

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

Kadri Onemar

**Eksoomi sekveneerimise rakendamine kliinilises praktikas 11
haigusjuhtumi näitel**

Bakalaureusetöö

Juhendajad: Margit Nõukas, M.Sc
prof Ants Kurg, Ph.D

TARTU 2014

SISUKORD

SISUKORD	2
LÜHENDID JA MÕISTED	4
SISSEJUHATUS	5
1 KIRJANDUSE ÜLEVAADE	6
1.1 DNA Sekveneerimine.....	6
1.1.1 Esimese põlvkonna sekveneerimine.....	6
1.1.2 Teise põlvkonna sekveneerimine	7
1.1.2.1 Roche/454 pürosekveneerimine	8
1.1.2.2 Applied Biosystems/SOLiD.....	10
1.1.2.3 Illumina/Solexa	11
1.2 Eksoomi sekveneerimine	13
1.2.1 Eksoomi väljapüüdmine.....	14
1.3 Mendeliaalsed haigused.....	16
1.3.1 Mendeliaalsete haiguste uurimine.....	16
1.3.1.1 Mendeliaalsete haiguste uurimine eksoomi sekveneerimisega.....	17
1.3.1.1.1 Uuritavate patsientide valiku strateegiad eksoomi sekveneerimisel....	18
1.3.2 Autosoom-retsessiivsete haiguste uurimine kogu-eksoomi sekveneerimisega.	20
1.3.3 Autosoom-dominantsete haiguste uurimine kogu-eksoomi sekveneerimisega	20
1.3.3.1 Eksoomi sekveneerimine ja <i>de novo</i> mutatsioonid	21
1.3.4 Eksoomi sekveneerimise puudused.....	22
2 EKSPERIMENTAALOSA	24
2.1 Töö eesmärgid	24
2.2 Materjal ja meetodika.....	25
2.2.1 Valimi kirjeldamine.....	25
2.2.2 Genoomse DNA eraldamine	25
2.2.3 Eksoomi sekveneerimine ja andmete analüüsimine.....	25
2.2.3.1 Eksoomi sekveneerimine.....	25
2.2.3.2 Eksoomi sekveneerimisandmete analüüs TÜ Eesti Geenivaramus.....	26
2.2.3.3 Andmete interpretatsioon	26
2.2.4 Eksoomi sekveneerimise tulemuste valideerimine Sangeri sekveneerimisega.	28
2.2.4.1 PCR reaktsiooni läbiviimine	28
2.2.4.2 Sekveneerimisreaktsioon.....	29
2.3 Tulemused	30

2.3.1	Patsient EST5	30
2.3.2	Patsient EST10	33
2.4	Arutelu	36
KOKKUVÕTE		39
SUMMARY		40
TÄNUAVALDUSED.....		41
KASUTATUD KIRJANDUS		42
KASUTATUD VEEBILEHED.....		48

LÜHENDID JA MÕISTED

ACMG	Ameerika Meditsiinigeneetika Kolledž (<i>American College of Medical Genetics and Genomics</i>)
ASD	autismispektri häired (<i>autism spectrum disorders</i>)
bp	aluspaar (<i>base pair</i>)
ID	intellektipuue (<i>intellectual disability</i>)
Mb	megabaas (<i>megabase</i>)
NGS	teise põlvkonna sekveneerimine (<i>next-generation sequencing</i>)
OMIM	Mendeliaalsete haiguste andmebaas (<i>Online Mendelian Inheritance in Man</i>)
WES	kogu-eksoomi sekveneerimine (<i>whole-exome sequencing</i>)
WGS	kogu-genoomi sekveneerimine (<i>whole-genome sequencing</i>)
ARID1B	<i>AT rich interactive domain 1B</i>
CHD7	<i>chromodomain helicase DNA-binding protein 7</i>
DHODH	<i>dihydroorotate dehydrogenase</i>
MLL2	<i>myeloid/lymphoid or mixed-lineage leukemia 2</i>
RUNX2	<i>runt-related transcription factor 2</i>
SETBP1	<i>SET binding protein 1</i>

SISSEJUHATUS

Viimastel aastatel on seoses tehnoloogia arengu ja uute sekveneerimismeetodite kasutuselevõtuga saanud võimalikuks järjestada ühe eksperimendi tulemusena kogu inimese genoomi või analüüsida ainult valke kodeerivaid järjestusi ehk eksoneid.

Eksoom moodustab umbes 1-2% kogu inimese genoomist (Ng *et al.*, 2009). Arvatakse, et enamik haiguspõhjuslikest mutatsioonidest võib paikneda just DNA kodeerivas alas, mistõttu on eksoomi sekveneerimisest saanud väga tõhus geenidiagnostika tööriist kliinilises praktikas.

Eksoomi sekveneerimine on olnud väga edukas just Mendeliaalsete haiguste geneetiliste põhjuste väljaselgitamisel. Eksoomi sekveneerimise tulemusena on võimalik lahendada keerulisi haigusjuhtumeid, mida seniste diagnostiliste võimalustega on olnud võimatu diagnoosida.

Kuigi eksoomi sekveneerimine on oluliselt kiirendanud ja lihtsustanud haigusseoseliste geenivariantide uurimist, esineb sellel analüüsimeetodil ka teatud puudusi. Vaatamata sellele võimaldab eksoomi sekveneerimine ebaselge etioloogiaga arenguhäiretega patsientide puhul selgitada välja häire põhjuse umbes 25% juhtudest (Yang *et al.*, 2013). Täpne diagnoos võimaldab pakkuda adekvaatset geneetilist nõustamist patsientide peredele, hinnata paremini haiguse kordusriske perekonnas ja rakendada võimalikku ravi.

Käesoleva töö eesmärgiks oli uurida võimalike haiguspõhjuslike geenivariantide esinemist ebaselge etioloogiaga arenguhäiretega patsientide valimis eksoomi sekveneerimise meetodi abil.

