

TARTU ÜLIKOOL

LOODUS- JA TÄPPISTEADUSTE VALDKOND

MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT

ja

MEDITSIIINITEADUSTE VALDKOND

BIO- JA SIIRDEMEDITSIINI INSTITUUT

Sekundaarsed geneetilised leiud viljatutel meestel ja nende kliiniline olulisus

Magistritöö

30 EAP

Kristiina Lillepea

Juhendajad:

Laura Kasak, PhD

inimese geneetika lektor, BSMI, TÜ

Maris Laan, PhD

inimese geneetika professor, BSMI, TÜ

TARTU 2021

INFOLEHT

„Sekundaarsed geneetilised leiud viljatutel meestel ja nende kliiniline olulisus“

Viljatuse põhjuseid ei suudeta leida kuni 20% azoospermiaga meestest. Idiopaatilise viljatuse diagnoosimiseks on üks kulutõhusamaid meetodeid eksoomi sekveneerimine, mille raames on patsientidel võimalik uurida ka sekundaarseid leide. Käesoleva töö peamiseks eesmärgiks oli tuvastada sekundaarsed leiud 85 geenis idiopaatilise mitte-obstruktiivse azoospermiaga meeste valimis. Lähtudes hüpoteesist, et viljatusel ja kasvajate arengul võib olla ühine geneetiline etioloogia, uuriti lisaks onkogeensete variantide rikastatust antud valimil. Töö käigus tuvastati 29 sekundaarset leidu 3,5% patsientidest, mis on proportsionaalne eelnevate teadustöödega. Onkogeensete variantide rikastatust ei leitud, kuid analüüsitud kasvajate juhtgeenide piiratud arvu ja väikse valimi tõttu võis seos jääda tuvastamata.

Märksõnad: sekundaarsed leiud, mitte-obstruktiivne azoospermia

CERCS klassifikaatorid: B220 geneetika, tsütogeneetika; B570 sünnitusabi, günekoloogia, androloogia, paljunemine, seksuaalsus; B790 kliiniline geneetika

“Secondary findings in infertile men and their clinical significance”

Up to 20% of azoospermia cases go undiagnosed. One of the most cost-effective methods to determine the genetic cause of idiopathic infertility is exome sequencing. With sequencing data it is also possible to determine other clinically important variants – secondary findings. The main aim of this thesis was to identify secondary findings in men with idiopathic non-obstructive azoospermia. In addition, based on the hypothesis that infertility and cancer share the same genetic etiology, enrichment of oncogenic variants was explored. 29 secondary findings were identified in 3.5% of the cohort, which is proportional to previously published works. No enrichment of oncogenic variants was found, however, because of the small sample size and limited number of oncogenes analyzed, the association might not have been identifiable.

Keywords: secondary findings, non-obstructive azoospermia

CERCS classification: B220 genetics, cytogenetics; B570 obstetrics, gynaecology, andrology, reproduction, sexuality; B790 clinical genetics

SISUKORD

INFOLEHT	2
SISUKORD.....	3
KASUTATUD LÜHENDID	5
SISSEJUHATUS.....	7
1 KIRJANDUSE ÜLEVAADE	8
1.1 Spermatogenees	8
1.2 Mehepoolne viljatus	9
1.2.1 Definiitsioon ja levimus	9
1.2.2 Azoospermia	11
1.3 Viljatus ja muud terviseriskid.....	13
1.4 Sekundaarsed leiud.....	16
1.4.1 Rahvusvahelised juhised sekundaarsete leidude tuvastamiseks.....	16
1.4.2 Sekundaarsete leidude esinemissagedus üldpopulatsioonis	22
1.4.3 Tagasiside andmine sekundaarsete leidude kohta	23
1.5 Geenivariantide patogeensuse hindamine.....	25
2 UURIMUS.....	28
2.1 Töö eesmärgid ja ülesehitus	28
2.2 Materjal ja metoodika	29
2.2.1 Valimi kirjeldus	29
2.2.2 Sekvenerimisandmed	31
2.2.3 Analüüsitavate geenide valik.....	31
2.2.4 Analüüsiks kasutatud andmebaasid ja tarkvara.....	31
2.2.5 Variantide annoteerimine ja prioritseerimine	33
2.2.6 Geenivariantide patogeensuse hindamine	33

2.2.7	Eesti patsientide patogeensete variantide valideerimine.....	34
2.2.8	Eksoonsete harvade variantide koormuse testimine.....	35
2.3	Tulemused.....	36
2.3.1	Analüüsitud geenide profiil	36
2.3.2	Geenivariantide filtreerimine	37
2.3.3	Variantide jagunemine geeniti	38
2.3.4	Variantide jagunemine valimi patsientide seas.....	40
2.3.5	Sekundaarsed leiud valimi patsientidel.....	40
2.3.6	Eksoonsete harvade variantide koormus	46
2.4	Arutelu	46
2.4.1	Eksoonne varieeruvus analüüsitud sekundaarsete leidudega seotud 85 geenis....	46
2.4.2	Sekundaarsete leidude profiil võrreldes teadustööde andmetega	48
2.4.3	Tuvastatud sekundaarsete leidude potentsiaalne seos viljatusega.....	52
2.4.4	Sekundaarsete leidude potentsiaalne lisandväärtus viljatute meeste käsitluses...	53
2.4.5	Eksoonsete harvade variantide rikastatus	54
	KOKKUVÕTE.....	56
	SUMMARY	58
	TÄNUSÕNAD.....	60
	KASUTATUD KIRJANDUS.....	61
	KASUTATUD VEEBIAADRESSID.....	70
	Lihtlitsents lõputöö elektrooniliseks avaldamiseks.....	71

KASUTATUD LÜHENDID

ACMG	Ameerika Meditsiinigeneetika ja Genoomika Kolleegium, <i>American College of Medical Genetics and Genomics</i>
ACMG SF v1.0	ACMG sekundaarsete leidude nimekiri, 1. versioon (56 geeni)
ACMG SF v2.0	ACMG sekundaarsete leidude nimekiri, 2. versioon (59 geeni)
ACMG SF v3.0	ACMG sekundaarsete leidude nimekiri, 3. versioon (73 geeni)
ANNOVAR	<i>ANNOtate VARIation</i>
CCI	Charlsoni komorbiidsusindeks, <i>Charlson comorbidity index</i>
CCMG	Kanada Meditsiinigeneetikute Kolleegium, <i>Canadian College of Medical Geneticists</i>
CDC	Haiguste Kontrolli ja Tõrje Keskus, <i>Centers of Disease Control and Prevention</i>
CBAVD	kaasasündinud bilateraalne seemnejuha puudumine, <i>congenital bilateral absence of the vas deferens</i>
ClinGen	<i>Clinical Genome Resource</i>
ESHG	Euroopa Inimesegeneetika Ühing, <i>European Society of Human Genetics</i>
FSH	folliikuleid stimuleeriv hormoon
GEMINI	<i>GEnetics of Male INFertility Initiative</i>
GHS	Geisingeri Tervishoiusüsteem, <i>Geisinger Health System</i>
GHS76	GHS sekundaarsete leidude nimekiri (76 geeni)
gnomAD	<i>Genome Aggregation Database</i>
HL	healoomuline
KMI	kehamassiindeks

KVS	kardiovaskulaarsüsteem
LH	luteiniseeriv hormoon
LoF	funktsioonikaoga, <i>loss-of-function</i>
MAF	minoorse alleeli sagedus, <i>minor allele frequency</i>
NFE	eurooplased, millest on välja arvatud soomlased, <i>non-Finnish Europeans</i>
NOA	mitte-obstruktiivne azoospermia, <i>non-obstructive azoospermia</i>
OA	obstruktiivne azoospermia
P	patogeenne
PCOS	polütsüstiline munasarjade sündroom, <i>polycystic ovary syndrome</i>
SHBG	suguhormoone siduv globuliin, <i>sex-hormone binding globulin</i>
SNV	ühenukleotiidne variatsioon, <i>single nucleotide variant</i>
TF	tüstiline fibroos
THL	tõenäoliselt healoomuline
TP	tõenäoliselt patogeenne
TRAPD	<i>Test Rare vAriants with Public Data</i>
VUS	teadmata patogeensusega variant, <i>variant of unknown significance</i>

SISSEJUHATUS

Umbes 15% reproduktiivses eas paaridest ei õnnestu ühe aasta jooksul regulaarse kaitsmata suguelu käigus rasestuda. Nendest juhtudest ligikaudu 50% on ainu- või kaaspõhjuseks mehepoolne viljatus, millest kõige raskem viljatuse vorm on azoospermia, s.t mehe ejakulaadis ei esine sperme (Jungwirth *et al.*, 2018). Tihti on selle põhjuseks häired spermatogeneesis ning see on defineeritud kui mitte-obstruktiivne azoospermia (NOA) (World Health Organization, 2010).

NOA juhtudest umbes veerand on diagnoositud teadaolevate geneetiliste põhjustega (Punab *et al.*, 2017). Kuna aga spermatogeneesis osaleb suur hulk geene, võib viljatuse geneetilisi põhjuseid olla palju ning seetõttu on NOA puhul tegu väga heterogeense haigusega, mille geneetiliste põhjuste välja selgitamiseks on parimad eksoomi või genoomi sekveneerimise uuringud.

Sekveneerimise käigus on võimalik lisaks patsiendi primaarse haigusega seotud geenivariantidele tuvastada ka muid kliiniliselt olulisi ja potentsiaalselt sekkumistõhusaid variante ehk sekundaarseid leide. Seetõttu on mitmed organisatsioonid, nagu ACMG (*American College of Medical Genetics and Genomics*) (Kalia *et al.*, 2017), GHS (*Geisinger Health System*) (Schwartz *et al.*, 2018) ja ClinGen (*Clinical Genome Resource*) (Webber *et al.*, 2018) välja töötanud nimekirjad geenidest, milles leiduvatest haigusseoselistest variantidest tuleks patsiente teavitada. Selleks, et patsientidele raporteerida vaid patogeenseid variante, on ACMG poolt välja töötatud ka juhised geenivariantide patogeensuse hindamiseks (Richards *et al.*, 2015).

Kuna eksoomi sekveneerimisandmete analüüs on kõige kulutõhusam meetod idiopaatilise viljatuse geneetiliste põhjuste väljaselgitamiseks, on NOA patsientidel võimalik uurida ka sekundaarseid leide, eriti arvestades, et varasemalt on viljatutel meestel tuvastatud ka kõrgem kaasuvate haiguste esinemissagedus (Del Giudice *et al.*, 2020b; Hanson *et al.*, 2016).

Käesoleva magistritöö eesmärgks oli ACMG, GHS ja ClinGen geeninimekirjade põhjal tuvastada sekundaarsed leiud idiopaatilise NOA-ga meestel ning võrrelda nende esinemissagedust üldpopulatsiooniga. Lisaks püstitati hüpotees, et raske spermatogeneesi häiretega patsientidel esineb kasvajatega seotud geenides varieeruvuste rikastatus võrreldes üldpopulatsiooniga.

Käesolev töö on teostatud Tartu Ülikooli bio- ja siirdemeditiini instituudis inimese geneetika uurimisgrupis.

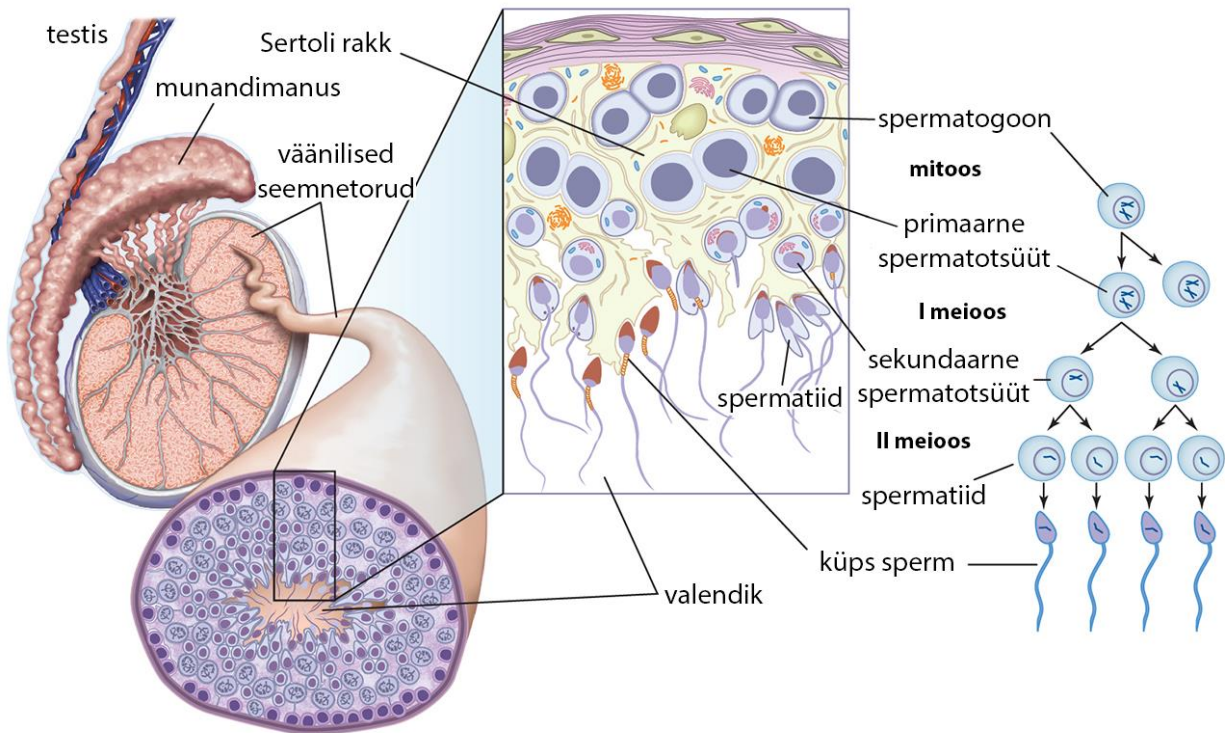
1 KIRJANDUSE ÜLEVAADE

1.1 Spermatogeneees

Inimorganism saab alguse küpse munaraku viljastumisel spermiga, mil mõlemalt vanemalt pärinev geneetiline materjal fuseerub ja moodustub diploidne rakk (Monroy, 2020). Haploidne sperm ise aga moodustub spermatogeneesi käigus, mille eesmärk on tagada igapäevane diferentseerunud spermide tootmine (Sutovsky ja Manandhar, 2006).

Spermatogeneees leiab aset testistes paiknevates väänilistes seemnetorudes ja kogu protsessi toetavad Sertoli rakud, varustades arenevaid sperme toitainetega (joonis 1). Spermatogeneees saab alguse diferentseerumata tüvirakkudest ehk spermatogoonidest, mis jagunevad asümmeetriliselt, uuendades tüvirakkude populatsiooni ning andes aluse primaarsetele spermatotsüütidele. Need rakud läbivad spermatogeneesi esimese meiootilise jagunemise ja moodustavad haploidsed sekundaarsed spermatotsüüdid. Sekundaarsed spermatotsüüdid läbivad teise meiootilise jagunemise ning sellega antakse algus spermatiididele. Spermatiidid läbivad spermiogeneesi, mille käigus moodustuvad küpsed spermid. Spermatogeneesi lõppfaasis transporditakse küpsed spermid läbi pika väänilise seemnetoru testise munandimanusesse, kus nad paiknevad kuni ejakulatsioonini (The Editors of Encyclopaedia Britannica, 2020). Kõik spermatogeneesi faasid toimuvad samaaegselt, tagades seeläbi pideva spermide tootmise (Sutovsky ja Manandhar, 2006).

Spermatogeneees on reguleeritud läbi hüpotalamuse-hüpopüüsi-sugunäärme hormoonide telje. Hüpotalamus toodab ja sekreteerib gonadotropiini vabastavat hormooni (GnRH, *gonadotropin-releasing hormone*), mis stimuleerib hüpopüüsi eesagarat vabastama folliikuleid stimuleerivat hormooni (FSH) ja luteiniseerivat hormooni (LH). FSH stimuleerib väänilistes seemnetorudes paiknevaid Sertoli rakke toetama spermatogeneesi. LH aga omakorda stimuleerib testistes olevaid Leydigi rakke tootma ja vabastama testosterooni (Choy ja Amory, 2020). Testosteroon on vajalik spermide produktsiooniks ja küpsemiseks ning sekundaarsete sugutunnuste arenguks ja funktsioneerimiseks (Oduwole *et al.*, 2018). Osa testosteroonist muudetakse hormoonideks östradiool ja inhibiin B, mis annavad hüpotalamusele ja hüpopüüsile negatiivset tagasisidet LH ja FSH sekretsiooni vähendamiseks (Choy ja Amory, 2020).



Joonis 1. Spermatogeneesis. Spermid saavad alguse spermatogoonidest, mis spermatogeneesi käigus läbivad mitmeid jagunemisi. Küpsed spermid transporditakse läbi vääniliste seemnetorude valendiku munandimanusesse. Kohandatud The Editors of Encyclopaedia Britannica (2020) järgi.

1.2 Mehepoolne viljatus

1.2.1 Definitsioon ja levimus

Viljatus on defineeritud kui haigus, mil paaril ei õnnestu ühe aasta jooksul regulaarse kaitsmata suguelu käigus rasestuda (Zegers-Hochschild *et al.*, 2009). Selle definitsiooni kohaselt on viljatud umbes 15% reproduktiivses eas paaridest, nendest 50% on ainu- või kaaspõhjuseks mehepoolne viljatus (Jungwirth *et al.*, 2018). Kesk- ja Ida-Euroopas on viljatuid mehi hinnanguliselt 8–12%, mujal Euroopas 7,5%, Austraalias 8–9% ning Põhja-Ameerikas 4,5–6% (Agarwal *et al.*, 2015).

Mehe viljakuse hindamiseks kasutatakse ühe meetodina seemnevedeliku analüüsi, milles hinnatakse mitmeid parameetreid. Maailma Terviseorganisatsioon (WHO, *World Health Organization*) on avaldanud standardiseeritud meetodi inimese seemnevedeliku hindamiseks ning antud juhiseid sisaldab ka normaalse viljakuse referentsväärtuseid (tabel 1) (World Health Organization, 2010).

Tabel 1. Normaalse seemnevedeliku referentsväärtused Maailma Terviseorganisatsiooni^a andmetel ja hormonaalse analüüsi referentsväärtused Tartu Ülikooli Kliinikumi^b andmetel.

Parameeter	Alampiirväärtus	Hormoon	Vanusegrupp	Referentsväärtus
Ejakulaadi maht	1,5 ml	FSH	≥18 a	1,5–12,4 U/L
Spermide kontsentratsioon	15 mln/ml	LH	≥18 a	1,7–8,6 U/L
Spermide koguarv	39 mln/ml	Testosteroon	19 a – 50 a	8,64–29,0 nmol/L
Liikuvus	40%		≥50 a	6,68–25,7 nmol/L
Progressiivne liikuvus	32%	Vaba	18 a – 50 a	>0,22 nmol/L
Normaalne morfoloogia	4%	testosteroon	≥50 a	>0,18 nmol/L

FSH – folliikuleid stimuleeriv hormoon. LH – luteiniseeriv hormoon.

^a World Health Organization (2010).

^b Tartu Ülikooli Kliinikum. Ühendlabor.

www.kliinikum.ee/yhendlabor/index.php?option=com_content&view=article&id=32&Itemid=44.

Kasutatud 25.05.2021.

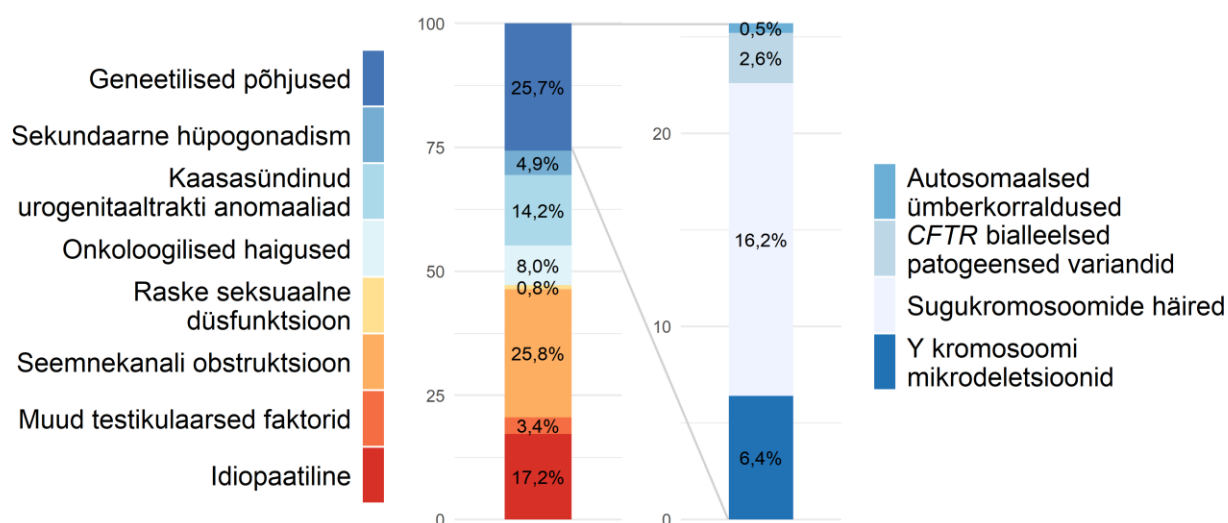
Lisaks seemnevedeliku analüüsile on viljatuse diagnoosimisel olulisel kohal ka hormoonide analüüs, nendest kõige tähtsamad on gonadotropiinid FSH ja LH ning testosteroon. Tihti esineb erinevate viljatuse fenotüüpide puhul kõrgenenud FSH ja LH tase koos või ilma madala testosterooni tasemega (Jungwirth *et al.*, 2018). Veel on viljatuse hindamisel olulised meditsiiniline ajalugu, füüsiline ülevaatus ja geneetiline testimine (Agarwal *et al.*, 2021).

Vastavalt ejakulaadis esinevale spermide arvule saab spermatogeneesi efektiivsuse jagada kolmeks: normozoospermia ehk normaalne spermide arv, oligozoospermia ehk spermide koguarv või kontsentratsioon on alla alampiirväärtuse ja azoospermia ehk spermide täielik puudumine ejakulaadis (World Health Organization, 2010).

1.2.2 Azoospermia

Kõige raskem mehe viljatuse vorm on azoospermia, s.t mehe seemnevedelikus ei esine sperme (World Health Organization, 2010). Ekstreemsele fenotübile vaatamata esineb azoospermia üldpopulatsioonis umbes ühel mehel 100-st (Practice Committee of the American Society for Reproductive Medicine in collaboration with the Society for Male Reproduction and Urology, 2018) ning viljakusprobleemidega meestest 10–20% (Olesen *et al.*, 2017; Punab *et al.*, 2017).

