

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOTEHNOLOOGIA ÕPPETOOL

Anette Kalnapenkis

**Naiste viljakusega seotud tunnuste geneetilised
korrelatsioonid**

Magistritöö geenitehnoloogia erialal (30 EAP)

Juhendajad: PhD Reedik Mägi

PhD Triin Laisk-Podar

TARTU 2016

Naiste viljakusega seotud tunnuste geneetilised korrelatsioonid

Lühikokkuvõte: Naiste reproduktiivne eluiga vältab enamasti aastakümneid, seda ilmestavad mitmed muutused, mille tulemusena naiste viljakuse hetkeseis muutub. Käesolevas töös leiti erinevate naiste viljakust peegeldavate tunnustega seotud geneetilised variandid geenides, millele on omane kõrge ekspressioon testistes ja naiste reproduktiivtraktis. Ühtlasi on mitmetel leitud geenidel kõrge ekspressioon kesknärvisüsteemis, millest tulenevalt on alust arvata, et need mõjutavad reproduktiivkäitumist. Antud magistritöö raames ei leitud erinevate naiste viljakusega seotud tunnuste vahel jagatud geneetilist tausta, kuid neid võivad siduda töös käsitlemata tunnused. Naiste viljakust peegeldavad tunnused omavad seost mitmete tunnustega epidemioloogilisel tasandil, kõige sagedamini on kirjeldatud seost kehamassiindeksiga. Magistritöös leiti, et kehamassiindeks on geneetilises korrelatsioonis nii vanusega menarhe ajal kui ka vanusega esimese raseduse ajal, mille võimalikud ühised bioloogilised komponendid on leptiini ja insuliiniga seotud geenirajad.

Märksõnad: naiste viljakus, GWAS, geneetilised korrelatsioonid

CERCS: B110 Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika

Genetic correlations of traits associated with female fertility

Abstract: Duration of female reproductive lifespan is in most cases decades. It is characterized by several changes that affect female fertility. In this study, genetic variants associated with different traits depicting female fertility were found, several of them were in genes that have high expression in testis and female reproductive tract. Several genes also have high expression in central nervous system. Thus, they could influence reproductive behavior. In this study, different traits depicting female fertility do not share genetic background, but they could be genetically binded by traits that were not included in master thesis. Traits depicting female fertility have been associated with different traits on epidemiological level, most frequently with body mass index. In this thesis, genetic correlations between age at menarche and body mass index and also between age at first pregnancy and body mass index were found. Candidates for shared biological component between female reproductive health and body mass index could be leptin and insulin.

Keywords: female fertility, GWAS, genetic correlations

CERCS: B110 Bioinformatics, medical informatics, biomathematics, biometrics

Sisukord

Kasutatud lühendid ja mõisted	4
Sissejuhatus	5
1. Kirjanduse ülevaade	6
1.1. Menarhe	6
1.2. Menstruaaltsükli iseloomustus.....	9
1.3. Menstruaaltsükli pikkus ja regulaarsus.....	10
1.4. Reproduktiivkäitumine	12
1.5. Raseduskatkemised	14
1.6. Ülegenoomsed assotsiatsiooniuuringud.....	15
1.7. LD skoori regressiooni ülevaade	16
2. Eksperimentaalosa	18
2.1. Töö eesmärk.....	18
2.2. Materjal ja metoodika	18
2.2.1. Valim ja uuritavad tunnused	18
2.2.2. Konsortsiumite assotsiatsiooniuuringute meta-analüüside koondstatistikud.....	20
2.2.3. Genotüpiseerimine	22
2.2.4. Kvaliteedikontroll.....	22
2.2.5. Imputatsioon.....	23
2.2.6. Ülegenoomne assotsiatsioonianalüüs.....	23
2.2.7. Meta-analüüs	23
2.2.8. LD skoori regressioon	24
2.3. Tulemused.....	25
2.3.1. Ülegenoomsete assotsiatsiooniuuringute meta-analüüsid.....	25
2.3.2. LD skoori regressioon	32
2.4. Arutelu	35
Kokkuvõte	42
Summary.....	43
Tänuõnad.....	45
Kasutatud kirjanduse loetelu	46
Kasutatud veebiaadressid	56
Lisa 1. Küsimustik.....	57
Lisa 2. Raseduskatkemiste välistusdiagnoosid.....	59

Kasutatud lühendid ja mõisted

CR	ingl <i>call rate</i> , edukalt genotüpiseeritud markerite protsentuaalne osakaal
EAF	ingl <i>effect allele frequency</i> , efektilleeli sagedus
EGG	ingl <i>Early Growth Genetics Consortium</i> , Varajase Kasvu Geneetika Konsortsium
FSH	folliikuleid stimuleeriv hormoon
GIANT	ingl <i>Genetic Investigation of Anthropometric Traits</i> , Antropomeetriliste Tunnuste Geneetilise Uurimise konsortsium
GLGC	ingl <i>Global Lipids Genetics Consortium</i> , Globaalne Lipiidide Geneetika Konsortsium
GnRH	ingl <i>gonadotropin-releasing hormone</i> , gonadotropiine vabastav hormoon
GRCh37	ingl <i>Genome Reference Consortium GRCh37</i> , inimese referentsgenoom GRCh37
GWAS	ingl <i>genome-wide association study</i> , ülegenoomne assotsiatsiooniuuring
h^2	päritavuskoeffitsient
hAMSC	ingl <i>human adipose tissue-derived mesenchymal stem cells</i> , inimese adipotsüütide mesenhümaalsed tüvirakud
HDL	ingl <i>high-density lipoprotein</i> , kõrge tihedusega lipoproteiin
HWE	ingl <i>Hardy-Weinberg equilibrium</i> , Hardy-Weinbergi tasakaalustatus
HPO	ingl <i>hypothalamic-pituitary-ovarian axis</i> , hüpotaalamuse-ajuripatsi-munasarjade telg
IBD	ingl <i>identity by descent</i> , päritolust tulenev identse DNA osakaal genoomis
IBS	ingl <i>identity by state</i> , sugulusest tulenev ühiste alleelide osakaal genoomis
KMI	kehamassiindeks
LD	ingl <i>linkage disequilibrium</i> , aheldatuse tasakaalutus
LDL	ingl <i>low-density lipoprotein</i> , madala tihedusega lipoproteiin
LH	luteiniseeriv hormoon
MAF	ingl <i>minor allele frequency</i> , harvema alleeli sagedus
Mb	ingl <i>mega base pairs</i> , miljon aluspaari
MCP-1	ingl <i>monocyte chemoattractant protein-1</i> , monotsüütide kemoatraktandi valk 1
NA	ingl <i>not available</i> , puuduv väärtus
PCOS	ingl <i>polycystic ovary syndrome</i> , polütsüstiliste munasarjade sündroom
r_g	geneetiline korrelatsioon
REML	ingl <i>restricted maximum likelihood</i> , piiratud suurima tõepära meetod
SNP	ingl <i>single nucleotide polymorphism</i> , ühenukleotiidne polümorfism
TPS	talje-puusa suhe

corpus luteum lad kollaskeha

pleiotroopsus geeni mitmene efekt - üksik geen mõjutab mitme tunnuse avaldumist

populatsiooni stratifikatsioon – erinevast põlvnemisest tulenevad alleelisageduste erinevused

Sissejuhatus

Naiste enamasti aastakümneid vältav reproduktiivne iga on piiritletud kahe sündmusega – menarhe ja menopausiga. Senised naiste reproduktiivtervist käsitlevad uuringud on peamiselt keskendunud sellesse ajavahemikku jäävatele patoloogilistele seisunditele ning normaalse reproduktiivtervise kontekstis on eelkõige uuritud menarhe ja menopausi geneetikat, mis on olemuselt üleminekuperioodid. Reproduktiivkäitumist mõjutavad eelkõige sotsiaalsed tegurid, mille seast on kõige olulisemal kohal indiviidi personaalsed valikud: kas ja millal luua perekond. Selle taustal on oma koht ka geneetikal ja sellega kaasnevatel bioloogilistel protsessidel, mis määravad reproduktiivse perioodi pikkuse ja edukuse. Sellest tulenevalt on oluline selgitada, mis põhjustab reproduktiivsete tunnuste varieeruvust normaalses vahemikus indiviiditi ning kuidas see on seotud reproduktiivse edukusega. Ülegenoomsete assotsiatsiooniuuringutega on võimalik leida tunnustega seotud uusi bioloogilisi radu. Leitud geenid on ühtlasi potentsiaalsed sihtmärgid biomarkeritele, mille abil on võimalik hinnata naise viljakuse seisundit konkreetsel ajahetkel, lähtuvalt geneetilisest ja bioloogilisest vaatenurgast. Geneetilise ühisosa leidmine võimaldab hinnata, kas erinevatel naiste viljakusega seotud tunnustel on jagatud geneetikast nii omavahel kui ka epidemioloogilisel tasandil seotud tunnustega, näiteks kehamassiindeksiga. Sellest tulenevalt on võimalik leida ühist bioloogilist komponenti.

Käesoleva töö üheks eesmärgiks oli leida geneetilisi variante, mis on seotud erinevate naiste viljakust peegeldavate tunnustega, kasutades selleks ülegenoomseid assotsiatsiooniuuringuid. Töö teiseks eesmärgiks oli hinnata naiste viljakusega seotud tunnuste omavahelist jagatud geneetikat ja vaadelda korrelatsioone epidemioloogilisel tasandil seotud tunnustega, leidmaks võimalikku ühist bioloogilist komponenti. Tunnuste vahelise geneetilise korrelatsiooni leidmisel kasutati LD skoori regressiooni meetodit. Antud magistr töö raames läbiviidud uuringud olid pilootprojektiks edasistele naiste reproduktiivtervist käsitlevatele projektidele.

1. Kirjanduse ülevaade

1.1. Menarhe

Naiste reproduktiivse eluea algust märgib esimene menstruatsioon ehk menarhe. Arenenud maades toimub tüdruku esimene menstruatsioon keskmiselt 12.–13. eluaastal (Whincup *et al.*, 2001; Anderson *et al.*, 2003). Eestis toimub tüdruku esimene menstruatsioon 2014. aasta seisuga keskmiselt vanuses 13,1 aastat (Lippus *et al.*, 2015). Lisaks individidevahelisele varieeruvusele mõjutab menarhet nii geograafiline asukoht kui ka etniline kuuluvus (WHO, 1986; Morabia & Costanza, 1998), nii saabub ekvaatori lähedal elavatel tüdrukutel menarhe varem kui suurematel laiuskraadidel elavatel tüdrukutel (Dossus *et al.*, 2013). Ameerika Ühendriikides läbiviidud uuringus leiti, et Euroopa päritolu tüdrukutel saabub menarhe keskmiselt mitmeid kuid hiljem kui Mehhiko ja Aafrika päritolu tüdrukutel (Chumlea *et al.*, 2003). Jaapani uuringus leiti, et naiste keskmine vanus menarhe ajal on 50 aasta vältel muutunud 1,6 aastat varasemaks (Hosokawa *et al.*, 2012). Selline trend on omane ka Eestile - kui 2004. aastal oli keskmine vanus menarhe ajal 13,2 aastat (Part *et al.*, 2007), siis juba 2014. aastal oli keskmine vanus menarhe ajal 13,1 aastat (Lippus *et al.*, 2015).

Frisch ja Revelle sõnastasid 1971. aastal esmakordselt „kriitilise kehakaalu hüpoteesi“, mille kohaselt naised jõuavad reproduktiivsesse eluikka alles teatud kehakaalu (47 kg) saavutamisel. Naistel, kes on kriitilise kaalu saavutanud, kiireneb samuti ainevahetus, mis omakorda põhjustab munasarjade ja emaka küpsemist ja sellest tulenevalt ka menarhet (Frisch & Revelle, 1971). Menstruatsiooni toimumiseks peab tüdrukutel keharasva olema vähemalt 17% ning menstruatsiooni püsimiseks ja regulaarseks muutumiseks peab keharasva olema 22% (Baker, 1985).

Menarhe ajastust mõjutab nii naise pikkus kui ka kehamass, mis mõlemad on omakorda tugeva geneetilise taustaga tunnused (Wood *et al.*, 2014; Locke *et al.*, 2015). Ühtlasi sõltuvad mõlemad tunnused toitumisest, seega sõltub menarhe ajastus nii geneetilistest kui ka keskkonnafaktoritest ning arvatavasti ka nende omavahelistest interaktsioonidest. Hetkel pole üheselt selge, kas suurenev kehamassiindeks (KMI) on varasema menarhe põhjuseks või varajane sugulise küpsuse saavutamine põhjustab keharasva hulga suurenemist ja keharasva jaotumist. Laste rasvumuse epideemia, mis algas 20. sajandi lõpul¹, võib mõjutada tüdrukutel puberteedi nihkumist varasemaks. Varane puberteet on positiivses seoses täiskasvanuea rasvumusseisunditega ja selle metaboolsete tagajärgedega edasises elus, kui on arvestatud lapsea kehamassiindeksit (Ng *et al.*, 2014). Selle põhjuseks võib olla lapsea

¹ <http://www.who.int/end-childhood-obesity/facts/en/>

kehamassiindeksi ning täiskasvanuea kehamassiindeksi tugev geneetiline korrelatsioon (Felix *et al.*, 2016). Varane menarhe on seotud erinevate haigusseisunditega, näiteks suurenenud riskiga II tüüpi diabeedile, mitmetele hormoontundlikele vähkkasvajatele ja kardio-metaboolsetele haigustele (Osuch *et al.*, 2010; Li *et al.*, 2007; Prentice & Viner, 2013). He jt uuringus leiti, et ühe aasta võrra hilisem menarhe vähendab II tüüpi diabeedi riski 10% (He *et al.*, 2010). Lisaks on varane menarhe oluline riskifaktor depressioonile (Kaltiala-Heino *et al.*, 2003) ja toitumishäiretele (Kaltiala-Heino *et al.*, 2001).

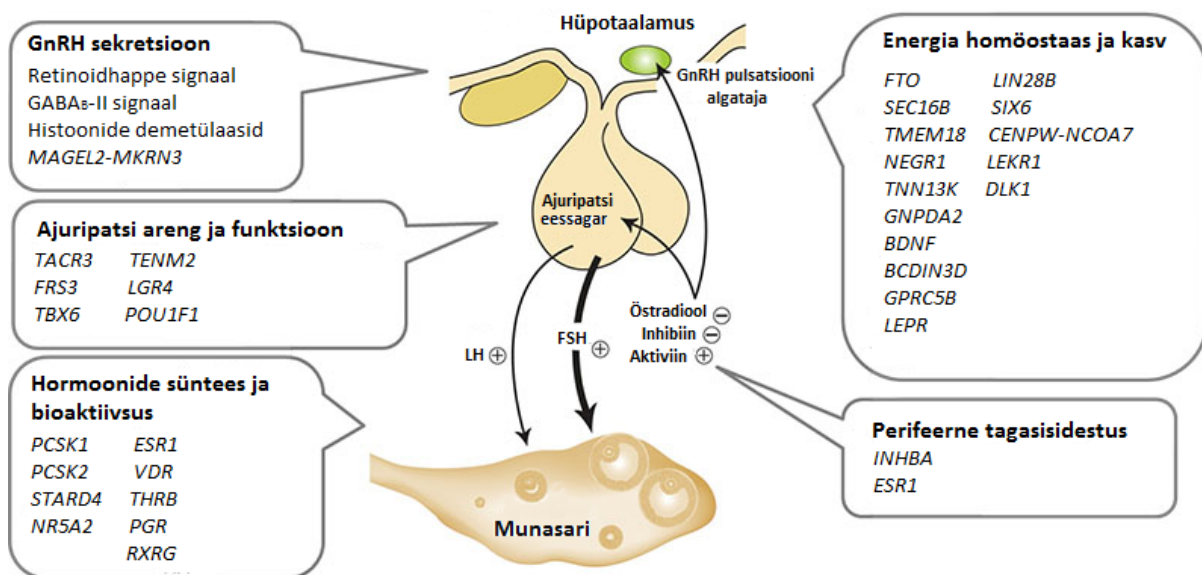
Ka sünnieelne kehaline areng võib mõjutada menarhe ajastust. Sünnikaalu ja menarhe seosest on raporteeritud vastuolulisi tulemusi: nii leiti Ühendkuningriigis läbiviidud uuringus, et väiksema sünnikaaluga naistel toimus menarhe oluliselt varem (Morris *et al.*, 2010), kuid Ameerika Ühendriikides läbiviidud uuringus ilmnes vastupidine seos, kus suurema sünnikaaluga naised saavutasid varem suguküpsuse (Wang *et al.*, 2012). Lisaks võib sünnikaalu ja menarhe seoses rolli mängida ka pikkus sündides. Tüdrukutel, keda sünnil kirjeldati pikkade ja saledatena (> 49 cm, < 3 kg), toimus menarhe ligikaudu 6 kuud varem kui tüdrukutel, keda sünnil kirjeldati lühikeste ja kergetena (< 49 cm, < 3 kg) (Adair, 2001).

Enamik menarhe päritavuse hinnanguid tuginevad kaksikute uuringutele. Hollandis läbiviidud uuringu kohaselt on individuaalsed erinevused menarhe ajastuses suuresti pärilikud, sealhulgas kirjeldavad aditiivsed geneetilised faktorid tunnuse varieeruvust vähemalt 70% ulatuses. Lisaks võib 1,5% ulatuses menarhe ea varieeruvuse põhjuseks olla genotüübikeskkonna vahelised interaktsioonid, mille korral tüdrukutel, kellel on geneetilised eeldused hiliseks menarheks, omavad keskkondlikud faktorid suuremat mõju menarhe ea kujunemisele (van den Berg & Boomsma, 2007).

Menarhe on kompleksne geneetiline tunnus, mille uurimiseks on läbi viidud mitmeid ülegenoomseid assotsiatsiooniuuringuid (*genome-wide association study* – GWAS) ning samuti paljusid kohorte hõlmavaid ülegenoomsete assotsiatsioonianalüüside meta-analüüse. Kuivõrd menarhe ja pikkus on omavahel epidemioloogiliselt seotud, ilmnes ka ülegenoomsetest uuringutest, et pikkusega seotud geneetilised variandid mõjutavad samuti menarhe geneetikat. Esimesed GWAS uuringud tuvastasid assotsiatsiooni signaali 6q21 lookuses *LIN28B* lähedal ja intergeenses piirkonnas 9q31.3 lookuses (He *et al.*, 2009). Kokku on GWAS uuringutes leitud üle 100 menarhe vanusega seotud lookuse ja signaali, mis seletavad ainult 3,6–6,1% menarhe ajastuse varieeruvusest (Elks *et al.*, 2010). Populatsiooni tasemel on GWAS-i sagedaste geneetiliste variantide puhul leitud, et üksiku variandi efektisuurus menarhe ajastusele varieerub vahemikus umbes 2 nädalat kuni 5 nädalat (Perry *et*

al., 2014). Kõige enam senistest leidudest mõjutavad menarhe ajastust populatsioonis madalama sagedusega esinevad geneetilised variandid *ALMS1* geenis, mille tulemusena lükkub menarhe umbes 3 kuud edasi (Lunetta *et al.*, 2015). Leitud geneetilised lookused on seotud mitmete bioloogiliste protsessidega, näiteks energia homöostaas, geenivaigistus ja hüpotaalamuse-ajuripatsi-munasarjade (*hypothalamic-pituitary-ovarian axis* – HPO) telje regulatsioon (Perry *et al.*, 2014)(Joonis 1).