1 KIRJANDUSE ÜLEVAADE

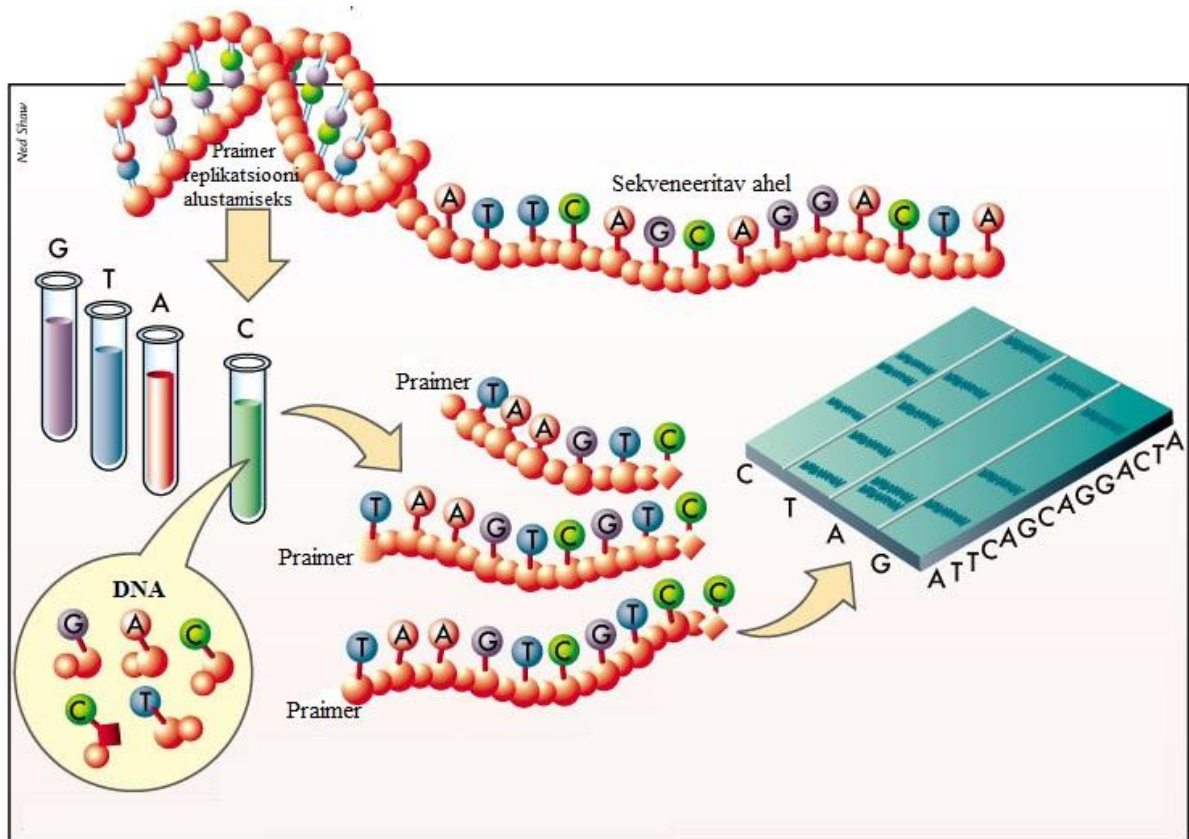
1.1 DNA sekveneerimine

Sekveneerimine ehk järjendamine on meetod, mis võimaldab määrata nukleiinhapete nukleotiidsed järjestust ja polüpeptiidide aminohappelist järjestust. DNA sekveneerimiseks nimetatakse DNA molekuli primaarse struktuuri nukleotiidsed järjestuse määramist. Kuigi sekveneerimise tehnoloogiad on aastate jooksul muutunud, on peamised põhimõtted jäänud samaks.

1.1.1 Esimese põlvkonna sekveneerimine

1977. aastal arendasid Maxam ja Gilbert mitte-ensümaatilise DNA sekveneerimise meetodi, mis oli tehniliselt keerukas ja kasutas ohtlikke kemikaale (Maxam and Gilbert, 1977). Samal aastal arendas Frederick Sanger ensümaatilise DNA sekveneerimise tehnoloogia, mida nimetatakse ka esimese põlvkonna sekveneerimise meetodiks. Ensümaatiline tehnoloogia põhineb ahela sünteesi peatamise meetodil, kasutades dideoksüribonukleotiide ehk terminaatoreid (ddNTP). Kuna dideoksüribonukleotiidil puudub positsioonis 3' hüdroksüülrühm, siis DNA polümeraas pole võimeline ahelat edasi sünteesima. Ensümaatilise reaktsiooni tulemusena toimub uuritava DNA järjestuse põhjal uute ahelate süntees ning tekivad erineva pikkusega DNA fragmendid, mille määrab ära reaktsioonis osalevate deoksüribonukleotiidide (dNTP) ja ddNTP suhe. Saadud reaktsiooni produktid eraldatakse geelelektroforeesil, et määrata DNA järjestus (Sanger *et al.*, 1977). Sanger sekveneerimise tööpõhimõte on esitatud joonisel 1.

Aastate jooksul on Sangeri sekveneerimismeetodit täiustatud. Näiteks on DNA märgistamine radioaktiivse isotoobiga asendatud fluorestsentsmärgisega ddNTP-ga (Smith *et al.*, 1986), sest fluorestsentsmärgis on inimese tervisele ohutum ning võimaldab ka DNA sekveneerimise automatiseerimist. 1990. aastal võetigi geelelektroforeesi asemel esmakordselt kasutusele automatiseeritud kapillaargeelektroforees (Luckey *et al.*, 1990). Sellised uuendused on muutnud DNA sekveneerimise Sangeri meetodiga kiiremaks ja efektiivsemaks, võimaldades tänapäeval sünteesida kuni 1000 aluspaariseid DNA fragmente (Shendure and Ji, 2008).