Azoospermia põhjuseid on mitmeid, sh geneetilised ja keskkonna faktorid, küll aga umbes 17% juhtudest ei suudeta põhjust leida, mispuhul on tegu idiopaatilise azoospermiaga. Eestis kogutud andmete põhjal on azoospermia põhjused kajastatud joonisel 2. Azoospermia jagatakse kaheks: obstruktiivne azoospermia (OA) ja mitte-obstruktiivne azoospermia (NOA, *non-obstructive azoospermia*).



Joonis 2. Azoospermia teadaolevate põhjuslike tegurite jaotus 388 Eesti meespatsiendi põhjal Punab *et al.* (2017) andmetel. Kõige sagedasemateks teadaolevateks põhjusteks on raske seksuaalne düsfunktsioon ja geneetilised faktorid, millest enim esineb sugukromosoomide häireid. Umbes 17% juhtudest jääb viljatuse põhjus välja selgitamata (idiopaatiline azoospermia).

1.2.2.1 *Obstruktiivne azoospermia*

OA puhul puuduvad seemnevedelikus spermid füüsilise obstruktsiooni tõttu ning see esineb pea 30% azoospermia haigetest (Punab *et al.*, 2017). Obstruktsioon võib olla kas kaasasündinud (nt bilateraalne seemnejuha puudumine) või omandatud (nt vasektoomia, infektsiooni või trauma tagajärjel) ning see võib asuda ühes või mitmes mehe reproduktiivsüsteemi osas, nt munandimanuses, seemne- või purskejuhas (Jungwirth *et al.*, 2018). Kaasasündinud obstruktsiooni kõige sagedasem põhjus on monogeenne muutus. Enamus juhtudest on põhjuseks bialleelsed patogeensed muutused geenis *CFTR*, mis sageli põhjustab kaasasündinud bilateraalset seemnejuha puudumist (CBAVD, *congenital bilateral absence of the vas deferens*), ning seetõttu on *CFTR* testimine sees ka kliinilistes juhistes (Bieth *et al.*, 2021; Jungwirth *et al.*, 2018). Üksikutel juhtumitel on tegu hiljuti kirjeldatud X-liitelise geeni *ADGRG2* muutustega (Pagin *et al.*, 2020).

CFTR bialleelsed haigusseoselised muutused põhjustavad lisaks ka lapseas algavat kroonilist kopsuhaigust ehk tsüstilist fibroosi (TF). CBAVD patsientidel, kel TF-i ei esine, võib siiski olla eelsoodumus hingamisteede nakkushaiguste raskekujulisele põdemisele ja kroonilisele pankreatiidile ehk kõhunäärme põletikule (Bombieri *et al.*, 2011). Lisaks on neil meestel kõrgem risk saada TF-i põdevaid lapsi, juhul kui ka nende partner on patogeense *CFTR* variandi kandja (Bieniek *et al.*, 2021).

1.2.2.2 *Mitte-obstruktiivne azoospermia*

NOA-ga on tegemist, kui spermide täielik puudumine seemnevedelikus on põhjustatud spermatogeneesi häiretest ning füüsilist obstruktsiooni ei esine. NOA on väga heterogeenne haigus, kuna sel on palju erinevaid põhjuseid, mil mõjutatud on erinevad protsessid nagu sugunäärmete diferentseerumine, hüpotalamuse-hüpofüüsi-sugunäärme telje funktsioon või spermatogenees (Tournaye *et al.*, 2017). Azoospermiaga meestest esineb NOA umbes 70% (Olesen *et al.*, 2017; Punab *et al.*, 2017).

NOA põhjused saab jagada pretestikulaarseteks ja testikulaarseteks. Pretestikulaarsete NOA põhjuste alla kuuluvad harva esinevad endokriinsüsteemi häired (nt Kallmanni sündroom) ja hüpofüüsiga seotud haigused, mispuhul on meestel FSH tasemed madalamad. Testikulaarsed NOA põhjused ehk primaarne testikulaarne puudulikkus võib olla kas kaasasündinud või elu

jooksul omandatud ning seda peegeldab FSH kõrge tase (Hwang *et al.*, 2018). Tüüpilised omandatud põhjused on varikotseele ehk seemneväädi veenilaiendid, orhiit ehk munandi põletik, keemiaravi ja radiatsioon ning trauma (Jungwirth *et al.*, 2018). Kaasasündinud põhjused on tavaliselt Y kromosoomi mikroleetsioonid (kuni 18% NOA juhtudest) ja kromosomaalsed häired, millest enim esineb 47,XXY karüotüüpi ehk Klinefelteri sündroomi (kuni 15% NOA juhtudest) (Krausz *et al.*, 2003; Punab *et al.*, 2017; Tüttelmann *et al.*, 2011). Kuigi NOA põhjuseid on mitmeid, jäävad kuni 20% azoospermia juhtudest idiopaatilisteks (Punab *et al.*, 2017; Tüttelmann *et al.*, 2011).

Lisaks kromosomaalsetele häiretele ja Y kromosoomi mikroleetsioonidele on NOA geneetilisteks põhjusteks ka monogeensed patogeensed variandid, mille tulemusel võib häiruda spermatogeneesi, sugunäärmete areng ja funktsioon või reproduktiivsüsteemi endokriinne regulatsioon. Mitmes sõltumatus uuringus kinnitatud NOA-ga seotud genee on tänaseks teada ainult paarikümne ringis ning nendest enamik põhjustavad spermatogeneesi maturatsiooni seiskumist (*FANCM*, *MEI1*, *MEIOB*, *STAG3*, *TEX11*, *TEX14*, *TEX15*, *NR5A1*, *SETX*, *WT1*) (Kasak ja Laan, 2021; Krausz *et al.*, 2020; Wyrwoll *et al.*, 2020). Nende geenide haigusseoselisi variante on aga kirjeldatud kogu maailmas vaid mõnesajal patsiendil.

Mehe genoomist on kuni 10% seotud reproduktiooniga (Kasman *et al.*, 2020), s.t üle 2000 geenis paiknevad geneetilised häired võivad mõjutada mehe viljakust. Seeläbi on NOA väga heterogeenne haigus, mille geneetiliste põhjuste välja selgitamiseks on kõige paremad eksoomi ja genoomi sekveneerimise uuringud. Seetõttu on NOA patsientidel võimalik uurida ka sekundaarseid leide, arvestades, et viljatutel meestel on tihti ka üldtervis kehvem kui viljakatel meestel.

1.3 Viljatus ja muud terviseriskid

Viljatutel meestel on hinnatud krooniliste ja muude haiguste esinemise sagedus kõrgemaks kui viljakatel meestel (Kasman *et al.*, 2020). Näiteks 38 a kestnud Taani uuringus leiti, et üleüldine hospitaliseerimise risk on kõrgem madalama sperma kontsentratsiooniga meestel ning kõrge sperma kontsentratsiooniga (195–200 mln/mL) mehed hospitaliseeritakse esmakordselt keskmiselt seitse aastat hiljem kui madala sperma kontsentratsiooniga (0–5 mln/mL) mehed (Latif

et al., 2017). Möödunud aastal ilmunud uuringus leiti, et azoospermiaga ja oligozoospermiaga meestel on kõrgeenenud surmarisk (riskitiheduste suhe (ingl *hazard ratio*) vastavalt 2,01 ja 1,17) (Del Giudice *et al.*, 2020a). Antud tulemuseni on jõudnud ka varasem töö, kus üle 380 000 meest hõlmanud uuringus leiti azoospermiaga meestel kõrgeenenud surmarisk (Glazer *et al.*, 2019). Madala spermide arvuga meestel on erinevad üldtervise näitajad kehvemad, nt kõrgem kehamassiindeks (KMI), vöö ümbermõõt, vererõhk, kõrgem LDL-kolesterool ja madalam HDL-kolesterool, lisaks on nende seas metaboolse sündroomi levimus kõrgem (Ferlin *et al.*, 2021).

Kaasuvate haiguste esinemist kirjeldatakse tihti läbi Charlsoni komorbiidsusindeksi (CCI, *Charlson comorbidity index*). Kõrgema CCI väärtusega meestel on kirjeldatud madalam sperma maht ja kontsentratsioon, spermide liikuvus ja koguarv ning testosterooni tase (Eisenberg *et al.*, 2015a; Ventimiglia *et al.*, 2015). Veel enam on CCI ≥ 1 meeste seas oligozoospermia ja NOA sagedus kõrgem (Ventimiglia *et al.*, 2015).

Mitmed uurimused on keskendunud vähktõve esinemissageduse ja levimuse uurimisele viljatutel meestel ning need andmed toetavad seost viljatuse ja erinevat tüüpi kasvajate vahel. 2015. a näitasid Eisenberg *et al.*, et viljatutel meestel on kõikide vähkide, eriti munandivähi ja mitte-Hodgkini lümfoomi risk kõrgem (Eisenberg *et al.*, 2015b). Kaks aastat varem leiti, et hinnanguliselt on viljatutel meestel 1,7 ja azoospermiaga meestel 2,9 korda tõenäolisem haigestuda vähki võrreldes üldpopulatsiooniga (Eisenberg *et al.*, 2013). Samas hilisemas Hanson *et al.* (2016) uuringus azoospermia korral kõrgeenenud vähiriski ei leitud, kuid sel puhul võis tegu olla väiksest valimist tuleneva artefaktiga. Küll aga leiti, et oligozoospermiaga meestel on munandivähki haigestumise risk üle kümne korra kõrgem. Lisaks leiti ka potentsiaalne seos sperma kvaliteedi ja melanoomi haigestumise vahel (Hanson *et al.*, 2016). Hilisemas Rootsi uuringus, mis hõlmas üle miljoni mehe, leiti, et tänu viljatusravile isaks saanud meestel oli statistiliselt kõrgem eesnäärmevähi risk võrreldes meestega, kes said isaks loomulikul viisil (riskitiheduste suhe 1,64). Veel enam haigestusid viljatusraviga rasestunud paaride mehed eesnäärmevähki varasemas eas (Al-Jebari *et al.*, 2019). Kasvajate ja mehepoolse viljatuse puhul on diskuteeritud ka võimaliku ühise geneetilise etioloogia üle ning tuvastatud on mitmeid kandidaatgeene, mis võiksid selgitada erinevates teadustöodes leitud ühist geneetilist riski (Nagirnaja *et al.*, 2018).

Peale onkoloogiliste haiguste on viljatust seostatud ka erinevate kardiovaskulaarsüsteemi (KVS) haigustega. 1737 Eesti viljakusprobleemidega meespatsientide seas on leitud kõrgem KVS haiguste ja ka muude krooniliste haiguste esinemise sagedus võrreldes viljakate meestega (Punab *et al.*, 2017). Seda tulemust kinnitavad ka muud uuringud, kus on leitud, et KVS haigustega meestel on kehvemad sperma parameetrid. Näiteks on KVS haigustega meestel spermide koguarv ja liikuvus madalam (Latif *et al.*, 2017), lisaks on hüperkolesteroleemiaga meestel kirjeldatud spermide morfoloogia muutuseid ja sperma mahu vähenemist (Schisterman *et al.*, 2014) ning hüpertensiooniga meestel üldist madalamat sperma kvaliteeti (Eisenberg *et al.*, 2015c). Vastavate haiguste sagedus viljatutel meestel on samuti kõrgem, nimelt esineb hüpertensiooni viljatute seas 10,7% võrra rohkem ning hüperlipideemiat, mille alla kuulub ka hüperkolesteroleemia, 3,2% võrra rohkem kui viljakatel meestel (Shiraishi ja Matsuyama, 2018). Veel enam on viljatute meeste puhul kirjeldatud ka kõrgemat riski haigestuda südame isheemiatõve (Eisenberg *et al.*, 2016). Madala sperma kontsentratsiooniga (<15 mln/mL) meestel on täheldatud kõrgemat KVS haiguste tõttu hospitaliseerimise sagedust võrreldes >40 mln/mL sperma kontsentratsiooniga meestega (Latif *et al.*, 2017).

Viljatutel meestel on kirjeldatud ka kõrgemat diabeeti haigestumise riski (Eisenberg *et al.*, 2016; Glazer *et al.*, 2017). Lisaks on leitud, et diabeeti põdevatel meestel on väiksem seemnevedeliku maht, vähenenud spermide liikuvus ja kõrgenenud FSH tase, samas kui LH ja testosterooni tasemed on korras (Pergialiotis *et al.*, 2016). Küll aga on leitud, et prediabeedi diagnoosiga meestel on madalamad kogu testosterooni ja suguhormoone siduva globuliini (SHBG, *sex-hormone binding globulin*) tasemed ning kõrgemad FSH ja östradiooli tasemed. Veel enam on prediabeediga meeste seas idiopaatilise NOA levimus kõrgem (Boeri *et al.*, 2019).

Mitmed avaldatud uuringud toetavad ülekaalu ja rasvumise ning infertiilsuse vahelist seost. 2013. a meta-analüüsis leiti, et ülekaaluliste ja rasvunud meeste seas on azoospermia ja oligozoospermia levimus kõrgem (Sermondade *et al.*, 2013). Kõrgema KMI-ga meestel on leitud madalam ejakulaadi maht, sperma kontsentratsioon, spermide arv ja liikuvus (Bieniek *et al.*, 2016; Ehala-Aleksejev ja Punab, 2015; Eisenberg *et al.*, 2014). Lisaks on kõrgema KMI-ga või metaboolse sündroomiga meestel täheldatud madalamad testosterooni, SHBG, inhibiin B ja anti-Mülleri hormooni tasemed ning kõrgem östradiooli tase (Bieniek *et al.*, 2016; Ehala-Aleksejev ja Punab, 2015; Ventimiglia *et al.*, 2016).

Kuna meeste sperma kontsentratsioon ja spermide koguarv on languses (Levine *et al.*, 2017), võib edasikestva allakäigu puhul mehepoolse viljatuse levimus suureneda. Sel puhul võib viljatus kui biomarker osutada tähtsaks meeste üldtervise kirjeldamisel.

1.4 Sekundaarsed leiud

1.4.1 Rahvusvahelised juhised sekundaarsete leidude tuvastamiseks

Eksoomi ja genoomi sekveneerimise kasutamine on meditsiinis tõusuteel. Peale spetsiifiliste leidude identifitseerimise võib sekveneerimisandmete analüüsil välja tulla ka muid kliiniliselt tähtsaid geneetilisi variante. Selliseid variante võib leida juhuslikult (juhuleiud) või sihilikult otsides (sekundaarsed leiud).

Juhu- ja sekundaarsed leiud on defineeritud kui geneetilised variandid, mis ei ole seotud testimise primaarse põhjusega, kuid mis on potentsiaalselt haigusseoselised ja sekkumistõhusad. 2013. a märtsis andis Ameerika Meditsiinigeneetika ja Genoomika Kolleegium (ACMG, *American College of Medical Genetics and Genomics*) välja algse nimekirja geenidest, milles leiduvatest patogeensetest ja tõenäoliselt patogeensetest geenivariantidest tuleks patsiente teavitada (Green *et al.*, 2013). Antud nimekiri sisaldas esialgu 56 geeni (ACMG SF v1.0), kuid hiljem on seda täiustatud ning nimekirja teine versioon sisaldab 59 geeni (ACMG SF v2.0) (Kalia *et al.*, 2017) ja kolmas versioon 73 geeni (ACMG SF v3.0) (Miller *et al.*, 2021a). Valitud on geenid, milles leiduvate patogeensete variantide tõttu võivad tekkida haigused, mis on kas ennetatavad, neile on olemas ravi või mille korral haiguse sümptomid võivad olla pikalt varjatud, kuid välise stiimuli mõjul võivad avalduda ja kiiresti progresseeruda (Green *et al.*, 2013). Seega on selliste haigusseoseliste sekundaarsete geneetiliste variantide identifitseerimine oluline, kuna tagasisidet andes on võimalik mõningaid hilisavalduvaid haigusi ennetavate meetmetega ära hoida, edasi lükata või leevendada.

ACMG alusel on oma geenide nimekirjad loonud ka Geisingeri Tervishoiusüsteem (GHS, *Geisinger Health System*) seoses projektiga MyCode (*MyCode Community Health Initiative*) ning ClinGen-i (*Clinical Genome Resource*) töögrupp. Geisingeri nimekirja kuulus 76 geeni (GHS76), millest 56 kattusid ACMG SF v2.0 nimekirja geenidega (Schwartz *et al.*, 2018). Küll aga on antud geeninimekiri hiljem üle vaadatud ning alles on jäetud vaid ACMG SF v2.0 nimekirja kuuluvad

geenid ning lisaks *HFE* geeni üks muutus (c.845G>A), mille homosügootseid leide tuleks raporteerida (Kelly *et al.*, 2021). ClinGen geeninimekirjas on kokku 22 geeni, millest kuus ei ole esindatud ACMG SF v2.0 ega GHS76 nimekirjades (Webber *et al.*, 2018).

ACMG SF v2.0 , GHS76 ja ClinGen nimekirjade peale on kokku 85 geeni, millest kolm on seotud autosoom-retsessiivse ja kolm X-liiteliste pärandumismustritega haigustega. Ülejäänud 79 geeni on autosoom-dominantsete pärandumisega haiguste geenid. Nimekirjadesse kuuluvad 85 geeni on võimalik jagada viide rühma (tabel 2):

1. Haiguste Kontrolli ja Tõrje Keskuse (CDC, *Centers of Disease Control and Prevention*) esimese taseme (ingl *tier 1*) haigused, milleks on pärilik rinna- ja munasarjavähk (*BRCA1* ja *BRCA2*), Lynch'i sündroom (*MLH1*, *MSH2*, *MSH6*, *PMS2*) ja perekondlik hüperkolesteroleemia (*APOB*, *LDLR*, *PCSK9*). Nende haigustega pered ei ole tihti teadlikud kõrgemast haiguse riskist, kuid varajane avastamine ja sekkumine vähendavad märgatavalt haigestumust ja suremust (CDC. Tier 1 Genomics Applications and their Importance to Public Health. Kasutatud 04.03.2021).
2. Kardiovaskulaarsete ja hematoloogiliste haigustega seotud geenid, nt *SCN5A* (Brugada sündroom, pika QT sündroom), *LMNA*, *MYBPC3*, *MYH7*, *TNNI3* (perekondlik hüpertroofiline/dilatatiivne kardiomiopaatia).
3. Kasvajatega seotud geenid, nt *RET* (endokriinse hulgikasvajate sündroom, perekondlik medullaarne kilpnäärme vähk), *WT1* (Wilmsi tuumor).
4. Endokrinoloogiliste haigustega seotud geenid, nt *HNF1A* (monogeenne diabeet).
5. Muud geenid, milles olevad patogeensed variandid võivad põhjustada adrenoleukodüstroofiat (*ABCD1*), dopamiinitundlikku düstooniast (*GCH1*), ornitiini transkarbamülaasi puudulikkust (*OTC*), maliigse hüpertermia eelsoodumust (*RYR1*) või Wilsoni haigust (*ATP7B*).

Nendest 26 geenist, mis olid esindatud vaid GHS76 ja/või ClinGen nimekirjas, on hiljutisse ACMG SF v3.0 lisatud neli geeni: *ACVRL1*, *BTD*, *ENG* ja *HNF1A* (Miller *et al.*, 2021a).

Tabel 2. ACMG SF v2.0¹, GHS76² ja ClinGen³ nimekirjade geenid ning nende vastavad fenotüübid.

Fenotüüp	Avaldumise iga	Geen	Pärandumise tüüp	Funktsioon	Allikas
CDC 1. taseme haigused					
Lynch'i sündroom	T	<i>MLH1</i>	AD	DNA reparatsioon	1, 2
		<i>MSH2</i>			
		<i>MSH6</i>			
		<i>PMS2</i>			
Perekondlik hüperkolesteroleemia	L/T	<i>APOB</i>	AD	Rasvade transport	1, 2, 3
		<i>LDLR</i>		Plasma kolesterooli taseme regulatsioon	
		<i>PCSK9</i>		Plasma kolesterooli taseme regulatsioon	
Pärilik rinna- ja munasarjavähk	T	<i>BRCA1</i>	AD	Tuumorsupressor, DNA reparatsioon	1, 2
		<i>BRCA2</i>		DNA reparatsioon	
Kardiovaskulaarsed ja hematoloogilised haigused					
Arütmogeenne parema vatsakese düsplaasia	L/T	<i>JUP</i>	AD	Tsütoskeleti struktuur, rakk-rakk adhesioon	2
Arütmogeenne parema vatsakese kardiomiopaatia	L/T	<i>DSC2</i>	AD	Desmosoomi komponent	1, 2
		<i>DSG2</i>		Desmosoomi komponent, transmembraanne kaltsiumit siduv glükoproteiin	
		<i>DSP</i>		Desmosoomi komponent	
		<i>PKP2</i>			
		<i>TMEM43</i>		Tuumamembraani struktuuri säilitamine	
Brugada sündroom	L/T	<i>CACNA1C</i>	AD	Osa kaltsiumi kanalist	2
		<i>CACNB2</i>		Katalüüs	
		<i>GPD1L</i>		Tsükliliste nukleotiidide poolt avatav kaaliumi kanal	
		<i>HCN4</i>		Osa kaltsiumi kanalist	
		<i>KCNE3</i>			
Brugada sündroom, kodade virvendusarütmia	L/T	<i>SCN1B</i> <i>SCN3B</i>	AD	Osa naatriumi kanalist	2

Ehlers-Danlosi sündroom (tüüp 4)	L/T	<i>COL3A1</i>	AD	Kortikaalse arengu regulatsioon pehmetes sidekudedes	1, 2
Katehoolaminergiline polümorfne ventrikulaarne tahhükardia	L/T	<i>RYR2</i>	AD	Lihaskontraktsioon	1, 2
Kodade virvendusarütmia	L/T	<i>KCNJ2</i>	AD	Aktsioonipotentsiaal	2
Loeys-Dietz'i sündroom	L	<i>SMAD3</i>	AD	Geenide aktiivsuse ja raku proliferatsiooni regulatsioon	1, 2, 3
		<i>TGFB2</i>		Angiogeneesi ja südame arengu regulatsioon	3
		<i>TGFB3</i>		Embrüogeneesi ja rakkude diferentseerumise regulatsioon	2, 3
		<i>TGFBR1</i>		Signaali transduktsioon	1, 2, 3
		<i>TGFBR2</i>			
Marfan'i sündroom, Loeys-Dietz'i sündroom, perekondlikud rinnaaordi aneurüsmid ja dissektsioonid	L/T	<i>FBN1</i>	AD	Kaltsiumit siduva mikrofibrilli struktuurne komponent	1, 2, 3
Perekondlik dilatatiivne kardiomüopaatia	L/T	<i>DES</i>	AD	Lihaste struktuur ja funktsioon	2
		<i>LMNA</i>		Tuuma stabiilsus, kromatiini struktuur ja geenide avaldumine	1, 2
Perekondlik hüpertroofiline kardiomüopaatia	L/T	<i>MYL2</i>	AD	Südamelihase kontraktsioon	1, 2
		<i>MYL3</i>			
		<i>PRKAG2</i>		Rakulise energia metabolismi regulatsioon	
Perekondlik hüpertroofiline kardiomüopaatia, perekondlik dilatatiivne kardiomüopaatia	L/T	<i>ACTC1</i>	AD	Lihaskudede kontraktiilsus	1, 2
		<i>MYBPC3</i>		Südamelihase kontraktsioon	
		<i>MYH7</i>		Müosiini koostisosa	
		<i>PLN</i>		Kaltsiumi regulatsioon südamelihases	2
		<i>TNNI3</i>		Troponiini subühik	1, 2
		<i>TNNT2</i>			
		<i>TPM1</i>		Kaltsium-sõltuva vöötlihase kontraktsiooni regulatsioon	
Pika QT sündroom	L/T	<i>KCNE1</i>	AD	Osa kaaliumi kanalist	2
		<i>KCNE2</i>			1, 2
		<i>KCNH2</i>			
		<i>KCNQ1</i>			
		<i>SNTA1</i>		Membranvalkude subtsellulaarse lokaliseerimise organiseerija	2