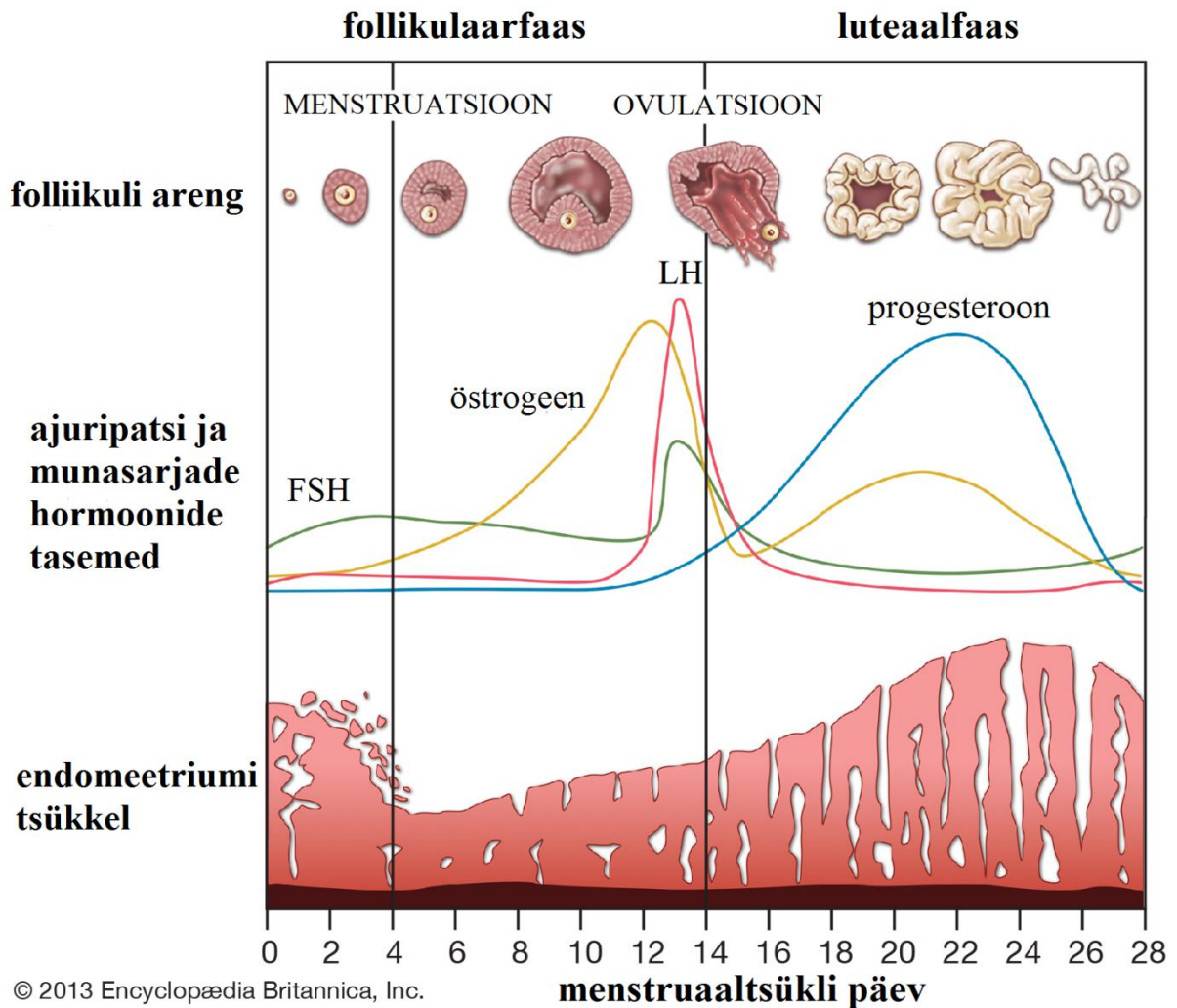
Tulenevalt menarhe ja KMI epidemioloogilisest korrelatsioonist on otsitud ka nende tunnuste geneetilist ühisosa. Pääaegu kõik KMI-d suurendavad GWAS leiud mõjutavad ka menarhe ajastust (Perry *et al.*, 2014; Speliotes *et al.*, 2010). Seda toetab ka Bulik-Sullivan jt uuringu leid, mille kohaselt menarhe aeg omab kõige enam geneetilist ühisosa erinevate antropomeetriliste tunnustega (Bulik-Sullivan *et al.*, 2015). HPO telg reguleerib munasarjade poolt sekreteerivate suguhormoonide tootmist ning mitmed monogeensed reproduktiivhäired tulenevad normaalse HPO telje regulatsiooni häirumisest. Sellest tulenevalt esineb ülekatet menarhe ajastust mõjutavate geneetiliste variantide ning monogeenseid reproduktiivhäireid põhjustavate geneetiliste piirkondade vahel. Menarhe ajastust mõjutavad geneetilised variandid paiknevad mitmetes steroidhormoonide retseptorvalke kodeerivates geenides, sealhulgas östrogeeni, progesterooni ning kilpnäärme hormoonide retseptorites (Perry *et al.*, 2014).



Joonis 1. Menarhega seostatud geenide võimalikud rollid hüpotaalamuse-ajuripatsi-munasarjade teljel ja teised võimalikud menarhet mõjutavad bioloogilised mehhanismid. Gonadotropiine vabastav hormoon (*gonadotropin-releasing hormone* – GnRH). (Perry *et al.*, 2014, kohandatud).

1.2. Menstruaaltsükli iseloomustus

Naiste normaalset reproduktiivset eluiga ilmestavad igakuised rütmilised muutused nii hormoonide sekretsioonis kui ka neile vastavad füsioloogilised muutused munasarjades ja teistes suguorganites (Joonis 2). Hüpotalamuses toodetakse GnRH, mille mõjul ajuripatsi gonadotroopsed rakud sekreteerivad luteiniseerivat hormooni (LH) ja folliikuleid stimuleerivat hormooni (FSH) kindla pulsatsiooniga. FSH on ajuripatsi hormoon, mis stimuleerib ootsüüte sisaldavate munasarjafolliiklite küpsemist. Munasari reageerib FSH ja LH taseme muutustele tsüklilisel viisil. Iga tsükkel kestab ligikaudu 28 päeva ning see on jagatud kaheks ligikaudu 14-päevaseks faasiks. Follikulaarfaas algab tsükli esimesel päeval, milleks loetakse menstruaaltsükli esimest päeva. Follikulaarfaasis toimub folliikulite kasvamine ning suureneb östradioli sekretsioon. Munasarja follikulaarfaasile vastab endomeetriumi proliferatiivne faas, mil östradioli toimel endomeetriumi epiteeli ja strooma rakud proliferatsioonivad (Bergeron & Ferenczy, 1988; Ferenczy & Bergeron, 1991). See on endomeetriumi küpsemise esimeseks faasiks ehk valmistumine embrüo implantatsiooniks, mis leiab aset menstruaaltsükli teises pooles. Ovulatsiooni, mis märgib munaraku vabanemist munasarjast, põhjustab östrogen-sõltuv LH tõus maksimumini menstruaaltsükli keskel. Ovulatsioon toimub LH maksimumi algusest 35–44 tunni jooksul (Tsafriri *et al.*, 1993). Pärast ovulatsiooni areneb lõhkenud folliikulist ja follikulaarrakkudest kollaskeha (*corpus luteum*), mis sekreteerib östradioli, progesterooni ja inhibiin A-d (Shikone *et al.*, 1997). Tsükli luteaalfaasis (vastab endomeetriumi sekretoorsele faasile) progesterooni toimel endomeetriumi epiteelrakud diferentseeruvad ja muutuvad sekretoorseteks rakkudeks (Bergeron & Ferenczy, 1988; Ferenczy & Bergeron, 1991) ning umbes nädal peale ovulatsiooni on endomeetrium valmis embrüo implantatsiooniks. Luteaalfaasi lõpus raseduse puudumisel kollaskeha taandareneb ning östradioli, progesterooni ja inhibiin A sekretsioon langeb (Shikone *et al.*, 1997). Endomeetrium irdub, põhjustades tsüklilist menstruaaltsükli (Ziegler *et al.*, 1992) ja kollaskeha taandarenemisest tuleneb FSH taseme tõus, mis võimaldab uue tsükli initsiatsiooni (Nestour *et al.*, 1993).



Joonis 2. Menstruaaltsükli üldine iseloomustus. (*Encyclopædia Britannica*, 2013², kohandatud).

1.3. Menstruaaltsükli pikkus ja regulaarsus

Menstruaaltsükli pikkus on väga varieeruv naiste reproduktiivtervist peegeldav tunnus, varieerudes oluliselt isegi samasse eagrüppi kuuluvate naiste vahel. Varieeruvus tuleneb erinevustest follikulaarfaasi (10–23 päeva) ja luteaalfaasi (7–19 päeva) kestvuses, näiteks ainult kümnel protsendil 28-päevase menstruaaltsükliga naistel on 14-päevane follikulaar- ja luteaalfaas (Harlow, 2000). Kõige suurem varieeruvus on follikulaarfaasi kestvuses, mis viljakuse langedes lüheneb 3–7 päeva. Folliikuli küpsemise liigse häirumise korral ei toimu ovulatsiooni. Sellele järgnevalt on menstruaaltsüklid ebaregulaarsed ning kestavad kaua (Waller *et al.*, 1998; Wilcox *et al.*, 2000).

Peamine informatsioon menstruaaltsükli pikkusest ning selle varieeruvusest reproduktiivse eluea vältel põhineb eelkõige 1960.–1970. aastatel läbiviidud neljal pikaajasel

² <http://www.britannica.com/science/menstrual-cycle>

menstruaalkalendrite uuringul (Treloar *et al.*, 1967; Chiazze *et al.*, 1968; Vollman, 1977; Matsumoto *et al.*, 1962). Uuringutest selgus, et populatsioonisisene varieeruvus menstruaaltsükli pikkuses on suurim pärast menarhet ja vahetult enne menopausi. Menarhe üleminekuperiood vältab ligikaudu 2–5 aastat ning menopausi üleminekuperiood kestab keskmiselt ligikaudu 6–10 aastat. Mõlemale perioodile on iseloomulikud suurenenud sagedusega nii väga pikad kui ka väga lühikesed menstruaaltsüklid. Tsükli pikkus varieerub 20–40 aastastel naistel vähem, kuid kahe aastakümne vältel menstruaaltsükkel keskmiselt lüheneb siiski paari päeva võrra (Harlow & Ephross, 1995). Ligikaudu 30% premenopausaalsetest naistest kogevad rohkem kui 14-päevast varieeruvust tsükli pikkustes aasta vältel. Kõige stabiilsem on menstruaaltsükli pikkus vanuses 25–35 aastat (Münster *et al.*, 1992). Lisaks vanusele on peamised menstruaaltsükli pikkust mõjutavateks teguriteks stress, suitsetamine ja kofeiini tarbimine (Marks *et al.*, 1994; Benowitz, 1990).

Menstruaaltsükli häired on amenorröa (primaarne ja sekundaarne), polümenorröa ja oligomenorröa. Amenorröa on menstruaaltsükli häire, mille korral menstruaaltsioonid puuduvad. Juhul kui menstruaaltsioonid pole kunagi olnud, on tegu primaarse vormiga, kuid menstruaaltsioonide puudumine 6 kuud ja enam on omane sekundaarsele vormile. Sekundaarne amenorröa esineb normaalse seisundina raseduse, imetamise või menopausi korral. Polümenorröa on menstruaaltsükli häire, mille korral tsükli intervall on lühem kui 23 päeva. Seda põhjustab folliikuli küpsemise häire, mil tsükli follikulaarfaas on patoloogiliselt lühike või tsükkel jääb anovulatoorseks. Sagedamini on seda mõne aasta jooksul pärast menarhet või enne menopausi. Oligomenorröa korral on menstruaaltsükkel tavalisest pikem (> 35 päeva) ja sageli ebaregulaarne (Ylikorkala & Kauppila, 2008).

Noorukieas on menstruaaltsiooni ebaregulaarsus sagedane nähtus, eriti 2–3 aastat pärast menarhet. Intervall esimesest menstruaaltsioonist teise menstruaaltsioonini võib olla pikk, kuid järgnevad tsüklid varieeruvad enamasti 21–35 päeva vahel ning vaid harva pole selles vahemikus. Menarhe järgselt läheb ligikaudu 5–7 aastat, mil hüpotaalamuse-ajuripatsimunasarjade telg saavutab küpsuse, mis tagab edasised regulaarsed menstruaaltsüklid. (Venturoli, 1992; Flug *et al.*, 1984). Menstruaaltsüklite regulaarsuse muutused kaasnevad ühtlasi naiste viljakuse langemisega, mis toimub üleminekul menopausi (Weiss, 2001).

Ebaregulaarsete menstruaaltsüklite põhjuseks võib olla polütsüstiliste munasarjade sündroom (*polycystic ovary syndrome* – PCOS), mis on sage hormonaalne häire, põhjustades väikeste tsüstide arengut munasarjades ning sellega kaasneva ebaregulaarseid menstruaaltsükleid. See on ühtlasi üheks peamiseks naisepoolse viljatuse põhjuseks. Süntomeid on võimalik

määrata juba noorukieas, kuid need on ajaliselt ülekattes sugulise küpsuse saavutamise perioodiga, mil menstruaaltsükkel on ebaregulaarne. Sellest tulenevalt diagnoositakse sündroomi harilikult alles täiskasvanueas (Welt & Carmina, 2013).

Menstruaaltsükli geneetilist regulatsiooni on hakatud põhjalikult uurima alles hiljuti. Ruth jt uuringus leiti, et menstruaaltsükli pikkust mõjutab *FSHB* lookus, mis kodeerib FSH hormooni β -subühikut. *FSHB* promootoris paiknev T-alleeli polümorfism alandab FSH taset ning selle tulemusena pikeneb menstruaaltsükkel umbes 1 päeva võrra. Lisaks leiti *FSHB* piirkonnast ühenukleotiidne insertioon, mille tulemusena menstruaaltsükkel samuti pikeneb päeva võrra (Ruth *et al.*, 2016).

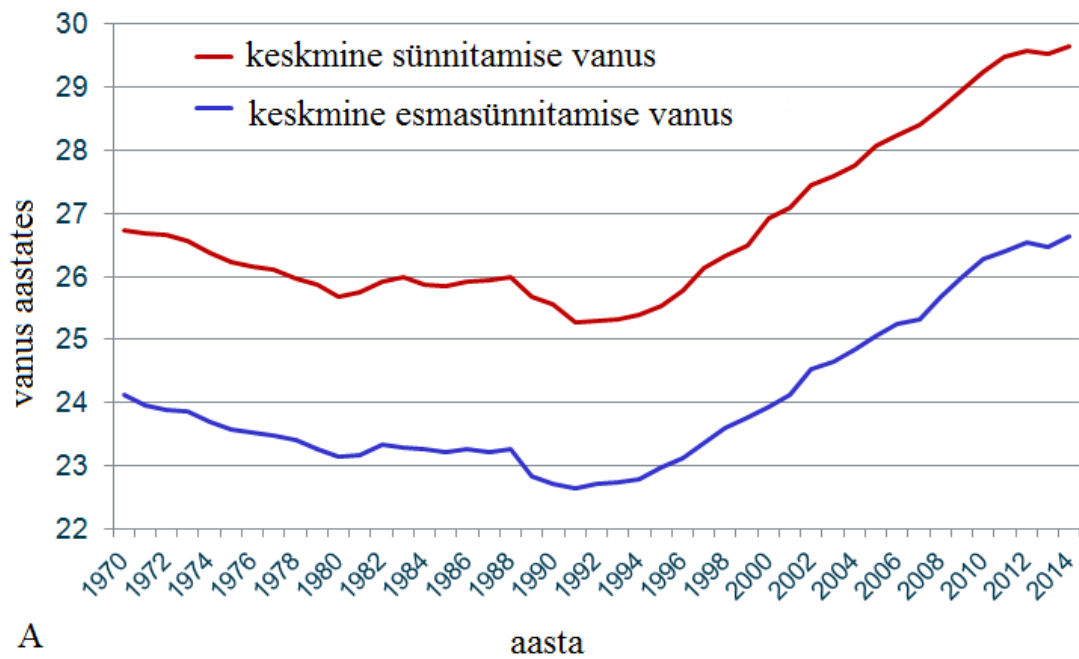
1.4. Reproduktiivkäitumine

Peaaegu kõik Euroopa riigid on jõudnud demograafilise ülemineku faasi, millele on iseloomulik madal sündimus ja kõrge oodatav eluiga. Arenenud riikides on viimastel aastakümnetel muutunud trendiks alustada pereplaneerimist aina hilisemas eas. Keskkondlikud faktorid, näiteks naiste haridusvõimaluste paranemine ja osalemine tööturul ning majanduslik ebakindlus mõjutavad tugevalt naiste viljakust peegeldavaid tunnuseid (Mills *et al.*, 2011). Tulenevalt antud muutustest on arenenud riikides esmasünnitajate vanus tõusnud^{3,4}, selline trend on omane samuti Eestile (Joonis 3A)(Tiit, 2015⁵). Arenenud riikides on kontratseptiivide laialdane kättesaadavus alates 1960.–1970. aastatest koos muutustega pereplaneerimishoiakutes mõjutanud sündinud laste arvu perekonna kohta. Euroopa Liidu 15 riigis oli 1965. aastal naisel keskmiselt 2,72 last, kuid juba 1975. aastaks oli keskmine laste arv langenud 1,96-ni ning nüüdseks on naisel keskmiselt alla 1,50 lapse (Leridon, 2006). Selline langev trend on omane ka Eestile, 2014. aastal oli naisel keskmiselt 1,54 last (Joonis 3B)(Tiit, 2015⁵). Võrreldes 20–24. aastaste naistega, on 25–29. aastaste naiste viljakus keskmiselt langenud 6%, 30–34. aastastel naistel 14%, 35–39. aastastel naistel 31% ning edaspidi jätkub veel järsem langev trend (Menken *et al.*, 1986). Esmasünnitajate vanuse tõus koos loomuliku vanusest tingitud viljakuse langusega võib põhjustada suurenenud viljatuse riski. Viljatust esineb arenenud maades kuni 16,7% paaridel ning vähem arenenud maades kuni 9,3% paaridel (Boivin *et al.*, 2007).

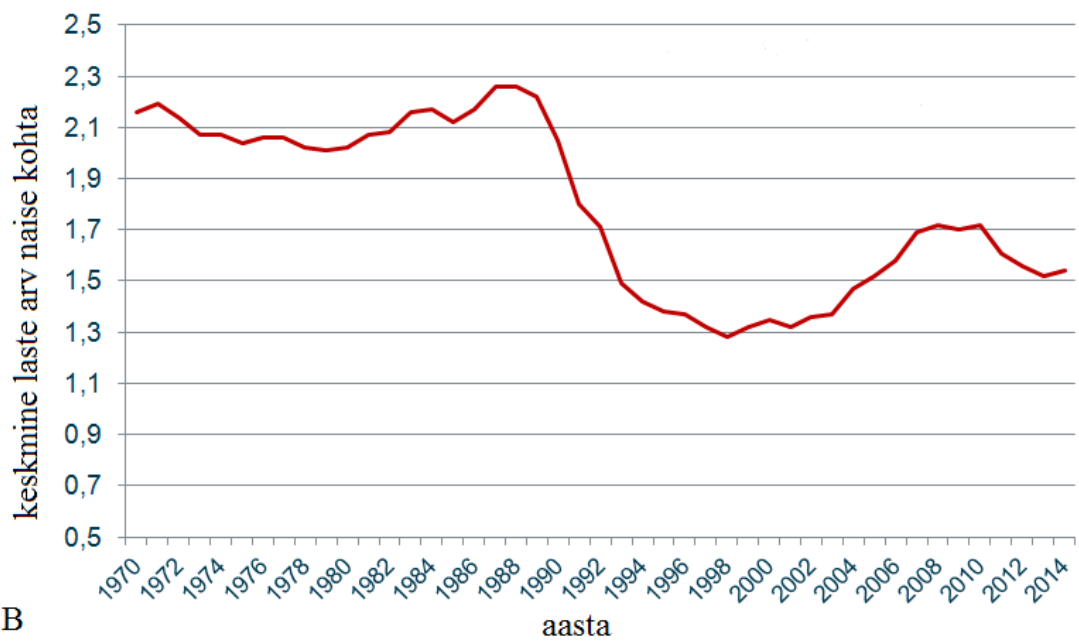
³ <http://www.cdc.gov/nchs/data/databriefs/db232.htm>

⁴ https://www.oecd.org/els/soc/SF_2_3_Age_mothers_childbirth.pdf

⁵ <https://www.stat.ee/311832>



A



B

Joonis 3. Reproduktiivkäitumist ilmestavad ajalised trendid. (A) Keskmine sünnitamise ja esmasünnitamise vanus Eestis aastatel 1970–2014. (B) Keskmine laste arv naise kohta Eestis aastatel 1970–2014. (Tiit, 2015⁵).

Eelnevalt on leitud, et bioloogilisel viljakusel on ka geneetiline komponent (Stolk *et al.*, 2009) ning kaksikute uuringu kohaselt võib viljakuskäitumise päritavus olla ligikaudu 40% (Kohler *et al.*, 1999). On leitud, et sagedased ühenukleotiidsed polümorfismid (*single nucleotide polymorphism* – SNP) mõjutavad 4–10% ulatuses sündinud laste arvu varieeruvust ja 5–15% ulatuses esmasünnitajate vanuse varieeruvust (Tropf *et al.*, 2015; Tropf *et al.*, 2016). Uuringutes leitud hinnangute erinevus võib tuleneda sellest, et geneetiliste efektide suurus on populatsioonide vahel väiksem kui populatsioonisiseselt, millest tulenevalt enamik viljakust

mõjutavatest geneetilistest variantidest on populatsioonispetsiifilised (Tropf *et al.*, 2016). Ühtlasi on näidatud, et esmasünnitajate vanuse ja laste arvu omavaheline geneetiline korrelatsioon on -0,62. Selle kohaselt on geenid, mis vastutavad hilisema esmasünnitamise vanuse eest, samuti seotud väiksema sündinud laste arvuga. Mõlemat tunnust võivad mõjutada samad geneetilised variandid või sündinud laste arvu mõjutavad geneetilised efektid tulenevad esmasünnitamise vanusest ning ka nende kombinatsioonist (Tropf *et al.*, 2015). Bangladeshi naiste seas läbiviidud raseduste ja laste arvuga seotud ülegenoomne uuring seletas 14,9% laste arvu päritavusest ning ainult 0,7% raseduste arvu päritavusest, kuid ülegenoomselt olulisi sagedasi geneetilisi variante ei tuvastatud. Olulisemad leitud seosed oli kromosoomides 4., 5. ja 6. eelkõige geenide vahelistes piirkondades paiknevate geneetiliste variantidega (Aschebrook-Kilfoy *et al.*, 2015). Hiljutises suuremahulises Day jt uuringus leiti, et neuronaalse raku adhesioonmolekuli kodeeriva geeni *CADM2* intronis paiknevad ühenukleotiidsed polümorfismid on seotud laste arvuga (Day *et al.*, 2016). Lisaks leiti östrogeeni retseptorit kodeerivas geenis *ESR1* paiknevate SNP-de seos varasema esmasünnitamise eaga ning suurema laste arvuga. *RBM6-SEMA3F* lookuse geneetiline variant oli seotud varasema esmasünnitamise vanusega, suurema laste arvuga ning ka väiksema tõenäosusega jääda lastetuks. Sellest tulenevalt on alust arvata, et reproduktiivne edukus on seotud nii kesksete protsessidega hüpotaalamuse ja ajuripatsi suguhormoonide signaliseerimisel kui ka neurokognitiivsete tunnustega (Day *et al.*, 2016).