Joonis 1. Sangeri sekvenerimine (Winnick, 2004) järgi. Üheaahelaline DNA segatakse praimeriga ja jaotatakse nelja tuubi, mis sisaldavad DNA polümeraasi, nelja deoksüribonukleotiidi (dGTP, dCTP, dATP, dTTP) ja ühte replikatsiooni terminaatorit (ddNTP). Iga reaktsioon kestab kuni ahelasse lisatakse terminaator-nukleotiid. Reaktsioonisegud kantakse geelile erinevatesse radadesse ja DNA fragmendid eraldatakse geelelektroforeesiga.

1.1.2 Teise põlvkonna sekvenerimine

Viimastel aastatel on seoses tehnoloogia arengu ja uute sekvenerimismeetodite kasutuselevõtuga toimunud pööre suuremahuliste inimese genoomi uuringute läbiviimiseks populatsiooni- ja meditsiinigeneetikas¹. Nüüd on võimalik järjestada ühe eksperimendi tulemusena kogu inimese genoomi (*WGS, whole-genome sequencing*) või analüüsida ainult valke kodeerivaid järjestusi ehk eksoneid (*WES, whole-exome sequencing*). Teise põlvkonna sekvenerimine on massiivselt paralleelne sekveneerimistehnoloogia, mis võimaldab tuvastada miljonite DNA matriitsahelate järgi sünteesitud fragmentide nukleotiidset järjestust samaaegselt (Majewski *et al.*, 2011), lugedes selleks nukleotiidide ahelasse lisamisel eralduvat signaali. Teise põlvkonna sekvenerimismeetodid (*NGS, next-generation sequencing*) on väga võimas tööriist, mille abil saab usaldusväärselt uurida inimese geneetilist varieeruvust ja seoseid erinevate haigustega.

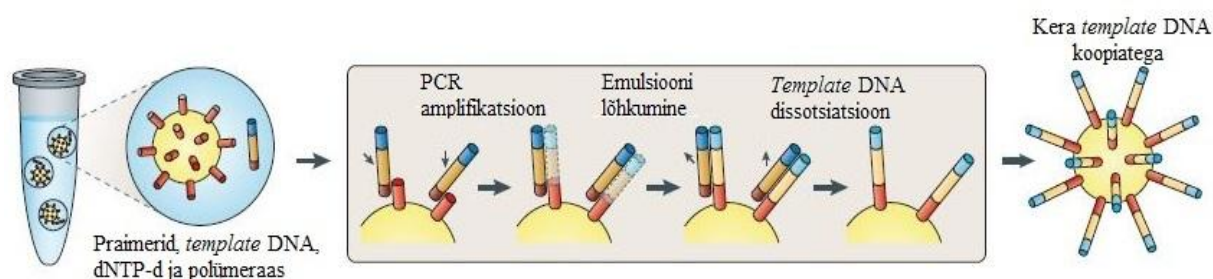
¹ Vaata <http://www.mu.ee/?PublicationId=de4ed9b5-3713-4e97-a33c-76270cf264e4>

2004. aastal valmis Inimese Genoomi Projekt (*HGP, Human Genome Project*), mille raames määrati inimese genoomi referentsjärjestus, mis koostati mitme inimese genoomi põhjal (International Human Genome Sequencing Consortium, 2004). Projekti valmimine andis tõe DNA sekveneerimistehnoloogiate edasi arendamiseks (Shendure *et al.*, 2004; Metzker, 2005). Kuigi esimene *NGS* masin valmis 2004. aastal (Morey *et al.*, 2013), siis kommertsiaalselt kättesaadavaks said *NGS* platvormid alles 2005. aastal (Margulies *et al.*, 2005). Tänapäeval on laialdaselt kasutusel kolm *NGS* sekveneerimistehnoloogiat, milleks on Roche/454 pürosekveneerimine, Illumina/Solexa ja Applied Biosystems/SOLiD (Mardis, 2008).

1.1.2.1 Roche/454 pürosekveneerimine

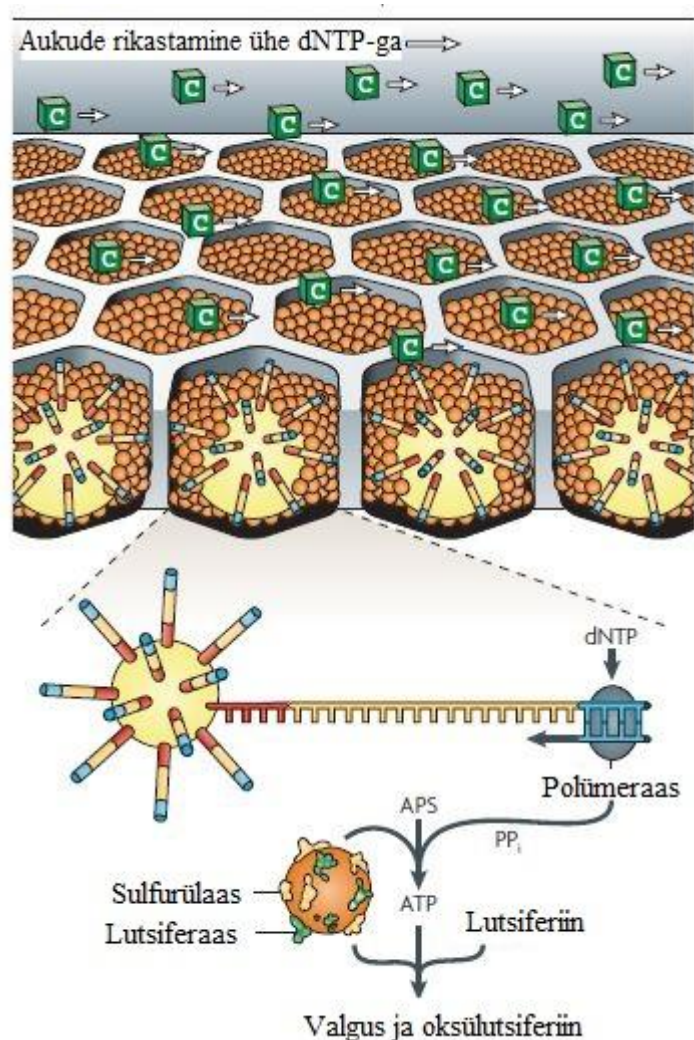
454 oli esimene kommertsiaalselt kättesaadav teise põlvkonna sekveneerimisplatvorm, mis kasutab pürosekveneerimise tehnoloogiat (Margulies *et al.*, 2005). Selle platvormiga sekveneeriti esimene inimese täisgenoom, milleks oli James D. Watsoni genoom (Wheeler *et al.*, 2008). Sekveneerimisele eelneb raamatukogu koostamine ja amplifikatsioon, mis toimub emulsioon polümeraasi ahelreaktsiooni ehk PCR-i meetodil (Mardis, 2008).