Pika QT sündroom, Brugada sündroom	L/T	SCN5A	AD	Osa naatriumi kanalist	1, 2
Pika QT sündroom, kodade virvendusarütmia	L/T	SCN4B	AD	Osa naatriumi kanalist	2
Pika QT sündroom, perekondlik hüpertroofiline kardiomiopaatia	L/T	CAV3	AD	Kaveooli komponent	2
Pärilik hemorraagiline teleangiiektaasia	L/T	ACVRL1	AD	TGF- β retseptor, vajalik angiogeneesis	2, 3
		ENG		TGF- β retseptorkompleksi komponent, angiogeneesi regulatsioon	2, 3
Pärilik rindkere aordi haigus	L/T	ACTA2	AD	Vaskulaarne kontraktiilsus ja vererõhu homöostaas	1, 2, 3
		MYH11		Müosiini koostisosa	
		MYLK		Südamelihase kontraktsioon	2, 3
Kasvajad					
Endokriinsete hulgikasvajate sündroom	L/T	MEN1	AD	Tuumorsupressor	1, 2
Endokriinsete hulgikasvajate sündroom, perekondlik medullaarne kilpnäärme vähk	L/T	RET	AD	Rakkude proliferatsioon, arenguline migratsioon ja diferentseerumine	1, 2, 3
Hüperparatüroidism-lõualuu kasvaja sündroom	T	CDC73	AD	Tuumorsupressor	3
Juveniilsete polüpooside sündroom	L/T	BMPR1A	AD	TGF- β signaalraja retseptor	1
		SMAD4		Tuumorsupressor, epiteelirakkude proliferatsiooni inhibitsioon	1, 2, 3
Li-Fraumeni sündroom	L/T	TP53	AD	Tuumorsupressor	1, 2
Neurofibromatoos	L/T	NF2	AD	Hippo/SWH signaalraja regulatsioon	1, 2
Perekondlik adenomatoosne polüpoos	L/T	APC	AD	Tuumorsupressor	1, 2
	T	MUTYH	AR	Oksüdatiivne DNA reparatsioon	
Peutz-Jeghers'i sündroom	L/T	STK11	AD	Tuumorsupressor	1, 3
PTEN hamartoomi tuumori sündroom	L/T	PTEN	AD	Tuumorsupressor	1, 2

Pärilik paraganglioom- feokromotsütoomi sündroom	L/T	<i>SDHAF2</i>	AD	Suksinaadi dehüdrogenaasi kokkupanek	1, 2
		<i>SDHB</i>		Elektronide liigutamine suksinaadilt ubikvinoonile	
		<i>SDHC</i>			
		<i>SDHD</i>			
Retinoblastoom	L	<i>RB1</i>	AD	Tuumorsupressor	1, 2
Tuberoosne skleroos	L	<i>TSC1</i>	AD	Tuumorsupressor	1, 2
		<i>TSC2</i>			
Von Hippel-Lindau sündroom	L/T	<i>VHL</i>	AD	Ubikvitineerumine ja sellele järgnev degradatsioon	1, 2, 3
Wilmsi tuumor	L	<i>WT1</i>	AD	Transkriptsioonifaktor, loote arenguline programmeerimine	1, 2
Endokrinoloogilised haigused					
Biotinidaasi puudulikkus	L	<i>BTD</i>	AR	Vaba biotiini regeneratsioon	3
Fabry haigus	L/T	<i>GLA</i>	XL	Glükolipiidide ja glükoproteiinide α -galaktosüülrühmade hüdroolüüs	1, 2
Monogeenne diabeet	L/T	<i>HNFA1</i>	AD	Transkriptsioonifaktor, maksa-spetsiifiliste geenide ekspressioon	3
Muud haigused					
Adrenoleukodüstroofia	L/T	<i>ABCD1</i>	XL	Väga pika ahelaga rasvhapete transport	3
Dopamiinitundlik düstoonia	L	<i>GCH1</i>	AD	Tetrahydrobiopterini (BH4) biosünteesi ensüüm	3
Ornitiini transkarbamülaasi puudulikkus	L/T	<i>OTC</i>	XL	Mitokondriaalse maatriksi ensüüm	1, 2
Maliigse hüpertermia eelsoodumus	L/T	<i>CACNA1S</i>	AD	Osa kaltsiumi kanalist	1, 2
		<i>RYR1</i>		Lihaskontraktsioon	
Wilsoni haigus	L	<i>ATP7B</i>	AR	Vase transport rakust välja	1

T – täiskasvanueas. L – lapseas. AD – autosoom-dominantne. AR – autosoom-retsessiivne. XL – X-liiteline.

¹ Kalia *et al.* (2017).

² Schwartz *et al.* (2018).

³ Webber *et al.* (2018).

1.4.2 Sekundaarsete leidude esinemissagedus üldpopulatsioonis

Mitmed uuringud on keskendunud erinevates valimites sekundaarsete leidude tuvastamisele, millest enamik on kasutanud ACMG poolt väljapakutud algset 56 või hilisemat 59 geeni sisaldavat nimekirja (tabel 3).

Tabel 3. Võrdlus sekundaarsete leidude uurimustest.

Uurimus	Nimekiri	Valimi suurus	Sekundaarsed leiud	Kandjad	
				n	%
Kelly <i>et al.</i> (2021)	ACMG SF v2.0 ^a + 1 HFE variant	130 048	980	3355	2,6
Gordon <i>et al.</i> (2020)	109 geeni ACMG SF v2.0	21 519	422	647	3,0
			398	549	2,6
Van Hout <i>et al.</i> (2020)	ACMG SF v2.0	49 960	548	992	2,0
Capalbo <i>et al.</i> (2019)	ACMG SF v2.0	14 125	218	327	2,3
Hart <i>et al.</i> (2019)	ACMG SF v1.0 ^b	6240	75	74	1,2
Thauvin-Robinet <i>et al.</i> (2019)	244 geeni ACMG SF v2.0	700	28	55	7,9
			13	13	1,9
Chen <i>et al.</i> (2018)	ACMG SF v2.0	421	11	12	2,9
Sapp <i>et al.</i> (2018)	ACMG SF v1.0	1197	14	18	1,5
Tang <i>et al.</i> (2018)	ACMG SF v2.0	954	21	24	2,5

^a Kalia *et al.* (2017).

^b Green *et al.* (2013).

Enim sekundaarsete leidudega üldpopulatsiooni indiviide (7,9%) on tuvastatud Thauvin-Robinet *et al.* (2019) uuringus, kus ACMG SF v2.0 listi geenidele oli lisaks uurimisel veel 185 geeni, kuid vaid ACMG gene arvestades leidis sekundaarseid variante antud uuringus 1,9% osalenutest, mis on kooskõlas teiste uuringutega (Thauvin-Robinet *et al.*, 2019). Geisingeri MyCode projekti raames on ACMG SF v2.0 + 1 listi kasutades tuvastatud sekkumistõhusaid leide 2,6% vaba-tahtlikest osalejatest (Kelly *et al.*, 2021). Selekteerimata valimites on veel ACMG SF v1.0 põhjal leitud sekundaarseid leide 1,2–1,5% (Hart *et al.*, 2019; Sapp *et al.*, 2018) ning ACMG SF v2.0

nimekirja kasutades 2,6% osalenutest (Gordon *et al.*, 2020). Gordon *et al.* (2020) uuringusse olid lisaks kaasatud ka 50 ACMG nimekirja mittekuuluvat geeni ning kõigi 109 geeni peale tuvastati sekundaarseid leide 3% valimist (Gordon *et al.*, 2020). Suurbritannia geenivaramu 49 960 indiviidi eksoomiantmete analüüsil tuvastati ACMG SF v2.0 nimekirja alusel sekundaarseid leide 2% (Van Hout *et al.*, 2020).

Eelselekteeritud uuringurühma kasutades on sekundaarseid leide ACMG SF v2.0 listi põhjal tuvastatud veidi alla 3% valimist (Capalbo *et al.*, 2019; Chen *et al.*, 2018; Tang *et al.*, 2018). Hirschsprungi tõve põdevate ja mittepõdevate inimeste seas läbiviidud uuringus tuvastati sekundaarseid leide 2,5% (Tang *et al.*, 2018) ning nii terveid, rasvunud kui kaasasündinud südamehaigustega lapsi kaasanud uuringus 2,9% (Chen *et al.*, 2018). Sugurakkude doonoreid ja kehavälist viljastamist kasutanud paaride hulgas on sekundaarseid leide raporteeritud 2,3% (Capalbo *et al.*, 2019).

Antud uuringute põhjal saab järeldada, et ACMG nimekirjade põhjal võib keskmine oodatav sekundaarsete leidudega patsientide hulk erinevates valimites olla 1–3%.

1.4.3 Tagasiside andmine sekundaarsete leidude kohta

Sekundaarsete leidude otsimisega kaasnevad küsimused, kas ja millistest leidudest tuleks patsiente teavitada. ACMG soovitude järgi tuleks patsiente teavitada sekundaarsetest leidudest, mis on sekkumistõhusad ja kõrge penetrantsusega, samas peab kõigile patsientidele jääma õigus antud informatsiooni mitte teada. Võimalus saada infot oma sekundaarsete leidude kohta tuleks tagada igale patsiendile nende vanusest sõltumata ning iga patsient peaks läbima testieelse ja -järgse geneetilise nõustamise (Miller *et al.*, 2021b). Küll aga soovitab Euroopa Inimesegeneetika Ühing (ESHG, *European Society of Human Genetics*) vältida patsiendi primaarse haigusega mitteseonduvate geenide analüüsimist, et ennetada sekundaarseid leide, ning enne testimist nõustada patsiente juhuleidude kohta, küsides nõusolekut nende raporteerimiseks (Matthijs *et al.*, 2016). Juhul kui otsida sekundaarseid leide, tuleks keskenduda hästi teadaolevatele kõrge penetrantsusega variantidele, mida saaks edukalt ja efektiivselt ennetada ja/või ravida. Seejuures tuleks arvestada ka konkreetse variandi penetrantsust konkreetses populatsioonis (de Wert *et al.*, 2021).

Sarnased soovitused on ka Kanada Meditsiinigeneetikute Kolleegiumi (CCMG, *Canadian College of Medical Geneticists*) poolt. CCMG ei kiida heaks sihilikku juhuleidude otsimist, kuid on siiski paika pannud selged soovitused sekundaarsete leidude raporteerimiseks (Boycott *et al.*, 2015; Hume *et al.*, 2019):

- Teovõimeliselt täiskasvanutelt tuleks enne testimist küsida informeeritud nõusolek.
- Lastel leitud kõrge penetrantsusega leidudest, mis on lapseas sekkumistõhusad, tuleks teavitada lapsevanemaid. Täiskasvanueas avalduvate haiguste kohta ei tohiks informatsiooni anda, v.a juhul kui lapsevanemad seda soovivad ja selle info jagamine aitaks potentsiaalselt ennetada lapsevanema või muu pereliikme rasket haigust.
- Teovõimetute täiskasvanute kõrge penetrantsusega sekkumistõhusatest variantidest tuleks teavitada nende seaduslikku esindajat, v.a kui kõnealune isik veel teovõimelisena soovis vastupidist.

Kuigi Euroopas puuduvad selged ja konkreetset soovitused või juhised sekundaarsete leidude raporteerimiseks ning üldine lähenemine on pigem konservatiivne ja ettevaatlik, leiab Prantsuse Prognostilise ja Personaalse Meditsiini Ühing (ingl *French Society of Predictive and Personalized Medicine*), et patsiente tuleks informeerida sekundaarsetest leidudest geenides, mis on seotud pärilike kasvaja eelsoodumusega. Samas tuleks enne igasugust geneetilist testimist patsiente sekundaarsete leidude osas nõustada ning võtta nende raporteerimiseks informeeritud nõusolek, jättes patsiendile õiguse sellest hiljem loobuda (Pujol *et al.*, 2018).

Viimastel aastatel on läbiviidud mitmeid uuringud, kus on välja selgitatud erineva taustaga inimeste soovid seoses geneetiliste leidude raporteerimisega. Mõne aasta taguses uuringus, kuhu olid kaasatud pea 7000 indiviidi 75 eri riigist, leiti, et 98% mitte-tervishoiutöötajatest tahaksid saada infot sekkumistõhusate sekundaarsete leidude kohta ning 73% lisaks ka mitte-sekkumistõhusate variantide kohta. Uuringusse kaasatud kliinilistel geneetikutel, geeniteadlastel ja tervishoiutöötajatel olid veidi konservatiivsemad arvamused, nimelt sooviksid neist 96% teada sekkumistõhusate ja 65% mitte-sekkumistõhusate leidudest (Middleton *et al.*, 2016). Perearstid on leidnud, et sekundaarsete leidude informeerimise peamine kasu on sekkumistõhusus ning informatsioon leidude kohta, mis pole sekkumistõhusad, on kasutu ja potentsiaalselt kahjulik ning võib patsientidel tekitada ärevust (Sebastian *et al.*, 2021).

Teismeliste seas läbiviidud küsitlustes on leitud, et kuni 83% noortest soovivad teada lisaks sekkumistõhusatele variantidele ka nendest variantidest, millega seotud haigused pole kas ravitavad või ennetatavad, tuues peamiseks põhjusteks, et see aitaks neil oma tulevikku planeerida (Hufnagel *et al.*, 2016; Pervola *et al.*, 2019). Mighton *et al.* (2020) uuringus leiti, et peamine põhjus, miks inimesed tahavad teada nii sekkumistõhusatest kui ka mitte-sekkumistõhusatest sekundaarsetest leidudest, on elukvaliteedi parandamine. Sama põhjuse tõid välja ka inimesed, kes ei soovinud üldse teada oma sekundaarsetest leidudest. Lisaks toodi välja, et sekundaarsete leidude kasulikkus seisneb võimaluses olla vaimselt ettevalmistunud, toetab tuleviku planeerimist ja tagab võimaluse jagada tulemusi pereliikmetega (Mighton *et al.*, 2020). Seni on uuringutes leitud, et enamik patsiente, kellel on tuvastatud ja raporteeritud sekundaarseid leide, ei kahetse oma valikut ning psühholoogilise stressi esinemine nende seas on madal (Hart *et al.*, 2019; Sapp *et al.*, 2018).

Antud uuringutest võib järeldada, et enamik inimesi sooviks neid puudutavastest sekkumistõhusatest variantidest teada, kuid igalt geeniuuringus osalejalt peaks enne geenianalüüsi erinevate geneetiliste leidude raporteerimise osas küsima informeeritud nõusoleku. Lisaks on siinkohal oluline aspekt ka patsientide nõustamine seoses nendel tuvastatud haigusseoseliste geneetiliste leidudega, kaasates vajadusel diskussiooni ka teised pereliikmed. Nõustamise järgselt saavad patsiendid teha informeeritud otsuseid testimisvõimaluste, tervise jälgimise, ravi ja pereplaneerimise osas.

1.5 Geenivariantide patogeensuse hindamine

Selleks, et patsiente teavitada vaid haigusseoselistest variantidest, on vajalik hinnata eksoomi või genoomi sekveneerimisel leitud variantide patogeensust. Eesmärgiga muuta kliiniliste geenivariantide hindamine ühtlasemaks ja läbinähtavamaks, avaldati 2015. a ACMG ja Molekulaarpatoloogia Assotsiatsiooni (AMP, *Association of Molecular Pathology*) koostöös soovitusel mendeliaarsete haiguste geneetiliste variantide patogeensuse hindamiseks. Antud juhend koosneb 16 patogeensuse ja 12 healoomulisuse hindamise kriteeriumist, mis võtavad arvesse populatsiooni, funktsionaalsuse, *in silico* programmide ennustuse, segregatsiooni jm andmeid. Väga tugevale patogeensusele viitab nullvariant geenis, kus funktsioonikadu on sage

haiguse tekkemehhanism, tugevale patogeensusele viitavad näiteks *de novo* variant (kui on olemas vanemate andmed), variandi statistiliselt kõrgem levimus haigetel võrreldes kontrollidega või kui sama aminohappelist muutust on varem patogeenseks hinnatud. Nõrgemad patogeensuse näitajad on näiteks variandi puudumine populatsiooni andmebaasides, paiknemine kõrge mutatsioonilise aktiivsusega (ingl *hot spot*) alas või asendusvariant geenis, kus on madal healoomuliste asendusvariantide esinemissagedus. Geenivariandi healoomulisusele viitavad näiteks kõrge minoorse alleeli esinemissagedus, kahjustava efekti puudumine funktsionaalsetes uuringutes või segregatsiooni puudumine haigusega. Iga variandi puhul tuleb hindajal otsustada, millised kriteeriumid on täidetud ning nende kombineerimisel jagatakse variandid ühte neljast rühmast: patogeenne (P), tõenäoliselt patogeenne (TP), healoomuline (HL) või tõenäoliselt healoomuline (THL). Kui kriteeriumite kombineerimise tulemus ei sobi ühtegi eelmainitud rühma, on tegu teadmata patogeensusega variandiga (VUS, *variant of unknown significance*) (Richards *et al.*, 2015).

Antud kriteeriumite hindamine võib aga kohati olla subjektiivne ja erinevate inimeste poolt erinevalt tõlgendatud. Seetõttu on välja töötatud mitmeid algoritme, mis lihtsustaksid ja ühtlustaksid meditsiinigeneetikute ja arstide tööd geenivariantide patogeensuse hindamisel. Sellised on näiteks ACMG juhistel põhinevad InterVar (Li ja Wang, 2017), VarSome (Kopanos *et al.*, 2019) ja Franklin (franklin.genoox.com). InterVar on bioinformaatiline käsureapõhine tarkvara geneetiliste variantide kliiniliseks interpreteerimiseks, mis hindab 18 ACMG kriteeriumi täitmist. Lisaks sellele on välja töötatud ka veebiserver wInterVar, mis võimaldab variantide automatiseeritud interpretatsiooni koos manuaalse kohendamisega (Li ja Wang, 2017).

Franklin on veebipõhine tööriist variantide patogeensuse interpreteerimiseks, mis kasutab tehisintellektil põhinevat variantide klassifikatsiooni mootorit (ingl *artificial intelligence-based variant classification engine*), et minimaliseerida ACMG juhendite tõlgendamise varieeruvust, tuginedes 17 ACMG kriteeriumile. Üks eesmärke Franklini puhul on kõigi haigusseoseliste variantide kinnipüüdmine, seega patogeensuse hindamisel võib juhtuda, et kriteeriumite spetsiifilisust langetatakse. Sellest tulenevalt antakse variantidele ka tugevus- ja usaldushinnang, et kasutaja saaks teha lõpliku informeeritud otsuse variandi haigusseoselisuse kohta (<https://franklin.genoox.com>).

VarSome'i puhul on tegu inimese geneetiliste variantide otsingumootoriga, kus variantide patogeensuse hindamine toimub automaatse variantide klassifikaatori abil, mis hindab haigusseoselisust vastavalt ACMG juhiste 21 kriteeriumile. Osadel juhtudel on lisaks arvestatud VarSome'i Teadusnõukogu (ingl *Scientific Advisory Board*) ja globaalse kogukonna ekspertarvamusi (Kopanos *et al.*, 2019). Lisaks automatiseeritud kriteeriumitele saab Franklinis ja VarSome'is läbi kasutajaliidese kriteeriume lisada (ja eemaldada), mis on vajalik eelkõige patsiendi fenotüüpi arvestavate parameetrite korral.

Eelmainitud tarkvarad sõltuvad suuresti teistest andmebaasidest, kust infot kogutakse. Üks omanäoline andmebaas, mis sõltub erinevate laborite, töögruppide ja ekspertpaneelide esitatud andmetest, on ClinVar. See on NCBI (*National Center for Biotechnology Information*) poolt hallatud avalik andmebaas inimese geneetilisest variantidest ja nende mõjust haiguste tekkele. ClinVar sai loodud peamise eesmärgiga olla andmebaas, kus kliinilise geneetika laborid saavad jagada variantide patogeensuse hinnanguid ja tõendeid, kuidas antud otsuseni jõuti. ClinVari kogutakse kokku info eri suuruses variantide kohta genoomi eri piirkondadest (Landrum *et al.*, 2018). Osadele geenivariantidest on oma haigusseoselisuse hinnangu andnud ka ekspertpaneelid, kes fokuseeruvad keerulistele variantidele, sh VUS-id ja vastuolulised (ingl *conflicting*) variandid (Landrum ja Kattman, 2018).

Hiljuti on oma geenivariantide klassifikatsioonimudeli välja töötanud ja avaldanud ka ESHG. Antud mudelit nimetatakse ABC süsteemiks ning seda on võimalik kasutada eraldiseisvalt või kombinatsioonis ACMG klassifikatsioonisüsteemiga. ABC süsteemi järgi antakse variandile esmalt funktsionaalne ja seejärel kliiniline hinnang ning nende kombineerimisel määratakse antud variandile klass A–F või klass 0, kui kliinilist hindamist ei peetud vajalikuks (Houge *et al.*, 2021).