1.5. Raseduskatkemised

Raseduskatkemised toimuvad 12–15% kliiniliselt tuvastatud raseduste korral (Zinaman *et al.*, 1996). Sagedus suureneb, kui arvestada ka varajasi raseduskatkemisi, mis definitsiooni kohaselt on toimunud ligikaudu 14 päeva pärast viljastumist. Need raseduskatkemised toimuvad ajaliselt järgmise menstruatsiooni lähedal, mistõttu veritsust võib pidada menstruatsioonile omaseks. Selliste varajaste raseduskatkemiste arvatav sagedus on 17–22% rasedustest (Ellish *et al.*, 1996). Raseduskatkemiste sagedus varieerub ka eagrupiti. Naistel vanuses 20–22 aastat on raseduskatkemiste esinemissagedus 10%, kuid 40–44. aastastel naistel katkevad juba üle poolte (51%) rasedustest (Nybo Andersen *et al.*, 2000). Peamiseks raseduskatkemise põhjuseks on loote kromosoomianomaaliad. Hinnanguliselt on ligikaudu 75% embrüotest 3. arengupäevaks vähemalt osaliselt aneuploidsed, omades nii meiotilist kui ka sügootilist päritolu vigu (Wells & Delhanty, 2000; Voullaire *et al.*, 2000), mille esinemissagedus tõuseb koos ema vanusega (Lamb *et al.*, 2005). Korduvate raseduse katkemiste korral on naisel toimunud kolm või enamat järjestikust raseduskatkemist enne 22. rasedusnädalat. Korduvate raseduskatkemiste sagedus on 1–3% ning selle põhjused on väga

mitmekesised (endokriinsed, immunoloogilised ja anatoomilised) või tihti hoopis ebaselged (Christiansen *et al.*, 2008). On leitud, et korduvate raseduskatkemistega naiste lähisugulastel on samuti suurem tõenäosus nii sporaadilisteks kui korduvateks raseduskatkemisteks, mis annab alust arvata, et raseduskatkemiste etioloogias on oma koht ka ema geneetikal (Christiansen *et al.*, 1990; Rull *et al.*, 2012). McCoy jt leidsid, et mitoetiliste vigade tekke üheks põhjuseks võib olla ema genoomis 4. kromosoomis q28.1 kuni q28.2 piirkond, kus paikneb *PLK4* geen, milles paikneva geneetilise variandi harvemini esineva alleeli korral suureneb mitoetiliste vigade määr. Selle tulemusena toimub rakutsükli anafaasis hiline mine ning hiline nud kromosoomi ei suudeta lisada tuuma koosseisu. Ühtlasi leiti, et antud genotüüpidega naistel võib minna keskmisest kauem aega eduka raseduse saavutamiseks (McCoy *et al.*, 2015). Harvemini esineva alleeli sagedus (*minor allele frequency* – MAF) on erinevates populatsioonides 20–45% (Durbin *et al.*, 2010).

1.6. Ülegenoomsed assotsiatsiooniuuringud

Ülegenoomsed assotsiatsiooniuuringud põhinevad aheldatuse tasakaalutuse (*linkage disequilibrium* – LD) printsiibil populatsiooni tasemel. LD on mittejuhuslik assotsiatsioon alleelide vahel, mis paiknevad erinevates lookustes. See põhineb evolutsioonilistel protsessidel nagu mutatsioonide teke, geenitriiv ja valik ning seda lahutab rekombinatsioon. Üldiselt on füüsiliselt lähestikku paiknevad lookused suuremas LD-s kui lookused, mis paiknevad üksteisest kromosoomil kaugemal (Daly *et al.*, 2001). Ülegenoomsed assotsiatsiooniuuringud võimaldavad uurida komplekstunnuste ja -haiguste geneetikat populatsiooni tasemel, võrreldes iga üksiku uuritava geneetilise variandi alleelisageduste erinevust juhtude ja kontrollide vahel, või siis seostatuna uuritava pideva tunnusega. GWAS-i suureks eeliseks on, et uuringud on disainilt hüpoteesivabad. GWAS-i käigus leitakse seoseid eelkõige sagedaste geneetiliste variantidega, kus harvema alleeli sagedus on uuritavas populatsioonis suurem kui üks protsent. Üksikuna on GWAS-i käigus leitud geneetiliste variantide efektid väikesed ja ka nende kumulatiivne efekt kirjeldab ainult osa tunnuse eeldatavast geneetilisest varieeruvusest (Visscher *et al.*, 2012).

Uuringu planeerimisel on vaja arvestada uuritava komplekstunnuse või -haiguse iseloomulike joontega. Kaheks peamiseks uuringute tüübiks on kategooriline (tihti binaarne juht-kontroll) ja kvantitatiivne. Juht-kontroll tüüpi uuringus on väga oluline, et juhtude määramisel ei esineks vigu klassifikatsioonis. Lisaks on vaja arvestada, et juhud ning kontrollid oleks sama etnilise taustaga, et vähendada valepositiivsete assotsiatsioonisignaalide arvu, mida põhjustavad erinevused alleelisagedustes, kui üks etniline grupp on üleesindatud kas juhu või

kontrollina (Pearson & Manolio, 2008). Analüütilised meetodid, näiteks peakomponentide analüüs, võimaldavad kohandada mitmeid mõõtmata populatsioonide erinevusi (populatsiooni stratifikatsioon). Leitakse omavahel korreleeruvad tunnused ja neist konstrueeritakse uued tunnused – põhikomponendid, mis omavahel definitsiooni järgi enam ei korreleeru (Stranger *et al.*, 2011).

1.7. LD skoori regressiooni ülevaade

Komplekstunnuste GWAS analüüsides on väga suur osa päritavusest seotud SNP-dega, mis ei saavuta seniste uuringu mahtude juures ülegenoomset olulisust. LD skoori regressioon on meetod, mis võimaldab leida tunnuste vahelisi geneetilisi korrelatsioone, et hinnata nende tunnuste omavahelisi geneetilisi seoseid, kasutades selleks GWAS andmestike koondstatistikuid. Geneetiliste korrelatsioonide hindamisel arvestatakse kõikide SNP-de efekte, sealhulgas ka selliste, mis ei saavutanud ülegenoomset olulisust. Eelnevad meetodid, mis hindasid tunnuste vahelisi geneetilisi korrelatsioone, lähtusid eeldusest, et on olemas ainult üks põhjuslik SNP ühe uuritava lookuse kohta, arvestamata võimalusega, et samas lookuses võib olla osalises LD-s ka mitmeid geneetilisi variante, mille seosed tunnusega on täielikult või osaliselt sõltumatud (Bulik-Sullivan *et al.*, 2015). Peamised kaks tehnikat GWAS andmestikest geneetiliste korrelatsioonide leidmiseks on piiratud suurima tõepära meetod (*restricted maximum likelihood* – REML)(Yang *et al.*, 2011) ja polügeensed skoorid (Dudbridge, 2013). Need meetodid kasutavad individuaalsete genotüüpide informatsiooni, kuid kuna mitmed suurimad GWAS analüüsid on läbiviidud meta-analüüsidenä, mille korral on saadaval ainult koondstatistikud ning enamasti puudub informatsioon individuaalse genotüübi tasemel, pole võimalik neid meetodeid väga suures mahus rakendada. Sellest tulenevalt on eelpool mainitud meetodeid rakendatud väheste fenotüüpide korral, sest indiviidide tasemel genotüübi informatsioon vajab informeeritud nõusolekuid (Bulik-Sullivan *et al.*, 2015).

LD skoori regressiooni meetodikal on mitmeid eeliseid. Näiteks ei vaja see indiviidi tasemel genotüübi andmeid, ning ei põhine mitte üksnes ülegenoomselt olulistel geneetilistel variantidel, vaid kasutab kõikide uuritavate variantide efekte. Selle meetodika korral ei põhjusta valimite ülekate geneetilise korrelatsiooni hindamisel nihet ning on arvutuslikult väga kiire. Meetod ei eelda samade indiviidide erinevate tunnuste mõõtmisi, millest tulenevalt saab läbi viia rohkem geneetiliste korrelatsioonide hindamisi, kui oli võimalik seniste meetoditega (Finucane *et al.*, 2015).

LD skoori regressiooni on võimalik rakendada korraga ainult kahe tunnuse omavahelise geneetilise korrelatsiooni leidmiseks, mis on ühtlasi ka meetodi peamine puudus. Geneetilist korrelatsiooni ei mõjuta keskkondlikud segajad vaid ainult geneetilised segajad, näiteks kui üks geen võib mõjutada mitme tunnuse avaldumist (pleiotroopsus). Kuigi geneetiliste korrelatsioonide hinnanguid ei kalluta juhtude ülesindatus, võivad seda teha teised valimi kõrvalekalded, näiteks vale klassifikatsioon (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013). LD skoori regressioon eeldab suurt valimit võrreldes indiviidi tasemel genotüübi andmeid kasutavate meetodikatega. Kuivõrd LD skoori regressiooni abil leitakse geneetilised korrelatsioonid LD referentspaneeli põhised, siis sellest tulenevalt on hetkel võimalik meetodit rakendada ainult ühtset etnilist päritolu populatsioonidele, mille jaoks on olemas sobilik LD referentspaneel (Bulik-Sullivan *et al.*, 2015). Geneetiliste korrelatsioonide hindamisel LD skoori regressiooniga ei ole seni uuritud positiivse valikulise paardumise mõju. Polügeensete tunnuste jaoks loodud meetodid nagu LD skoori regressioon ja REML on kõige efektiivsemad, kui neid rakendatakse tunnustele, mis on polügeense geneetilise arhitektuuriga. Tunnustel, mille päritavus sõltub märkimisväärsel hulgal ainult kindlatest SNP-dest, võib ainult nendel geneetilistel variantidel põhinev analüüs olla parem valik. Hetkel ei suuda LD skoori regressioon optimaalselt kasutada geneetilise korrelatsiooni leidmiseks väga suure efektiga geneetilisi variante (Finucane *et al.*, 2015).

2. Eksperimentaalosa

2.1. Töö eesmärk

Käesoleva magistritöö eesmärgiks oli ülegenoomse assotsiatsioonianalüüsi abil leida uusi geneetilisi seoseid naiste viljakusega seotud erinevate tunnustega. Töö teiseks eesmärgiks oli LD skoori regressiooni meetodit kasutades hinnata naiste viljakusega seotud tunnuste omavahelist jagatud geneetikat ning lisaks vaadelda korrelatsioone epidemioloogilisel tasandil seotud tunnustega, näiteks kehamassiindeksi ja erinevate kolesterooli tasemetega, et leida tunnuste vahelist jagatud bioloogilist komponenti.

2.2. Materjal ja metoodika

2.2.1. Valim ja uuritavad tunnused

Tartu Ülikooli Eesti Geenivaramu on Eestis elavaid täisealisi vabatahtlikke indiviide koondav populatsioonipõhine biobank, millega on liitunud ligi 52 000 inimest. See moodustab ligi 5% Eesti täiskasvanud elanikkonnast, olles ühtlasi suurim epidemioloogiline kohort Balti riikides. Geenidoonorid on värvatud üle Eesti koostöös perearstide, haigla eriarstide ja medõdedega. Osalejad annetasid vereproovid, mida kasutati DNA, valgete vererakkude ja vereplasma kogumiseks. Lisaks vastasid indiviidid küsimustikule, mis hõlmas isiklikke ja geneoloogilisi andmeid, kliinilisi diagnoose vastavalt ICD-10⁶ klassifikatsioonile ja elustiili informatsiooni, samuti sisaldab küsimustik eraldi naiste reproduktiivtervist puudutavat osa (Lisa 1)(Leitsalu *et al.*, 2014). Antud uurimuse käigus kasutati 9702 naise andmeid (naiste keskmine vanus oli 44,0 (std. hälve 18,9) aastat ja keskmine KMI oli 25,6 (std. hälve 5,9)), kellel oli olemas info vastavate fenotüüpide kohta ja kes olid genotüpiseeritud ülegenoomsete genotüpiseerimiskiipidega. Käesoleva töö raames teostavate uuringute läbiviimiseks on geenidoonorid andnud informeeritud nõusoleku ning on olemas Tartu Ülikooli inimuuringute eetika komitee kooskõlastus (nr. 234/T-12).

Naiste viljakuse geneetika uurimiseks oli vaatluse all 7 erinevat tunnust: vanus menarhe ajal, menstruaaltsükli pikkus ja regulaarsus, raseduste esinemine, vanus esimese raseduse ajal, raseduste arv ning raseduskatkemiste esinemine (Tabel 1).

⁶ <http://rhk.sm.ee/>

Tabel 1. Uuritavate tunnuste valimi suurus koos tunnuse keskmise väärtusega. Tabelis on iga tunnuse põhiselt naiste arv valimis. Juht-kontroll tüüpi uuringu korral on välja toodud juhtude ja kontrollide arv ning juhtude protsentuaalne osakaal. Mõõdetava väärtusega tunnustel on esitatud tunnuse keskmine ning standardhälve. Juhtudena on esitatud naised, kellel on vastavalt regulaarne menstruaaltsükkel, raseduste esinemine või raseduskatkemiste esinemine.

Uuritav tunnus	Valimi suurus	Tunnuse keskmine (std.hälve) või %
Vanus menarhe ajal	5723	13,5 (1,4) aastat
Menstruaaltsükli pikkus	4375	26,6 (2,7) päeva
Menstruaaltsükli regulaarsus	(juhud) 4376 (kontrollid) 231	95%
Raseduste esinemine	(juhud) 544 (kontrollid) 1548	26%
Vanus esimese raseduse ajal	4741	22,2 (4,1) aastat
Raseduste arv	2794	2,7 (2,6)
Raseduskatkemiste esinemine	(juhud) 1371 (kontrollid) 7441	16%

Iga tunnuse korral kasutati valimi filtreerimisel järgmisi kriteeriume:

Menarhe

Menarhe ehk esimese menstruatsiooni toimumine märgib reproduktiivse eluea algust. Uuritava tunnusena kasutati vanust menarhe ajal täisaastates. Väga varane või hiline menarhe on mõjutatud mitmetest teguritest, mis on indiviiditi väga varieeruvad ning võivad olla seotud mõne haigusseisundiga. Sellest tulenevalt jäeti menarhe vanuse uuringust välja naised, kelle vanus esimese menstruatsiooni ajal oli ≤ 9 või ≥ 17 aastat.

Menstruaaltsükli pikkus ja regulaarsus

Menstruaaltsükli pikkus ja regulaarsus varieerub naiste reproduktiivse eluea jooksul ning neid mõjutavad mitmed antropomeetrilised tunnused, näiteks kehamassiindeks. Menstruaaltsükli pikkusena oli defineeritud naiste menstruaaltsükli pikkus nende 25.–35. eluaastal, mil menstruaaltsükkel on väljakujunenud sagedusega (Münster *et al.*, 1992). Pärast 35. eluaastat toimub naiste viljakuse järsk langus ning ka muutused menstruaaltsükli pikkuses. Menstruaaltsükli pikkus on indiviiditi väga erinev ning sellest tulenevalt on tunnus küsimustikus jagatud kategooriateks. Antud töös teisendati need kategooriad ümber menstruaaltsükli pikkuseks päevades, kasutades antud kategooria keskmist väärtust, et tunnust

saaks analüüsida kvantitatiivsena (Lisa 1). Antud uurimuses kasutati indiviidide andmeid, kelle vastused tsükli pikkuse kohta liigitusid järgmistesse kategooriatesse: 1) tsükli pikkus \leq 21 päeva, 2) tsükli pikkus 21–24 päeva (keskmise 22,5 päeva), 3) tsükli pikkus 25–29 päeva (keskmise 27 päeva), 4) tsükli pikkus 30–35 päeva (keskmise 32,5 päeva) ning 5) tsükli pikkus $>$ 35 päeva.

Menstruaaltsükli regulaarsust on uurimuses hinnatud binaarse tunnusega: regulaarne/ebaregulaarne.

Raseduste esinemine

Raseduste esinemise valimis võrreldi omavahel naisi, kes on olnud elu jooksul rasedad, nendega, kes pole teadaolevalt seda olnud. Selle aluseks olid vastavad vastused küsimustikus. Valimis kasutati ainult menopausis naisi, sest reproduktiivses eas naised on hetkeseisust sõltumata võimelised veel rasestuma.

Vanus esimese raseduse ajal

Esimese raseduse aja valimisse olid kaasatud kõik naised, kellel on olnud vähemalt üks rasedus ja vastavad andmed olemas.

Raseduste arv

Raseduste arvu valimis olid naised, kellel on menopaus ja kellel on olnud rasedusi ning naised, kellel on olemas info menopausi kohta või on vanusega \geq 55 aastat ning neil pole rasedusi esinenud.

Raseduskatkemiste esinemine

Raseduse katkemiseks on palju põhjuseid, sellest tulenevalt on antud uurimuses käsitletava tunnuse korral välistusdiagnoosideks naiste endi kromosoomidefektid, erinevad endokriinsed ja immunoloogilised haigused, lisaks anatoomilistest ja geneetilistest eripäradest tulenevad haigusseisundid (täpne nimekiri haiguskoodidest, mille alusel on naisi uuringust välistatud, on Lisas 2). Antud uuringus võrreldi naisi, kellel oli üks või rohkem raseduskatkemist, naistega, kellel oli rasedusi, ent teadaolevalt polnud raseduskatkemisi.

2.2.2. Konsortsiumite assotsiatsiooniuuringute meta-analüüside koondstatistikud

Antropomeetriliste Tunnuste Geneetilise Uurimise konsortsium (*Genetic Investigation of ANthropometric Traits – GIANT*) on rahvusvaheline koostööprojekt, mille eesmärgiks on tuvastada geneetilisi lookusi, mis mõjutavad inimeste keha suurust ja kuju, sealhulgas pikkust

ning rasvumuse määra. GIANT konsortsium toimib koostöös erinevate töögruppide, asutuste, riikide ning uuringute vahel. Peamiseks lähenemisviisiks on ülegenoomsete assotsiatsiooniuuringute andmestikega või teiste suuremahuliste geneetiliste andmestikega läbiviidavad meta-analüüsid. GIANT konsortsium uurib antropomeetrilisi tunnuseid nagu KMI ja pikkus ning tunnuseid, mis on seotud talje ümbermõõduga, näiteks talje-puusa suhe (TPS). GIANT konsortsium on seni tuvastanud antropomeetriliste tunnustega seotud sagedasi geneetilisi variante sadades lookustes⁷. Antud magistritöös on kasutatud GIANT konsortsiumi kehamassiindeksi GWAS meta-analüüsi andmestikku, mille käigus 249 796 indiviidi andmete analüüsil leiti 18 uut kehamassiindeksiga seotud geneetilist lookust. GIANT konsortsiumi uuringu esimeses etapis viidi läbi kehamassiindeksi GWAS meta-analüüsid, mille korral kasutati 2,8 miljonit HapMap2 referentsil imputeeritud ja genotüpiseeritud geneetilist markerit 123 865 indiviidil 46 uuringust. GIANT konsortsiumi uuringu teises etapis viidi läbi assotsiatsioonianalüüsid 42 kehamassiindeksiga seotud geneetilise variandiga 125 931 indiviidil (Speliotes *et al.*, 2010). Samuti on käesolevas magistritöös kasutatud kehamassiindeksile kohandatud talje-puusa suhte GWAS meta-analüüsi andmestikku, mille korral kasutati HapMap2 referentsil imputeeritud geneetilisi variante lisaks 77 167 Euroopa päritolu indiviidi andmetele 32 uuringust (Heid *et al.*, 2010).

Globaalne Lipiidide Geneetika Konsortsium (*Global Lipids Genetics Consortium – GLGC*) on 2006. aastal diabeedi geneetikaga seotud ülegenoomse assotsiatsiooniuuringuga algatatud koostööprojekt, mille eesmärgiks on selgitada plasma lipiidide geneetikat, sealhulgas madala tihedusega lipoproteiinide (*low-density lipoprotein – LDL*) ja kõrge tihedusega lipoproteiinide (*high-density lipoprotein – HDL*) geneetikat, mis seletaks LDL kolesterooli taseme, HDL kolesterooli taseme ja triglütseriidide taseme varieeruvust indiviiditi. GLGC toimib koostöös >200 uurija ning > 80 institutsiooniga üle maailma⁸. Antud uuringus on kasutatud LDL kolesterooli taseme ja HDL kolesterooli taseme GWAS meta-analüüsi andmestikke, mis põhinevad 188 577 Euroopa päritolu ja 7898 mitte-Euroopa päritolu indiviidi ülegenoomsetel genotüpiseerimisandmetel ja HapMap2 referentsil imputatsioonil (Willer *et al.*, 2013). Mõlemad GWAS meta-analüüsi andmestikud hõlmavad erinevate rasside andmeid, mis pole LD skoori regressiooni jaoks ideaalne, kuid on siiski antud fenotüüpide jaoks parimad andmestikud.