Genoomne DNA lõigatakse fragmentideks ja segatakse keradega, millel asuvad oligonukleotiidsed järjestused on komplementaarsed fragmentidele lisatud järjestustega (joonis 2). Iga keraga seotakse ainult üks fragment. DNA-kerade kompleksid eraldatakse emulsiooni tilkadesse, kus asuvad PCR reaktsiooniks vajalikud reagentid. Emulsiooni tilgad toimivad kui mikroreaktorid. Toimub amplifikatsioon ja kera kaetakse tuhandate DNA järjestustega. Peale amplifikatsiooni lõhutakse emulsioon, mille käigus eraldatakse kerale kinnitumata fragmendid. Saadakse kera, mis on kaetud üheaheelaliste DNA fragmentidega (Mardis, 2008).



Joonis 2. Emulsioon PCR (Metzker, 2010) järgi. DNA fragment isoleeritakse mikroreaktorisse, mis on ümbritsetud õlifaasiga. Mikroreaktor sisaldab praimeritega kaetud kerakesi ja PCR reaktsiooniks vajalikke reagente. Mikroreaktorites toimub PCR reaktsioonid, mille järel on terve kera kaetud *template* DNA koopiatega. Emulsioon lõhutakse ja *template* DNA lagundatakse.

Peale amplifitseerimist toimub järjestuste sekveneerimine. DNA fragmentidega kaetud kerad asetatakse pikotiiterplaadile nii, et igasse auku kinnitub ainult üks kera (joonis 3). See tingib iga kera kindla asukoha, mistõttu on võimalik iga sekveneerimisreaktsiooni jälgida CCD (*charge-coupled device*) kaameraga. Pikotiiterplaadile lisatakse kerad, millele on kinnitatud sulfüülaasid ja lutsiferaasid. Plaat rikastatakse korduvalt nelja erineva dNTP-ga ja järjestusse sobiva dNTP korral seotakse ta DNA polümeraasi poolt. Iga nukleotiidi ühendamine ahelasse vabastab pürofosfaadi. Sulfüülaas seob pürofosfaadi adenosiin-5'-fosfosulfaadiga (APS), moodustades ATP, mis on vajalik valguse tootmiseks lutsiferaasiga. Valguse intensiivsus sõltub seotud nukleotiidide arvust ehk mida rohkem seotakse nukleotiide ühe reaktsiooni käigus, seda tugevam on saadav signaal (Mardis, 2008; Morey *et al.*, 2013).

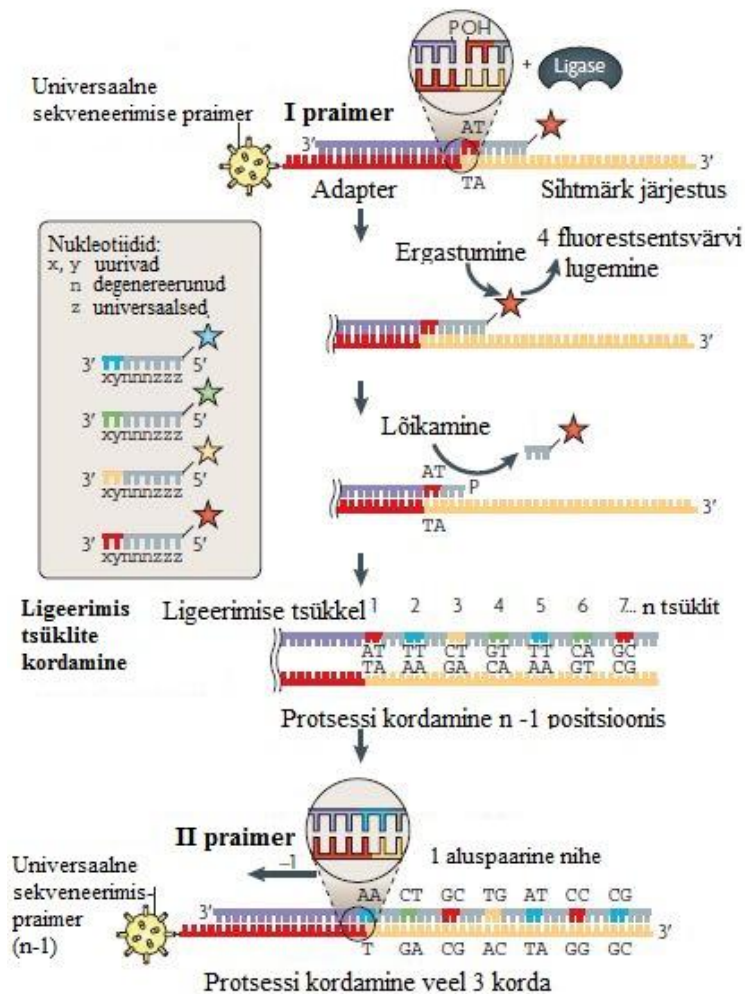


Joonis 3. Pürosekveneerimine (Metzker, 2010) järgi. Igas augus asub üks DNA fragmentidega kaetud kera. Augud rikastatakse kordamööda erinevate nukleotiididega. Õige nukleotiidi seondumisel eraldub pürofosfaat ja järgnevatel reaktsioonidel eraldub valgus.