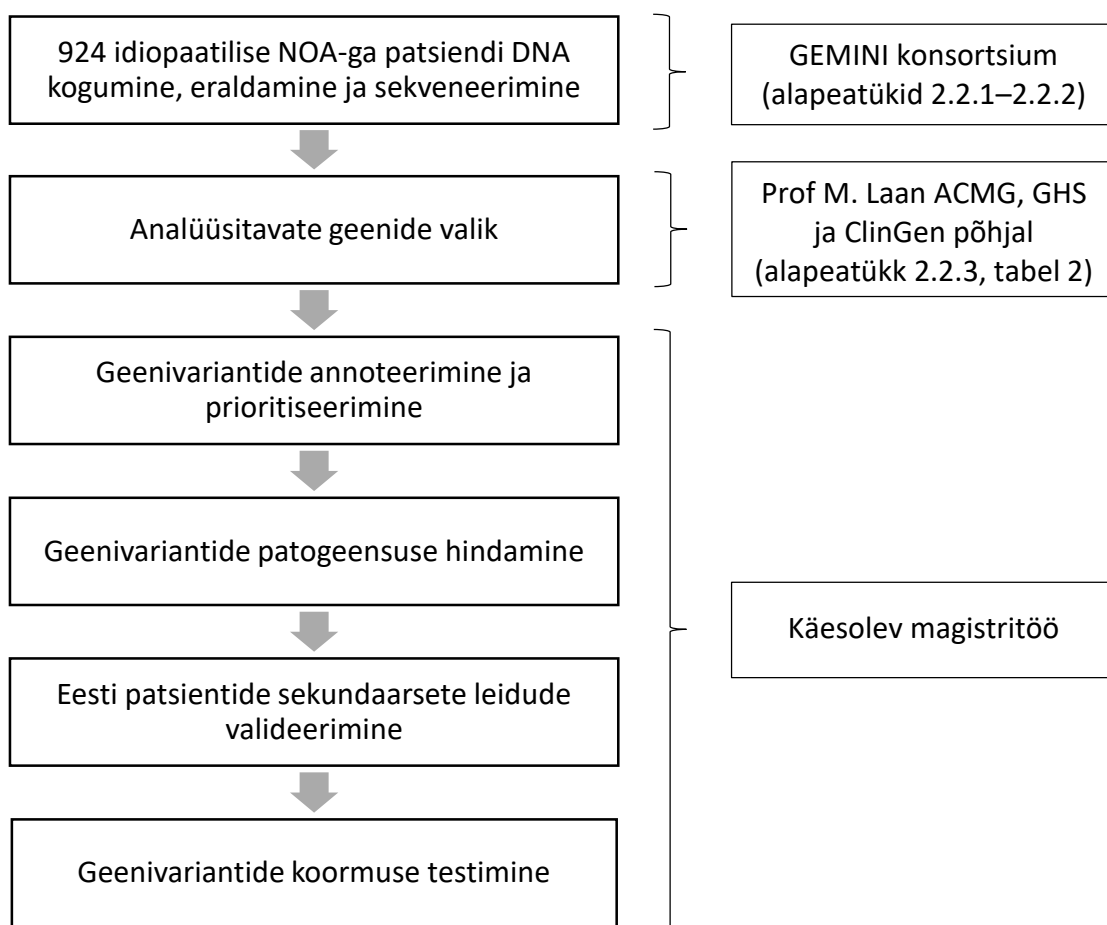
2 UURIMUS

2.1 Töö eesmärgid ja ülesehitus

Käesoleva töö eesmärgid olid:

- 1) tuvastada kliiniliselt olulised sekundaarsed leiud idiopaatilise mitte-obstruktiivse azoospermiaga meeste valimis;
- 2) analüüsida, kas idiopaatilise mitte-obstruktiivse azoospermiaga meestel esineb võimalikke sekkumistõhusaid geenivariante rohkem kui üldpopulatsioonis;
- 3) testida hüpoteesi, kas raskete spermatogeneesi häirete ja kasvajate geneetiline etioloogia on omavahel ülekattes;
- 4) võimaldada patsientidele lähitulevikus haigusriskide osas tagasisidet.

Antud magistr töö raames tehtud ja sellele eelnenud tööd iseloomustab joonis 3.



Joonis 3. Ülevaade käesoleva magistr töö raames tehtud ja sellele eelnenud tööst.

2.2 Materjal ja meetoodika

2.2.1 Valimi kirjeldus

Käesolevas töös kasutati GEMINI (*GENetics of Male Infertility Initiative*) konsortsiumi 924 idiopaatilise NOA-ga patsiendi täiseksoomi sekveneerimisel saadud andmeid. GEMINI puhul on tegu rahvusvahelise konsortsiumiga, kuhu on kaasatud üheksa riigi 20 keskust ning üle 1500 patsiendi ja 2000 kontrolli. Patsiendid on peamiselt kas NOA või oligozoospermia diagnoosiga ning konsortsiumi teadustegevus keskendub mehepoolse viljatuse geneetiliste põhjuste identifitseerimisele ja kirjeldamisele (gemini.conradlab.org).

Käesolevas magistritöös analüüsitud 924 patsienti olid kogutud 11 keskuses (tabel 4) ning neil oli diagnoositud azoospermia (n=915, spermide puudumine ejakulaadis) või raske oligozoo-/krüptozoospermia (n=9, spermide kontsentratsioon $< \sim 1$ mln/mL) ning kategoriseeritud kui NOA vastavalt füüsilisele ülevaatusele (munandimaht), endokriinsetele näitajatele (FSH, LH ja testosteroon) ning histoloogilistele leidudele. Enamik patsiente olid Euroopa päritolu (91%), ülejäänud Lõuna-Aasia (4%), Aafrika (3%) ja Ida-Aasia (2%) päritolu. Meeste keskmine vanus oli 35,5 aastat. Keskmine seerumi LH kontsentratsioon oli 8,1 IU/L, FSH kontsentratsioon 20,4 IU/L ja kogu testosterooni kontsentratsioon 14,7 nmol/L. Kõikidest keskustest on olemas eetikakomitee luba ning kõigilt meestelt on võetud kirjalik nõusolek uuringus osalemiseks.

Eestist on GEMINI konsortsiumis esindatud nii Tartu Ülikool (prof Maris Laan, BSMI inimese geneetika uurimisrühm) kui ka SA Tartu Ülikooli Kliinikumi Meestekliinik (dr Margus Punab). Eesti patsientide keskmine vanus oli 35,5 aastat, keskmine LH kontsentratsioon 7,8 IU/L, FSH 22,0 IU/L ja testosteroon 16,2 nmol/L.

Tartu Ülikooli inimuuringute eetikakomitee poolt on välja antud heakskiit 74/54 (viimane muudatusettepanek 288/M-13).

Tabel 4. Käesoleva magistr töö valimi jaotus GEMINI projektis osalevate keskuste vahel (n=924).

Keskus	Patsientide arv	Keskuse koordinaator(id)
Ameerika Ühendriigid	311	
Cornell (Weill Cornelli Meditsiini- kolledž)	205	Peter N. Schlegel
Stanford (<i>Stanford Medicine</i>)	10	Michael L. Eisenberg
Utah (Utah' Ülikool)	62	Kenneth I. Aston, Douglas T. Carrell, James M. Hotaling, Timothy G. Jenkins
WashU (Washingtoni Ülikool)	34	Emly S. Jungheim, Kenan R. Omurtag
Austraalia (Monashi Ülikool)	88	Rob McLachlan
Eesti (Tartu Ülikool ja Tartu Ülikooli Kliinikum)	82	Maris Laan, Margus Punab
Hispaania (<i>Instituto de Investigaciones Biomédicas Sant Pau</i>)	46	Daniel Moreno Mendoza, Csilla Krausz
Kanada (Toronto Ülikool ja Mount Sinai haigla)	35	Keith A. Jarvi
Portugal	275	
Lissabon (<i>Instituto Nacional de Saude Doutor Ricardo Jorge</i>)	147	João Gonçalves
Porto (Porto Ülikool)	128	Alexandra M. Lopes, Susana Seixas, Filipa Carvalho, Susana Fernandes, Alberto Barros
Taani (Kopenhaageni Ülikool)	87	Ewa Rajpert-De Meyts, Niels Jørgensen, Kristian Almstrup

2.2.2 Sekvenerimisandmed

GEMINI täiseksoomi sekvenerimine teostati Ameerika Ühendriikides Washingtoni Ülikoolis McDonnelli genoomi instituudis (www.genome.wustl.edu), kasutades samas laboris välja töötatud ja valideeritud (ingl *in-house*) eksoomi sekvenerimise reagenti Illumina HiSeq 4000 platvormil. Lugemite katvuseks saadi keskmiselt 80x. Sekveneritud lugemid töödeldi ja joondati vastu GRCh38 inimese referentsgenoomi, kasutades tööriistu BWA-MEM (v0.7.17) (Li ja Durbin 2009), Genome Analysis Toolkit (v3.6.0) (McKenna *et al.* 2010) ja Picard (v2.10.0) (broadinstitute.github.io/picard). Baasinformaatilise osa teostas dr Liina Nagirnaja (Oregon Health & Science University, Portland, USA) (Nagirnaja *et al.*, 2021, avaldamisele saadetud).

2.2.3 Analüüsitavate geenide valik

Töös analüüsitud 85 geeni olid nimetatud sekkumistõhusateks ACMG (Kalia *et al.*, 2017), GHS (Schwartz *et al.*, 2018) ja/või ClinGen (Webber *et al.*, 2018) ekspertide poolt (tabel 2). ACMG nimekirjadest kasutati antud töös ACMG SF v2.0 ning GHS puhul esmast 76 geeniga nimekirja.

2.2.4 Analüüsiks kasutatud andmebaasid ja tarkvara

Geneetiliste variantide analüüsi erinevates etappides kasutati mitmeid andmebaase ja tarkvarasid, mis on välja toodud tabelis 5. Üldiseks analüüsiks kasutati programmeerimiskeelt R (v3.6.3) (R Core Team, 2020), Linux terminali ja käsureapõhist programmi BCFtools (v1.12) (Li *et al.*, 2009). Jooniste tegemiseks kasutati R paketti ggplot2 (v3.3.0) (Wickham, 2016).

Tabel 5. Käesolevas töös kasutatud andmebaaside ja tarkvarade lühikirjeldus.

Andmebaas/tarkvara	Kirjeldus
ANNOVAR	Tarkvara geenivariantide anoteerimiseks, kasutades erinevaid andmebaase, sh RefGene, ensGene, dbSNP, gnomAD, ExAc, ClinVar, InterVar (annovar.openbioinformatics.org).
ClinVar	Andmebaas, mis sisaldab infot inimese geenivariantide kohta ning nende seoseid haigustega (clinvar.com).
dbSNP (avsnp150)	NCBI alla kuuluv erinevate liikide geenivariantide andmebaas (www.ncbi.nlm.nih.gov/snp).
ensGene	Ensembli poolt välja töötatud vahend kodeerivate ja mittekodeerivate geeniregioonide tuvastamiseks (genome.ucsc.edu/cgi-bin/hgTrackUi?db=danRer10&g=ensGene).
ExAc	Populatsiooniandmebaas, mis sisaldab 60 706 isiku eksoomiandmeid. Kuulub nüüd gnomAD-i alla (exac.broadinstitute.org).
FinchTV	Programm Sanger sekveneerimise andmete illustreerimiseks (digitalworldbiology.com/FinchTV).
Franklin	Geenivariantide otsingumootor (franklin.genoox.com).
GenomeTester	Tarkvara praaimerite unikaalsuse kontrollimiseks (bioinfo.ut.ee/genometester).
gnomAD	Populatsiooniandmebaas, v2.1 sisaldab 125 748 eksoomi ja 15 708 genoomi andmeid (gnomad.broadinstitute.org).
InterVar	Geenivariantide kliiniline interpreteerimine (wintervar.wglab.org).
MGI	Hiire geenide, geenifunktsioonide, fenotüüpide ja hiiremudelite andmebaas (www.informatics.jax.org).
OMIM	Andmebaas, mis hõlmab inimese gene ja geneetilisi haiguseid, fokuseerides geneetiliste variantide ja fenotüübi molekulaarsetele seostele (www.omim.org).
Primer3	Tarkvara praaimerite disainimiseks (bioinfo.ut.ee/primer3).
Primer-BLAST	Tarkvara praaimerite disainimiseks ja unikaalsuse kontrollimine (www.ncbi.nlm.nih.gov/tools/primer-blast).
RefGene	Teadaolevate inimese genoomi kodeerivate ja mittekodeerivate geenide kirjeldamine, kasutades NCBI RNA referentsjärjestuste andmebaasi (RefSeq) (varianttools.sourceforge.net/Annotation/RefGene).
VarSome	Geenivariantide otsingumootor (varsome.com).

2.2.5 Variantide annoteerimine ja prioritseerimine

Kogu eksoomi sekveneerimisel saadud andmed uuritavate 85 geeni kohta annoteeriti ANNOVAR-iga (*ANNOtate VARIation*, v20191024). ANNOVAR on tarkvara geneetiliste variantide interpreteerimiseks ja prioritseerimiseks, kus on võimalik rakendada geeni-, regiooni- või filtripõhist annoteerimist (Wang *et al.*, 2010).

Annoteeritud andmetest filtreeriti sisse variandid:

- mille genotüüpide katvus valimis oli vähemalt 90%;
- mis asusid eksonites või ekson-intron piiril (splaiss-variandid);
- mis ei olnud sünonüümsed ühenukleotiidsed variatsioonid (SNV, *single nucleotide variant*) refGene (varianttools.sourceforge.net/Annotation/RefGene) ja ensGene (genome.ucsc.edu/cgi-bin/hgTrackUi?db=danRer10&g=ensGene) ennustuste põhjal;
- mis olid uued ehk andmebaasides esindamata või harvad variandid ehk mille minoorse alleeli esinemissagedus (MAF, *minor allele frequency*) üldpopulatsioonis $\leq 1\%$.

Geenivariantide MAF võeti vastavalt andmebaasidele gnomAD (*Genome Aggregation Database*, v2.1.1) (Karczewski *et al.*, 2020) ja ExAc (*Exome Aggregation Consortium*, v03) (Karczewski *et al.*, 2017).

2.2.6 Geenivariantide patogeensuse hindamine

Edasiseks uurimiseks kasutati automatiseeritud ClinVari (v20200316) (Landrum *et al.*, 2018) ja InterVari (v20180118) (Li ja Wang, 2017) haigusseosisuse hinnanguid. Nende põhjal välistati variandid, mille mõlemad hinnangud olid VUS (teadmata patogeensusega variant) või millel vähemalt üks hinnang oli HL (healoomuline) või THL (tõenäoliselt healoomuline). Alles jäid:

- Potentsiaalsed patogeensed variandid – vähemalt üks TP (tõenäoliselt patogeenne) või P (patogeenne) hinnang ning teine hinnang polnud HL või THL.
- Hindamata patogeensusega variandid – hinnangud puudusid.

Allesjäänud variantide haigusseosisuse kinnitamiseks leiti juurde hinnangud VarSome'ist (Kopanos *et al.*, 2019) ja Franklinist (franklin.geenox.com) ning tehti tehti manuaalne ClinVari

väljavõtte, kuna ClinVari andmebaasi uuendatakse pidevalt. Lisaks anti hinnang ACMG juhiste järgi. Saadud andmete põhjal pandi paika kriteeriumid määramaks variandi lõplikku hinnangut:

- Kui variandile oli olemas ClinVaris ekspertpaneeli hinnang, jäi lõplik hinnang vastavalt sellele.
- Kui variandile leidis nii THL/HL ja TP/P hinnanguid, jäi lõplikuks hinnanguks vastuoluline.
- Ülejäänud variantide lõplik hinnang jäi vastavalt ACMG hinnangule.

2.2.7 Eesti patsientide patogeensete variantide valideerimine

Eesti NOA patsientidel tuvastatud sekundaarsetele leidudele viidi läbi valideerimine Sanger sekveneerimisega. Selle jaoks disainiti praimerid programmiga Primer3 (v4.1.0) (Kõressaar *et al.*, 2018) ning unikaalsuse kontrollimiseks kasutati programme Primer-BLAST (Ye *et al.*, 2012) ja GenomeTester (v1.3) (Andreson *et al.*, 2006). Kasutatud praimerite järjestused on välja toodud tabelis 6.

Tabel 6. Valideerimiseks kasutatud praimerite nukleotiidsed järjestused.

Geen	cDNA muutus	Produkti pikkus	Praimeri järjestus (F/R)
<i>BRCA2</i>	c.1796_1800del	424 bp	TCTTGCAGTAAAGCAGGCAAT/ TGCTTCAAACCTGGGCTGAAC
<i>BRCA2</i>	c.3645_3646insTTTA	407 bp	ACACCTAGCCAAAAGGCAGA/ TTCACAGCTTTTTGCAGAGC

F – juhtiva ahela järjestus. R – mahajääva ahela järjestus.

Praimerite töötavuse kontrolliks viidi läbi PCR, mille jaoks kasutatud programmid on välja toodud tabelis 7. PCR produktide geelelektroforees viidi läbi 1% TBE geelil. Sanger sekveneerimine teostati Eesti Biokeskuse tuumiklabori poolt. Sekveneerimise andmeid analüüsiti programmiga FinchTV (v1.4).

Tabel 7. Valideerimiseks valitud geenivariantide PCR programmid.

Geen	cDNA muutus	Programm			
BRCA2 ^a	c.3645_3646 insTTTA ^a	Algne denaturatsioon	95°C	15 min	
		Denaturatsioon	95°C	45 s	35 tsüklit
BRCA2 ^b	c.1796_1800 delCTTAT ^b	Praimerite seondumine	51 ^a /53 ^b °C	30 s	
		Ekstensioon	72°C	45 s	
		Inkubatsioon	72°C	10 min	
		Säilitamine	4°C	∞	

2.2.8 Eksoonsete harvade variantide koormuse testimine

Uuritavas GEMINI valimis (n=924) esinevate harvade geenivariantide koormuse testimiseks võrreldes gnomAD andmebaasiga kasutati tarkvara TRAPD (*Test Rare vAriants with Public Data*). TRAPD puhul on tegu tarkvarapaketiga, mis on mõeldud geenipõhiseks koormuse võrdluseks avalike andmebaasidega, arvutades rikastatuse Fisheri täpse testi põhjal (Guo *et al.*, 2018). Käesolevas töös valiti võrdlusvalimiks gnomAD-i eurooplaste (millest on välja arvatud soomlased, NFE, *non-Finnish Europeans*) eksoomandmed (n=56 885).

Kuna TRAPD ei toeta X-liiteliste geenide testimist, välistati antud pärandumistüübiga geenid analüüsist (n=3).

Lähtudes Guo *et al.* (2018) tööst, pandi paika kriteeriumid testitavatele variantidele:

- GEMINI ja gnomAD-i positsioonide >10x katvus vähemalt 90% patsientidest kummaski valimis (kontroll teostatud TÜ genoomika instituudi spetsialisti Maarja Lepametsa poolt).
- Variantide populatsioonide maksimaalne MAF ≤1% gnomAD (v2.1.1) andmebaasi järgi.
- Sünonüümsete variantide testimise abil kinnitati sobiv QD (*quality by depth*) skooride lävend.

QD skoori lävendi kinnitamiseks viidi läbi testi kalibreerimine vastavalt mõlemas valimis esinevatele eksoonsete variantide QD skooridele. Selle jaoks arvutati GEMINI valimi QD skooridele vastavad 85, 80, 75, 70 ja 65 protsentiilid ja gnomAD-i valimis 95 ja 90 protsentiilid ning eksoonsete sünonüümsete harvade variantide peal kombineeriti erinevate protsentiilide

paare, tehes kindlaks parimad tingimused tasakaalustatud koormuse testi jaoks. Antud kalibreerimistesti alusel sätestati ka statistiline olulisusnivoo.

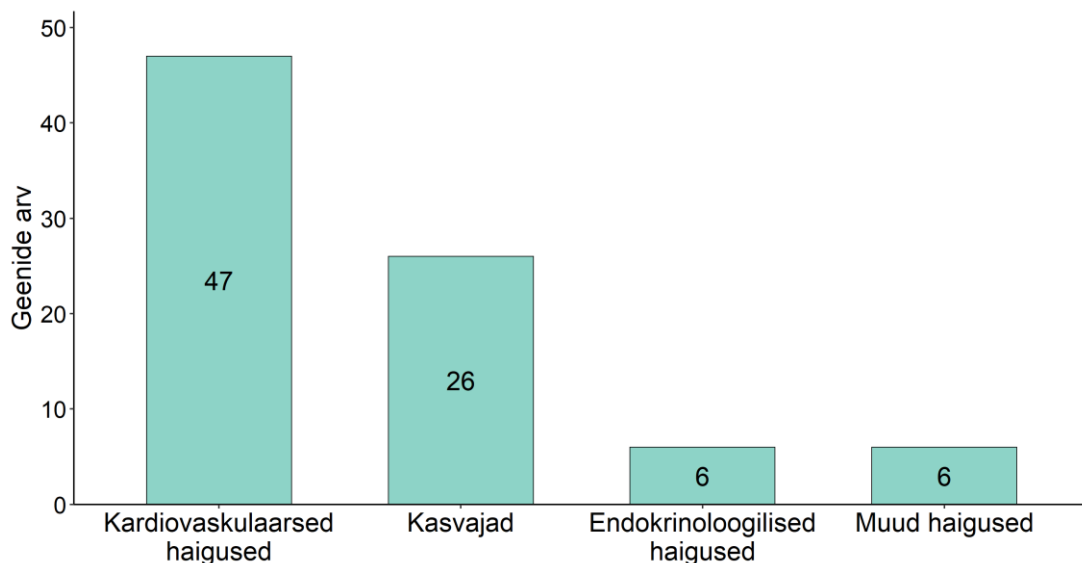
Seejärel viidi läbi analüüs harvade (MAF $\leq 1\%$) eksoonsete lugemisraami muutvate indelide, asendus- ja stoppvariantide ning splaiss-variantide peal. GEMINI QD skoori filtriks valiti 65% (QD $\geq 8,981$) ja gnomAD-i valimil 95% (QD $\geq 11,383$). Olulisusnivoo valiti vastava sünonüümse kalibreerimistesti (65% vs. 95%) alusel, milleks oli $2,21 \times 10^{-5}$.

Kuna GEMINI ja gnomAD-i valimite suurused on väga erinevad, oli oluline konservatiivne analüüs, et vältida juhuslikke p-väärtuseid, mis tulenevad valimi juhuslikest erinevustest sekveneeritud alades ja nende toorandmete käsitlest.