Varajase Kasvu Geneetika Konsortsium (*Early Growth Genetics Consortium – EGG konsortsium*) on koostööprojekt, mille eesmärgiks on erinevate uuringute GWAS andmestike

⁷ https://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium

⁸ <http://www.kathiresanlab.org/collaborators/global-lipids-genetic-consortium/>

kombineerimisel tuvastada inimgenoomist lookuseid, mis on seotud varajase kasvu erinevate tunnustega. Antud uuringus on kasutatud sünnikaalu GWAS meta-analüüsi andmestikku, milles on kombineeritud 26 836 indiviidi genotüpiseerimisandmed 18 populatsiooni-põhisest uuringust Euroopas. Andmestikus on kokku > 2 miljoni genotüpiseeritud ja HapMap projekti referentsil (*release 27*) imputeeritud geneetilist varianti (Horikoshi *et al.*, 2013).

2.2.3. Genotüpiseerimine

Antud projekti raames ei viidud läbi genotüpiseerimist, vaid kasutati olemasolevaid genotüpiseerimisandmeid, mis on saadud kasutades Illumina HumanCNV370-DUO, HumanOmniExpress ja HumanCoreExome-12v1.0 kiipe. Neist kõige suurema katvusega on HumanOmniExpress, millel on enam kui 715 000 markerit. HumanCNV370-DUO kiibil on 370 000 markerit ning HumanCoreExome-12v1.0 kiibil on üle 500 000 markeri⁹. Genotüpiseerimine teostati TÜ Eesti Geenivaramu Tuumiklabori Genotüpiseerimiskeskuses vastavalt tootja protokollile.

2.2.4. Kvaliteedikontroll

Genotüpiseerimise järgselt läbis uuritav valim kvaliteedikontrolli järgnevate parameetrite ja näitajate osas:

- 1) edukalt genotüpiseeritud markerite protsentuaalne osakaal (*call rate* – CR) iga indiviidi ja geneetilise variandi kohta
- 2) harvem esineva alleeli sagedus
- 3) Hardy-Weinbergi tasakaalustatus (*Hardy-Weinberg equilibrium* – HWE)
- 4) individidevaheline sugulus
- 5) genotüübi- ja fenotüübipõhise soo ebakõla

Edasisest analüüsist jäeti välja isikud, kelle CR oli alla 95% ning inimesed, kelle genotüübis ilmnenud sugu ei vastanud fenotüübilisele soole. Individidevaheline sugulus leiti ühiste alleelide osatähtsuse (*identity by state* – IBS) hindamise järgi genoomis, millest tulenevalt ennustati iga sugulaspaari päritolult identse DNA (*identity by descent* – IBD) osatähtsus. Olukorras, mil IBD osatähtsus kahe indiviidi vahel oli $\geq 10\%$ genoomist, eemaldati edasistest analüüsides üks sugulastest. Lisaks eemaldati SNP-d, millel oli CR < 95%, MAF < 1% ning HWE testi p-väärtus < 10^{-6} . Andmete kvaliteedikontrollil ja filtreerimisel kasutati programmi PLINK vs 1.08 (Purcell *et al.*, 2007).

⁹ www.illumina.com

2.2.5. Imputatsioon

Imputeerimine on statistiline meetod, mis võimaldab haplotüüpide referentspaneeli põhjal ennustada genotüpiseerimata jäänud SNP-de genotüüpide tõenäosusi, suurendades seega analüüsitava markerite arvu ning tõstes uuringu võimsust (Marchini *et al.*, 2007). Seda võimaldab puuduvate SNP-ide LD kiibil olevate markeritega. LD korral on geneetilised variandid erinevates lookustes omavahel aheldunud tugevamalt kui võiks juhuslikkuse korral eeldada. Täpseks imputeerimiseks peavad referentspaneeli andmestik ja uuritav andmestik olema geneetiliselt lähedased ja ühise LD struktuuriga (Halperin & Stephan, 2009).

Peale kvaliteedikontrolli jagati genotüpiseerimisandmeid kromosoomide kaupa eraldi failidesse, kasutades programmi ShapeIT v2 (Delaneau *et al.*, 2013). Faasimisel leiti võimalikud vanematelt vanematelt pärandunud haplotüübid. Seejärel imputeeriti kiipidel puuduolevad variandid, kasutades programmi Impute v2.2.2 (Howie *et al.*, 2009) 5 miljoni aluspaariliste (*mega base pairs* – Mb) genomilõikude kaupa. Referentsina kasutati 1000 Genoomi Projekti raames sekveneeritud haplotüüpe, vastavalt autosoomide korral *1000 Genomes Phase 3 integrated variant set, Oct 2014* ning X-kromosoomi korral *1000 Genomes Phase 1, March 2012*. Imputeerimise viimases etapis tõsteti kromosoomide tükid taas kokku.

Andmete kvaliteedikontrolli ja imputeerimise viis läbi Evelin Mihailov, välja arvatud Illumina HumanCoreExome-12v1.0 X-kromosoomi analüüs, mille teostas töö autor.

2.2.6. Ülegenoomne assotsiatsioonianalüüs

Geneetiliste markerite ja fenotüübi vaheliseks assotsiatsioonianalüüsiks kasutati pidevate tunnuste korral lineaarset regressiooni ning binaarsete tunnuste korral logistilist regressiooni, mõlemal korral kasutati aditiivset geneetilist mudelit programmis SNPTTEST v2.5 (Marchini *et al.*, 2007). Analüüsides kasutati kovariaatidena proovide suguluste maatriksi põhjal leitud nelja esimest peakomponenti, kahandamaks võimalikku populatsiooni stratifikatsiooni. Lisaks kasutati kovariaadina vanuse asemel sünniaastat, et kohandada ka võimalikele seksuaal- ja reproduktiivkäitumist mõjutavatele ajalistele trendidele. Menstruaaltsükli pikkuse korral oli kovariaadina arvestatud ka seda mõjutav kehamassiindeks. Arvutused viidi läbi Tartu Ülikooli teadusarvutuste keskus¹⁰.

2.2.7. Meta-analüüs

Meta-analüüsis ühildati kolme kiibi assotsiatsioonianalüüsides tulemused, mille tulemusena suurenes analüüsi võimsus ja vähenes valepositiivsete leidude arv. Selleks kasutati GWAMA

¹⁰ <http://www.hpc.ut.ee>

programmi (Mägi & Morris, 2010). Sisendfaile filtreeriti imputatsiooni kvaliteedi suhtes ($info \geq 0,4$) ja populatsiooni stratifikatsiooni korrigeerimiseks kasutati genomset kontrolli (*Genomic control ON*). Meta-analüüsi järgselt filtreeriti edasiseks analüüsiks markerid, mis olid peale filtreerimist olemas vähemalt kahel kiibil ning mille heterogeensus (Cochrani Q-testi p-väärtus) oli > 0.05 ning harvem esineva alleeli sagedus $> 1\%$. Mitmesest testimisest tulenevate vigade vältimiseks loeti ülegenoomselt statistiliselt oluliseks markereid, mille p-väärtus oli $< 5 \times 10^{-8}$ (Pe'er *et al.*, 2008).

Assotsiatsiooniuuringu tulemused visualiseeriti Manhattan joonistena. Mitmed valitud tulemused visualiseeriti ka regionaaljooniste abil. Positsioonid on kujutatud inimese referentsgenoomist hg19/GRCh37 lähtuvalt. LD väärtused on lähtuvalt 1000 Genoomi Nov 2014 EUR referentsist. Joonistel halli värviga esitatud markeritel puuduvad referentsis LD väärtused. Regionaaljoonistel on y-teljel negatiivne kümnendlogaritm p-väärtusest ning x-teljel on esitatud 1 Mb ulatuses ümbritsev lookus. Lisaks on joonistel esitatud rekombinatsioonimäär ja antud regioonis paiknevate geenide nimed (Pruim *et al.*, 2011).

2.2.8. LD skoori regressioon

LD skoori regressioon võimaldab leida geneetilisi korrelatsioone, mille põhjal saab hinnata, kas tunnuste epidemioloogiliste seoste aluseks on ka jagatud geneetika. Tunnuste vahelise jagatud geneetika leidmiseks kasutati LDSC (LD Score) v1.0.0 programmi, mis hindab tunnuste päritavust ja tunnuste vahelist geneetilist korrelatsiooni, kasutades GWAS koondstatistikuid (Bulik-Sullivan *et al.*, 2015). GWAS koondstatistikud sisaldavad informatsiooni antud valimiga läbiviidud ülegenoomses assotsiatsioonianalüüsis kasutatud kõikide SNP-ide efektisuuruste, standardhälvete, alleelisageduste ja p-väärtuste kohta. Antud meetod toetub faktile, et GWAS käigus leitud kindla SNP-i efektisuurused hõlmavad kõikide SNP-ide efekti, mis on antud SNP-iga aheldatuse tasakaalutuses. Meetod võimaldab tunnuste võrdlemisel kasutada ka geneetilisi variante, mis polnud ülegenoomselt statistiliselt olulised. Lisaks ei nõua koondstatistikute informatsioonil põhinev meetod indiviidi tasemel genotüpiseerimisandmeid, mille kättesaamine on tihti informeeritud nõusolekute tõttu raskendatud (Finucane *et al.* 2015).

Esimese sammuna konverteeriti koondstatistikute fail .sumstats formaati, kasutades `munge_sumstats.py` programmi, mis on LDSC (LD Score) v1.0.0 kaasnev laiend. Koondstatistikute fail peab sisaldama SNP-i identifitseerimiseks vajalikku infot (näiteks rs-number), valimi suurust, z-skoori (efektisuurus jagatud standardveaga) või z-skoori kujule viidavaid andmeid, efekti alleeli ning teist alleeli. Antud programm ei võimalda LD skoori

regressiooni rakendada sugukromosoomidele, seetõttu kasutati analüüsis ainult autosomaalseid kromosome. Meetodit pole võimalik rakendada insertioonidele ja deletsioonidele, mistõttu need on eemaldatud GWAS andmestikest. Teise etapina arvutatakse päritavuskoeffitsient (h^2) ning geneetiline korrelatsioon (r_g). Kahe tunnuse vahelise geneetilise korrelatsiooni leidmisel on oluline, et mõlema tunnuse GWAS andmestikud oleks ühtset etnilist päritolu. LD skoori regressioon võimaldab arvutada geneetilist korrelatsiooni näiteks kahe Euroopa päritolu GWAS andmestiku vahel või kahe Aasta päritolu GWAS andmestiku vahel, kuid programm ei ole suuteline leidma geneetilist korrelatsiooni Euroopa ning Aasia päritolu GWAS andmestiku vahel. Lisaks ei võimalda programm kasutada korrelatsiooni leidmisel kahe mitme kontinendi põhiseid meta-analüüside andmeid. Sellisel juhul leitakse igale kontinendile eraldi geneetiline korrelatsioon ning hiljem arvutatakse tulemuste ühine keskmine. LD skoori regressiooni korral on oluline valimi suurus – usaldusväärsete tulemuste jaoks on vajalik valimi suurus ligikaudu 5000 indiidivi. Lisaks võib standardhälve suurened, kui valimis on vähem kui 600 000 SNP-i. Meetodit ei ole võimalik rakendada andmetele, mis pärinevad spetsiifilistelt genotüpiseerimiskiipidelt (näiteks Illumina Cardio-MetaboChip)¹¹. Tulemused on visualiseeritud R tarkvarapaketi (v3.1.1)¹².

2.3. Tulemused

2.3.1. Ülegenoomsete assotsiatsiooniuuringute meta-analüüsid

Antud töö raames viidi läbi ülegenoomsed assotsiatsiooniuuringud, leidmaks geneetilisi variante, mis on seotud naiste viljakust iseloomustavate tunnustega. Geneetiliste variantide positsioonid on lähtuvalt inimese referentsgenoomist GRCh37 (*Genome Reference Consortium GRCh37, Feb 2009*). Meta-analüüsi käigus ühildati kolme kasutatud genotüpiseerimiskiibi assotsiatsioonianalüüside tulemused. Uuringus vaadeldi 7 erinevat tunnust: vanus menarhe ajal, menstruaaltsükli pikkus, menstruaaltsükli regulaarsus, raseduste esinemine, vanus esimese raseduse ajal, raseduste arv ning raseduskatkemiste esinemine. Ülegenoomsete assotsiatsiooniuuringute meta-analüüside käigus leitud ülegenoomselt olulised markerid on esitatud Tabelis 2. Ülegenoomselt olulisi (p -väärtus $< 5 \times 10^{-8}$) markereid leiti neljale tunnusele: menstruaaltsükli pikkus, menstruaaltsükli regulaarsus, raseduste esinemine ja raseduskatkemiste esinemine. Kokku leiti meta-analüüsides 9 ülegenoomselt olulist markerit. Kolmel tunnusel – vanusel menarhe ajal, vanusel esimese raseduse ajal ja raseduste arvul, ülegenoomselt olulisi markereid antud uuringu raames ei tuvastatud. Seetõttu on välja toodud antud tunnustega seotud markereid, mille p -väärtus oli väiksem kui 5×10^{-7} .

¹¹ <https://github.com/bulik/ldsc/>

¹² <https://www.r-project.org/>

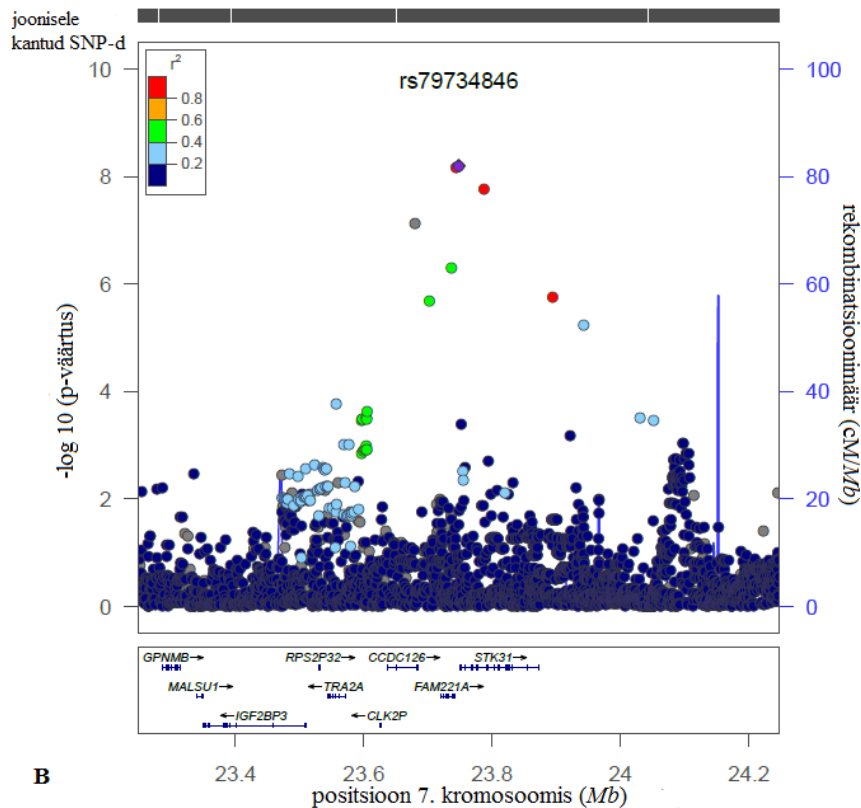
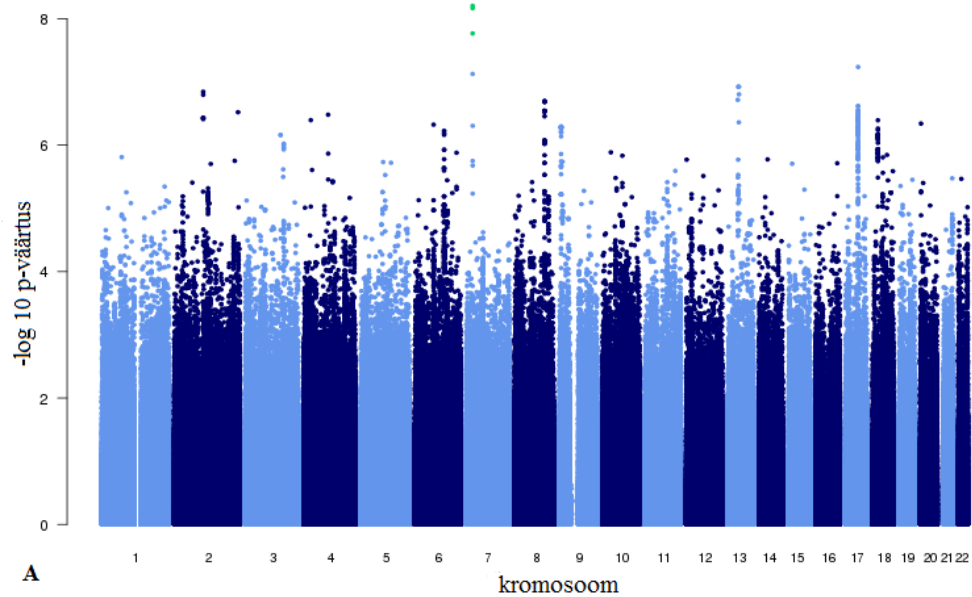
Menarhe vanuse assotsiatsiooniuuringu käigus ei leitud vaadeldava tunnusega ühtegi üleüldiselt statistiliselt olulist seost. Statistiliselt mitteolulised, ent madalaimad p-väärtused olid geneetilise variandiga rs34231612 (p-väärtus $1,62 \times 10^{-7}$), mis asub geeni *CNTN6* lähedal 3. kromosoomis ja geneetilise variandiga rs56376587 (p-väärtus $3,08 \times 10^{-7}$), mis paikneb 18. kromosoomis geeni *NFATC1* intronis.

Menstruaaltsükli pikkusega assotsieerus geneetiline variant rs147328053 (p-väärtus $1,91 \times 10^{-8}$) 3. kromosoomis *SGOL1* geeni lähedal. Tegemist on deletsiooniga, mille sagedus on populatsioonis ligikaudu 3%.

Menstruaaltsükli regulaarsuse assotsiatsiooniuuringus leiti üks üleüldiselt assotsieerunud lookus 7. kromosoomis, mis hõlmab gene *STK31* ja *FAM221A* (Joonis 4A). Leitud kolme markeri seast oli statistiliselt kõige olulisem *STK31* geenis paiknev rs79734846 (p-väärtus $6,27 \times 10^{-9}$) (Joonis 4B). Teisteks leitud geneetilisteks variantideks olid rs74524971 (p-väärtus $6,80 \times 10^{-9}$), mis asub geenis *FAM221A* ning geneetiline variant rs74904363 ($1,71 \times 10^{-8}$), mis asub geenis *STK31*.

Tabel 2. Ülegenoomse assotsiatsiooniuringute meta-analüüside käigus leitud ülegenoomselt olulised markerid. Efektialleeli sagedus (*effect allele frequency* – EAF). Inimese referentsgenoom GRCh37 (*Genome Reference Consortium GRCh37, Feb 2009*).