1.1.2.2 Applied Biosystems/SOLiD

SOLiD (Sequencing by Oligo Ligation Detection) platvormil toimub sekveneerimine ligeerimise teel ja sarnaselt 454 pürosekveneerimisplatvormile kasutab SOLiD fragmentide amplifitseerimiseks emulsioon PCR-i (joonis 2). Amplifitseeritud fragmendid kinnitatakse paramagneetilistele keradele ja peale emulsiooni lõhkumist immobiliseeritakse fragmente kandvad kerad klaaspinnale (Shendure and Ji, 2008).

Erinevalt 454 pürosekveneerimisplatvormist, toimub SOLiD platvormil sekveneerimine DNA ligaasiga mitte polümeraasiga. Adapterjärjestusele komplementaarsed universaalsed praimerid hübridiseeritakse fragmentidele, mis asuvad kera pinnal (joonis 4). Iga sekveneerimise tsükkel sisaldab fluorestsentsmärgisega oktameeride ligeerimist, kus fluorestsentsmärgis on seotud viimase kolme nukleotiidiga. Sekveneerimine toimub kahe aluspaari kaupa, milleks on oktameeri kaks esimest nukleotiidi (joonisel 4 tähistatud kui x ja y). Selliselt sekveneeritakse ühte nukleotiidi kaks korda, mistõttu on võimalik eristada tekkinud vigu polümorfismidest. Oktameeri ligeerimisel ergastatakse fluorestsentsmärgis ja loetakse eralduv kiirgus, mille järel lõigatakse kolm viimast nukleotiidi ülejäänud viie nukleotiidi küljest lahti. Lõikamine vabastab fosfaatrühma, mistõttu on võimalik alustada uut tsükli. Ühe praimeriga toimub 10 tsükli, mille järel see praimer degradeeritakse ja järjestusega seotakse uus praimer. Uus praimer asub teises positsioonis (n-1), mis nihutab oktameeri seondumissaidi ühe aluspaari võrra vasakule. DNA fragmentide sekveneerimiseks kasutatakse viite erinevas positsioonis praimerit. Üks praimer võimaldab sekveneerida iga viiendat nukleotiidi ja viie erineva praimeriga kasutamisel sekveneeritakse matriits ahelale komplementaarne järjestus (Shendure and Ji, 2008; Metzker, 2010).



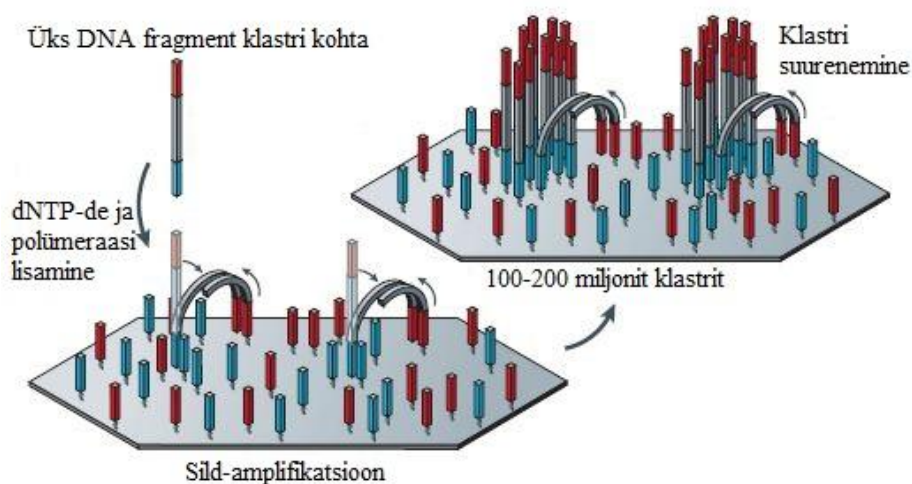
Joonis 4. Ligeerimise teel sekveneerimine (Metzker, 2010) järgi. Adapterjärjestusele hübridiseeritakse universaalne praimer. Praimeri külge ligeeritakse sobivate aluspaaridega oktameer, mille järel loetakse fluorestsents ja eemaldatakse kolm viimast nukleotiidi. Sekveneerimine viiakse läbi viie praimeriga, mis asuvad erinevates positsioonides.

1.1.2.3 Illumina/Solexa

Solexa platvorm on tänapäeval enim kasutatud teise põlvkonna sekveneerimisplatvorm. Solexa tehnoloogia põhineb sekveneerimisel sünteesi käigus põhimõttel ja fragmentide amplifikatsioon toimub silla-PCR meetodil (Morey *et al.*, 2013).