2.3 Tulemused

2.3.1 Analüüsitud geenide profiil

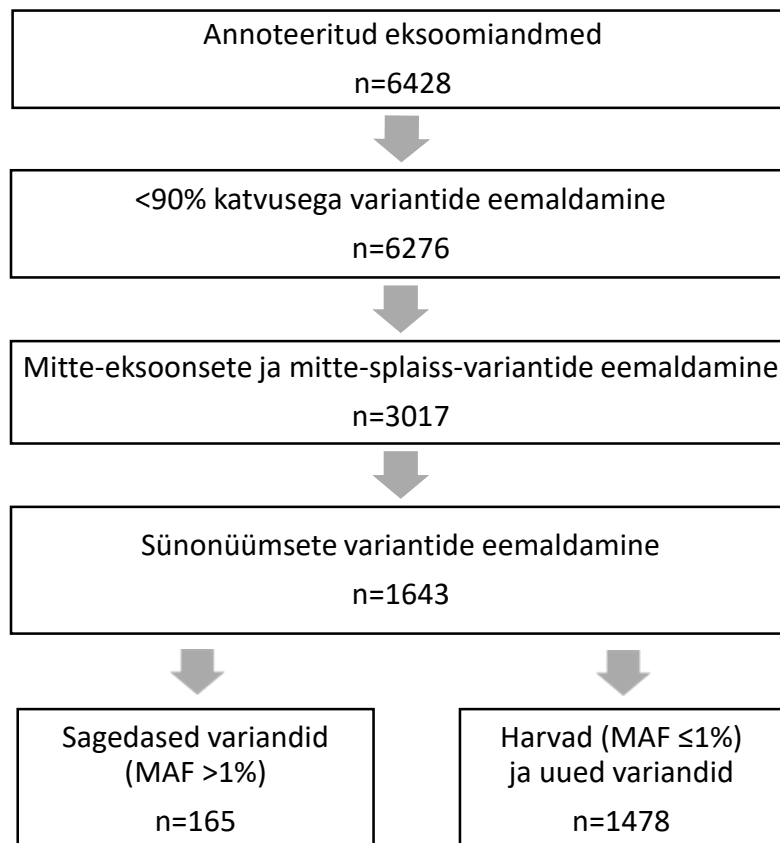
Uurimistöökäigus analüüsiti ACMG SF v2.0, GHS76 ja ClinGen nimekirjade alusel 85 geenis asuvaid varieeruvusi (tabel 2, joonis 4). Antud nimekirjades esines nii autosoom-retsessiivse kui ka X-liitelise pärandumismustriga gene kolm, ülejäänud 79 geeni olid autosoom-dominantse pärandumisega. Enamik geenidest (55%) olid seotud KVS haigustega, seejärel kasvajatega (31%) ning endokrinoloogiliste (7%) ja muude haigustega (7%) seotud geenid.



Joonis 4. ACMG SF v2.0, GHS76 ja ClinGeni nimekirjade geenide (n=85) jaotumine vastavalt haiguste rühmadele.

2.3.2 Geenivariantide filtreerimine

Annoteeritud eksoomi sekveneerimisandmetes leidus geenivariante kokku 6428. Varieeruvusi esines kõigis väljavalitud geenides, v.a geen *PLN*. Huvipakkuvate ehk potentsiaalselt patogeensete geenivariantide leidmiseks viidi läbi kõigi geenivariantide filtreerimine vastavalt uuringus sätestatud tingimustele (vt alapeatükid 2.2.4 ja 2.2.5, joonis 5).



Joonis 5. Geenivariantide filtreerimise protsess. MAF – minoorse alleeli sagedus.

Pärast kõigi variantide filtreerimist alles jäänud erineva sagedusega huvipakkuvate variantide (n=1643) jagunemine MAF järgi on välja toodud tabelis 8. Valdava enamuse variantide (96,5%) puhul oli tegu mittesünonüümsete ehk asendusvariantidega. Ennustatavalt funktsioonikaoga (LoF, *loss-of-function*) variandid, mille alla kuulusid splaiis-saidis asuvad variandid, lugemisraami muutvad insertioonid ja deletsioonid ning stopp- ehk nonsenssvariandid, moodustasid kõigist variantidest 2,2%. Ülejäänud variantide alla kuulusid lugemisraami mitte-muutvad indelid ning start- või stoppkoodoni kadu tekitavad variandid (1,3%).

Tabel 8. Antud töö valimi patsientidel esinevad geenivariandid (n=1643) vastavalt minoorse alleeli esinemissagedusele (MAF).

	Sagedased variandid (MAF >1%)		Harvad (MAF ≤1%) ja uued variandid		Kokku	
	n	%	n	%	n	%
Asendusvariandid	161	97,6	1425	96,4	1586	96,5
LoF variandid	2	1,2	34	2,3	36	2,2
Muud variandid*	2	1,2	19	1,3	21	1,3
Kokku	165		1478		1643	

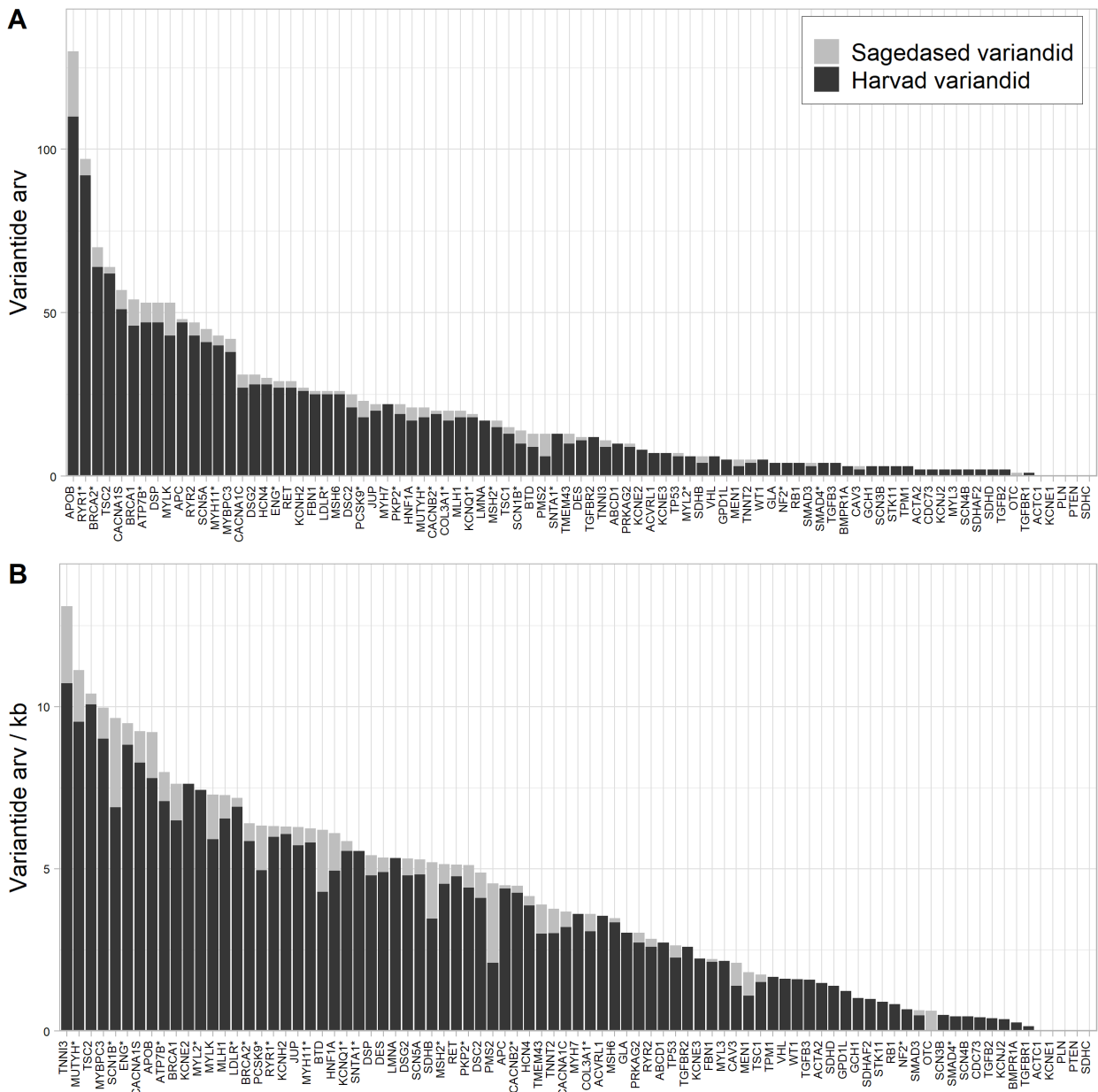
LoF – funktsioonikaoga variandid (splaiss-variandid, lugemisraami muutvad indelid ning stopp-variandid).

* Lugemisraami mitte-muutvad indelid ning start- või stoppkoodoni kadu tekitavad variandid.

Kahe sagedase LoF variandi puhul oli tegu lugemisraami muutva insertiooniga. Esimese geenivariandi puhul oli tegu *MYLK* geeni ekson-intron piiril oleva ühenukleotiidses insertiooniga (c.4289-10dup, rs200371896), mille sagedus üldpopulatsioonis on 3–4% ning mida antud töö valimis leidis 65 patsiendil, nendest neli homosügootses seisundis. Teise insertiooni puhul oli tegu *DSC2* geenis oleva kahenukleotiidses insertiooniga (c.2686_2687dup, rs200056085), mille sagedus genoomis on veidi üle 1%. Selle insertiooniga patsiente leidis antud töö valimis 16, kellest kõik olid heterosügootid. Kuigi tegu on lugemisraami muutvate insertioonidega ja seega LoF variantidega, on ClinVari patogeensuse hinnang mõlemale variandile HL/THL.

2.3.3 Variantide jagunemine geeniti

Geenivariantide filtreerimise järgselt jäi alles 1643 varianti (joonis 6). Enim leidis variante geenis *APOB* (n=133) ning vähim geenides *OTC* ja *TGFBR1* (n=1). Varieeruvusi ei esinenud geenides *ACTC1*, *KCNE1*, *PLN*, *PTEN* ja *SDHC*. Vastavalt geeni peamise transkripti pikkusele leidis enim variante geenis *TNNI3* (13,1 varianti 1 kb kohta), vähim geenis *TGFBR1* (0,2 varianti 1 kb kohta).



Joonis 6. Variantide ($n=1643$) jagunemine geeniti (A) ja 1 kb transkripti kohta (B). Harvad variandid defineeritud kui minoorse alleeli sagedus $\leq 1\%$ või andmebaasides esindamata variandid. * Funktsioonikaoga variante (splaiss-variante, lugemisraami muutvaid indeleid ja/või stopp-variante) sisaldav geen.

2.3.4 Variantide jagunemine valimi patsientide seas

Pärast sagedaste variantide (MAF >1%) välja filtreerimist jäi alles 1478 harva või varem andmebaasides esindamata varianti. Selliseid variante esines keskmiselt 3,12 patsiendi kohta. Enim esines ühel patsiendil 14 heterosügootset varianti ja kaks homosügootset varianti (tabel 9). Asendusvariante leidis ühel patsiendil kõige rohkem 12, ühelgi patsiendil ei tuvastatud rohkem kui üks LoF variant.

Tabel 9. Harvade (minoorse alleeli sagedus ≤1%) ja uute variantide (n=1478) esinemine valimi 924 patsiendil.

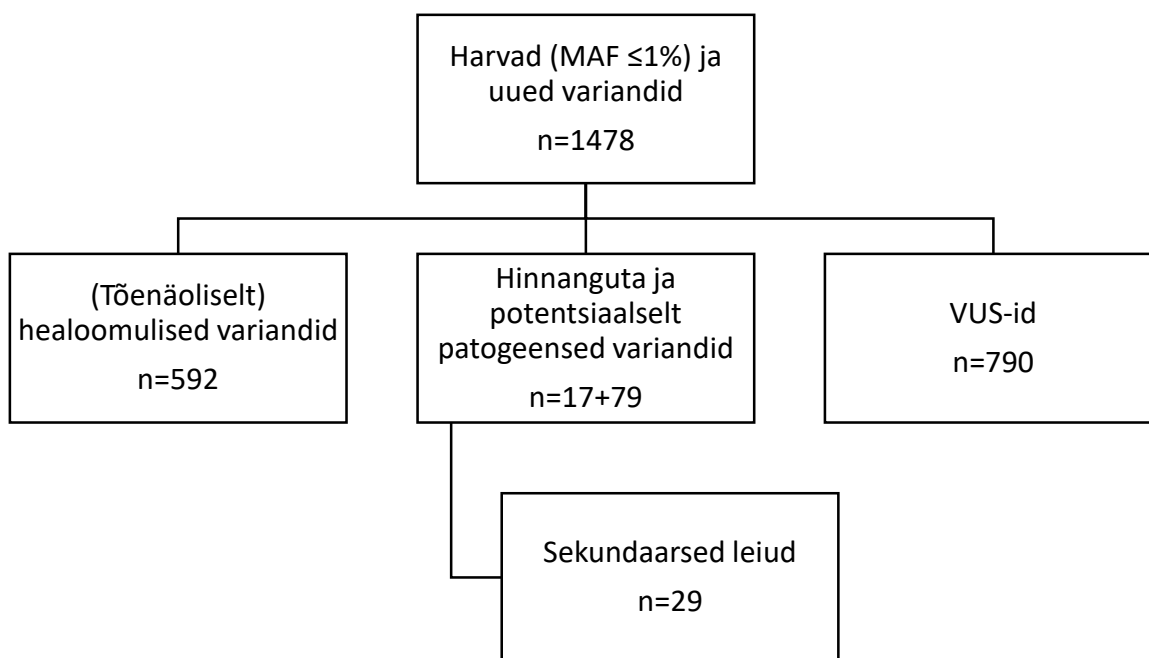
	Heterosügootsed				Homosügootsed			
	AV	LoF	Muud	Kõik	AV	LoF	Muud	Kõik
Keskmine	2,92	0,03	0,08	3,04	0,05	0,03	0,00	0,05
Mediaan	3,00	0,00	0,00	3,00	0,00	0,00	0,00	0,00
Max	12,00	1,00	2,00	14,00	2,00	1,00	0,00	2,00

AV – asendusvariandid. LoF – funktsioonikaoga variandid (splaiss-variandid, lugemisraami muutvad indelid ning stopp-variandid). Muud – lugemisraami mitte-muutvad indelid ning start- või stoppkoodoni kadu tekitavad variandid.

2.3.5 Sekundaarsed leiud valimi patsientidel

Harvade ja uute variantide (n=1478) haigusseoselisuse hindamiseks kasutati töös sätestatud kriteeriume (vt alapeatükk 2.2.5), mille järgi hinnati haigusseoselisus 96 geenivariandil (joonis 7). Sekundaarsete leidudena defineeriti 29 geenivarianti, nendest 14 TP ja 15 P varianti 21 eri geenis (tabel 10). Leitud geenivariandid esinesid kokku 32 patsiendil 924-st ehk 3,5%. 27 varianti asusid autosoom-dominantse pärandumisega geenides (nendest üks homosügootses olekus), üks hemisügootne variant tuvastati X-liitelises *ABCD1* geenis ning üks homosügootne variant autosoom-retsessiivse pärandumisega *MUTYH* geenis. Ühelgi patsiendil ei esinenud üle ühe sekundaarse leiu.

Tuvastatud sekundaarsetest leidudest ei kuulu ACMG SF v2.0 geenide nimekirja *ABCD1* ja *HNF1A* (küll aga on viimane lisatud ACMG SF v3.0 nimekirja). Seega ACMG SF v2.0 nimekirja järgi leidis 27 TP/P varianti 30 idiopaatilise NOA-ga patsiendil ehk 3,2% valimist.



Joonis 7. Geenivariantide patogeensuse hindamise protsess. 1478 geenivarianti jagati kolme gruppi, millest hinnanguta (n=17) ja potentsiaalselt patogeensetele (n=79) tehti haigusseoselisuse hindamine ning tuvastati 29 sekundaarset leidu. MAF – minoorse alleeli sagedus. VUS – teadmata patogeensusega variant.

Tabel 10. Tuvastatud sekundaarsed leiud (n=29) 924 idiopaatilise mitte-obstruktiivse azoospermiaga patsiendil.

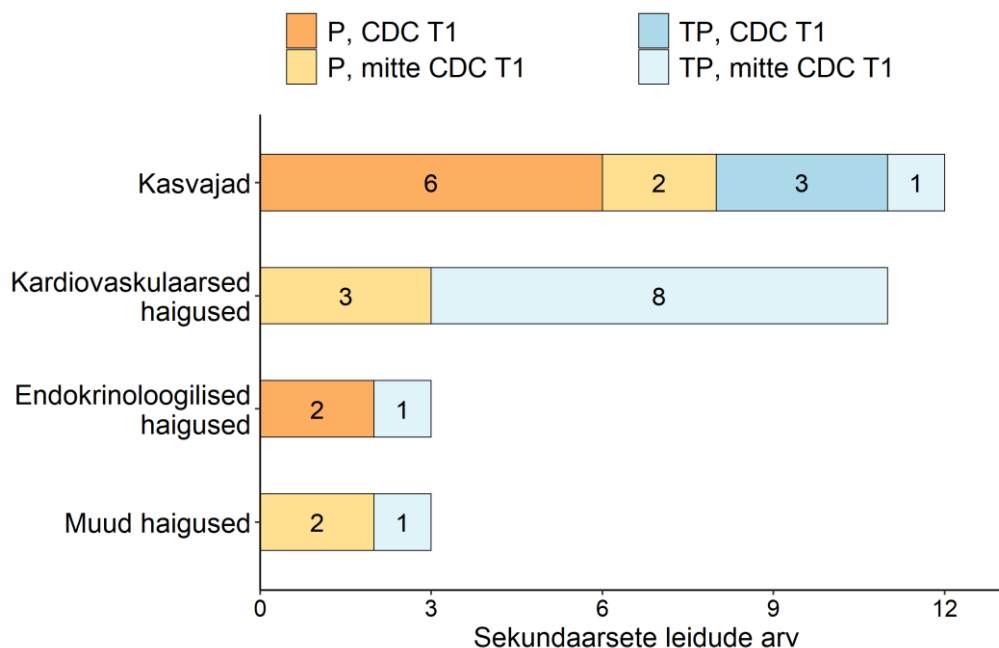
Geen	cDNA muutus, valgumuutus	rs kood	MAF gnomAD-is (%)	Hinnang	Patsientide arv
<i>ABCD1</i>	c.854G>A, p.Arg285His	rs782635828	0,0014	TP	1
<i>BRCA1</i>	c.2299delA, p.Ser767fs*25	rs80357786	-	P	1
<i>BRCA2</i>	c.1670T>G, p.Leu557*	rs80358452	-	P	1
<i>BRCA2</i>	c.1796_1800delCTTAT, p.Ser599*	rs276174813	0,0004	P	1
<i>BRCA2</i>	c.5576_5579delTTAA, p.Ile1859fs*3	rs80359520	0,0016	P	1
<i>BRCA2</i>	c.3645_3646insTTTA, p.R1217Ifs*2	-	-	TP	1
<i>BRCA2</i>	c.3680_3681delTG, p.Leu1227fs*5	rs80359395	0,0032	P	1

<i>BRCA2</i>	c.8488-1G>A, -	rs397507404	-	P	1*
<i>FBN1</i>	c.287G>C, p.Arg96Thr	rs794728291	-	TP	3
<i>HNF1A</i>	c.467C>A, p.Thr156Lys	rs150513055	0,0012	TP	1
<i>KCNQ1</i>	c.187C>T, p.Arg63Trp	rs199473662	0,0004	TP	1
<i>LDLR</i>	c.1216C>T, p.Arg406Trp	rs121908043	0,0018	P	1
<i>LDLR</i>	c.1775G>A, p.Gly592Glu	rs137929307	0,0057	P	1
<i>LMNA</i>	c.1718C>T, p.Ser573Leu	rs60890628	0,0140	TP	1
<i>MSH6</i>	c.3226C>T, p.Arg1076Cys	rs63750617	0,0095	TP	2
<i>MUTYH</i>	c.1187G>A, p.Gly396Asp	rs36053993	0,3059	P	1*
<i>MYBPC3</i>	c.1484G>A, p.Arg495Gln	rs200411226	0,0024	TP	1
<i>MYH11</i>	c.5083-1G>A, -	-	-	P	1
<i>MYH7</i>	c.2167C>T, p.Arg723Cys	rs121913630	0,0012	P	1
<i>PKP2</i>	c.2014-1G>C, -	rs193922674	0,0032	P	1
<i>PMS2</i>	c.137G>T, p.Ser46Ile	rs121434629	0,0173	TP	1
<i>RET</i>	c.2370G>T, p.Leu790Phe	rs75030001	0,0020	P	1
<i>RYR1</i>	c.12317delA, p.Ile4107fs*15	rs754572007	0,0040	P	1
<i>RYR1</i>	c.10347+1G>A, -	rs111436401	0,0004	P	1
<i>SCN5A</i>	c.361C>T, p.Arg121Trp	rs199473556	-	TP	1
<i>SCN5A</i>	c.4877G>A, p.Arg1626His	rs199473283	0,0042	TP	1

<i>TGFBR2</i>	c.1067G>A, p.Arg356Gln	rs727504292	0,0018	TP	1
<i>TNNI3</i>	c.586G>A, p.Asp196Asn	rs104894727	0,0007	TP	1
<i>WT1</i>	c.455delG, p.G152Afs*11	-	-	TP	1

MAF – minoorse alleeli sagedus. TP – tõenäoliselt patogeenne. P – patogeenne. * – homosügoot.

Enim TP ja P variante tuvastati geenis *BRCA2* (n=6, nendest üks homosügootse kandjaga), mis on CDC esimese taseme geen ning on seotud päriliku rinna- ja munasarjavähiga (joonis 8). CDC esimese taseme geenidest tuvastati TP/P variante lisaks geenides *BRCA1* (n=1, pärilik rinna- ja munasarjavähk), *LDLR* (n=2, perekondlik hüperkolesteroleemia) ning *MSH6* ja *PMS2* (n=1, Lynch'i sündroom). Nende haiguste varajane tuvastamine ja sekkumine langetab märgatavalt haigestumust ja suremust. Kokku kuulusid CDC esimese taseme geenidesse 38% tuvastatud sekundaarsetest leidudest, millest enamus olid P hinnanguga (73%, 8/11).



Joonis 8. Sekundaarsete leidude (n=29) jagunemine vastavalt haiguse kategooriale. P – patogeenne, TP – tõenäoliselt patogeenne, CDC T1 – CDC esimese taseme haiguste geenides olevad variandid.

Vastavalt haigusgruppidele tuvastati sekundaarseid leide enim (41%, 12/29) kasvajate geenides, nendest kaks kolmandikku olid P leiud. Üks sekundaarne leid vähem (39%, 11/29) leiti KVS haiguste geenides, millest pea kolmandik olid P hinnanguga. Kolm leidu (10%, 3/29) tuvastati nii endokrinoloogiliste kui ka muude haiguste gruppides, millest mõlemal juhul kaks olid P leiud.

Marfan'i ja Loeys-Dietz'i sündroomi ning perekondlikku rinnaaordi aneurüsme ja dissektsioone põhjustavas geenis *FBN1* tuvastatud TP muutus c.287G>C (p.Arg96Thr) esines valimis kolmel patsiendil Portugali kahest keskusest, seetõttu võib antud variandi puhul tegu olla populatsiooni-spetsiifilise variandiga. Kuna andmebaasides pole seda varianti varem kirjeldatud ja patsientide kohta detailne suguvõsa anamnees puudus, on keeruline hinnata, kas tegemist on populatsiooni-spetsiifilise variandiga või kaugete sugulusastmetega indiviididega.