Uuritav tunnus	Tunnuse tüüp	Marker	Kromosoom ja positsioon	Lähim geen	Efektialleel/ teine alleel	EAF	Efekt (standardhälve)	P-väärtus
Menstruaaltsükli pikkus	Kvantitatiivne	rs147328053	3:20570696	<i>SGOL1</i>	G/GA	0,16	0,18 (0,03)	$1,91 \times 10^{-8}$
		rs74524971	7:23743366	<i>FAM221A</i>	C/G	0,04	-1,02 (0,18)	$6,80 \times 10^{-9}$
Menstruaaltsükli regulaarsus	Binaarne	rs79734846	7:23747984	<i>STK31</i>	T/G	0,04	-1,02 (0,18)	$6,27 \times 10^{-9}$
		rs74904363	7:23787549	<i>STK31</i>	T/G	0,04	-1,00 (0,18)	$1,71 \times 10^{-8}$
		rs150079937	2:230404792	<i>DNER</i>	C/T	0,03	-1,64 (0,29)	$2,71 \times 10^{-8}$
Raseduste esinemine	Binaarne	rs200713680	2:230428227	<i>DNER</i>	G/A	0,04	-1,69 (0,29)	$8,35 \times 10^{-9}$
		chr3:100797937:I	3:100797937	<i>ABI3BP</i>	CTA/C	0,39	-0,73 (0,13)	$1,68 \times 10^{-8}$
Raseduskatkemiste esinemine	Binaarne	rs79309044	4:109669412	<i>AGXT2L1</i>	C/G	0,02	0,71 (0,12)	$3,49 \times 10^{-8}$
		rs115007121	4:109708225	<i>AGXT2L1</i>	A/G	0,03	0,67 (0,12)	$2,94 \times 10^{-8}$



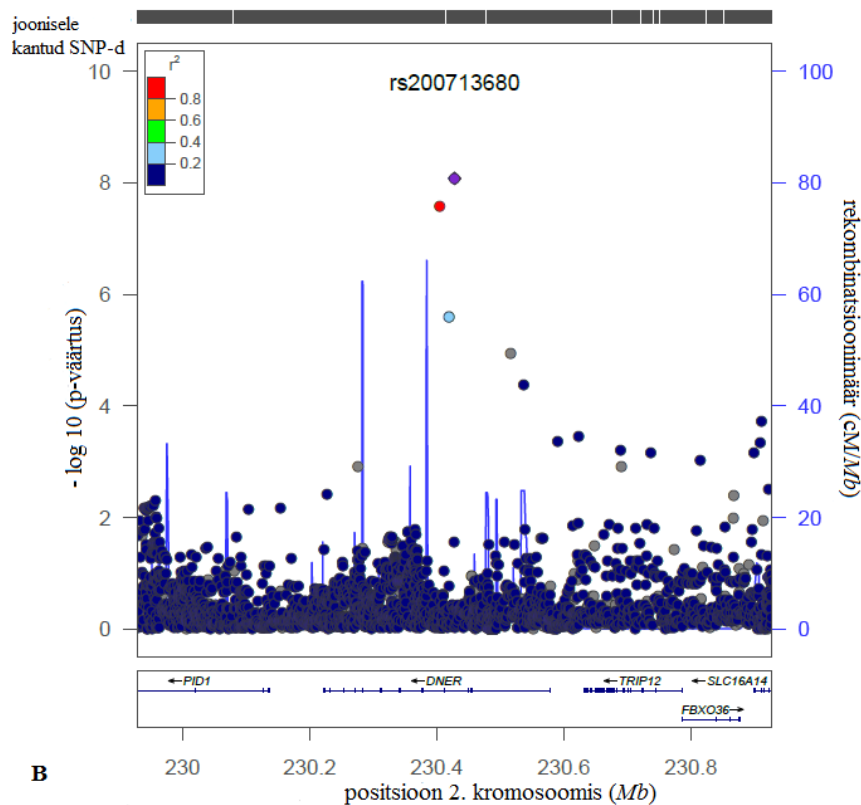
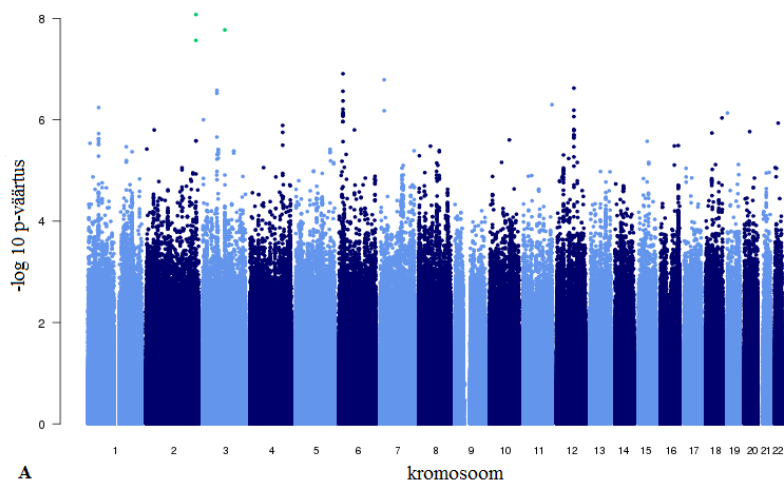
Joonis 4. Menstruaaltsükli regulaarsuse GWAS tulemused. (A) Ülegenoomse assotsiatsiooniuringu tulemused visualiseerituna Manhattan joonisel. Y-teljel on negatiivne kümnendlogaritm assotsiatsiooni p-väärtusest. X-teljel on genoomsed asukohad. Rohelised täpid on ülegenoomselt statistiliselt olulised assotsiatsioonid, antud juhul 7. kromosoomis. (B) Markeri rs79734846 regionaaljoonis. Erineva värvusega on kujutatud markerite LD (r^2) markeriga rs79734846 (lilla romb). X-teljel on markerite positsioonid ja geenide asukohad genoomis. Vasakpoolisel y-teljel on negatiivne kümnendlogaritm p-väärtusest ja parempoolisel y-teljel kujutab sinine joon rekombinatsioonimäära.

Raseduste esinemise assotsiatsioonianalüüsis leiti kolm ülegenoomselt olulist seost (Joonis 5A). Kõige olulisem leitud geneetiline variant rs200713680 (p-väärtus $8,35 \times 10^{-9}$) asub 2. kromosoomis *DNER* geenis (Joonis 5B), kus paikneb ka geneetiline variant rs150079937 (p-väärtus $2,71 \times 10^{-8}$). Lisaks oli ülegenoomselt oluline seos 3. kromosoomis positsioonil chr3:100797937 paikneva insertiooniga (p-väärtus $1,68 \times 10^{-8}$).

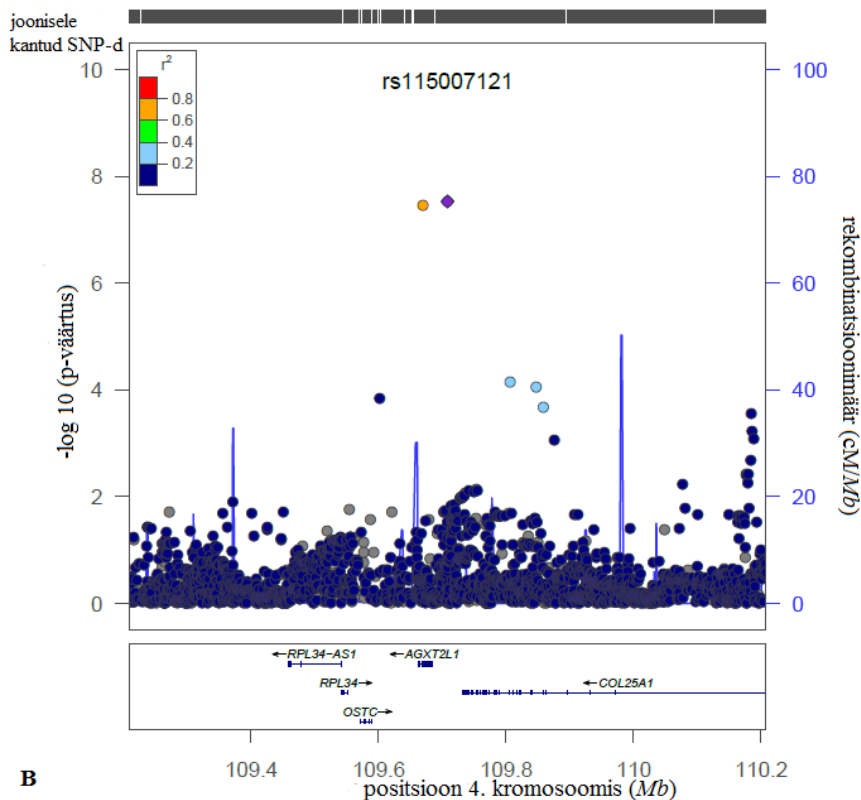
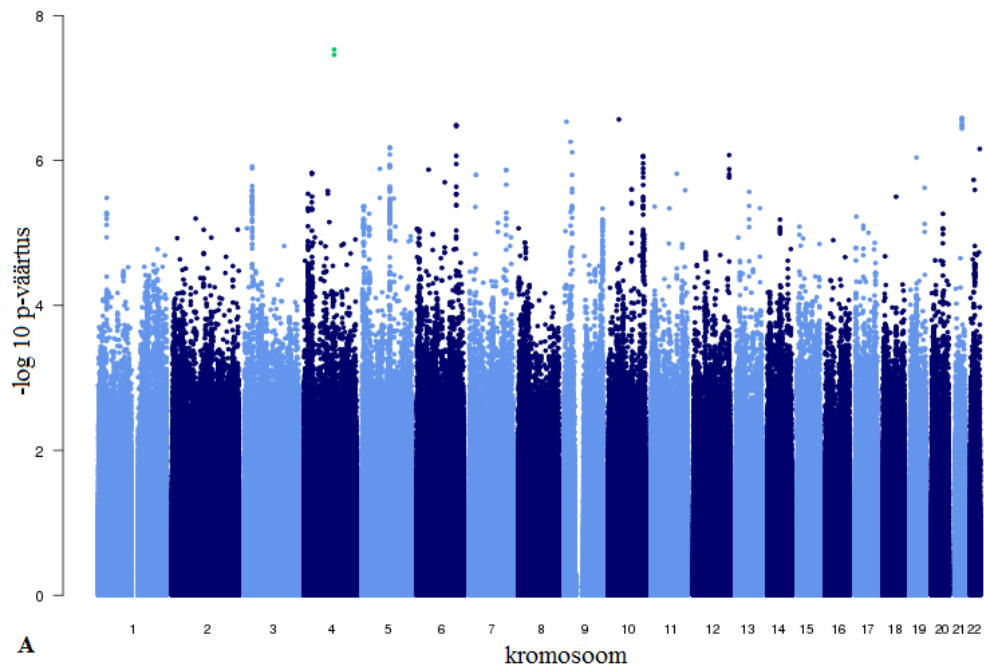
Vanus esimese raseduse ajal assotsiatsiooniuringus ei leitud ülegenoomselt olulisi seoseid. Ülegenoomselt statistiliselt mitteolulised, kuid madalaima p-väärtusega seosed olid 5. kromosoomis 5q15 lookusega. Kõige tugevam leitud seos oli markeriga rs2351012 (p-väärtus $2,52 \times 10^{-7}$), mis asub geeni *ERAPI* intronis.

Raseduste arvu assotsiatsioonianalüüsis ei leitud ülegenoomselt olulisi seoseid. Uuringus leiti kõige tugevam seos geneetilise variandiga rs117885050 (p-väärtus $1,06 \times 10^{-7}$), mis asub 17. kromosoomis geenide *SLC16A11* ja *CLEC10A* vahelises piirkonnas. Tunnusega leiti veel seoseid 3p.13 lookuses *PDZRN3* geenis paiknevate geneetiliste variantidega. Piirkonnast leitud assotsiatsioonidest oli kõige olulisem seos geneetilise variandiga rs9818960 (p-väärtus $2,73 \times 10^{-7}$), mis asub *PDZRN3* intronis. Samas genoomses piirkonnas asuvad ühtlasi leitud markerid rs11926413 (p-väärtus $3,75 \times 10^{-7}$) ja rs4677314 (p-väärtus $2,79 \times 10^{-7}$).

Raseduskatkemise esinemise assotsiatsiooniuringus leiti kaks ülegenoomselt olulist seost 4q25 lookuses (Joonis 6A). Leitud geneetiline variant rs79309044 (p-väärtus $3,49 \times 10^{-8}$) paikneb *AGXT2LI* geeni intronis ning rs11500712 (p-väärtus $2,94 \times 10^{-8}$) (Joonis 6B) asub geeni *COL25A1* lähedal.



Joonis 5. Raseduste esinemise GWAS tulemused. (A) Ülegenoomse assotsiatsiooniuringu tulemused visualiseerituna Manhattan joonisel. Y-teljel on negatiivne kümnendlogaritm assotsiatsiooni p-väärtusest. X-teljel on geoomsed asukohad. Rohelised täpid on ülegenoomselt statistiliselt olulised assotsiatsioonid. (B) Markeri rs200713680 regionaaljoonis. Erineva värvusega on kujutatud markerite LD (r^2) markeriga rs200713680 (lilla romb). X-teljel on markerite positsioonid ja geenide asukohad geoomis. Vasakpoolsel y-teljel on negatiivne kümnendlogaritm p-väärtusest ja parempoolsel y-teljel kujutab sinine joon rekombinatsioonimäära.



Joonis 6. Raseduskatkemiste esinemise GWAS tulemused. (A) Ülegenoomse assotsiatsiooniuringu tulemused visualiseerituna Manhattan joonisel. Y-teljel on negatiivne kümnendlogaritm assotsiatsiooni p-väärtusest. X-teljel on genoomsed asukohad. Rohelised täpid on ülegenoomselt statistiliselt olulised assotsiatsioonid. (B) Marker **rs115007121 regionaaljoonis. Erineva värvusega on kujutatud markerite LD (r^2) markeriga rs115007121 (lilla romb). X-teljel on markerite positsioonid ja geenide asukohad genoomis. Vasakpoolsel y-teljel on negatiivne kümnendlogaritm p-väärtusest ja parempoolsel y-teljel kujutab sinine joon rekombinatsioonimäära.**

2.3.2. LD skoori regressioon

Kõikidele uuritavatele tunnustele leiti päritavuse hinnangud, mida kajastab päritavuskoeffitsiendi h^2 väärtus ning selle standardhälve (Tabel 3). Neljal tunnusel – menstruaaltsükli pikkusel, menstruaaltsükli regulaarsusel, raseduste esinemisel ja raseduste arvul ei õnnestunud päritavuskoeffitsienti usaldusväärset hinnata liiga väikse valimi tõttu. Kolmel tunnusel – vanusel menarhe ajal, vanusel esimese raseduse ajal ja raseduskatkemiste esinemisel leiti päritavuse hinnangud, mis varieerusid 20% ringis.

Tabel 3. Uuritavate tunnuste päritavuskoeffitsiendi h^2 väärtused ning standardhälve. Puuduv väärtus (*not available* – NA).

Uuritav tunnus	Päritavuskoeffitsient h^2 (standardhälve)
Vanus menarhe ajal	0,27 (0,06)
Menstruaaltsükli pikkus	NA
Menstruaaltsükli regulaarsus	NA
Raseduste esinemine	NA
Vanus esimese raseduse ajal	0,24 (0,07)
Raseduste arv	NA
Raseduskatkemiste esinemine	0,17 (0,06)

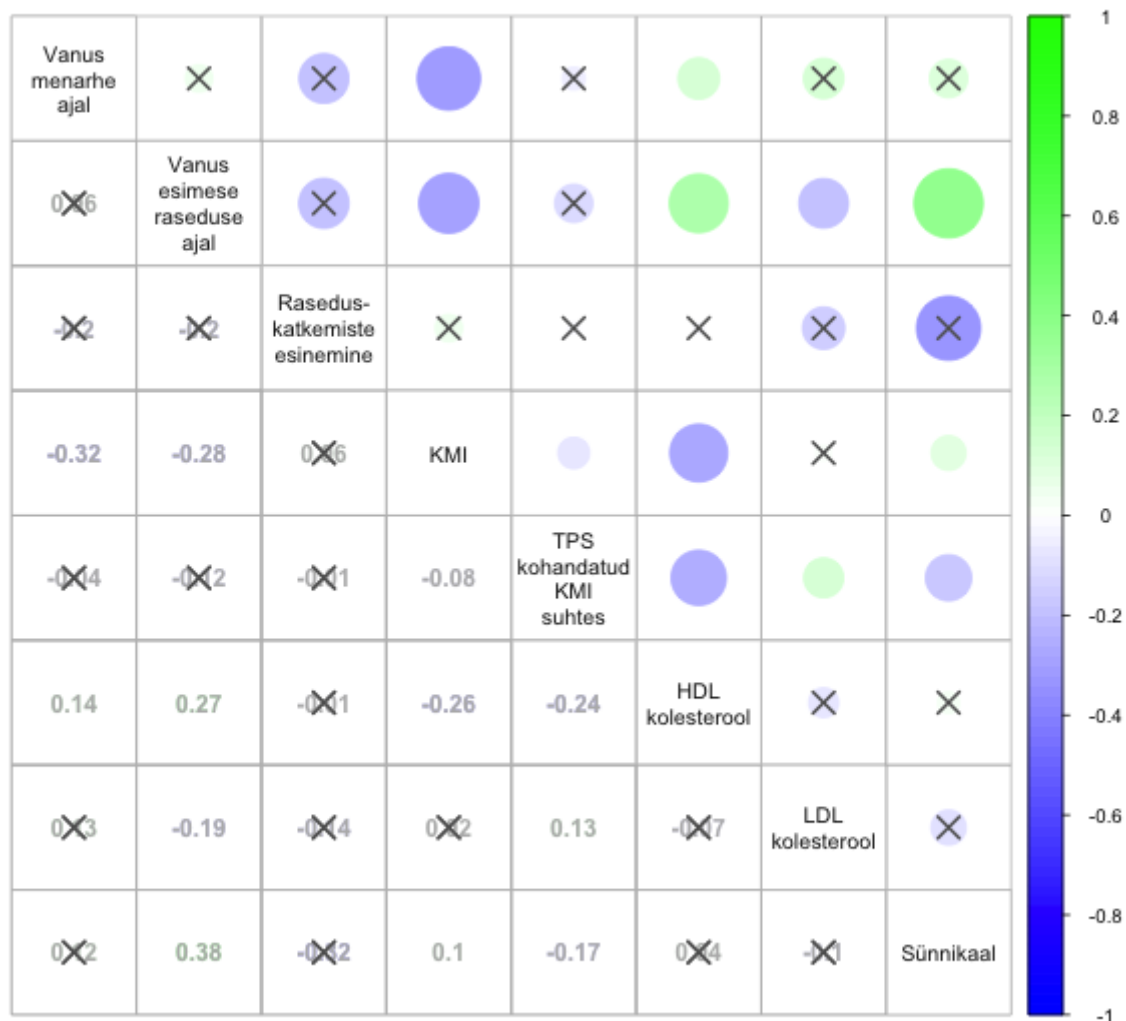
Geneetiliste korrelatsioonide leidmiseks sai kasutada ainult tunnuseid, mille valimi suurus oli piisav päritavuskoeffitsiendi arvutamiseks: vanus menarhe ajal, vanus esimese raseduse ajal ja raseduskatkemiste esinemine. Lisaks kasutati geneetiliste korrelatsioonide leidmisel vabalt kättesaadavaid GIANT konsortsiumi kehamassiindeksi ning kehamassiindeksile kohandatud talje-puusa suhte GWAS andmestikke, GLGC konsortsiumi HDL ja LDL kolesterooli GWAS andmestikke ning EGG konsortsiumi sünnikaalu GWAS andmestikku. Kõikide LD skoori regressiooniks kasutatud tunnuste vahelised korrelatsioonid on esitatud Joonisel 7. Raseduskatkemiste esinemisel puudusid nominaalselt olulised geneetilised korrelatsioonid teiste vaadeldavate tunnustega.

Vanus menarhe ajal oli negatiivses geneetilises korrelatsioonis kehamassiindeksiga (geneetiline korrelatsioon -0,32; p-väärtus $1,28 \times 10^{-6}$). Geneetilised eeldused kõrgemaks kehamassiindeksiks on seotud ka varasema menstruaaltsükli ajastusega. Ühtlasi oli tegu menarhega kõige tugevamalt seotud tunnusega uuritavate geneetiliste korrelatsioonide hulgas. Vanus menarhe ajal oli nõrgas positiivses geneetilises korrelatsioonis HDL kolesterooli tasemega (geneetiline korrelatsioon 0,14; p-väärtus $3,03 \times 10^{-2}$). Naistel, kellel on geneetilised eeldused kõrgemaks HDL kolesterooli tasemeks, algab menstruaaltsükkel hiljem või

vastupidiselt – naised, kellel on hiljem menarhe, on geneetiliste eeldustega kõrgemaks HDL kolesterooli tasemeks.

Vanus esimese raseduse ajal ja kehamassiindeks olid mõõdukas negatiivses geneetilises korrelatsioonis (geneetiline korrelatsioon $-0,28$; p -väärtus 3×10^{-4}), millest tulenevalt naised, kellel on geneetilised eeldused kõrgemaks kehamassiindeksiks ka rasestuvad varasemas eas kui naised, kellel selleks geneetilised eeldused puuduvad. Vanus esimese raseduse ajal oli mõõdukas positiivses geneetilises korrelatsioonis HDL kolesterooli tasemega (geneetiline korrelatsioon $0,27$; p -väärtus $4,80 \times 10^{-3}$). Geneetilised eeldused kõrgemaks HDL kolesterooli tasemeks on eeldusteks ka hilisemaks esimeseks raseduseks. Erinevalt seosest HDL kolesterooli tasemega, oli vanus esimese raseduse ajal ning LDL kolesterooli tase nõrgas negatiivses geneetilises korrelatsioonis (geneetiline korrelatsioon $-0,19$; p -väärtus $4,56 \times 10^{-2}$), millest tulenevalt naised, kellel on geneetilised eeldused kõrgemaks LDL kolesterooli tasemeks, on esimese raseduse ajal nooremad kui naised, kellel puuduvad vastavad geneetilised eeldused. Vanus esimese raseduse ajal oli mõõdukas positiivses geneetilises korrelatsioonis sünnikaaluga (geneetiline korrelatsioon $0,38$; p -väärtus $4,80 \times 10^{-2}$). Naised, kellel oli geneetilisest eeldustest tulenevalt kõrgem sünnikaal, on eeldus ka hilisemaks raseduseks.

Geneetilist korrelatsiooni ei leitud vanusel menarhe ajal ning vanusel esimese raseduse ajal. Vanus menarhe ajal kui ka vanus esimese raseduse ajal omasid positiivset geneetilist korrelatsiooni HDL kolesterooli tasemega, kuid vanusel esimese raseduse ajal oli leitud korrelatsioon oluliselt tugevam. Vanusel esimese raseduse ajal oli negatiivne korrelatsioon LDL kolesterooli tasemega, kuid vanusel menarhe ajal ja LDL kolesterooli tasemel oluline geneetiline korrelatsioon puudus. Vanus esimese raseduse ajal oli tugevas positiivses geneetilises korrelatsioonis sünnikaaluga, kuid menarhe ja sünnikaalu vahel antud uuringu raames geneetilist korrelatsiooni ei leitud.