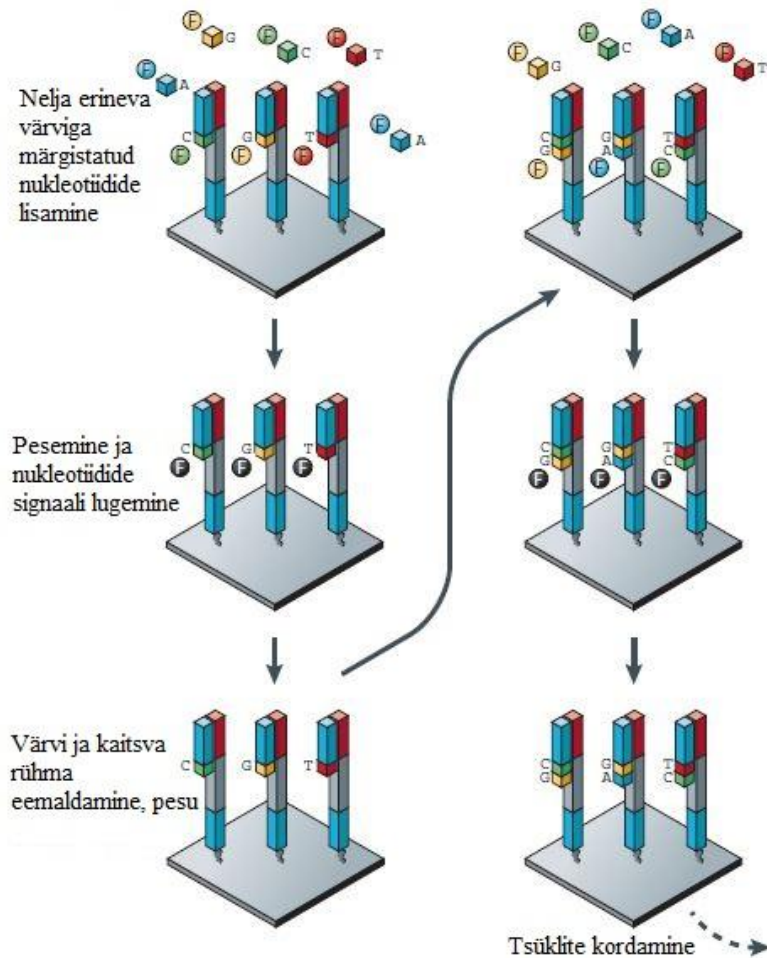
Sild-amplifikatsioon võimaldab samaaegselt amplifitseerida suurt hulka DNA fragmente, kasutades tahket oligonukleotiididega kaetud pinda, mida kutsutakse ka *flow cell*'ks (joonis 5) (Morey *et al.*, 2013). Sild-amplifitseerimisel seotakse päri- ja vastassuunalised praimerid painduva *linker*'i kaudu tahkele pinnale nii, et sünteesitavad fragmendid jäävad immobiliseerituks ja on klasterdunud kindlale kohale kiibil (Shendure and Ji, 2008). Oligonukleotiidid kinnitatakse *linker*'ile 5' otsapidi, jättes 3' otsa polümeraasi jaoks vabaks. Pinnale seotud oligonukleotiidid on komplementaarsed järjestustele, mis liideti DNA fragmentidele raamatukogu koostamisel. Genoomse DNA fragmendid denatureeritakse ja

kinnitatakse *flow cell* pinnale. DNA polümeraas sünteesib kinnitunud fragmendile komplementaarse ahela, mille käigus moodustub kaheahelaline DNA. Kaheahelalise DNA denatureerimisel paindub üheaahelaline fragment ja hübridiseerub lähedalasuva praimeriga, moodustades silla. Moodustunud sillalt sünteesitakse komplementaarne ahel. Eelnevaid etappe korratakse mitmeid kordi, tootes igast fragmendist umbes tuhat koopiat. Peale sild-amplifikatsiooni denatureeritakse klastrid, et saada sekveneerimiseks vajalikud ühe-ahelalised fragmendid (Morey *et al.*, 2013).



Joonis 5. Sild-amplifikatsioon (Metzker, 2010) järgi. Tahkele pinnale kantakse 5 μ g DNA fragmente. Üheaahelalised fragmendid kinnitatakse kovalentselt oligonukleotiididele ja lisatakse DNA polümeraas ning dNTP-d. Toimuvad järjestikused tsüklid, kus sünteesitakse komplementaarne ahel ja see omakorda denatureeritakse. Tahke faasi amplifikatsioon võimaldab toota 100-200 miljonit ruumiliselt eraldatud *template*-DNA klastrit.

Sekveneerimine toimub pöördterminaatoritega tsükli meetodil, mis toimub *flow cell*'i tahkel pinnal. Enne sekveneerimist lisatakse huvipakkuvatele järjestustele sekveneerimispraimer. Esmalt seob DNA polümeraas ühe fluorestsentsmärgisega nukleotiidi, mis on komplementaarne *template* ahelaga. Nukleotiidi lisamine ahelasse takistab järgnevate nukleotiidide lisamist ja süntees peatub ajutiselt. Samaaegselt toimub miljonite ahelate süntees, kuid ühte ahelasse seotakse tsükli jooksul ainult üks nukleotiid. Seondumata nukleotiidid eemaldatakse ja loetakse ahelasse seondunud nukleotiidilt signaal. Järgneb lõhustumisfaas, millega eemaldatakse 3' inhibeeriv rühm ja fluorestsentsvärv. Toimub uus pesu ja algab järgmise nukleotiidi seondumine (Metzker, 2010; Morey *et al.*, 2013).