Lynch'i sündroomi põhjustavas *MSH6* geenis tuvastatud variant c.3226C>T (p.Arg1076Cys) esines kahel erinevates keskustes (Hispaania, Portugal) uuringusse kaasatud patsientidel, kel omavahel teadaolevat sugulussidet ei olnud (tabel 11).

Tabel 11. Sekundaarsete leidude (n=29) jaotus valimi keskuste vahel.

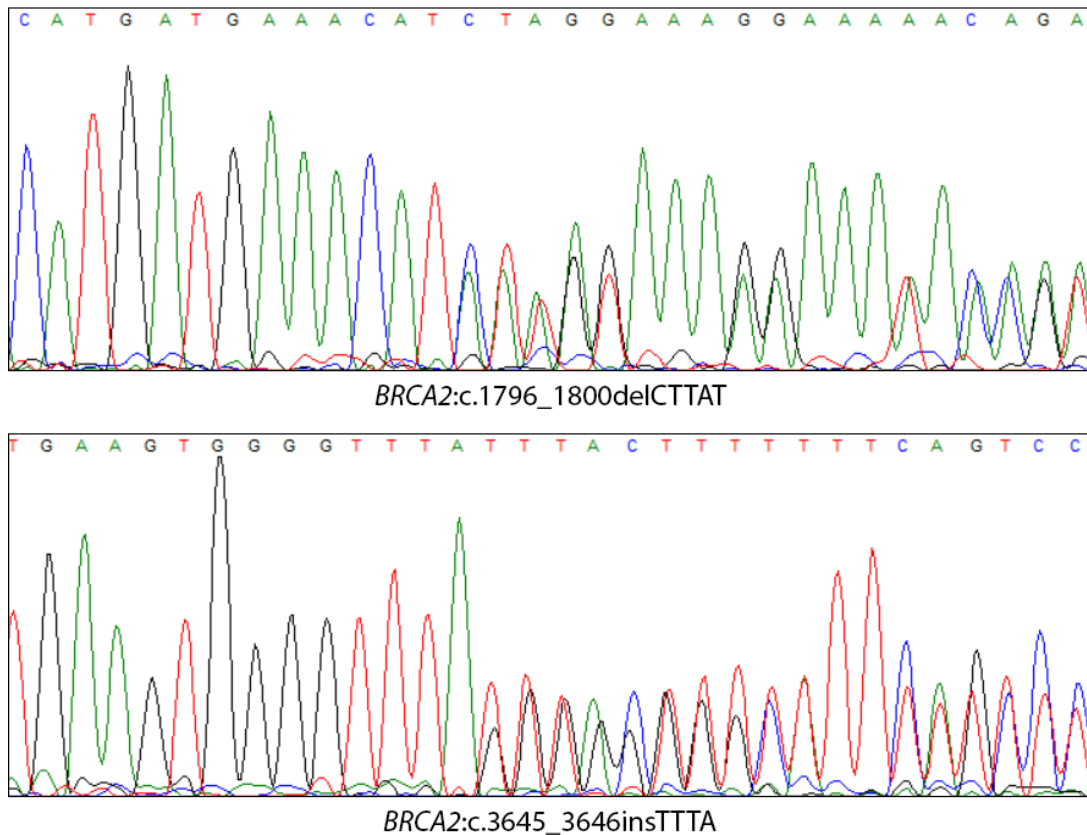
Keskus	Patsientide arv	Sekundaarsed leiud	Kandjad	
			n	%
Ameerika Ühendriigid	311	11	11	3,5
Cornell	205	6	6	2,9
Stanford	10	0	0	0
Utah	62	4	4	6,5
WashU	34	1	1	2,9
Austraalia	88	2	2	2,3
Eesti	82	2	2	2,4
Hispaania	46	1*	1	2,2
Kanada	35	0	0	0
Portugal (Porto + Lissabon)	275	13*	15	5,5
Taani	87	1	1	1,1

* Ühine variant.

Erinevates valimi keskustes esines sekundaarsete leidudega patsiente keskmiselt 2,6%. Enim sekundaarsete leidudega patsiente tuvastati Utah ja Portugali keskustes. Ühtegi sekundaarse leiuga patsienti ei tuvastatud väikse analüüsitud patsientide arvuga Stanfordini ja Kanada keskustes (vastavalt n=10 ja 35).

2.3.5.1 Eesti patsientide sekundaarsed leiud

Eesti 82-st NOA patsiendist tuvastati sekundaarseid leide kahel. Mõlemal puhul asus tuvastatud heterosügootne variant *BRCA2* geenis. Nendele variantidele viidi läbi sõltumatu meetodiga eksperimentaalne valideerimine Sangeri sekveneerimisega (joonis 9). Mõlemad variandid kinnitusid.



Joonis 9. Eesti patsientide *BRCA2* geenis tuvastatud sekundaarsete leidude sekveneerimispildid.

2.3.6 Eksoonsete harvade variantide koormus

Eesmärgina teha kindlaks, kas spermatogeneesi häirete ja kasvajate geneetiline etioloogia on omavahel ülekattes, viidi valimis esinevatele harvadele asendus- ja LoF variantidele läbi koormuse testimine. Analüüs viidi läbi vastavalt alapeatükis 2.2.8 sätestatud tingimustele.

Analüüsi filtritele vastavaid variante esines 65 geenis, millest kolm (*RET*, *ENG* ja *SCN5A*) andsid statistiliselt olulise ($p < 2,2 \times 10^{-5}$) tulemuse (tabel 12). *RET* puhul on tegu onkogeeniga, mis on sekundaarsete leidude nimekirja valitud, kuna patogeensed muutused selles geenis võivad põhjustada endokriinsete hulgikasvajate sündroomi või perekondlikku medullaarse kilpnäärme vähki. *ENG* ja *SCN5A* on KVS haiguste geenid, põhjustades vastavalt pärilikku hemorraagilist teleangiiektaasiat ja pika QT või Brugada sündroomi.

Tabel 12. Koormuse testi statistiliselt olulised tulemused käesoleva töö valimi (GEMINI) ja gnomAD andmebaasi eksoomiandmete vahel. Koormuse test põhines Fisheri täpsel testil.

Geen	GEMINI (n=924)			gnomAD NFE (n=56 885)			p-väärtus	OR
	HET	HOM	%	HET	HOM	%		
<i>ENG</i>	20	0	2,2	338	1	0,6	$1,8 \times 10^{-6}$	3,7
<i>RET</i>	16	0	1,7	201	0	0,4	$5,3 \times 10^{-7}$	5,0
<i>SCN5A</i>	13	0	1,4	143	0	0,3	$1,6 \times 10^{-6}$	5,7

NFE – eurooplased, millest on välja arvatud soomlased. HET – heterosügoodid. HOM – homosügoodid. OR – šansisuhe.

2.4 Arutelu

2.4.1 Eksoonne varieeruvus analüüsitud sekundaarsete leidudega seotud 85 geenis

Käesoleva töö käigus uuriti 924 idiopaatilise NOA patsiendi 85 geeni hõlmavaid eksoomiandmeid. Geenide nimekiri pandi kokku ACMG SF v2.0 (Kalia *et al.*, 2017), GHS76 (Schwartz *et al.*, 2018) ja ClinGen (Webber *et al.*, 2018) nimekirjade baasil, mis hõlmavad haigusgeene, milles leiduvatest patogeensetest variantidest tuleks patsiente teavitada.

Annoteeritud sekveneerimisandmetes leidis geenivariante kokku 6428. Ainsana ei esinenud varieeruvusi geenis *PLN*. *PLN* puhul on tegu küllaltki lühikese geeniga, mille transkripti pikkus on 2001 bp ja valgu pikkus 52 aminohapet. Kokku leidub gnomAD-is selles geenis vaid 19 asendus-,

neli LoF ja viis sünonüümset varianti, mistõttu on antud geenis varieeruvuste puudumine oodatav.

Pärast eksoomiandmete filtreerimist, mil välja jäeti <90% valimist kaetud, mitte-eksoonsed ja mitte-splaiiss-variandid ning sünonüümsed SNV-d, ei esinenud filtritele vastupidavaid varieeruvusi lisaks geenides *ACTC1*, *KCNE1*, *PTEN* ja *SDHC*. gnomAD andmetel esineb asendus- ja LoF variante *ACTC1* geenis kokku 47 (kõrgeim MAF eksoomiandmete põhjal 0,004%), *KCNE1* geenis 105 (60%), *PLN* geenis 13 (0,003%), *PTEN* geenis 82 (0,07%) ja *SDHC* geenis 121 (0,5%). Nendest geenidest on ainsana kõrge alleelisagedusega asendus- ja LoF variante sisaldav *KCNE1*, mistõttu oleks ka antud töös oodanud selle geeni varieeruvusi. Teiste nelja geeni varieeruvuste puudumine on oodatav.

Filtritele vastupidavatest variantidest (n=1643) leidis enim asendusvariante (96,5%). LoF variandid moodustasid kõigist variantidest 2,2%. Ülejäänud 1,3% olid lugemisraami mitte-muutvad indelid ning start- või stoppkoodoni kadu tekitavad variandid.

LoF variandid on sagedamini patogeensed (Pagel *et al.*, 2017; Richards *et al.*, 2015), mistõttu on arusaadav nende vähene esinemine võrreldes asendusvariantidega. Näiteks geen *TNN*, mille produktiks on inimese teadaolevalt pikim valk titiin, leidub gnomAD andmetel asendus- ja LoF variante kokku veidi üle 20 000, nendest 96% on asendusvariandid. Inimese teadaolevalt pikim geen on *DMD*, mille asendus- ja LoF variantidest moodustavad viimased vaid 2%. Seega on igati ootuspärane, et käesoleva töö valimis moodustavad väga suure osa asendusvariandid.

Eelmainitud LoF variantidest kahe puhul oli tegu sagedaste (MAF <1%) lugemisraami muutvate insertioonidega, mille haigusseoselisuse hinnang ClinVar andmetel oli THL/HL. Traditsiooniliselt on LoF variante hinnatud harvadeks ja kõrge tõenäosusega patogeenseteks. Samas on mitmetes uuringutes tuvastatud tervetel inimestel mitmeid LoF variante, mis ei oma mõju fenotübile (Narasimhan *et al.*, 2016), ning on isegi leitud, et osad LoF variandid võivad kasulikud olla. Näiteks LoF variandid geen *SLC30A8* vähendavad 2. tüüpi diabeedi riski 65% (Flannick *et al.*, 2014).

2.4.2 Sekundaarsete leidude profiil võrreldes teadustööde andmetega

Edasi uuriti vaid harvu (MAF $\leq 1\%$) ja andmebaasides varem esindamata variante (n=1478). Manuaalne haigusseoselisuse hindamine viidi läbi 96 geenivariandil ning sekundaarsed leiud tuvastati kogu analüüsitud ACMG SF v2.0, GHS76 ja ClinGen geeninimekirju arvestades 3,5% (n=32, tabel 13) ning vaid ACMG SF v2.0 nimekirja geene arvestades 3,2% (n=30) valimi patsientidest, mis on lähedane enamike tööde andmetele. Sekkumistõhusate sekundaarsete geneetiliste leidude osakaal erinevates teadustöödes on 1,2–7,9% ning vaid ACMG SF v1.0 ja 2.0 nimekirju arvestades 1,2–2,9% (tabel 3). Käesolevas töös tuvastatud leidude proportsioon on seega ootuspärane tulemus ja kattub varasemate valdavalt üldpopulatsioonis läbiviidud uuringutega.

MUTYH puhul on tegu onkogeeniga, milles bialleelsed patogeensed muutused võivad olla perekondliku adenomatoosse polüpoosi põhjuseks. Võrdluses olevatest töödest tuvastati homosügootseid indiviide peale käesoleva töö veel vaid Gordon *et al.* (2020) uuringus, kuhu oli kaasatud veidi üle 20 000 indiviidi (Gordon *et al.*, 2020). Mõlemates töödes leiti üks homosügootset varianti kandev indiviid, samas tuleb silmas pidada, et käesoleva töö valimi suurus oli pea 24 korda väiksem. Käesolevas uurimuses tuvastatud muutuse (c.1187G>A, p.Gly396Asp) MAF on gnomAD-is 0,3%, samas Gordon *et al.* (2020) töös leitud geenivariant (c.536A>G, p.Tyr179Cys) pole gnomAD andmebaasis esindatud. Sama olukord on ka geeniga *TGFBR2* (Loeys-Dietz'i sündroom), kus mõlemates töödes tuvastati üks selle geeni haigusseoselise variandi kandja. Käesoleva töö variandi (c.1067G>A, p.Arg356Gln) alleelisagedus gnomAD eksoomiandmete põhjal on 0,002% ning Gordon *et al.* (2020) töös tuvastatud varianti (c.1582C>T, p.Arg528Cys) pole jällegi gnomAD-is esindatud. Seega käesolevas töös tuvastatud variandid on tõenäoliselt kõrgema esinemissagedusega ning nende leidmine väiksema valimi hulgas pole ootamatu.

Tabel 13. Erinevates valimites tuvastatud sekundaarsete leidude esinemissagedus võrreldes käesoleva töö ja selles esindatud geenidega.

Geen	ACMG SF v1.0/2.0	Käesolev töö n=924 ACMG SF v2.0, GHS76, ClinGen	Gordon <i>et al.</i> (2020) n= 21 915 106 geeni (k.a ACMG SF v2.0)	Capalbo <i>et al.</i> (2019) n=14 125 ACMG SF v2.0	Hart <i>et al.</i> (2019) n=6240 ACMG SF v1.0	Thauvin-Robinet <i>et al.</i> (2019) n=700 244 geeni (k.a ACMG SF v2.0)	Chen <i>et al.</i> (2018) n=421 ACMG SF v2.0	Tang <i>et al.</i> (2018) n=954 ACMG SF v2.0
<i>ABCD1</i>	-	1/924	n.a	n.a	n.a	n.a	n.a	n.a
<i>BRCA1</i>	+	1/924	1/535	1/706	1/624	0	0	1/954
<i>BRCA2</i>	+	1/154	1/274	1/248	1/567	1/175	1/211	1/136
<i>FBN1</i>	+	1/308 ^a	1/1826	1/942	1/6240	0	1/421	0
<i>HNF1A</i>	-	1/924	1/7305	n.a	n.a	n.a	n.a	n.a
<i>KCNQ1</i>	+	1/924	1/877	1/1009	1/1560	1/233	0	0
<i>LDLR</i>	+	1/462	1/318	1/642	1/1560	0	1/421	1/954
<i>LMNA</i>	+	1/924	1/4383	1/14 125	0	0	0	0
<i>MSH6</i>	+	1/462 ^b	1/1565	1/1284	1/2080	0	0	0
<i>MUTYH*</i>	+	1/924	1/21 915	0	0	0	0	0
<i>MYBPC3</i>	+	1/924	1/487	1/673	1/2080	0	0	1/318
<i>MYH11</i>	+	1/924	0	0	0	0	0	1/954
<i>MYH7</i>	+	1/924	1/812	1/1009	1/2080	1/700	1/140	1/954
<i>PKP2</i>	+	1/924	1/1153	1/1413	1/6240	0	0	1/318

<i>PMS2</i>	+	1/924	1/877	1/1009	1/6240	0	1/211	1/954
<i>RET</i>	+	1/924	1/1565	1/2354	1/6240	0	0	0
<i>RYR1</i>	+	1/462	1/1289	1/642	1/780	0	0	0
<i>SCN5A</i>	+	1/462	1/812	1/2018	1/693	1/350	0	0
<i>TGFBR2</i>	+	1/924	1/21 915	0	0	0	0	0
<i>TNNI3</i>	+	1/924	1/7305	1/2825	1/6240	0	0	0
<i>WT1</i>	+	1/924	1/10 958	0	0	0	0	0

n.a – teadustöös polnud antud geen uurimisel.

* Näidatud homosügootide arv.

^a Ühine variant kolmel patsiendil.

^b Ühine variant kahel patsiendil.

Lisaks leiti peale käesoleva töö võrdluses olevatest töödest vaid Gordon *et al.* (2020) uuringus sekundaarseid leide *WT1* geenis. Käesolevas töös tuvastati *WT1* geenis üks ühenukleotiidne deletsioon (c.455delG, p.G152Afs*11) ning Gordon *et al.* (2020) töös kaks splaiss-varianti (c.1001+1G>A ja c.1432+5G>A) (Gordon *et al.*, 2020). Kõik kolm varianti on gnomAD andmebaasis esindamata. Antud töö patsiendil on see leid potentsiaalselt tema NOA põhjuseks, mis võib selgitada selle tuvastamist idiopaatilise NOA-ga meeste valimis ning miks üheski Gordon *et al.* (2020) uuringust väiksema valimiga töödes pole selles geenis sekundaarseid leide tuvastatud.

Käesolevas töös tuvastati üks *LMNA* geenis sekundaarset leidu kandev patsient (MAF 0,014%). Antud geenis on võrdluses olevatest töödest veel TP/P variante tuvastatud vaid kahes suurima valimiga uuringutes. Capalbo *et al.* (2019) töös tuvastati samuti üks leid (MAF 0,004%) ning Gordon *et al.* (2020) töös viis leidu (gnomAD-is esindamata) (Capalbo *et al.*, 2019; Gordon *et al.*, 2020). Käesoleva töö variant on seega tõenäoliselt kõige sagedasem.

Antud tööle sarnaste valimitega uuringutes ei tuvastatud sekundaarseid leide kaheksas geenis (8/19, 42%), milles antud valimi patsientidel leiti sekundaarseid leide (Chen *et al.*, 2018; Tang *et al.*, 2018; Thauvin-Robinet *et al.*, 2019). Neli nendest geenidest on ülalpool mainitud (*LMNA*, *MUTYH*, *TGFBR2*, *WT1*). Kolme geeni (*MSH6*, *RYR1*, *TNNI3*) puhul tuvastati TP/P variantide kandjaid kõigis kolmes suurima valimitega töödes (Capalbo *et al.*, 2019; Gordon *et al.*, 2020; Hart *et al.*, 2019). Tõenäoliselt jäi nii käesolevas töös kui ka teistes sarnaste valimi suurustega töödes nendes geenides sekundaarsete leidude kandjad lihtsalt valimi väiksuse tõttu tuvastamata.

Gordon *et al.* (2020) töös tuvastati *RET* geenis seitse haigusseoselist varianti 14 indiviidil (Gordon *et al.*, 2020), Hart *et al.* (2019) uuringus üks variant ühel kandjal ning Capalbo *et al.* (2019) töös üks variant kuuel kandjal (Capalbo *et al.*, 2019; Gordon *et al.*, 2020; Hart *et al.*, 2019). Gordon *et al.* (2020) uuringus seitsmel indiviidil ja Capalbo *et al.* (2019) töös kuuel indiviidil tuvastatud *RET* variant on ühine ning gnomAD eksoomiandmete järgi on MAF 0,012%. Teistest kuuest Gordon *et al.* (2020) töös leitud variandist on gnomAD-is esindatud kolm (MAF 0,001–0,002%), mis on käesolevas töös tuvastatud *RET* asendusvariandiga sarnase alleelisagedusega. Hart *et al.* (2019) töös tuvastatud varieeruvus oli veidi madalama alleelisagedusega.

Kõikide uuritud geenide kohta aga ei tuvastatud idiopaatilise NOA-ga meestel rohkem sekundaarseid leide kui varasemalt publitseeritud teadustöödes. Lisaks tasub mainimist, et teistes

teadustöodes leiti sekundaarsete leidude kandjaid mitmetes geenides, milles käesolevas töös haigusseoselisi geenivariante ei tuvastatud.

2.4.3 Tuvastatud sekundaarsete leidude potentsiaalne seos viljatusega

Sekundaarseid leide tuvastati käesolevas töös 21 geenis, nendest seitsme puhul on MGI (*Mouse Genome Informatics*) andmebaasis (Bult *et al.*, 2019) hiiremudelitel kirjeldatud kas azoospermiat või spermatogeneesi häireid. Nendeks geenideks on *BRCA1*, *BRCA2*, *HNF1A*, *LMNA*, *PMS2*, *RET* ja *WT1*. Sekundaarsete leidudega meestest esines antud geenides TP/P varieeruvusi 12 patsiendil ehk 38%.

BRCA1, *BRCA2* ja *PMS2* geenide produktid on seotud DNA reparatsiooniga. DNA parandamise defektide tõttu võivad häiruda nii meioosi kui mitoosi protsessid ja seeläbi ka spermatogenees. Antud töö käigus tuvastati ühel patsiendil *BRCA2* geenis homosügootne splaiss-variant c.8488-1G>A. Selle patsiendi puhul on antud variant defineeritud kui primaarne NOA põhjus (Nagirnaja *et al.*, 2021, avaldamisele saadetud).

HNF1A geeniproduktiks on küll maksaspetsiifiline transkriptsioonifaktor, kuid geen on ekspresseeritud ka munasarjades ja munandites. Varasemalt on kirjeldatud potentsiaalne seos *HNF1A* ja naistel esineva polütsüstiliste munasarjade sündroomi (PCOS, *polycystic ovary syndrome*) vahel (Castillo-Higuera *et al.*, 2020; Lv *et al.*, 2017). PCOS-iga naistel on sageli kõrgenenud LH ja alanenud FSH tasemed. Antud hormoonid on tähtsad ka mehe reproduktiivsüsteemis, täpsemalt spermatogeneesis, kus LH stimuleerib testosterooni produktsiooni ja FSH Sertoli rakke (Choy ja Amory, 2020). PCOS fenotüüpi on kirjeldatud ka *LMNA* patogeensete variantidega naistel (Vantyghem *et al.*, 2008). Seetõttu võib arvata, et antud geenid võivad potentsiaalselt olla seotud ka mehepoolse viljatusega.

RET ja *WT1* on olulised loote arengus. Hiirtel on näidatud, et *RET* valguproduktid on vajalikud diferentseerumata spermatogoniaalsete tüvirakkude säilitamiseks (Naughton *et al.*, 2006). *WT1* on tähtis meessoos determinatsioonil ja idurakkude arengul (Seabra *et al.*, 2015). Raske oligozoospermia ja NOA-ga patsientidel on *WT1* geenis tuvastatud erinevaid asendusvariante, millel on potentsiaalne seos viljatusega (Seabra *et al.*, 2015; Xu *et al.*, 2017). Käesolevas töös tuvastati ühel patsiendil *WT1* geenis heterosügootne ühenukleotiidne deletsioon c.455delG

(p.G152Afs*11), mis on tõenäoliselt patsiendi primaarne NOA põhjus. Antud patsiendi kohta on ka teada, et tal on diagnoositud krüptorhism ja Wilmsi tuumor (Donald F. Conrad, Portland, USA, kirjalik kommunikatsioon).