Joonis 7. Geneetilised korrelatsioonid 8 ülegenoomse assotsiatsiooniuringu andmetike vahel. Rohelisega on märgitud positiivsed geneetilised korrelatsioonid, sinisega on märgitud negatiivsed geneetilised korrelatsioonid. Tunnustest moodustuvast diagonaalist jäävad ülispoole leitud korrelatsioonid ringidena kujutatuna: mida tugevam on seos, seda suurem on ring. Diagonaalist allapoole jäävad leitud seoste r_g väärtused numbritena. Ristiga on märgitud seosed, mis ei ole statistiliselt olulised (nominaalne nivoo p -väärtus $< 0,05$). Peale mitmese testimise korrektsiooni jäävad statistiliselt oluliseks järgnevad seosed: vanus menarhe ajal ja KMI; vanus esimese raseduse ajal ja KMI.

2.4. Arutelu

Antud magistritöö üheks eesmärgiks oli leida geneetilisi seoseid naiste viljakust peegeldavate erinevate tunnustega. Naiste viljakusega seotud tunnusteks olid – vanus menarhe ajal, menstruaaltsükli pikkus, menstruaaltsükli regulaarsus, raseduste esinemine, vanus esimese raseduse ajal, raseduste arv ja raseduskatkemiste esinemine. Selleks viidi läbi iga uuritava tunnusega ülegenoomne assotsiatsiooniuuring. Naiste viljakusega seotud tunnustega leiti ülegenoomse assotsiatsioonianalüüsi käigus ülegenoomselt olulisi geneetilisi seoseid nelja tunnusega – menstruaaltsükli pikkusega, menstruaaltsükli regulaarsusega, raseduste esinemisega ja raseduskatkemiste esinemisega. Järgnevalt vaadeldi naiste viljakusega seotud tunnuste omavahelist jagatud geneetikat, kasutades selleks LD skoori regressiooni meetodit. Esmalt leiti igal tunnusele päritavushinnangud, kuid tulenevalt valimi suurusest, suudeti ainult kolmele tunnusele leida päritavuskoeffitsient. Nendeks tunnusteks olid vanus menarhe ajal, vanus esimese raseduse ajal ja raseduskatkemiste esinemine. Naiste viljakusega seotud tunnustel puudusid omavahelised geneetilised korrelatsioonid. Lisaks uuriti antud magistritöös, kas naiste viljakusega seotud tunnustel on korrelatsioone epidemioloogilisel tasandil seotud tunnustega. Geneetilised korrelatsioonid epidemioloogilisel tasandil seotud tunnustega leiti vanusele menarhe ajal ja vanusele esimese raseduse ajal.

Menarhe ajaga ei leitud ülegenoomses assotsiatsiooniuuringus ülegenoomselt olulisi seoseid, kuid olulisemad seosed olid geneetiliste variantidega 3. ja 18. kromosoomis. Geneetilised variandid 3. kromosoomis paiknevad *CNTN6* geeni lähedal, mis on seotud raku adhesiooni ja neuronaalse arenguga (Bouyain & Watkins, 2010; Hu *et al.*, 2015). Naiste reproduktiivtervise kontekstis on Day jt varasemalt leidnud seose neuronaalse raku adhesioonmolekuli kodeeriva geeni *CADM2* intronis paiknevate SNP-de ja laste arvu vahel (Day *et al.*, 2016). Geneetilised variandid 18. kromosoomis paiknevad *NFATC1* geeni intronis, mille poolt kodeeritavat valku on eelnevalt seostatud immuunvastuses vajalike geenide transkriptide indutseerimises (Kaminuma *et al.*, 2008). Leitud piirkonnad ei ühtinud seni kõige suuremahulisema menarhe aja ülegenoomse assotsiatsiooniuuringu kõige olulisemate tulemustega (Perry *et al.*, 2014). Peamine põhjus selleks on valimi suurus, kuigi menarhe aja valimis oli ligi 6000 naist, jääb see suuruselt alla uuringutele, millesse on kaasatud andmed kümnetest uuringutest.

Menarhe aeg oli negatiivses geneetilises korrelatsioonis kehamassiindeksiga, mis kinnitab ka epidemioloogilistes uuringutes leitud seoseid. Peters jt uuringus leiti, et naistel, kellel on menarhe enne 12. eluaastat, on keskmine kehamassiindeks 29,0. Naistel, kellel on menarhe

alles pärast 14. eluaastat, on keskmine kehamassiindeks 26,5 (Peters *et al.*, 2016). Menarhe aja ning täiskasvanuea kehamassiindeksi vaheline geneetiline seos võib tuleneda eelkõige lapsea kehamassiindeksist, kuivõrd hiljutises uuringus on leitud, et lapsea kehamassiindeks ning täiskasvanuea kehamassiindeks on tugevas geneetilises korrelatsioonis (Felix *et al.*, 2016).

Geneetilised lookused, mis on seotud keharasva paiknemisega tuhara ning reite piirkonnas võiksid mõjutada menarhet, kuivõrd nendes piirkondades paiknevat rasva on võimalik vabamalt mobiliseerida kui kõhupiirkonna rasvavarusid reproduktiivseks talitluseks (Snijder *et al.*, 2003). Uuringu tulemus viitab sellele, et menarhe aeg ei sõltu keharasvade paiknemist vaid üldisest rasvumusseisundist. Menarhe ning kehamassiindeksi bioloogiliseks ühisosaks võiks olla leptiin. Keharasvadest sekreteeritud leptiin on söögiisu reguleeriv hormoon (Friedman & Halaas, 1998), mille mõju reproduktiivsele süsteemile on pigem kaudne. Näriliste katsetel leiti, et eesaju neuronite konditsionaalse leptiini retseptori deletsiooni korral hilines närilistel puberteet ja selle tulemusena olid närilised viljatud, samuti oli blokeeritud östradiolist tulenev LH taseme tõus. GnRH neuronite leptiini retseptori deletsiooni korral ainult isastel närilistel puberteet hilines, kuid edasisi reproduktiivseid probleeme ei ilmnenud (Quennell *et al.*, 2009). Samuti võib bioloogiliseks ühenduslüliks olla insuliin. *In vivo* hiirte uuringus näidati, et suurenevad insuliini kontsentratsioonid stimuleerivad LH sekretsiooni ning *in vitro* mudelis näidati, et otseselt insuliini doosist sõltuvalt toimus GnRH-i sekretsioonis muutus (Burcelin *et al.*, 2003).

Epidemioloogilistes uuringutes on varajast menarhet seostatud ka südameveresoonekonnahaiguste suurenenud riskiga (Prentice & Viner, 2013). Antud magistritöös leitud menarhe ja HDL kolesterooli positiivne geneetiline korrelatsioon ühtib varasema Bulik-Sullivan jt kahe tunnuse vahelise geneetilise korrelatsiooniga (Bulik-Sullivan *et al.*, 2015). Varasemalt on leitud menarhe ajaga seotud geneetilisi variante steroidhormoonide sünteesis osalevatest geenidest või nende lähedalt. Perry jt uuringus leiti, et mittekodeerivas RNA-s paiknev geneetiline variant mõjutab lipiidide transpordis ja metabolismis osaleva (Calderon-Dominguez *et al.*, 2014) *STARD4* geeni ekspressioonitaset (Perry *et al.*, 2014). *STARD4* kuulub valguperekonda, mis osaleb steroidhormoonide sünteesis regulaatorina, konverteerides kolesterooli pregnenolooniks, mis on androgeenide ja östrogeenide prekursoriks (Lin *et al.*, 1995).

Antud magistritöös leiti menstruaaltsükli pikkusega olulisim seos *SGOL1* geeni, mis on seotud mitoosis tütarchromatiidide korrektse segregatsiooniga (Yamagishi *et al.*, 2008), lähedal paikneva piirkonnaga. Varasemalt on assotsiatsiooniuuringutes leitud *SGOL1* intergeensest piirkonnast geneetilised variandid rs13070421¹³ ja rs648238¹⁴, millel on seos monotsüütide kemoatraktandi valguga-1 (*monocyte chemoattractant protein-1* – MCP-1). Lisaks on eelnevates uuringutes leitud 3p24.3 lookusest marker rs1395103, millel on leitud seos psühhiaatriliste haigustega, sealhulgas skisofreenia ja toitumishäiretega (Wang *et al.*, 2010). *SGOL1* ja pseudogeeni *VENTXP7* intergeensest alast on leitud geneetilised variandid rs4292253¹⁵, rs1523351¹⁶ ja rs10440098¹⁷, millel on assotsiatsioonid südame isheemiatõvega. Samuti on intergeensest regioonist leitud marker rs1403651, millel on seos madala tihedusega lipoproteiinide kolesterooliga (Kathiresan *et al.*, 2007). Naise reproduktiivtervise kontekstis ei ole *SGOL1* ja *VENTXP7* geene varem kirjeldatud. Sellest tulenevalt on vaja edasisi uuringuid, et selgitada, kas ja kuidas need võiksid olla seotud menstruaaltsükli pikkusega. Magistritöös leitud menstruaaltsükli pikkusega seotud tulemused ei ühti seni ainsas publitseeritud menstruaaltsükli pikkuse ülegenoomse assotsiatsiooniuuringus leitud 11. kromosoomis geeni *FSHB* piirkonnaga, milles paikneva geneetilise variandi seos ei osutunud magistritöö analüüsidest oluliseks (p-väärtus > 0,05). Peamine põhjus selleks on uuritava tunnuse väike valim antud magistritöös, võrreldes Ruth jt uuringuga, millesse oli kaasatud üle 9000 naise andmed (Ruth *et al.*, 2016).

Menstruaaltsükli regulaarsusega leiti ülegenoomselt olulised seosed 7. kromosoomi piirkonnas, kus paiknevad geenid *FAM221A* ja *STK31*. *FAM221A* geenil on kõrge ekspressioon testistes ning seda ekspresseeritakse ka naiste reproduktiivtraktis¹⁸. *STK31* geenil on tugev ekspressiooni rikastus testistes¹⁹. *STK31* avaldab mõju mitmete kasvajate korral, omades bioloogilist funktsiooni raku migratsiooni suurenemisel kui ka rakkude apoptoosi mahasurumisel (Kuo *et al.*, 2014). Yin jt uuringu kohaselt omab *STK31* potentsiaalset rolli emakaelavähi prognoosil, diagnoosil kui ka ravil (Yin *et al.*, 2016). Lisaks on varasemalt assotsiatsiooniuuringus leitud geenide *FAM221A* ja *STK31* intergeensest alast geneetilise variandi rs6974500 ja rauataseme vaheline seos (Tanaka *et al.*, 2010).

¹³ <http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=13070421#pgGAP>

¹⁴ <http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=648238#pgGAP>

¹⁵ <http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=4292253#pgGAP>

¹⁶ <http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=1523351#pgGAP>

¹⁷ <http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=10440098#pgGAP>

¹⁸ <http://www.proteinatlas.org/ENSG00000188732-FAM221A/tissue>

¹⁹ <http://www.proteinatlas.org/ENSG00000196335-STK31/tissue>

Raseduste esinemisega leiti olulisemad seosed *DNER* geeni piirkonnast. *DNER* on ehituslikult sarnane Notch signaaliraja retseptoriga ja selle ligandi Deltaga ning seda ekspresseeritakse nii arenevas kui ka küpses kesknärvisüsteemis (Eiraku *et al.*, 2002), neerupealsetes ning teataval määral ka munajuhas ja endomeetriumis²⁰. Varasemates assotsiatsiooniuuringutes on leitud *DNER* geenis oleva markeri rs10197959 seos autoimmuunhaigustega (Lauc *et al.*, 2013). Ühtlasi on varasemalt seostatud *DNER* geenis paiknevaid geneetilisi variante II tüüpi diabeediga (Deng *et al.*, 2015; Hanson *et al.*, 2014). See võib tuleneda *DNER* funktsioonist adipogeneesis. Park jt leidsid, et *DNER* inhibeerimise tulemusena suurenes adipotsüütide küpsemise määr inimese adipotsüütide mesenhümaalsetes tüvirakkudes (*human adipose tissue-derived mesenchymal stem cells – hAMSC*)(Park *et al.*, 2010). Eelnevalt on *DNER* geeni intronis paikneval variandil rs10519389 ning atriaalse natriureetilise faktori, mis on südamelihaskude poolt sekreteeritav hormoon, vahel leitud assotsiatsioone (Benjamin *et al.*, 2007).

Vanusega esimese raseduse ajal ei leitud ülegenoomselt olulisi seoseid, kuid olulisim leitud seos oli 5q15 lookuses *ERAPI* geeni intronis paikneva geneetilise variandiga. *ERAPI* valg lõikab peptiide lühemaks, et need oleksid optimaalse pikkusega peamise koosobivuskompleksi klass I molekulidega seondumiseks (Saric *et al.*, 2002). Varasemates uuringutes on leitud esmasünnitamise vanusega seotud geneetilisi variante *ESR1* ja *RBM6-SEMA3F* geenidest või nende lähedalt, mis antud magistritöö raames olulisust ei saavutanud: *ESR1* kõige olulisem geneetiline variant puudus esimese raseduse vanuse analüüsist ja *RBM6-SEMA3F* geneetiline variant oli p-väärtusega $> 0,05$ esimese raseduse vanuse uuringus. Peamine põhjus tulemuste erinevuseks tuleneb valimi suurusest, Day jt uuring hõlmas enam kui 60 000 naise andmeid (Day *et al.*, 2016).

Vanus esimese raseduse ajal oli negatiivses geneetilises korrelatsioonis kehamassiindeksiga. Seos ei tulene menarhe vanuse ning kehamassiindeksi geneetilisest korrelatsioonist, kuivõrd menarhe vanusel ja vanusel esimese raseduse ajal puudub jagatud geneetika. Seni on esmasünnitamise vanusega seotud geneetilised variandid leitud ka östrogeeni retseptori *ESR1* geenist või selle lähedalt (Day *et al.*, 2016). Goulart jt uuringus leiti, et *ESR1* geenis paiknevate polümorfismide tulemusena on naistel vähenenud risk ülekaalulisusele (Goulart *et al.*, 2009).

²⁰ <http://www.proteinatlas.org/ENSG00000187957-DNER/tissue>

Vanus esimese raseduse ajal oli positiivses geneetilises korrelatsioonis HDL kolesterooli tasemega, kuid negatiivses geneetilises korrelatsioonis LDL kolesterooli tasemega. Arvatavalt on vanus esimese raseduse ajal ja HDL kolesterooli taseme seos menarhest sõltumatu. Kuigi menarhe aja ja HDL kolesterooli taseme vahel oli samuti positiivne geneetiline korrelatsioon, siis menarhe ajal ning vanusel esimese raseduse ajal puudus omavaheline geneetiline korrelatsioon. Varast esmasünnitamist ning suurt laste arvu on epidemioloogilistes uuringutes seostatud kardiovaskulaarhaigustega, mis võib tuleneda raseduse järgsest väikesest HDL kolesterooli taseme langusest (Lewis *et al.*, 1996). Sünnitusjärgsele perioodile omane HDL kolesterooli taseme langus püsib mitmeid aastaid ning see ei tulene kaalutõusust, suurenenud tsentraalsest rasvumisest ega käitumuslikest muutustest (Gunderson *et al.*, 2004). Arvatavalt on rasedusele omaste lipiidide tasemete muutuste aluseks ka suguhormoonide tasemete muutused. LDL kolesterooli tase on negatiivses seoses progesterooni kontsentratsiooniga ning HDL kolesterooli tase on positiivses seoses nii östradioli kui ka progesterooni kontsentratsioonidega (Knopp *et al.*, 1986). Seosed ei võimalda küll tuvastada põhjuseid ja tagajärgi. LDL kolesterooli ja progesterooni seos võib tuleneda platsenta suurenenud LDL kolesterooli kasutusest progesterooni sekretsiooni vältel, progesterooni sekretsiooni suurenedes LDL kolesterooli kontsentratsioonid vähenevad (Knopp *et al.*, 1998). Emapoolne kolesterool, LDL kolesterooli kujul, on platsentas peamiseks substraadiks progesterooni biosünteesil (Simpson & MacDonald, 1981; Winkel *et al.*, 1980).

Vanus esimese raseduse ajal oli positiivses geneetilises korrelatsioonis sünnikaaluga. Barkeri hüpoteesi kohaselt loote kahjustunud toitumise tagajärjel sünnib madala sünnikaaluga imik, kes omab edasises elus sellest tulenevalt suurenenud riski erinevatele haigusseisunditele (Barker, 1995). Senised uuringud pole käsitletud, kuidas naiste sünnikaal võib mõjutada edasises elus reproduktiivset edukust.

Raseduste arvu ülegenoomses assotsiatsiooniuringus oli olulisim seos geneetilise variandiga, mis paikneb 17. kromosoomis geenide *SLC16A11* ja *CLEC10A* vahelises piirkonnas. *SLC16A11* geenil on rikastatud ekspressioon munajuhas²¹. Varasemates uuringutes on *SLC16A11* geenil paiknevaid polümorfisme seostatud II tüüpi diabeedi suurenenud riskiga (Lara-Riegos *et al.*, 2015; SIGMA Type 2 Diabetes Consortium *et al.*, 2014). *CLEC10A* geenil on rikastatud ekspressioon testistes²². Lisaks leiti seoseid 3. kromosoomis *PDZRN3* geeni intronis paiknevate geneetiliste variantidega. *PDZRN3* geenil on kõrge ekspressioon

²¹ <http://www.proteinatlas.org/ENSG00000174326-SLC16A11/tissue>

²² <http://www.proteinatlas.org/ENSG00000132514-CLEC10A/tissue>

munsarjades kui ka endomeetriumis ning kesknärvisüsteemis²³. *PDZRN3* geen on oluline vaskulaarses morfogeneesis kui ka adipotsüütide, osteoblastide ja müoblastide diferentseerumise reguleerimisel (Honda *et al.*, 2013; Sewduth *et al.*, 2014). Antud magistritöö raseduste arvu uuringu olulisemad leiud ei ühti Bangladeshis naiste uuringu tulemustega, mille kohaselt raseduste ja laste arvuga võimalikud seotud geneetilised variandid paiknevad 4., 5. ja 6. kromosoomi piirkondades (Aschebrook-Kilfoy *et al.*, 2015). Tulemuste mitte kattumise peamiseks põhjuseks võib olla kahes uuringus osalenud naiste erinev etniline kuuluvus. Lisaks ei ühtinud magistritöö uuringu tulemused Day jt laste arvu uuringuga, mille korral olulisemad piirkonnad olid 3. kromosoomis *CADM2* geenis ja 6. kromosoomis *ESR1* geenis (Day *et al.*, 2016). Day jt uuringu *ESR1* geeni kõige olulisem marker oli magistritöö raseduste arvu uuringus p-väärtusega 0,01. Kahe uuringu tulemuste erinevuse peamiseks põhjuseks on vahe valimite suuruses.

Kuivõrd rasedus on mitme osapoolega sündmus (ema, isa, laps/platsenta), siis sellest tulenevalt on biopanga andmetel põhineva antud uuringu korral võimalik keskenduda ainult ema andmetele. Sellest tulenevalt ei saa raseduskatkemise põhjusena välistada isa haigusi, loote haigusseisundeid, infektsioone ning teisi põhjuseid, mis ema haigusloos ei kajastu. Antud magistritöös leiti ülegenoomselt olulised raseduskatkemistega seotud geneetilised variandid 4. kromosoomis geeni *AGXT2L1* lähedalt, mis osaleb fosfolipiidide metabolismis (Schiroli *et al.*, 2013). Ühtlasi on *AGXT2L1* tugevalt ekspresseeritud suurajukoos kui ka hipokampuses²⁴. Levy jt uuringus leiti antud piirkonnast geneetiline variant rs2575655 geenis *OSTC*, mis on seotud vererõhuga (Levy *et al.*, 2007). Lookuses 4q25 paikneb geneetiline marker rs10029851, mida on seostatud amüotroofse lateraalskleroosiga Ida-Aasia haniinlaste populatsioonis (Xie *et al.*, 2014). Lisaks on piirkonnast varasemalt naistel leitud geneetilise markeri rs4956211 ja süsteemse erütematoosse luupuse vaheline seos (Harley *et al.*, 2008). Seni on kirjeldatud ema genoomist tulenevalt raseduskatkemise põhjusena loote aneuploidiat põhjustavaid geneetilisi variante 4q28 lookuses *PLK4* geenis (McCoy *et al.*, 2015), mis antud magistritöös olulisust ei saavutanud.

Kuigi antud magistritöös ei leitud naiste viljakust peegeldavate tunnuste omavahelisi geneetilisi korrelatsioone, on siiski alust arvata, et mingisugune geneetiline ühisosa esineb läbi teiste tunnuste. Day jt uuringus leiti, et menarhe aeg on positiivses geneetilises korrelatsioonis vanusega esimese seksuaalvahekorra ajal. Samuti oli vanus esimese

²³ <http://www.proteinatlas.org/ENSG00000121440-PDZRN3/tissue>

²⁴ <http://www.proteinatlas.org/ENSG00000164089-ETNPPL/tissue>

seksuaalvahekorra ajal positiivses geneetilises korrelatsioonis esmasünnitamise vanusega. Lisaks leiti, et vanus esimese seksuaalvahekorra ajal on negatiivses geneetilises korrelatsioonis laste arvuga. Leitud korrelatsioonide aluseks on riskikäitumisega seotud geneetilised lookused, millest osalt tuleneb ka vanus esimese seksuaalvahekorra ajal (Day *et al.*, 2016).