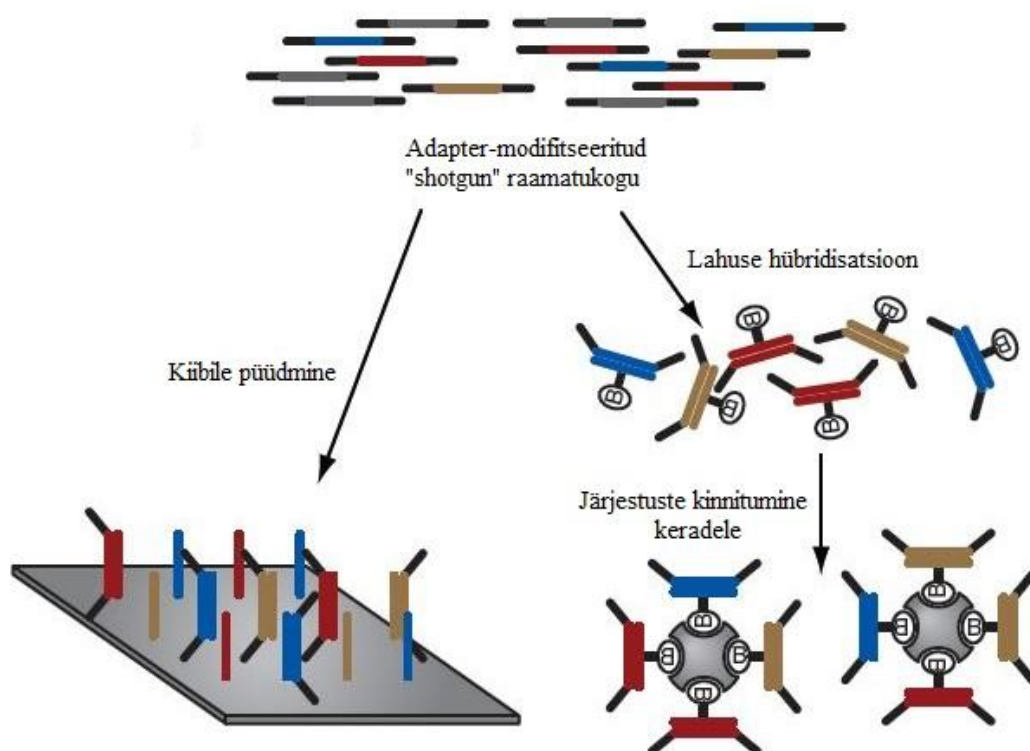
Joonis 6. Illumina/Solexa sekveneerimistehnoloogia (Metzker, 2010) järgi. Samaegselt toimub mitmesse ahelasse nukleotiidide lisamine. Seotud nukleotiididel on 3' ots blokeeritud, mis takistab järgnevate nukleotiidide lisamist ahelasse. Peale nukleotiidi lisamist loetakse fluorestsents signaal, mille järel vabaneb 3' otsast inhibeerib rühm ja fluorestsentsvärv. Ahelasse lisatakse järgmine nukleotiid.

1.2 Eksoomi sekveneerimine

Eksoomi sekveneerimine on inimese genoomi kodeeriva osa DNA nukleotiidne järjestamine. Eksoom koosneb umbes 23 000 kodeerivast geenist (Gonzaga-Jauregui *et al.*, 2012) ja 180 000-200 000 eksonist (Ng *et al.*, 2009; Ku *et al.*, 2012). Kuigi eksoom moodustab umbes 1-2% inimese genoomist (Ng *et al.*, 2009), on enamik Mendeliaalseid haigusi põhjustatud mutatsioonidest just valke kodeerivates alades ehk eksonites. Arvatakse, et ligikaudu 85% haiguspõhjuslikest mutatsioonidest võib paikneda just DNA kodeerivas alas (Choi *et al.*, 2009; Lalonde *et al.*, 2010). Just seetõttu on eksoomi sekveneerimisest saanud väga tõhus geenidiagnostika tööriist kliinilises praktikas.

1.2.1 Eksoomi väljapüüdmine

Et sekveneerida ainult inimese kodeerivaid alasid, on vaja eksonid eelnevalt välja püüda. Enne NGS tehnoloogiate turule jõudmist toimus huvipakkuvate järjestuste väljapüüdmine PCR meetodiga. Kuna eksoomi suurus on ~30 Mb (Gonzaga-Jauregui *et al.*, 2012), siis pole selline meetod otstarbekas väga suurte uuringute läbiviimiseks ühe eksperimendiga (Albert *et al.*, 2007). Tänapäeval kasutatakse peamiselt kahte eksoomi väljapüüdmise meetodit, milleks on kiibipõhine ja lahusepõhine meetod (joonis 7) (Mamanova *et al.*, 2010). Selliste meetoditega on võimalik püüda rohkem sihtmärke ja saadud katvus on homogensem kui PCR meetodiga. (Morey *et al.*, 2013).



Joonis 7. Kiibi- ja lahusepõhine eksoomi väljapüüdmine (Mamanova *et al.*, 2010) järgi. Genoomse DNA raamatukogu fragmendid hübridiseeritakse sihtmärk spetsiifiliste oligonukleotiididega mikrokiibi pinnal või lahuses. Kinnitumata DNA pestakse ära ja sihtmärk DNA loputatakse välja, amplifitseeritakse ja sekveneeritakse.

Kiibipõhisel meetodil hübridiseeritakse raamatukogu fragmendid immobiliseeritud oligonukleotiididega, mis esindavad sihtmärk regioone, kiibi pinnale. Mittespetsiifilised hübriidid eemaldatakse pesuga ja sihtmärk DNA loputatakse välja. Viimaste uuendustega on kiibile kinnitatud 2,1 miljonit oligonukleotiidi, mis võimaldab ühe eksperimendiga väljapüüda kuni 34 Mb. Kuigi kiibipõhisel meetodil on teatud eelised PCR-i meetodi ees, on üheks suureks puuduseks vajaminev DNA kogus. Olenemata väljapüütava piirkonna suuruselt, on

tavaliselt kiibipõhisel meetodil vaja 10-15 µg DNA-d (Mamanova *et al.*, 2010; Morey *et al.*, 2013).

Lahusepõhise meetodi peamine tööpõhimõte on kiibipõhise meetodiga sarnane - disainitakse oligonukleotiidid, mis on spetsiifilised huvipakkuvatele järjestustele. Kuid erinevalt kiibipõhisest meetodist, kinnituvad lahusepõhisel meetodil oligonukleotiidsed järjestused sihtmärk fragmentidele lahuses ja alles seejärel keradele. Kinnitumata fragmendid eemaldatakse pesuga ja sihtmärk DNA loputatakse välja (Mamanova *et al.*, 2010). Kahe meetodi erinevuseks on ka see, et kiibipõhine meetod kasutab oligonukleotiidide suhtes liigselt DNA fragmente, kuid lahusepõhisel meetodil oligonukleotiide ülehulga. Seetõttu vajab lahusepõhine meetod väiksemat (3 µg) raamatukogu hulka (Gnirke *et al.*, 2009).