Sekundaarsetest leidudest tuvastati kaks varianti rohkem kui ühel patsiendil. *MSH6* geenis tuvastatud asendusvariant c.3226C>T (p.Arg1076Cys) esines kahel eri riigi patsiendil ning geenis *FBN1* tuvastatud asendusvariant c.287G>C (p.Arg96Thr) kolmel Portugali keskuste patsiendil. Kuna antud valimis tuvastati neid kahte varianti mitmel korral, võib nende leidude puhul samuti olla potentsiaalne seos viljatusega, küll aga pole neid gene varasemalt viljatusega seostatud. Samas *MSH6* puhul on tegu DNA reparatsiooni geeniga, mispuhul võib teoreetiliselt spermatogeneesis olla häirunud.

Peale *BRCA2* ja *WT1* geenides välja toodud primaarsete leidude võivad ülejäänud hiiremudelites viljatusega seotud geenides tuvastatud variandid olla potentsiaalsed spermatogeneesi häire tekke kaaspanustajad, tõstes põhivariandi penetrantsust. Seda ainult käesoleva töö andmete põhjal veel kinnitada ei saa ning vajalik on teostada lisauuringuid.

2.4.4 Sekundaarsete leidude potentsiaalne lisandväärtus viljatute meeste käsitluses

Eksoomi sekveneerimine on kulutõhus viis idiopaatilise viljatuse geneetiliste põhjuste välja selgitamiseks, kuid ka mitmete teiste kaasuvate haiguste diagnoosimiseks. Mitmetes teadustöodes on mehepoolse viljatuse puhul leitud kõrgem kaasuvate haiguste esinemissagedus ja kõrgem suremus (Del Giudice *et al.*, 2020b, 2020a; Glazer *et al.*, 2019). Seega ei peaks viljatu mehe ravi ja käsitlus fokuseeruma ainult reproduktiivsele tervisele, kuid hõlmama ka üldtervise hindamist, kuna kaasuvatel haigustel võib olla mõju mehe viljatusele ja ka oodatavale elueale. Sekundaarsete leidude tuvastamisel ja raporteerimisel on võimalus seda protsessi edendada ning seeläbi võimaldada viljatutele meestele varajast kaasuvate haiguste ennetust, diagnoosi ja ravi.

Lisaks võivad mitmed haigused ja viljatus olla põhjustatud ühiste geenide ja/või molekulaarsete radade häirumise tõttu (Tarín *et al.*, 2015). Seega tänu sekundaarsete leidude tuvastamisele võib juhtuda, et idiopaatiline viljatus võib leida molekulaarse diagnoosi. Käesolevas töös tuvastatud kahe variandi (*BRCA2*:c.8488-1G>A ja *WT1*:c.455delG) puhul on tõenäoliselt tegu varasemalt idiopaatilise viljatusega patsientide primaarsete NOA põhjustega.

Käesoleva töö käigus tuvastati lisaks nendele kahele variandile veel 27 sekundaarset leidu. Kõik leiud on ka edastatud uuringu juhi prof M. Laane poolt vastavatele rahvusvahelistele kliinilistele keskustele. Tuvastatud leiud on kas kinnitatud Sanger sekveneerimisega või on veel valideerimisel kas GEMINI koordineeriva partneri (Portland, USA) või vastava keskuse molekulaardiagnostika laborites. Sekundaarsete leidude edastamine GEMINI keskustes sõltub kaasatud patsientide informeeritud nõusoleku vormist ja patsiendi raviarsti otsusest. Hilisemas vanuses avalduva haiguse suhtes kõrgenenud riskiga patsientidele saab pakkuda nõustamist, et vajadusel alustada süstemaatilise jälgimise ja/või ennetava raviga, et haigust edasi lükata või vähendada haigusest tulenevaid komplikatsioone.

Eesti patsientidest tuvastati sekundaarseid leide kahel ning mõlemad variandid on käesoleva töö raames eksperimentaalselt kinnitatud Sanger sekveneerimisega (joonis 9). Ühel patsientidest tuvastati deletsioon c.1796_1800delCTTAT (p.Ser599*). Antud patsiendil on diagnoositud leiomüosarkoom, s.o harvaesinev silelihaskoes arenev vähkkasvaja (dr Margus Punab, TÜK Meestekliinik, suuline kommunikatsioon). Leiomüosarkoomi puhul on tihti tegu muutustega homoloogilise rekombinatsiooni rajal, sh *BRCA2* geen (Chudasama *et al.*, 2018). Erinevates uuringutes on *BRCA2* muutuseid tuvastatud kuni 10% leiomüosarkoomi jt sarkoomidega patsientidel, kuid seda suuremas osas naissoost valimites (Hensley *et al.*, 2020; Rosenbaum *et al.*, 2019; Seligson *et al.*, 2019).

Teisel Eesti patsiendil tuvastati insertioon c.3645_3646insTTTA (p.R1217Ifs*2). Antud patsiendi kohta puuduvad reproduktiivtervist mittepuudutavad andmed.

2.4.5 Eksoonsete harvade variantide rikastatus

Kirjanduses on püstitatud hüpotees, et spermatogeneesi häiretel ja kasvajate eelsoodumuse vahel on geneetiline seos (Nagirnaja *et al.*, 2018). Selle hüpoteesi kinnitamiseks viidi valimis esinevatele harvadele asendus- ja LoF variantidele läbi koormuse testimine vastu gnomAD andmebaasi NFE eksoomiandmeid. Antud koormuse testi käigus tuvastati idiopaatilise NOA-ga patsientide valimis geenivariantide rikastatus analüüsitud 32 kasvajate juhtgeenidest ainult onkogeenis *RET*.

Geeni *RET* geen-fenotüüp seosed OMIM andmebaasis (Online Mendelian Inheritance in Man, OMIM®. Kasutatud 24.05.2021) on kaasasündinud tsentraalne hüpoventilatsiooni sündroom (MIM#209880), medullaarne kilpnäärme vähk (MIM#155240), endokriinsete hulgikasvajate sündroom (MIM#171400, #162300) ja feokromotsütoom (MIM#171300). MGI andmebaasis on hiiremudelitel täheldatud *RET* geeniga seoses reproduktiivsüsteemi häireid, täpsemalt on kirjeldatud vääniliste seemnetorude degradatsiooni, krüptorhismi, spermatogeneesi häireid, ebanormaalset seemnejuha morfoloogiat ning viljatust. Kuigi kirjanduse andmetel ei põhjusta *RET* geeni muutused inimesel azoospermiat, püstitavad käesoleva uurimistöö andmed hüpoteesi nende võimalikust panustavast rollist spermatogeneesi häire kujunemisel.

Teadaolevalt on kasvajate juhtgeene 299 (Bailey *et al.*, 2018), samas sekundaarsete leidude nimekirjades oli nendest geenidest esindatud väike hulk (21/299, 7%), mistõttu ei saanud käesolevas töös spermatogeneesi häirete ja kasvajate geneetilise etioloogia seost usaldusväärselt testida.

Statistiliselt oluline ($p < 2,21 \times 10^{-5}$) rikastatus harvade variante koormuse suhtes tuvastati veel geenides *ENG* ja *SCN5A*. *ENG* geeni puhul on OMIM andmebaasi järgi ainus geen-fenotüüp seos päriliku hemorraagilise teleangiiektaasiaga (MIM#187300), mille tõttu on antud geen ka sekundaarsete leidude nimekirjas. *SCN5A* geen-fenotüüp seosed autosoom-dominantse pärandumisega haigustest on perekondlik kodade virvendusarütmia (MIM#614022), Brugada sündroom (MIM#601144), dilatatiivne kardiomiopaatia (MIM#601154), progresseeruv ja mitteprogresseeruv südame seiskumine (MIM#113900) ning pika QT sündroom (MIM#603830). Antud geene pole küll viljatusega otseselt seostatud, kuid nagu kirjanduse ülevaates mainitud, esineb viljatutel meestel mitmeid muid haiguseid, sh KVS süsteemi haiguseid rohkem kui viljakatel meestel, mis võib seletada antud analüüsis tuvastatud rikastatust. Pigem võib aga olla nendes geenides tuvastatud harvade variantide rikastatus juhuslik ja põhjustatud väiksest valimist.

Kuigi idiopaatilise NOA-ga patsientidel leiti kolmes geenis harvade variantide kandluse osas statistiliselt oluline rikastatus võrreldes gnomAD NFE üldpopulatsiooni valimiga, oli vaid ühe puhul tegu kasvajatega seotud geeniga. Seega geenivariantide rikastatus tuvastati antud valimis 4,8% analüüsitud kasvajate juhtgeenidest. See kinnitab, et idiopaatilise NOA-ga patsiendid ei eristu kasvajatega seotud geenide varieeruvuse määras ja profiilis üldpopulatsioonist.

KOKKUVÕTE

Käesoleva magistritöö esimeses peatükis anti ülevaade mehepoolsest viljatusest, spermatogeneesist ning viljatuse seosest üldtervisega. Lisaks kirjeldati sekundaarsete leidude olemust, esinemissagedust üldpopulatsioonis, tagasisidestamist ja geenivariantide patogeensuse hindamist vastavalt Ameerika Meditsiinigeneetika ja Genoomika Kolleegiumi (ACMG) juhistele.

Töö eesmärkideks oli tuvastada 85 geenist koosneva nimekirja alusel kliiniliselt olulised sekundaarsed leiud GEMINI konsortsiumi idiopaatilise mitte-obstruktiivse azoospermiaga (NOA) meeste valimis (n=924) ning teha kindlaks, kas idiopaatilise NOA-ga meestel esineb võimalikke sekkumistõhusaid geenivariante rohkem kui üldpopulatsioonis. Lisaks püstitati hüpotees, et raske spermatogeneesi häiretega patsientidel esineb onkogeenides rikastatus võrreldes üldpopulatsiooniga.

Haigusseoselisuse hindamine viidi läbi 96 geenivariandil. Antud töö käigus tuvastati 29 sekundaarset leidu 21 geenis 32 idiopaatilise NOA-ga patsiendil (3,5%) ning vaid ACMG nimekirja arvestades 27 leidu 30 patsiendil (3,2%). Antud tulemused sarnanevad varasematele teadustöödele, kus on tuvastatud sekundaarseid leide üldpopulatsioonil umbes 1–3%. Käesolevas töös tuvastatud leidude proportsioon on seega ootuspärane tulemus.

Ühe *BRCA2* leiu puhul oli tegu homosügootse variandiga ning antud leid on suure tõenäosusega patsiendi primaarne NOA põhjus. Lisaks tuvastati ühel patsiendil ühenukleotiidne deletsioon *WT1* geenis, mis võib samuti olla potentsiaalselt patsiendi primaarne leid.

Kasvajatega seotud geenides esinevate variantide rikastatuse uurimiseks teostati harvade geenivariantide koormuse test. Võrdluseks kasutati gnomAD eksoomiandmeid. Analüüsi käigus tuvastati rikastatus kardiovaskulaarsüsteemi haigustega seotud geenides *ENG* ja *SCN5A* ning vaid ühes onkogeenis *RET*. Hiiremudelitel on näidatud *RET* geeni seos spermatogeneesi häirete ja viljatusega. Kuigi kirjanduse andmetel ei põhjusta *RET* geeni muutused inimesel azoospermiat, püstitavad käesoleva uurimistöö andmed hüpoteesi nende võimalikust panustavast rollist spermatogeneesi häire kujunemisel.

Geenivariantide rikastatus tuvastati antud valimis kasvajate juhtgeenidest 4,8%. See näitab, et idiopaatilise NOA-ga patsiendid ei eristu kasvajate geenide varieeruvuse määras ja profiilis üldpopulatsioonist. Samas oli sekundaarsete leidude nimekirjades vähe kasvajatega seotud geene, mistõttu võis raske spermatogeneesi häirete ja kasvajate ühine geneetiline etioloogia jääda tuvastamata.

Androloogia praktikas on sekundaarsete leidude uurimine olulise väärtusega nii patsientide enda tervise ja pereplaneerimise kui ka nende pereliikmetega seoses. Seeläbi võib lisaks tuvastada ka potentsiaalseid viljatusele kaasperanustajaid.

Secondary findings in infertile men and their clinical significance

Kristiina Lillepea

SUMMARY

Infertility is defined as the inability to achieve pregnancy after one year, despite having regular unprotected intercourse. It affects approximately 15% of couples worldwide and about half of the cases are due to male-factor infertility. Azoospermia, which is the absence of spermatozoa in the ejaculate, can be observed in about 1% of men. In most cases it is caused by spermatogenic failure, termed non-obstructive azoospermia (NOA).

Known genetic causes (e.g. Klinefelter syndrome and Y chromosome microdeletions) account for up to 25% of NOA cases and up to 20% still remain idiopathic. NOA is a very heterogeneous condition, since up to 10% of the male genome is considered to be involved in reproduction. Because of this, exome or genome sequencing is the most cost-effective method to reach a molecular diagnosis for NOA patients.

During exome/genome sequencing it is also possible to determine clinically important and potentially actionable genetic variants that are not associated with the patient's primary condition, termed secondary findings (SFs). American College of Medical Genetics and Genomics (ACMG) has published a minimum list of genes that should be evaluated in individuals undergoing clinical exome/genome sequencing based on the medical actionability of the associated condition. This list has three versions, the newest version having been published just this May. Similar gene lists have also been published by Geisinger Health System and ClinGen. Since it is strongly recommended to only report clinically actionable variants, ACMG and the Association for Molecular Pathology have jointly developed guidelines for the assessment of evidence in clinical variant interpretation.

Because exome sequencing is a good way to potentially determine the genetic cause of male-factor infertility and considering infertile men have been described to have impaired overall health, it is convenient to also look for SFs in patients with NOA.

The aim of this master's thesis was to identify medically actionable genetic variants in men with idiopathic NOA (n=924) and determine whether the prevalence of SFs is higher than in the general

population. In addition, it was hypothesized that men with severe spermatogenic failure have an enrichment of oncogenic variants.

Manual curation for pathogenicity was carried out on 96 genetic variants. 29 actionable findings were identified, with 15 classified as pathogenic and 14 as likely pathogenic. SFs were located in 21 different genes and identified in 32 patients (3.5%). Of these, 27 findings in 30 patients were variants in genes recommended by the ACMG (SF v2.0). This prevalence is similar to previous publications, where SFs have been identified in about 1–3% of the general population.

One patient was homozygous for a pathogenic splice-site variant in *BRCA2*. It is highly likely that this SF is actually the primary cause of NOA in this patient. One single nucleotide deletion was identified in *WT1*, which could also potentially be the primary cause of the patient's NOA.

A burden test was performed to determine whether men with severe spermatogenic failure have an enrichment of genetic variants found in cancer-related genes. gnomAD exome sequencing data of non-Finnish Europeans (n=56,885) was used as a control cohort. Three genes showed significant enrichment ($p < 2.2 \times 10^{-5}$), of which only one was an oncogene (*RET*, $p = 5.3 \times 10^{-7}$; odds ratio, OR=5.0). Two others were genes associated with cardiovascular conditions (*ENG*, $p = 1.8 \times 10^{-6}$; OR=3.7 and *SCN5A*, $p = 1.6 \times 10^{-6}$; OR=5.7). Therefore, the results suggest that men with severe spermatogenic failure do not have an enrichment of oncogenic variants. However, the amount of tested cancer-related genes was limited, which is why an association between spermatogenic failure and cancer cannot be excluded.

In mouse models, *RET* has previously been associated with abnormal spermatogenesis and infertility. Although literature data does not support the association of *RET* and azoospermia in humans, the presented data in this thesis proposes the hypothesis of its potential co-contributing role in the development of spermatogenic failure.

In clinical practice, SFs have significant value for not just patients themselves regarding their health and family planning, but also for their relatives. Through SFs, it is also possible to determine potential co-contributing factors for infertility.

TÄNUSÕNAD

Suurimad tänud minu juhendajatele Laura Kasak ja Maris Laan suurepärase juhendamise ja heade nõuannete eest töö valmimisel. Samuti soovin tänada Maarja Lepametsa abi eest eksoomiandmete katvuse analüüsil. Tänuõnad kuuluvad ka minu perekonnale, elukaaslasele ja sõpradele, kes mind õpingute ja töö valmimise ajal toetasid.

KASUTATUD KIRJANDUS

Agarwal, A., Mulgund, A., Hamada, A., ja Chyatte, M.R. (2015). A unique view on male infertility around the globe. *Reprod. Biol. Endocrinol.* *13*.

Agarwal, A., Baskaran, S., Parekh, N., Cho, C.L., Henkel, R., Vij, S., Arafa, M., Panner Selvam, M.K., ja Shah, R. (2021). Male infertility. *Lancet* *397*, 319–333.

Al-Jebari, Y., Elenkov, A., Wirestrand, E., Schütz, I., Giwercman, A., ja Lundberg Giwercman, Y. (2019). Risk of prostate cancer for men fathering through assisted reproduction: Nationwide population based register study. *BMJ* *366*.

Andreson, R., Reppo, E., Kaplinski, L., ja Remm, M. (2006). GENOMEMASKER package for designing unique genomic PCR primers. *BMC Bioinformatics* *7*, 172.

Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., *et al.* (2018). Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* *173*, 371-385.e18.

Bieniek, J.M., Kashanian, J.A., Deibert, C.M., Grober, E.D., Lo, K.C., Brannigan, R.E., Sandlow, J.I., ja Jarvi, K.A. (2016). Influence of increasing body mass index on semen and reproductive hormonal parameters in a multi-institutional cohort of subfertile men. *Fertil. Steril.* *106*, 1070–1075.

Bieniek, J.M., Lapin, C.D., ja Jarvi, K.A. (2021). Genetics of CFTR and male infertility. *Transl Androl Urol* *10*, 1391–1400.

Bieth, E., Hamdi, S.M., ja Mieusset, R. (2021). Genetics of the congenital absence of the vas deferens. *Hum. Genet.* *140*, 59–76.

Boeri, L., Capogrosso, P., Ventimiglia, E., Pederzoli, F., Frego, N., Cazzaniga, W., Chierigo, F., Alfano, M., Piemonti, L., Viganò, P., *et al.* (2019). Undiagnosed prediabetes is highly prevalent in primary infertile men – results from a cross-sectional study. *BJU Int.* *123*, 1070–1077.

Bombieri, C., Claustres, M., De Boeck, K., Derichs, N., Dodge, J., Girodon, E., Sermet, I., Schwarz, M., Tzetis, M., Wilschanski, M., *et al.* (2011). Recommendations for the classification of diseases as CFTR-related disorders. *J. Cyst. Fibros.* *10*, S86–S102.

Boycott, K., Hartley, T., Adam, S., Bernier, F., Chong, K., Fernandez, B.A., Friedman, J.M., Geraghty, M.T., Hume, S., Knoppers, B.M., *et al.* (2015). The clinical application of genome-wide sequencing for monogenic diseases in Canada: Position statement of the Canadian College of medical geneticists. *J. Med. Genet.* *52*, 431–437.

Bult, C.J., Blake, J.A., Smith, C.L., Kadin, J.A., Richardson, J.E., Anagnostopoulos, A., Asabor, R., Baldarelli, R.M., Beal, J.S., Bello, S.M., *et al.* (2019). Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* *47*, D801–D806.

Capalbo, A., Valero, R.A., Jimenez-Almazan, J., Pardo, P.M., Fabiani, M., Jiménez, D., Simon, C., ja Rodriguez, J.M. (2019). Optimizing clinical exome design and parallel gene-testing for recessive genetic conditions in preconception carrier screening: Translational research genomic data from 14,125 exomes. *PLoS Genet.* *15*.

- Castillo-Higuera, T., Alarcón-Granados, M.C., Marin-Suarez, J., Moreno-Ortiz, H., Esteban-Pérez, C.I., Ferrebuz-Cardozo, A.J., Forero-Castro, M., ja Camargo-Villalba, G. (2020). A Comprehensive Overview of Common Polymorphic Variants in Genes Related to Polycystic Ovary Syndrome. *Reprod. Sci.* 1–14.
- Chen, W., Li, W., Ma, Y., Zhang, Y., Han, B., Liu, X., Zhao, K., Zhang, M., Mi, J., Fu, Y., *et al.* (2018). Secondary findings in 421 whole exome-sequenced Chinese children. *Hum. Genomics* 12, 42.
- Choy, J.T., ja Amory, J.K. (2020). Nonsurgical management of oligozoospermia. *J. Clin. Endocrinol. Metab.* 105, 1–14.
- Chudasama, P., Mughal, S.S., Sanders, M.A., Hübschmann, D., Chung, I., Deeg, K.I., Wong, S.H., Rabe, S., Hlevnjak, M., Zapatka, M., *et al.* (2018). Integrative genomic and transcriptomic analysis of leiomyosarcoma. *Nat. Commun.* 9.
- Ehala-Aleksejev, K., ja Punab, M. (2015). The different surrogate measures of adiposity in relation to semen quality and serum reproductive hormone levels among Estonian fertile men. *Andrology* 3, 225–234.
- Eisenberg, M.L., Betts, P., Herder, D., Lamb, D.J., ja Lipshultz, L.I. (2013). Increased risk of cancer among azoospermic men. *Fertil. Steril.* 100, 681–685.e1.
- Eisenberg, M.L., Kim, S., Chen, Z., Sundaram, R., Schisterman, E.F., ja Buck Louis, G.M. (2014). The relationship between male BMI and waist circumference on semen quality: Data from the LIFE study. *Hum. Reprod.* 29, 193–200.
- Eisenberg, M.L., Li, S., Behr, B., Pera, R.R., ja Cullen, M.R. (2015a). Relationship between semen production and medical comorbidity. *Fertil. Steril.* 103, 66–71.
- Eisenberg, M.L., Li, S., Brooks, J.D., Cullen, M.R., ja Baker, L.C. (2015b). Increased risk of cancer in infertile men: Analysis of U.S. claims data. *J. Urol.* 193, 1596–1601.
- Eisenberg, M.L., Chen, Z., Ye, A., ja Buck Louis, G.M. (2015c). Relationship between physical occupational exposures and health on semen quality: Data from the Longitudinal Investigation of Fertility and the Environment (LIFE) study. *Fertil. Steril.* 103, 1271–1277.
- Eisenberg, M.L., Li, S., Cullen, M.R., ja Baker, L.C. (2016). Increased risk of incident chronic medical conditions in infertile men: Analysis of United States claims data. *Fertil. Steril.* 105, 629–636.
- Ferlin, A., Garolla, A., Ghezzi, M., Selice, R., Palego, P., Caretta, N., Di Mambro, A., Valente, U., De Rocco Ponce, M., Dipresa, S., *et al.* (2021). Sperm Count and Hypogonadism as Markers of General Male Health. *Eur. Urol. Focus* 7, 205–213.
- Flannick, J., Thorleifsson, G., Beer, N.L., Jacobs, S.B.R., Grarup, N., Burt, N.P., Mahajan, A., Fuchsberger, C., Atzmon, G., Benediktsson, R., *et al.* (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat. Genet.* 46, 357–363.
- Del Giudice, F., Kasman, A.M., Li, S., Belladelli, F., Ferro, M., de Cobelli, O., De Berardinis, E., Busetto, G.M., ja Eisenberg, M.L. (2020a). Increased Mortality Among Men Diagnosed With Impaired Fertility: Analysis of US Claims Data. *Urology* 147, 143–149.