Antud töö tulemuste analüüs näitab, et magistritöös leitud naiste viljakusega seotud geneetilised variandid paiknevad geenides, millele on omane kõrge ekspressioon testistes või naiste reproduktiivtraktis. Lisaks olid mitmed geenid seotud adipotsüütide morfogeneesiga. Mitmetel leitud geenidel on kõrge ekspressioon kesknärvisüsteemis, millest tulenevalt on alust arvata, et geneetilised variandid võivad mõjutada käitumist üldisemalt ja sealhulgas ka reproduktiivkäitumist. Antud magistritöös leitud geneetilised variandid ei ühtinud seniste naiste reproduktiivtervist käsitlevate ülegenoomsete uuringute tulemustega, mille põhjuseks on võib olla antud uuringu raames käsitletud tunnuste väiksed valimid. Tulenevalt väiksest valimi suurusest õnnestus leida päritavuse hinnangud ainult kolmele tunnusele, mis antud magistritöö tulemuste kohaselt jagatud geneetikat ei oma. Naiste viljakust kirjeldavate tunnustega on mitmetel epidemioloogilisel tasandil seotud tunnustel jagatud geneetika, kõige olulisemad seosed olid kehamassiindeksiga, ning sellest tulenevalt võib eeldada, et kahel tunnusel on siduv bioloogiline komponent. Tulemuste kinnitamiseks viiakse antud tunnustega läbi analüüsid suuremate valimitega.

Kokkuvõte

Naiste normaalse reproduktiivtervise kontekstis läbiviidud uuringud on keskendunud eelkõige kahele sündmusele, mis piiritlevad reproduktiivset eluiga – menarhe ja menopaus. Sellest tulenevalt on vajalik leida, mis mõjutab naiste viljakust kogu reproduktiivperioodi vältel.

Käesoleva töö raames teostati 9702 naise genotüpiseerimisandmetel, mida oli täiendatud 1000 Genoomi Projekti referentsil põhineva imputatsiooniga, ülegenoomsed assotsiatsiooniuuringud, leidmaks geneetilisi variante, mis on seotud erinevate naiste viljakust peegeldavate tunnustega. Ülegenoomselt olulised geneetilised variandid leiti neljale tunnusele – menstruaaltsükli pikkus, menstruaaltsükli regulaarsus, raseduste esinemine ja raseduskatkemiste esinemine. Kõikidele leitud ülegenoomselt olulistele geneetilistele variantidele, sealhulgas ka olulisuse piiril olevatele, on ühisjooneks paiknemine geenides, millel on kõrge ekspressioon testistes ja reproduktiivtraktis (*SGOL1*, *FAM221A*, *STK31*, *DNER*, *SLC16A11*, *PDZRN3*). Ühtlasi on mitmed geenid kõrgelt ekspresseeritud kesknärvisüsteemis (*CNTN6*, *DNER*, *PDZRN3*, *AGXT2L1*), mistõttu need võiksid mõjutada käitumist üldisemalt ning sealhulgas ka reproduktiivkäitumist.

Naiste viljakusega seotud erinevate tunnuste jagatud geneetika leidmiseks kasutati LD skoori regressiooni meetodit, mis esmalt hindab tunnuse päritavust ja seejärel leiab kahe tunnuse geneetilise korrelatsiooni. Tulenevalt väikesest valimist, õnnestus päritavuse hinnangud leida kolmele tunnustele – vanus menarhe ajal, vanus esimese raseduse ajal ja raseduskatkemiste esinemine. Tunnustele ei leitud omavahelist geneetilist korrelatsiooni, mistõttu võib eeldada, et neil geneetiline ühisosa puudub või see tuleneb läbi tunnuste, mida antud magistritöös ei vaadeldud.

Varasemalt on naiste viljakusega seotud tunnused epidemioloogilisel tasandil seostatud erinevate tunnustega, nendest seast kõige sagedamini kehamassiindeksiga, kuid on ebaselge, kas seosed on ka geneetilisel tasandil ning omavad ühist bioloogilist komponenti. Kõige olulisemad geneetilised korrelatsioonid olid menarhe vanuse ja kehamassiindeksi vahel ning esimese raseduse vanuse ning kehamassiindeksi vahel. Naiste reproduktiivtervise kontekstis on oluline rasvumise määr, mitte keharasvade paiknemine. Naiste reproduktiivtervise ja kehamassiindeksi jagatud bioloogilise komponendi võimalikud kandidaadid on leptiin ja insuliin.

Genetic correlations of traits associated with female fertility

Anette Kalnapenkis

Summary

Female reproductive lifespan is bounded by two events – menarche and menopause. Previous studies on female reproductive health have mainly focused on pathologic conditions during female reproductive lifespan. In the context of normal reproductive health, studies have mostly only focused on genetics of menarche and menopause, while both of them are transitional periods. Reproductive behavior is mostly influenced by social factors, most important of them being personal reproductive choices, if and when to start a family. But it is influenced by genetics and biological processes that determine length and success of reproductive lifespan.

During this study genome-wide association analyses were conducted using genotyping data of 9702 women and data based on 1000 Genome Project reference imputation in order to find genetic variants associated with different traits depicting female fertility. Genetic variants for four traits – menstrual cycle length, menstrual cycle regularity, ever been pregnant and ever had a miscarriage. Genetic variants, even those that did not reach genome-wide significance, were in genes that are highly expressed in testis and female reproductive tract (*FAM221A*, *STK31*, *DNER*, *SLC16A11*, *PDZRN3*). Several genes are also highly expressed in central nervous system (*CNTN6*, *DNER*, *PDZRN3*, *AGXT2L1*). Thus, they could influence overall behavior and including reproductive behavior.

In order to find shared genetic background between different traits depicting female fertility LD score regression method was used. Firstly, LD score regression estimates heritability of traits and then finds genetic correlation between two trait of interest. Estimates of heritability were found for three traits – age at menarche, age at first pregnancy and ever had a miscarriage. No genetic correlations were found between traits depicting female fertility. Thus, it is possible that they don't have shared genetic background or it comes through traits that were not included in this study.

Previously traits associated with female fertility have been correlated with different traits on epidemiological level, most often with body mass index. But it is unclear if correlations exist on genetic level. Most significant genetic correlations were found between age at menarche and body mass index also between age at first pregnancy and body mass index. In the context of female reproductive health the extent of body fat is more important than distribution of

body fat. Candidates for shared biological component between female reproductive health and body mass index could be leptin and insulin.

Tänuõnad

Kõige enam sooviksin tänada oma juhendajaid, Reedik Mägi ja Triin Laisk-Podarit, kes võimaldasid nii huvitaval teemal magistritööd kirjutada, olles seejuures alati abivalmid ja toetavad.

Lisaks tänaksin oma perekonda ja lähedasi aktiivse kaasaelamise ning toetuse eest töö valmimisel.

Kasutatud kirjanduse loetelu

- Adair, L. S. (2001). Size at birth predicts age at menarche. *Pediatrics*, *107*(4), E59.
- Anderson, S. E., Dallal, G. E., & Must, A. (2003). Relative weight and race influence average age at menarche: results from two nationally representative surveys of US girls studied 25 years apart. *Pediatrics*, *111*(4 Pt 1), 844–50.
- Aschebrook-Kilfoy, B., Argos, M., Pierce, B. L., Tong, L., Jasmine, F., Roy, S., ... Ahsan, H. (2015). Genome-Wide Association Study of Parity in Bangladeshi Women. *Plos One*, *10*(3), e0118488.
- Baker, E. R. (1985). Body weight and the initiation of puberty. *Clinical Obstetrics and Gynecology*, *28*(3), 573–9.
- Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ (Clinical Research Ed.)*, *311*(6998), 171–4.
- Benjamin, E. J., Dupuis, J., Larson, M. G., Lunetta, K. L., Booth, S. L., Govindaraju, D. R., ... Vasani, R. S. (2007). Genome-wide association with select biomarker traits in the Framingham Heart Study. *BMC Medical Genetics*, *8 Suppl 1*, S11.
- Benowitz, N. L. (1990). Clinical pharmacology of caffeine. *Annual Review of Medicine*, *41*, 277–88.
- van den Berg, S. M., & Boomsma, D. I. (2007). The Familial Clustering of Age at Menarche in Extended Twin Families. *Behavior Genetics*, *37*(5), 661–667.
- Bergeron C, Ferenczy A, S. G. (1988). Immunocytochemical study of progesterone receptors in the human endometrium during the menstrual cycle. *Lab Invest*, *59*, 862–869.
- Boivin, J., Bunting, L., Collins, J. A., & Nygren, K. G. (2007). International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Human Reproduction (Oxford, England)*, *22*(6), 1506–12.
- Bouyain, S., & Watkins, D. J. (2010). The protein tyrosine phosphatases PTPRZ and PTPRG bind to distinct members of the contactin family of neural recognition molecules. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(6), 2443–8.
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., ... Neale. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, *47*(11), 1236–41.
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Consortium, S. W. G. of the P. G., ... Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet, advance on*(3), 291–295.
- Burcelin, R., Thorens, B., Glauser, M., Gaillard, R. C., & Pralong, F. P. (2003). Gonadotropin-releasing hormone secretion from hypothalamic neurons: stimulation by insulin and potentiation by leptin. *Endocrinology*, *144*(10), 4484–91.
- Calderon-Dominguez, M., Gil, G., Medina, M. A., Pandak, W. M., & Rodríguez-Agudo, D. (2014). The StarD4 subfamily of steroidogenic acute regulatory-related lipid transfer (START) domain proteins: new players in cholesterol metabolism. *The International Journal*

of Biochemistry & Cell Biology, 49, 64–8.

Chiazze, L., Brayer, F. T., Macisco, J. J., Parker, M. P., Duffy, B. J., Geist, S. H., ... Haman, J. O. (1968). The Length and Variability of the Human Menstrual Cycle. *JAMA: The Journal of the American Medical Association*, 203(6), 377.

Christiansen, O. B., Mathiesen, O., Lauritsen, J. G., & Grunnet, N. (1990). Idiopathic recurrent spontaneous abortion. Evidence of a familial predisposition. *Acta Obstetrica et Gynecologica Scandinavica*, 69(7-8), 597–601.

Christiansen, O. B., Steffensen, R., Nielsen, H. S., & Varming, K. (2008). Multifactorial etiology of recurrent miscarriage and its scientific and clinical implications. *Gynecologic and Obstetric Investigation*, 66(4), 257–67.

Chumlea, W. C., Schubert, C. M., Roche, A. F., Kulin, H. E., Lee, P. A., Himes, J. H., & Sun, S. S. (2003). Age at menarche and racial comparisons in US girls. *Pediatrics*, 111(1), 110–3.

Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee, S. H., Ripke, S., Neale, B. M., Faraone, S. V., Purcell, S. M., ... International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, 45(9), 984–94.

Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., & Lander, E. S. (2001). High-resolution haplotype structure in the human genome. *Nature Genetics*, 29(2), 229–232.

Day, F. R., Helgason, H., Chasman, D. I., Rose, L. M., Loh, P.-R., Scott, R. A., ... Perry, J. R. B. (2016). Physical and neurobehavioral determinants of reproductive onset and success. *Nature Genetics*, (April). doi:10.1038/ng.3551

Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, 10(1), 5–6.

Deng, Z., Shen, J., Ye, J., Shu, Q., Zhao, J., Fang, M., & Zhang, T. (2015). Association between single nucleotide polymorphisms of delta/notch-like epidermal growth factor (EGF)-related receptor (DNER) and Delta-like 1 Ligand (DLL 1) with the risk of type 2 diabetes mellitus in a Chinese Han population. *Cell Biochemistry and Biophysics*, 71(1), 331–5.

Dossus, L., Kvaskoff, M., Bijon, A., Engel, P., Verdebout, J., Fervers, B., ... Mesrine, S. (2013). Latitude and ultraviolet radiation dose in the birthplace in relation to menarcheal age in a large cohort of French women. *International Journal of Epidemiology*, 42(2), 590–600.

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, 9(3), e1003348.

Durbin, R. M., Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., ... McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061–1073.

Eiraku, M., Hirata, Y., Takeshima, H., Hirano, T., & Kengaku, M. (2002). Delta/Notch-like Epidermal Growth Factor (EGF)-related Receptor, a Novel EGF-like Repeat-containing Protein Targeted to Dendrites of Developing and Adult Central Nervous System Neurons. *Journal of Biological Chemistry*, 277(28), 25400–25407.

Elks, C. E., Perry, J. R. B., Sulem, P., Chasman, D. I., Franceschini, N., He, C., ... Murray, A.

- (2010). Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nature Genetics*, 42(12), 1077–1085.
- Ellish, N. J., Saboda, K., O'Connor, J., Nasca, P. C., Stanek, E. J., & Boyle, C. (1996). A prospective study of early pregnancy loss. *Human Reproduction (Oxford, England)*, 11(2), 406–12.
- Felix, J. F., Bradfield, J. P., Monnereau, C., van der Valk, R. J. P., Stergiakouli, E., Chesni, A., ... Bone Mineral Density in Childhood Study BMDCS. (2016). Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Human Molecular Genetics*, 25(2), 389–403.
- Ferenczy A, B. C. (1991). Histology of the human endometrium: from birth to senescence. *Annals of the New York Academy of Sciences*, 622, 6–27.
- Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., ... Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, 47(11), 1228–35.
- Flug, D., Largo, R. H., & Prader, A. (1984). Menstrual patterns in adolescent Swiss girls: a longitudinal study. *Annals of Human Biology*, 11(6), 495–508.
- Friedman, J. M., & Halaas, J. L. (1998). Leptin and the regulation of body weight in mammals. *Nature*, 395(6704), 763–770.
- Frisch, R. E., & Revelle, R. (1971). Height and weight at menarche and a hypothesis of menarche. *Archives of Disease in Childhood*, 46(249), 695–701.
- Goulart, A. C., Zee, R. Y. L., & Rexrode, K. M. (2009). Estrogen receptor 1 gene polymorphisms and decreased risk of obesity in women. *Metabolism: Clinical and Experimental*, 58(6), 759–64.
- Gunderson, E. P., Lewis, C. E., Murtaugh, M. A., Quesenberry, C. P., Smith West, D., & Sidney, S. (2004). Long-term plasma lipid changes associated with a first birth: the Coronary Artery Risk Development in Young Adults study. *American Journal of Epidemiology*, 159(11), 1028–39.
- Halperin, E., & Stephan, D. a. (2009). SNP imputation in association studies. *Nature Biotechnology*, 27(4), 349–351.
- Hanson, R. L., Muller, Y. L., Kobes, S., Guo, T., Bian, L., Ossowski, V., ... Baier, L. J. (2014). A genome-wide association study in American Indians implicates DNER as a susceptibility locus for type 2 diabetes. *Diabetes*, 63(1), 369–76.
- Harley, J., & Alarcón-Riquelme, M. (2008). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nature Genetics*, 40(2), 204–210.
- Harlow, S. D. (2000). 9 – Menstruation and Menstrual Disorders: The Epidemiology of Menstruation and Menstrual Dysfunction. In *Women and Health* (pp. 99–113).
- Harlow, S. D., & Ephross, S.A. (1995). Epidemiology of menstruation and its relevance to women's health. *Epidemiologic Reviews*, 17(2), 265–286.

- He, C., Kraft, P., Chen, C., Buring, J. E., Paré, G., Hankinson, S. E., ... Chasman, D. I. (2009). Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nature Genetics*, *41*(6), 724–728.
- He, C., Zhang, C., Hunter, D. J., Hankinson, S. E., Buck Louis, G. M., Hediger, M. L., & Hu, F. B. (2010). Age at menarche and risk of type 2 diabetes: results from 2 large prospective cohort studies. *American Journal of Epidemiology*, *171*(3), 334–44.
- Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., ... Lindgren, C. M. (2010). Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature Genetics*, *42*(11), 949–60.
- Honda, T., Ishii, A., & Inui, M. (2013). Regulation of adipocyte differentiation of 3T3-L1 cells by PDZRN3. *American Journal of Physiology. Cell Physiology*, *304*(11), C1091–7.
- Horikoshi, M., Yaghoobkar, H., Mook-Kanamori, D. O., Sovio, U., Taal, H. R., Hennig, B. J., ... Early Growth Genetics (EGG) Consortium. (2013). New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nature Genetics*, *45*(1), 76–82.
- Hosokawa, M., Imazeki, S., Mizunuma, H., Kubota, T., & Hayashi, K. (2012). Secular trends in age at menarche and time to establish regular menstrual cycling in Japanese women born between 1930 and 1985. *BMC Women's Health*, *12*(1), 19.
- Howie, B. N., Donnelly, P., & Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genetics*, *5*(6), e1000529.
- Hu, J., Liao, J., Sathanoori, M., Kochmar, S., Sebastian, J., Yatsenko, S. A., ... Schachner, M. (2015). CNTN6 copy number variations in 14 patients: a possible candidate gene for neurodevelopmental and neuropsychiatric disorders. *Journal of Neurodevelopmental Disorders*, *7*(1), 26.
- Kaltiala-Heino, R., Kosunen, E., & Rimpelä, M. (2003). Pubertal timing, sexual behaviour and self-reported depression in middle adolescence. *Journal of Adolescence*, *26*(5), 531–45.
- Kaltiala-Heino, R., Rimpelä, M., Rissanen, A., & Rantanen, P. (2001). Early puberty and early sexual activity are associated with bulimic-type eating pathology in middle adolescence. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *28*(4), 346–52.
- Kaminuma, O., Kitamura, F., Kitamura, N., Hiroi, T., Miyoshi, H., Miyawaki, A., & Miyatake, S. (2008). Differential Contribution of NFATc2 and NFATc1 to TNF- Gene Expression in T Cells. *The Journal of Immunology*, *180*(1), 319–326.
- Kathiresan, S., Manning, A. K., Demissie, S., D'Agostino, R. B., Surti, A., Guiducci, C., ... Cupples, L. A. (2007). A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Medical Genetics*, *8 Suppl 1*, S17.
- Knopp, R. H., Bonet, B., & Zhu, X. (1998). Lipid Metabolism in Pregnancy. In *Principles of Perinatal—Neonatal Metabolism* (pp. 221–258). New York, NY: Springer New York.
- Knopp, R. H., Warth, M. R., Charles, D., Childs, M., Li, J. R., Mabuchi, H., & Van Allen, M.

- I. (1986). Lipoprotein metabolism in pregnancy, fat transport to the fetus, and the effects of diabetes. *Biology of the Neonate*, 50(6), 297–317.
- Kohler, H.-P., Rodgers, J. L., & Christensen, K. (1999). Is Fertility Behavior in Our Genes? Findings from a Danish Twin Study. *Population and Development Review*, 25(2), 253–288.
- Kuo, P.-L., Huang, Y.-L., Hsieh, C. C.-J., Lee, J.-C., Lin, B.-W., Hung, L.-Y., ... Ullrich, A. (2014). STK31 Is a Cell-Cycle Regulated Protein That Contributes to the Tumorigenicity of Epithelial Cancer Cells. *PLoS ONE*, 9(3), e93303.
- Lamb, N. E., Yu, K., Shaffer, J., Feingold, E., & Sherman, S. L. (2005). Association between maternal age and meiotic recombination for trisomy 21. *American Journal of Human Genetics*, 76(1), 91–9.
- Lara-Riegos, J. C., Ortiz-López, M. G., Peña-Espinoza, B. I., Montúfar-Robles, I., Peña-Rico, M. A., Sánchez-Pozos, K., ... Menjivar, M. (2015). Diabetes susceptibility in Mayas: Evidence for the involvement of polymorphisms in HHEX, HNF4 α , KCNJ11, PPAR γ , CDKN2A/2B, SLC30A8, CDC123/CAMK1D, TCF7L2, ABCA1 and SLC16A11 genes. *Gene*, 565(1), 68–75.
- Lauc, G., Huffman, J. E., Pučić, M., Zgaga, L., Adamczyk, B., Mužinić, A., ... Rudan, I. (2013). Loci Associated with N-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. *PLoS Genetics*, 9(1).
- Leitsalu, L., Haller, T., Esko, T., Tammesoo, M. L., Alavere, H., Snieder, H., ... Metspalu, A. (2014). Cohort profile: Estonian biobank of the Estonian genome center, university of Tartu. *International Journal of Epidemiology*, 44(4), 1137–1147.
- Leridon, H. (2006). Demographic effects of the introduction of steroid contraception in developed countries. *Human Reproduction Update*, 12(5), 603–16.
- Levy, D., Larson, M. G., Benjamin, E. J., Newton-Cheh, C., Wang, T. J., Hwang, S.-J., ... Mitchell, G. F. (2007). Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. *BMC Medical Genetics*, 8 Suppl 1, S3.
- Lewis, C. E., Funkhouser, E., Raczynski, J. M., Sidney, S., Bild, D. E., & Howard, B. V. (1996). Adverse effect of pregnancy on high density lipoprotein (HDL) cholesterol in young adult women. The CARDIA Study. Coronary Artery Risk Development in Young Adults. *Am J Epidemiol*, 144(3), 247–254.
- Li, C. I., Littman, A. J., & White, E. (2007). Relationship between age maximum height is attained, age at menarche, and age at first full-term birth and breast cancer risk. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 16(10), 2144–9.
- Lin, D., Sugawara, T., Strauss, J. F., Clark, B. J., Stocco, D. M., Saenger, P., ... Miller, W. L. (1995). Role of steroidogenic acute regulatory protein in adrenal and gonadal steroidogenesis. *Science (New York, N.Y.)*, 267(5205), 1828–31.
- Lippus, H., Laanpere, M., Part, K., Ringmets, I., Rahu, M., Haldre K., Allvee, K., Karro, H. (2015). Eesti Naiste Tervis 2014: seksuaal- ja reproduktiivtervis, tervisekäitumine, hoiakud ja tervishoiuteenuste kasutamine. Uurimisaruanne. Tartu: Tartu Ülikooli Naistekliinik.

- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., ... Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, *518*(7538), 197–206.
- Lunetta, K. L., Day, F. R., Sulem, P., Ruth, K. S., Tung, J. Y., Hinds, D. A., ... Perry, J. R. B. (2015). Rare coding variants and X-linked loci associated with age at menarche. *Nature Communications*, *6*, 7756.
- Marchini, J., Howie, B., Myers, S., McVean, G., & Donnelly, P. (2007a). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genetics*, *39*(7), 906–13.
- Marks, J. L., Hair, C. S., Klock, S. C., Ginsburg, B. E., & Pomerleau, C. S. (1994). Effects of menstrual phase on intake of nicotine, caffeine, and alcohol and nonprescribed drugs in women with late luteal phase dysphoric disorder. *Journal of Substance Abuse*, *6*(2), 235–243.
- Matsumoto, S., Mogami, Y., Ohkuri, S. (1962). Statistical studies on menstruation; a criticism on the definition of normal menstruation. *Gunma J Med Sci*, *11*, 294–318.
- McCoy, R. C., Demko, Z., Ryan, A., Banjevic, M., Hill, M., Sigurjonsson, S., ... Petrov, D. (2015). Common variants spanning PLK4 are associated with mitotic-origin aneuploidy in human embryos. *Science*, *348*(6231), 235–238.
- Menken, J., Trussell, J., & Larsen, U. (1986). Age and infertility. *Science (New York, N.Y.)*, *233*(4771), 1389–94.
- Mills, M., Rindfuss, R. R., McDonald, P., te Velde, E., & ESHRE Reproduction and Society Task Force. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human Reproduction Update*, *17*(6), 848–60.
- Morabia, A., & Costanza, M. C. (1998). International variability in ages at menarche, first livebirth, and menopause. World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. *American Journal of Epidemiology*, *148*(12), 1195–205.
- Morris, D. H., Jones, M. E., Schoemaker, M. J., Ashworth, A., & Swerdlow, A. J. (2010). Determinants of age at menarche in the UK: analyses from the Breakthrough Generations Study. *British Journal of Cancer*, *103*(11), 1760–4.
- Mägi, R., & Morris, A. P. (2010). GWAMA: Software for genome-wide association meta-analysis. *BMC Bioinformatics*, *11*(ii), 288.
- Münster, K., Schmidt, L., & Helm, P. (1992). Length and variation in the menstrual cycle--a cross-sectional study from a Danish county. *British Journal of Obstetrics and Gynaecology*, *99*(5), 422–9.
- le Nestour, E., Marraoui, J., Lahlou, N., Roger, M., de Ziegler, D., Bouchard., P. (1993). Role of estradiol in the rise in follicle-stimulating hormone levels during the luteal-follicular transition. *The Journal of Clinical Endocrinology & Metabolism*, *77*(2), 439–442.
- Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., ... Gakidou, E. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet (London, England)*, *384*(9945), 766–81.

- Nybo Andersen, A. M., Wohlfahrt, J., Christens, P., Olsen, J., & Melbye, M. (2000). Maternal age and fetal loss: population based register linkage study. *BMJ (Clinical Research Ed.)*, *320*(7251), 1708–12.
- Osuch, J. R., Karmaus, W., Hoekman, P., Mudd, L., Zhang, J., Haan, P., & Mikucki, D. (2010). Association of age at menarche with adult leg length and trunk height: Speculations in relation to breast cancer risk. *Annals of Human Biology*, *37*(1), 76–85.
- Park, J.-R., Jung, J.-W., Seo, M.-S., Kang, S.-K., Lee, Y.-S., & Kang, K.-S. (2010). DNER modulates adipogenesis of human adipose tissue-derived mesenchymal stem cells via regulation of cell proliferation. *Cell Proliferation*, *43*(1), 19–28.
- Part, K., Laanpere, M., Rahu, K., Haldre, K., Rahu, M., Karro, H. (2007). Eesti Naiste Tervis: seksuaal- ja reproduktiivtervis, tervisekäitumine, hoiakud ja tervishoiuteenuste kasutamine. Uurimisaruanne. Tartu.
- Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology*, *32*(4), 381–5.
- Pearson, T., & Manolio, T. (2008). How to Interpret a Genome-wide Association Study. *JAMA*, *299*(11), 1335–44.
- Perry, J. R. B., Day, F., Elks, C. E., Sulem, P., Thompson, D. J., Ferreira, T., ... Ong, K. K. (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature*, *514*(7520), 92–7.
- Peters, S. A. E., Huxley, R. R., & Woodward, M. (2016). Women's reproductive health factors and body adiposity: findings from the UK Biobank. *International Journal of Obesity*, *40*(5), 803–808.
- Prentice, P., & Viner, R. M. (2013). Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. *International Journal of Obesity*, *37*(8), 1036–1043.
- Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., ... Frishman, D. (2011). LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics*, *27*(13), 2336–2337.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–75.
- Quennell, J. H., Mulligan, A. C., Tups, A., Liu, X., Phipps, S. J., Kemp, C. J., ... Anderson, G. M. (2009). Leptin Indirectly Regulates Gonadotropin-Releasing Hormone Neuronal Function. *Endocrinology*, *150*(6), 2805–2812.
- Rull, K., Nagirnaja, L., & Laan, M. (2012). Genetics of recurrent miscarriage: challenges, current knowledge, future directions. *Frontiers in Genetics*, *3*, 34.
- Ruth, K. S., Beaumont, R. N., Tyrrell, J., Jones, S. E., Tuke, M. A., Yaghoobkar, H., ... Murray, A. (2016). Genetic evidence that lower circulating FSH levels lengthen menstrual cycle, increase age at menopause and impact female reproductive health. *Human Reproduction (Oxford, England)*, *31*(2), 473–81.

- Saric, T., Chang, S.-C., Hattori, A., York, I. A., Markant, S., Rock, K. L., ... Goldberg, A. L. (2002). An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nature Immunology*, *3*(12), 1169–76.
- Schirotti, D., Cirrincione, S., Donini, S., & Peracchi, A. (2013). Strict reaction and substrate specificity of AGXT2L1, the human O-phosphoethanolamine phospho-lyase. *IUBMB Life*, *65*(7), 645–50.
- Sewduth, R. N., Jaspard-Vinassa, B., Peghaire, C., Guillabert, A., Franzl, N., Larrieu-Lahargue, F., ... Dupl a, C. (2014). The ubiquitin ligase PDZRN3 is required for vascular morphogenesis through Wnt/planar cell polarity signalling. *Nature Communications*, *5*, 4832.
- Shikone, T., Kokawa, K., Yamoto, M., & Nakano, R. (1997). Apoptosis of human ovary and uterine endometrium during the menstrual cycle. *Hormone Research*, *48 Suppl 3*, 27–34.
- SIGMA Type 2 Diabetes Consortium, Williams, A. L., Jacobs, S. B. R., Moreno-Mac as, H., Huerta-Chagoya, A., Churchhouse, C., ... Altshuler, D. (2014). Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature*, *506*(7486), 97–101.
- Simpson, E. R., & MacDonald, P. C. (1981). Endocrine physiology of the placenta. *Annual Review of Physiology*, *43*, 163–88.
- Snijder, M. B., Dekker, J. M., Visser, M., Yudkin, J. S., Stehouwer, C. D. A., Bouter, L. M., ... Seidell, J. C. (2003). Larger thigh and hip circumferences are associated with better glucose tolerance: the Hoorn study. *Obesity Research*, *11*(1), 104–11.
- Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., ... Kaakinen, M. (2010). Association analyses of 249,796 individuals reveal eighteen new loci associated with body mass index. *Nature Genetics*, *42*(11), 937–948.
- Stolk, L., Zhai, G., van Meurs, J. B. J., Verbiest, M. M. P. J., Visser, J. A., Estrada, K., ... Uitterlinden, A. G. (2009). Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nature Genetics*, *41*(6), 645–7.
- Stranger, B. E., Stahl, E. a., & Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, *187*(2), 367–83.
- Ziegler, D. De, Bergeron, C., Cornel, C., Medalie, D. A., Massai, M. R., Milgrom, E., ... Bouchard, P. (1992). Effects of Luteal Estradiol Transformation Gonadotropins. *The Journal of Clinical Endocrinology & Metabolism*, *74*(2), 322–331.
- Zinaman, M. J., Clegg, E. D., Brown, C. C., O'Connor, J., & Selevan, S. G. (1996). Estimates of human fertility and pregnancy loss. *Fertility and Sterility*, *65*(3), 503–9.
- Tanaka, T., Roy, C. N., Yao, W., Matteini, A., Semba, R. D., Arking, D., ... The, N. (2010). Brief report A genome-wide association analysis of serum iron concentrations. *Quest*, *115*(1), 94–96.
- Treloar, A. E., Boynton, R. E., Behn, B. G., & Brown, B. W. (1967). Variation of the human menstrual cycle through reproductive life. *Int J Fertil. Jan-Mar*, *12*(1 Pt 2): 77–126.
- Tropf, F. C., Stulp, G., Barban, N., Visscher, P. M., Yang, J., Snieder, H., & Mills, M. C. (2015). Human fertility, molecular genetics, and natural selection in modern societies. *PLoS*

One, 10(6), e0126821.

Tropf, F. C., Verweij, R. M., van der Most, P. J., Stulp, G., Bakshi, A., Briley, D. A., ... Mills, M. C. (2016). *Mega-analysis of 31,396 individuals from 6 countries uncovers strong gene-environment interaction for human fertility*. *bioRxiv*. Cold Spring Harbor Labs Journals. <http://dx.doi.org/10.1101/049163>

Tsafri A, Chun SY, Reich. R. (1993). Follicular rupture and ovulation. In *Adashi EY, Leung PCK Eds. The Ovary* (pp. 227–224). New York: Raven Press.

Venturoli, S. (1992). Longitudinal evaluation of the different gonadotropin pulsatile patterns in anovulatory cycles of young girls. *Journal of Clinical Endocrinology & Metabolism*, 74(4), 836–841.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five Years of GWAS Discovery. *The American Journal of Human Genetics*, 90(1), 7–24.

Vollman, R. F. (1977). The menstrual cycle. *Major Problems in Obstetrics and Gynecology*, 7, 1–193.

Voullaire, L., Slater, H., Williamson, R., & Wilton, L. (2000). Chromosome analysis of blastomeres from human embryos by using comparative genomic hybridization. *Human Genetics*, 106(2), 210–7.

Waller, K., Swan, S. H., Windham, G. C., Fenster, L., Elkin, E. P., & Lasley, B. L. (1998). Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. *American Journal of Epidemiology*, 147(11), 1071–80.

Wang, K. S., Liu, X. F., & Aragam, N. (2010). A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophrenia Research*, 124(1-3), 192–199.

Wang, Y., Dinse, G. E., & Rogan, W. J. (2012). Birth weight, early weight gain and pubertal maturation: a longitudinal study. *Pediatric Obesity*, 7(2), 101–9.

Weiss, G. (2001). Menstrual irregularities and the perimenopause. *Journal of the Society for Gynecologic Investigation*, 8(1 Suppl Proceedings), S65–6.

Wells, D., & Delhanty, J. D. (2000). Comprehensive chromosomal analysis of human preimplantation embryos using whole genome amplification and single cell comparative genomic hybridization. *Molecular Human Reproduction*, 6(11), 1055–62.

Welt, C. K., & Carmina, E. (2013). Clinical review: Lifecycle of polycystic ovary syndrome (PCOS): from in utero to menopause. *The Journal of Clinical Endocrinology and Metabolism*, 98(12), 4629–38.

Whincup, P. H., Gilg, J. A., Odoki, K., Taylor, S. J., & Cook, D. G. (2001). Age of menarche in contemporary British teenagers: survey of girls born between 1982 and 1986. *BMJ (Clinical Research Ed.)*, 322(7294), 1095–6.

WHO - World Health Organization. (1986). World Health Organization Multicenter Study on Menstrual and Ovulatory Patterns in Adolescent Girls. *J Adolesc Health Care*, (1), 229–235.

Wilcox, A. J., Dunson, D., & Baird, D. D. (2000). The timing of the "fertile window" in the

menstrual cycle: day specific estimates from a prospective study. *BMJ (Clinical Research Ed.)*, 321(7271), 1259–62.

Willer, C. J., Schmidt, E. M., Sengupta, S., Peloso, G. M., Gustafsson, S., Kanoni, S., ... Abecasis, G. R. (2013). Discovery and refinement of loci associated with lipid levels. *Nature Genetics*, 45(11), 1274–83.

Winkel, C. A., Snyder, J. M., MacDonald, P. C., & Simpson, E. R. (1980). Regulation of cholesterol and progesterone synthesis in human placental cells in culture by serum lipoproteins. *Endocrinology*, 106(4), 1054–60.

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., ... Frayling, T. M. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46(11), 1173–86.

Xie, T., Deng, L., Mei, P., Zhou, Y., Wang, B., Zhang, J., ... Xu, R. (2014). A genome-wide association study combining pathway analysis for typical sporadic amyotrophic lateral sclerosis in Chinese Han populations. *Neurobiology of Aging*, 35(7), 1778.e9–1778.e23.

Yamagishi, Y., Sakuno, T., Shimura, M., & Watanabe, Y. (2008). Heterochromatin links to centromeric protection by recruiting shugoshin. *Nature*, 455(7210), 251–5.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88(1), 76–82.

Yin, F.-F., Wang, N., Bi, X.-N., Yu, X., Xu, X.-H., Wang, Y.-L., ... Pappu, S. (2016). Serine/threonine kinases 31(STK31) may be a novel cellular target gene for the HPV16 oncogene E7 with potential as a DNA hypomethylation biomarker in cervical cancer. *Virology Journal*, 13(1), 60.

Ylikorkala, O., Kauppila, A. (2008). II Hormonaalsed häired ja muud haigused p. 62, 79. *Sünnitusabi ja günekoloogia*. Tallinn: AS Medicina.

Kasutatud veebiaadressid

<http://www.who.int/end-childhood-obesity/facts/en/>

<http://www.britannica.com/science/menstrual-cycle> (Encyclopædia Britannica, 2013)

<http://www.cdc.gov/nchs/data/databriefs/db232.htm>

https://www.oecd.org/els/soc/SF_2_3_Age_mothers_childbirth.pdf

<https://www.stat.ee/311832> (Ene Margit Tiit, ettekanne „Rahvastikustatistika. Sünnimus ja iive“, loeng matemaatikaõpetajatele. 3. detsember 2015)

<http://rhk.sm.ee/>

https://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium

<http://www.kathiresanlab.org/collaborators/global-lipids-genetic-consortium/>

<http://www.illumina.com>

<http://www.hpc.ut.ee>

<https://github.com/bulik/ldsc/>

<https://www.r-project.org/>

<http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=13070421#pgGAP>

<http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=648238#pgGAP>

<http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=4292253#pgGAP>

<http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=1523351#pgGAP>

<http://www.ncbi.nlm.nih.gov/gap/phegeni?tab=2&rs=10440098#pgGAP>

<http://www.proteinatlas.org/ENSG00000188732-FAM221A/tissue>

<http://www.proteinatlas.org/ENSG00000196335-STK31/tissue>

<http://www.proteinatlas.org/ENSG00000187957-DNER/tissue>

<http://www.proteinatlas.org/ENSG00000174326-SLC16A11/tissue>

<http://www.proteinatlas.org/ENSG00000132514-CLEC10A/tissue>

<http://www.proteinatlas.org/ENSG00000121440-PDZRN3/tissue>

<http://www.proteinatlas.org/ENSG00000164089-ETNPPL/tissue>

Lisa 1. Küsimustik

Kui vanalt algasid Teil menstruatsioonid.....

Kui pikk oli Teie keskmine menstruaaltsükkel vanuses 25–35

ei tea

pole menstruaaltsükleid

ebaregulaarsed

≤ 20 päeva

21–24 päeva

25–29 päeva

30–35 päeva

>35 päeva

Kui kiiresti peale esimest menstruatsiooni muutusid menstruaaltsüklid regulaarseteks

kohe või kuni kuue kuu jooksul

aasta möödudes

2 a

3 a

4 a

5 või enam aastaga

peale esimest rasedust

kogu aeg ebaregulaarsed

Kas Teil esineb praegu menstruatsioone.....

Kui jah, siis mitu menstruaaltsüklit on Teil olnud viimase 12 kuu jooksul

1–3

4–5

6–9

10 või rohkem

ei tea

Kui ei, siis mis vanuses..... ja mis põhjusel menstruatsioonid lõppesid

menopaus

olen praegu rase

muu

Muud põhjused

menopaus

olen praegu rase

rasestumisvastased vahendid

günekoloogiline operatsioon

ravimid

madal kehakaal

tippspordi tegemine

kaasuv haigus

lapse rinnaga toitmine

Mitu rasedust Teil on olnud.....

Mitmel juhul on tegemist olnud

elussünniga.....
iseenesliku raseduse katkemisega.....
tehisliku raseduse katkemisega.....
emakavälise rasedusega.....
surnultsünniga.....

Kui vana Te olite oma esimese raseduse ajal.....

Kas Teil on olnud kunstlikku viljastamist.....

Kui jah, siis mitu.....

Mitu neist on lõppenud elussünniga.....

Kas Te olete kunagi tarvitanud hormonaalseid rasestumisvastaseid vahendeid.....

Mis selle vahendi nimi on.....

Kui vanalt Te nende tarvitamist alustasite.....

Kas nende kasutamisel on olnud pause

ei
kuni 6 kuud
6–12 kuud
1–2 aastat
üle 2 aasta
ei tea

Kui Te praegu neid vahendeid ei kasuta, siis kui vanalt Te lõpetasite nende kasutamise.....

Küsimused menopausi kohta

Kas Te olete kasutanud hormoonravimeid menopausi tõttu.....

Kui vanalt Te neid tarvitama hakkasite.....

Kas nende kasutamisel on olnud pause.....

ei
kuni 6 kuud
6–12 kuud
1–2 aastat
üle 2 aasta
ei tea

Kas Te kasutate selliseid vahendeid ka praegu.....

Mis selle vahendi nimi on.....

Kui Te praegu neid vahendeid ei kasuta, siis kui vanalt Te lõpetasite nende kasutamise.....

Lisa 2. Raseduskatkemiste välistusdiagnoosid

C73 (kilpnäärme pahaloomuline kasvaja), C74 (neerupealise pahaloomuline kasvaja), C75.1 (ajuripatsi pahaloomuline kasvaja), D34 (kilpnäärme healoomuline kasvaja), D35.0 (neerupealise healoomuline kasvaja), D35.2 (ajuripatsi healoomuline kasvaja), D68.5 (primaarne trombofiilia), D68.6 (muud trombofiilia haigusseisundid, sealhulgas antifosfolipiidsündroom), E02 (subkliiniline joodivaegushüpotüreoos), E03 (muu hüpotüreoos), E05 (tür[e]otoksikoos [hüpertüreoos]), E06 (tür[e]oidiit e kilpnäärmepõletik), E07 (kilpnäärme muud haigusseisundid), E22 (hüpofüüsi hüpofunktsioon e vaegtalitus e liigtalitus), E23 (hüpofüüsi hüpofunktsioon e vaegtalitus ja muud väärtalitlused), E24 (Cushingi sündroom), E25 (adrenogenitaalsed häired), E26 (hüperaldosteronism), E27 (neerupealiste muud haigusseisundid), E28.2 polütsüstilise munasarja sündroom, M32 (süsteemne erütematoosluupus e -söötraig), M35.0 (kuiv sündroom [Sjögren]), Q50-Q56 (suguelundite kaasasündinud väärendid, sealhulgas Q50 munasarjade ja -juhade ning emakalaisideme kaasasündinud väärendid, Q51 emaka ja emakakaela kaasasündinud väärendid, Q52 naissuguelundite muud kaasasündinud väärendid, Q56 määratlemata sugu ja pseudohermafroditism), Q90-Q99 mujal klassifitseerimata kromosoomianomaaliad (sealhulgas Q90 Downi sündroom, Q91 Edwardsi sündroom ja Patau sündroom, Q92 autosoomide mujal klassifitseerimata muud trisoomiad ja osalised trisoomiad, Q93 mujal klassifitseerimata monosoomiad ja autosoomide deletsioon, Q95 mujal klassifitseerimata tasakaalustatud rekombinatsioonid ja struktuurimarkerid, Q96 Turneri sündroom, Q97 mujal klassifitseerimata naisfenotüübi muud sugukromosoomide anomaaliad, Q98 mujal klassifitseerimata meesfenotüübi muud sugukromosoomide anomaaliad, Q99 mujal klassifitseerimata muud kromosoomianomaaliad).

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, **Anette Kalnapenkis**

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „**Naiste viljakusega seotud tunnuste geneetilised korrelatsioonid**“, mille juhendajateks on **Reedik Mägi** ja **Triin Laisk-Podar**,
 - 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 kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 27. mail 2016