- Del Giudice, F., Kasman, A.M., Ferro, M., Sciarra, A., De Berardinis, E., Belladelli, F., Salonia, A., ja Eisenberg, M.L. (2020b). Clinical correlation among male infertility and overall male health: A systematic review of the literature. *Investig Clin Urol* *61*, 355–371.
- Glazer, C.H., Bonde, J.P., Giwercman, A., Vassard, D., Pinborg, A., Schmidt, L., ja Vaclavik Bräuner, E. (2017). Risk of diabetes according to male factor infertility: A register-based cohort study. *Hum. Reprod.* *32*, 1474–1481.
- Glazer, C.H., Eisenberg, M.L., Tøttenborg, S.S., Giwercman, A., Flachs, E.M., Bräuner, E.V., Vassard, D., Pinborg, A., Schmidt, L., ja Bonde, J.P. (2019). Male factor infertility and risk of death: A nationwide record-linkage study. *Hum. Reprod.* *34*, 2266–2273.
- Gordon, A.S., Zouk, H., Venner, E., Eng, C.M., Funke, B.H., Amendola, L.M., Carrell, D.S., Chisholm, R.L., Chung, W.K., Denny, J.C., *et al.* (2020). Frequency of genomic secondary findings among 21,915 eMERGE network participants. *Genet. Med.* *22*, 1470–1477.
- Green, R.C., Berg, J.S., Grody, W.W., Kalia, S.S., Korf, B.R., Martin, C.L., McGuire, A.L., Nussbaum, R.L., O’Daniel, J.M., Ormond, K.E., *et al.* (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet. Med.* *15*, 565–574.
- Guo, M.H., Plummer, L., Chan, Y.M., Hirschhorn, J.N., ja Lippincott, M.F. (2018). Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data. *Am. J. Hum. Genet.* *103*, 522–534.
- Hanson, H.A., Anderson, R.E., Aston, K.I., Carrell, D.T., Smith, K.R., ja Hotaling, J.M. (2016). Subfertility increases risk of testicular cancer: evidence from population-based semen samples. *Fertil. Steril.* *105*, 322–328.e1.
- Hart, M.R., Biesecker, B.B., Blout, C.L., Christensen, K.D., Amendola, L.M., Bergstrom, K.L., Biswas, S., Bowling, K.M., Brothers, K.B., Conlin, L.K., *et al.* (2019). Secondary findings from clinical genomic sequencing: prevalence, patient perspectives, family history assessment, and health-care costs from a multisite study. *Genet. Med.* *21*, 1100–1110.
- Hensley, M.L., Chavan, S.S., Solit, D.B., Murali, R., Soslow, R., Chiang, S., Jungbluth, A.A., Bandlamudi, C., Srinivasan, P., Tap, W.D., *et al.* (2020). Genomic Landscape of Uterine Sarcomas Defined through Prospective Clinical Sequencing. *Clin. Cancer Res.* *26*, 3881–3888.
- Houge, G., Laner, A., Cirak, S., de Leeuw, N., Scheffer, H., ja den Dunnen, J.T. (2021). Stepwise ABC system for classification of any type of genetic variant. *Eur. J. Hum. Genet.* *22*, 1956–1966.
- Van Hout, C. V., Tachmazidou, I., Backman, J.D., Hoffman, J.D., Liu, D., Pandey, A.K., Gonzaga-Jauregui, C., Khalid, S., Ye, B., Banerjee, N., *et al.* (2020). Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature* *586*, 749–756.
- Hufnagel, S.B., Martin, L.J., Cassidy, A., Hopkin, R.J., ja Antommaria, A.H.M. (2016). Adolescents’ preferences regarding disclosure of incidental findings in genomic sequencing that are not medically actionable in childhood. *Am. J. Med. Genet. Part A* *170*, 2083–2088.
- Hume, S., Nelson, T.N., Speevak, M., McCreedy, E., Agatep, R., Feilotter, H., Parboosingh, J., Stavropoulos, D.J., Taylor, S., ja Stockley, T.L. (2019). CCMG practice guideline: Laboratory guidelines for next-generation sequencing. *J. Med. Genet.* *56*, 792–800.

- Hwang, K., Smith, J.F., Coward, R.M., Penzias, A., Bendikson, K., Butts, S., Coutifaris, C., Falcone, T., Fossum, G., Gitlin, S., *et al.* (2018). Evaluation of the azoospermic male: a committee opinion. *Fertil. Steril.* *109*, 777–782.
- Jungwirth, A., Diemer, T., Kopa, Z., Krausz, C., Minhas, S., ja Tournaye, H. (2018). EAU Guidelines on Male Infertility.
- Kalia, S.S., Adelman, K., Bale, S.J., Chung, W.K., Eng, C., Evans, J.P., Herman, G.E., Hufnagel, S.B., Klein, T.E., Korf, B.R., *et al.* (2017). Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): A policy statement of the American College of Medical Genetics and Genomics. *Genet. Med.* *19*, 249–255.
- Karczewski, K.J., Weisburd, B., Thomas, B., Solomonson, M., Ruderfer, D.M., Kavanagh, D., Hamamsy, T., Lek, M., Samocha, K.E., Cummings, B.B., *et al.* (2017). The ExAC browser: Displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* *45*, D840–D845.
- Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., *et al.* (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* *581*, 434–443.
- Kasak, L., ja Laan, M. (2021). Monogenic causes of non-obstructive azoospermia: challenges, established knowledge, limitations and perspectives. *Hum. Genet.* *140*, 135–154.
- Kasman, A.M., Del Giudice, F., ja Eisenberg, M.L. (2020). New insights to guide patient care: the bidirectional relationship between male infertility and male health. *Fertil. Steril.* *113*, 469–477.
- Kelly, M.A., Leader, J.B., Wain, K.E., Bodian, D., Oetjens, M.T., Ledbetter, D.H., Martin, C.L., ja Strande, N.T. (2021). Leveraging population-based exome screening to impact clinical care: The evolution of variant assessment in the Geisinger MyCode research project. *Am. J. Med. Genet. Part C Semin. Med. Genet.* *187*, 83–94.
- Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C.E., Albarca Aguilera, M., Meyer, R., ja Massouras, A. (2019). VarSome: the human genomic variant search engine. *Bioinformatics* *35*, 1978–1980.
- Köressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R., ja Remm, M. (2018). Primer3-masker: Integrating masking of template sequence with primer design software. *Bioinformatics* *34*, 1937–1938.
- Krausz, C., Forti, G., ja McElreavey, K. (2003). The Y chromosome and male fertility and infertility. *Int. J. Androl.* *26*, 70–75.
- Krausz, C., Riera-Escamilla, A., Moreno-Mendoza, D., Holleman, K., Cioppi, F., Algaba, F., Pybus, M., Friedrich, C., Wyrwoll, M.J., Casamonti, E., *et al.* (2020). Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. *Genet. Med.* *22*, 1956–1966.
- Landrum, M.J., ja Kattman, B.L. (2018). ClinVar at five years: Delivering on the promise. *Hum. Mutat.* *39*, 1623–1630.

- Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., *et al.* (2018). ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* *46*, D1062–D1067.
- Latif, T., Jensen, T.K., Mehlsen, J., Holmboe, S.A., Brinth, L., Pors, K., Skouby, S.O., Jørgensen, N., ja Lindahl-Jacobsen, R. (2017). Semen Quality as a Predictor of Subsequent Morbidity: A Danish Cohort Study of 4,712 Men with Long-Term Follow-up. *Am. J. Epidemiol.* *186*, 910–917.
- Levine, H., Jørgensen, N., Martino-Andrade, A., Mendiola, J., Weksler-Derri, D., Mindlis, I., Pinotti, R., ja Swan, S.H. (2017). Temporal trends in sperm count: A systematic review and meta-regression analysis. *Hum. Reprod. Update* *23*, 646–659.
- Li, H., ja Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* *25*, 1754–1760.
- Li, Q., ja Wang, K. (2017). InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am. J. Hum. Genet.* *100*, 267–280.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., ja Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* *25*, 2078–2079.
- Lv, Y., Sun, C., Tian, Y., Zhao, S., Bian, Y., Cheng, L., Sun, M., Liu, H. Bin, Zhao, H., ja Ma, J. (2017). Association study of HNF1A in women with polycystic ovary syndrome. *J. Assist. Reprod. Genet.* *34*, 677–682.
- Matthijs, G., Souche, E., Alders, M., Corveleyn, A., Eck, S., Feenstra, I., Race, V., Sistermans, E., Sturm, M., Weiss, M., *et al.* (2016). Guidelines for diagnostic next-generation sequencing. *Eur. J. Hum. Genet.* *24*, 2–5.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303.
- Middleton, A., Morley, K.I., Bragin, E., Firth, H. V., Hurles, M.E., Wright, C.F., ja Parker, M. (2016). Attitudes of nearly 7000 health professionals, genomic researchers and publics toward the return of incidental results from sequencing research. *Eur. J. Hum. Genet.* *24*, 21–29.
- Mighton, C., Carlsson, L., Clausen, M., Casalino, S., Shickh, S., McCuaig, L., Joshi, E., Panchal, S., Semotiuk, K., Ott, K., *et al.* (2020). Quality of life drives patients' preferences for secondary findings from genomic sequencing. *Eur. J. Hum. Genet.* *28*, 1178–1186.
- Miller, D.T., Lee, K., Chung, W.K., Gordon, A.S., Herman, G.E., Klein, T.E., Stewart, D.R., Amendola, L.M., Adelman, K., Bale, S.J., *et al.* (2021a). ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.*
- Miller, D.T., Lee, K., Gordon, A.S., Adelman, K., Bale, S.J., Chung, W.K., Gollob, M.H., Harrison, S.M., Herman, G.E., Hershberger, R.E., *et al.* (2021b). Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2021 update: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet. Med.*

Monroy, A. (2020). Fertilization.

Nagirnaja, L., Aston, K.I., ja Conrad, D.F. (2018). Genetic intersection of male infertility and cancer. *Fertil. Steril.* 109, 20–26.

Nagirnaja, L., Lopes, A.M., Charng, W.-L., Miller, B., Stakaitis, R., Golubickaite, I., Stendahl, A., Luan, T., Friedrich, C., Mahyari, E., *et al.* (2021). Diverse Monogenic Subforms of Human Spermatogenic Failure.

Narasimhan, V.M., Hunt, K.A., Mason, D., Baker, C.L., Karczewski, K.J., Barnes, M.R., Barnett, A.H., Bates, C., Bellary, S., Bockett, N.A., *et al.* (2016). Health and population effects of rare gene knockouts in adult humans with related parents. *Science* (80-.). 352, 474–477.

Naughton, C.K., Jain, S., Strickland, A.M., Gupta, A., ja Milbrandt, J. (2006). Glial cell-line derived neurotrophic factor-mediated RET signaling regulates spermatogonial stem cell fate. *Biol. Reprod.* 74, 314–321.

Oduwole, O.O., Peltoketo, H., ja Huhtaniemi, I.T. (2018). Role of follicle-stimulating hormone in spermatogenesis. *Front. Endocrinol. (Lausanne)*. 9, 763.

Olesen, I.A., Andersson, A.M., Aksglaede, L., Skakkebaek, N.E., Rajpert-de Meyts, E., Joergensen, N., ja Juul, A. (2017). Clinical, genetic, biochemical, and testicular biopsy findings among 1,213 men evaluated for infertility. *Fertil. Steril.* 107, 74-82.e7.

Pagel, K.A., Pejaver, V., Lin, G.N., Nam, H.J., Mort, M., Cooper, D.N., Sebat, J., Iakoucheva, L.M., Mooney, S.D., ja Radivojac, P. (2017). When loss-of-function is loss of function: Assessing mutational signatures and impact of loss-of-function genetic variants. In *Bioinformatics*, lk i389–i398.

Pagin, A., Bergougnoux, A., Girodon, E., Reboul, M.P., Willoquaux, C., Kesteloot, M., Raynal, C., Bienvenu, T., Humbert, M., Lalau, G., *et al.* (2020). Novel ADGRG2 truncating variants in patients with X-linked congenital absence of vas deferens. *Andrology* 8, 618–624.

Pergialiotis, V., Prodromidou, A., Frountzas, M., Korou, L.M., Vlachos, G.D., ja Perrea, D. (2016). Diabetes mellitus and functional sperm characteristics: A meta-analysis of observational studies. *J. Diabetes Complications* 30, 1167–1176.

Pervola, J., Myers, M.F., McGowan, M.L., ja Prows, C.A. (2019). Giving adolescents a voice: the types of genetic information adolescents choose to learn and why. *Genet. Med.* 21, 965–971.

Practice Committee of the American Society for Reproductive Medicine in collaboration with the Society for Male Reproduction and Urology (2018). Evaluation of the azoospermic male: a committee opinion. *Fertil. Steril.* 109, 777–782.

Pujol, P., Vande Perre, P., Faivre, L., Sanlaville, D., Corsini, C., Baertschi, B., Anahory, M., Vaur, D., Olschwang, S., Soufir, N., *et al.* (2018). Guidelines for reporting secondary findings of genome sequencing in cancer genes: the SFMPP recommendations. *Eur. J. Hum. Genet.* 26, 1732–1742.

Punab, M., Poolamets, O., Paju, P., Vihljajev, V., Pomm, K., Ladva, R., Korrovits, P., ja Laan, M. (2017). Causes of male infertility: A 9-year prospective monocentre study on 1737 patients with reduced total sperm counts. *Hum. Reprod.* 32, 18–31.

- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., *et al.* (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* *17*, 405–424.
- Rosenbaum, E., Jonsson, P., Seier, K., Chi, P., Dickson, M.A., Gounder, M.M., Kelly, C.M., Keohan, M.L., Qin, L.-X., Hensley, M.L., *et al.* (2019). DNA damage response pathway alterations and clinical outcome in leiomyosarcoma. *J. Clin. Oncol.* *37*, 11048–11048.
- Sapp, J.C., Johnston, J.J., Driscoll, K., Heidlebaugh, A.R., Miren Sagardia, A., Dogbe, D.N., Umstead, K.L., Turbitt, E., Alevizos, I., Baron, J., *et al.* (2018). Evaluation of Recipients of Positive and Negative Secondary Findings Evaluations in a Hybrid CLIA-Research Sequencing Pilot. *Am. J. Hum. Genet.* *103*, 358–366.
- Schisterman, E.F., Mumford, S.L., Chen, Z., Browne, R.W., Boyd Barr, D., Kim, S., ja Buck Louis, G.M. (2014). Lipid concentrations and semen quality: The LIFE study. *Andrology* *2*, 408–415.
- Schwartz, M.L.B., McCormick, C.Z., Lazzeri, A.L., Lindbuchler, D.M., Hallquist, M.L.G., Manickam, K., Buchanan, A.H., Rahm, A.K., Giovanni, M.A., Frisbie, L., *et al.* (2018). A Model for Genome-First Care: Returning Secondary Genomic Findings to Participants and Their Healthcare Providers in a Large Research Cohort. *Am. J. Hum. Genet.* *103*, 328–337.
- Seabra, C.M., Quental, S., Lima, A.C., Carvalho, F., Gonçalves, J., Fernandes, S., Pereira, I., Silva, J., Marques, P.I., Sousa, M., *et al.* (2015). The mutational spectrum of WT1 in male infertility. *J. Urol.* *193*, 1709–1715.
- Sebastian, A., Carroll, J.C., Vanstone, M., Clausen, M., Kodida, R., Reble, E., Mighton, C., Shickh, S., Aronson, M., Eisen, A., *et al.* (2021). Widening the lens of actionability: A qualitative study of primary care providers' views and experiences of managing secondary genomic findings. *Eur. J. Hum. Genet.* 1–9.
- Seligson, N.D., Kautto, E.A., Passen, E.N., Stets, C., Toland, A.E., Millis, S.Z., Meyer, C.F., Hays, J.L., ja Chen, J.L. (2019). BRCA1/2 Functional Loss Defines a Targetable Subset in Leiomyosarcoma. *Oncologist* *24*, 973–979.
- Sermondade, N., Faure, C., Fezeu, L., Shayeb, A.G., Bonde, J.P., Jensen, T.K., Van Wely, M., Cao, J., Martini, A.C., Eskandar, M., *et al.* (2013). BMI in relation to sperm count: An updated systematic review and collaborative meta-analysis. *Hum. Reprod. Update* *19*, 221–231.
- Shiraishi, K., ja Matsuyama, H. (2018). Effects of medical comorbidity on male infertility and comorbidity treatment on spermatogenesis. *Fertil. Steril.* *110*, 1006-1011.e2.
- Sutovsky, P., ja Manandhar, G. (2006). Mammalian spermatogenesis and sperm structure: anatomical and compartmental analysis. In *The Sperm Cell: Production, Maturation, Fertilization, Regeneration*, C.J. de Jonge, ja C. Barratt, toim (Cambridge), lk 1–30.
- Tang, C.S. man, Dattani, S., So, M. ting, Cherny, S.S., Tam, P.K.H., Sham, P.C., ja Garcia-Barcelo, M.M. (2018). Actionable secondary findings from whole-genome sequencing of 954 East Asians. *Hum. Genet.* *137*, 31–37.

Tarín, J.J., García-Pérez, M.A., Hamatani, T., ja Cano, A. (2015). Infertility etiologies are genetically and clinically linked with other diseases in single meta-diseases. *Reprod. Biol. Endocrinol.* *13*.

Thauvin-Robinet, C., Thevenon, J., Nambot, S., Delanne, J., Kuentz, P., Bruel, A.L., Chassagne, A., Cretin, E., Pelissier, A., Peyron, C., *et al.* (2019). Secondary actionable findings identified by exome sequencing: expected impact on the organisation of care from the study of 700 consecutive tests. *Eur. J. Hum. Genet.* *27*, 1197–1214.

The Editors of Encyclopaedia Britannica (2020). Spermatogenesis.

Tournaye, H., Krausz, C., ja Oates, R.D. (2017). Novel concepts in the aetiology of male reproductive impairment. *Lancet Diabetes Endocrinol.* *5*, 544–553.

Tüttelmann, F., Werny, F., Cooper, T.G., Kliesch, S., Simoni, M., ja Nieschlag, E. (2011). Clinical experience with azoospermia: Aetiology and chances for spermatozoa detection upon biopsy. *Int. J. Androl.* *34*, 291–298.

Vantyghem, M.C., Vincent-Desplanques, D., Defrance-Faivre, F., Capeau, J., Fermon, C., Valat, A.S., Lascols, O., Hecart, A.C., Pigny, P., Delemer, B., *et al.* (2008). Fertility and obstetrical complications in women with LMNA-related familial partial lipodystrophy. *J. Clin. Endocrinol. Metab.* *93*, 2223–2229.

Ventimiglia, E., Capogrosso, P., Boeri, L., Serino, A., Colicchia, M., Ippolito, S., Scano, R., Papaleo, E., Damiano, R., Montorsi, F., *et al.* (2015). Infertility as a proxy of general male health: Results of a cross-sectional survey. *Fertil. Steril.* *104*, 48–55.

Ventimiglia, E., Capogrosso, P., Colicchia, M., Boeri, L., Serino, A., Castagna, G., Clementi, M.C., La Croce, G., Regina, C., Bianchi, M., *et al.* (2016). Metabolic syndrome in white European men presenting for primary couple's infertility: investigation of the clinical and reproductive burden. *Andrology* *4*, 944–951.

Wang, K., Li, M., ja Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* *38*.

Webber, E.M., Hunter, J.E., Biesecker, L.G., Buchanan, A.H., Clarke, E. V., Currey, E., Dagan-Rosenfeld, O., Lee, K., Lindor, N.M., Martin, C.L., *et al.* (2018). Evidence-based assessments of clinical actionability in the context of secondary findings: Updates from ClinGen's Actionability Working Group. *Hum. Mutat.* *39*, 1677–1685.

de Wert, G., Dondorp, W., Clarke, A., Dequeker, E.M.C., Cordier, C., Deans, Z., van El, C.G., Fellmann, F., Hastings, R., Hentze, S., *et al.* (2021). Opportunistic genomic screening. Recommendations of the European Society of Human Genetics. *Eur. J. Hum. Genet.* *29*, 365–377.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York).

World Health Organization (2010). *Laboratory manual for the examination and processing of human semen* (Geneva).

Wyrwoll, M.J., Temel, Ş.G., Nagirnaja, L., Oud, M.S., Lopes, A.M., van der Heijden, G.W., Heald, J.S., Rotte, N., Wistuba, J., Wöste, M., *et al.* (2020). Bi-allelic Mutations in M1AP Are a Frequent Cause of Meiotic Arrest and Severely Impaired Spermatogenesis Leading to Male Infertility. *Am. J. Hum. Genet.* *107*, 342–351.

Xu, J., Jiang, L., Yu, W., Guo, H., Zhang, H., Wei, D., Liang, L., Feng, K., Song, X., Liu, Q., *et al.* (2017). A novel functional variant in Wilms' Tumor 1 (WT1) is associated with idiopathic non-obstructive azoospermia. *Mol. Reprod. Dev.* *84*, 222–228.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., ja Madden, T.L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* *13*, 134.

Zegers-Hochschild, F., Adamson, G.D., de Mouzon, J., Ishihara, O., Mansour, R., Nygren, K., Sullivan, E., ja Vanderpoel, S. (2009). International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertil. Steril.* *92*, 1520–1524.

KASUTATUD VEEBIAADRESSID

<https://www.cdc.gov/genomics/implementation/toolkit/tier1.htm>, kasutatud 04.03.2021.

<http://clinvar.com>, viimati kasutatud 24.03.2021.

<https://franklin.genoox.com>, viimati kasutatud 24.03.2021.

<https://gemini.conradlab.org>, kasutatud 13.01.2021.

<https://gnomad.broadinstitute.org>, viimati kasutatud 31.05.2021.

<http://www.informatics.jax.org>, kasutatud 24.05.2021.

https://www.kliinikum.ee/yhendlabor/index.php?option=com_content&view=article&id=32&Itemid=44, kasutatud 25.05.2021.

<https://www.omim.org>, kasutatud 24.05.2021.

<https://varsome.com>, viimati kasutatud 24.03.2021.

Lihtlitsents lõputöö elektrooniliseks avaldamiseks

Mina, Kristiina Lillepea,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Sekundaarsed geneetilised leiud viljatutel meestel ja nende kliiniline olulisus“, mille juhendajad on Laura Kasak ja Maris Laan, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, alates **03.06.2024** kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kristiina Lillepea

03.06.2021