1.3 Mendeliaalsed haigused

Mendeliaalsed haigused on monogeensed haigused, mis päranduvad Mendeli seaduste alusel. Mendeliaalsed haigused on harvaesinevad, esinedes üle maailma umbes 40 kuni 82 inimesel 1000 sünni kohta². OMIM (*Online Mendelian Inheritance in Man*) andmebaasi järgi on avastatud üle 6000 Mendeliaalse haiguse, kuid umbes pooltel on käesoleval ajal haiguse põhjus teadmata (Zhang, 2014). Selle põhjusteks võivad olla näiteks väga vähe uuritavaid patsiente või perekondasid, kellel on avaldunud sarnane fenotüübiliselt äratuntav tunnuse muster, või lookuste heterogeensus. Sellest tulenevalt ei pruugi näiteks traditsioonilised lookusspetsiifilised uurimismeetodid olla edukad haiguse põhjuse väljaselgitamisel (Lalonde *et al.*, 2010; Ng *et al.*, 2010).

Mendeliaalsete haiguste geneetilised põhjused võivad olla väga heterogeensed. Geneetiliselt heterogeensed on sellised haigused, mis on põhjustatud mutatsioonidest erinevates geenides (Gilissen *et al.*, 2011). Üheks selliseks näiteks on retsessiivselt päranduv Fanconi aneemia (OMIM #227650), mis põhjustab organsüsteemide arenguanomaaliad, füüsilisi ebanormaalsusi ja luuüdi häireid ning esineb 1/160 000 sünni kohta. Käesolevaks ajaks on teada, et Fanconi aneemia võib olla põhjustatud mutatsioonidest vähemalt 15 erinevas geenis, sealhulgas 80-90% juhtudest on põhjustatud mutatsioonidest FANCC geeniperikonna geenides, nagu *FANCA* (*fanconi anemia, complementation group A*), *FANCC* (*fanconi anemia, complementation group C*) ja *FANCG* (*fanconi anemia, complementation group G*)³.

1.3.1 Mendeliaalsete haiguste uurimine

Mendeliaalsete haiguste põhjuseid on uuritud aastakümneid, kasutades selleks erinevaid meetodeid, nagu kandidaatgeenide uuringud ja aheldatuse analüüs positsioonilise kloonimisega. Enne 1980. aastat, kui pakuti välja ülegenoomne aheldatuse analüüs, ei olnud ühtset meetodit Mendeliaalsete haiguste uurimiseks. Inimese geneetilise aheldatuse kaart, meetodid ja algoritmid tagasid esimest korda usaldusväärse tööriista, mis aitas mõista fenotüübi-genotüübi vahelisi seoseid (Botstein *et al.*, 1980). 1986. aastal klooniti esimene geen, milleks oli *CYBB* (*cytochrome b(-245), beta subunit*), mis on seotud X-liitelise kroonilise granulomatoosiga. On oluline rõhutada, et *CYBB* geen klooniti referentsjärjestust teadmata, toetudes ainult kromosomaalse kaardi positsioonile. (Royer-Pokora *et al.*, 1986).

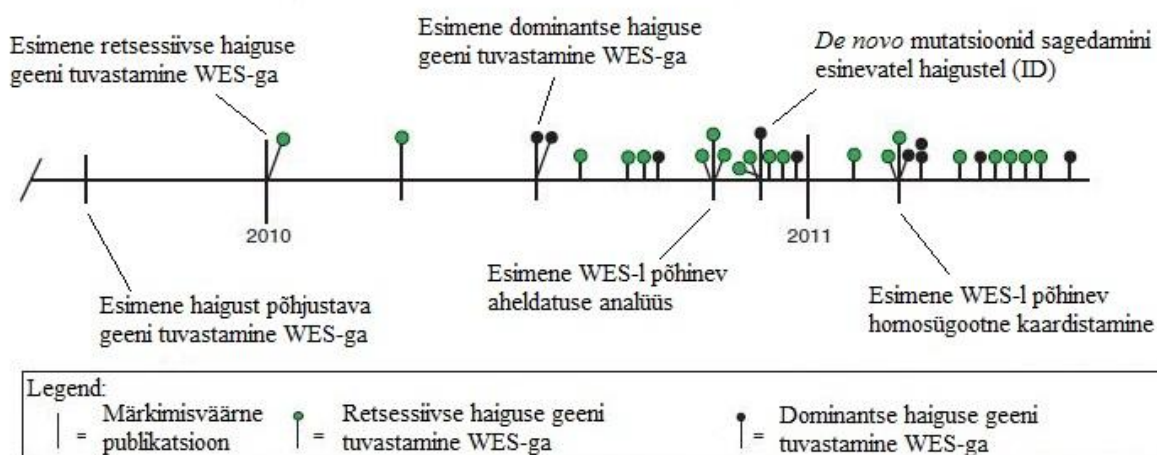
² Vaata <http://www.marchofdimes.com/materials/global-report-on-birth-defects-the-hidden-toll-of-dying-and-disabled-children-executive-summary.pdf>

³ Vaata <http://ghr.nlm.nih.gov/condition/fanconi-anemia>

Kuigi NGS tehnoloogiad on oluliselt kiirendanud ja lihtsustanud haigusseoseliste geenivariantide uurimist, pole tänapäeval geenidiagnostika tööriistana kuhugi kadunud ka näiteks esimese põlvkonna sekveneerimise tehnoloogial põhinev Sangeri sekveneerimine.

1.3.1.1 Mendeliaalsete haiguste uurimine eksoomi sekveneerimisega

Kui Sangeri sekveneerimine võimaldab analüüsida ainult teatud kindlaid geenijärjestusi, siis WES-i abil saab hüpoteesivabalt uurida kogu genoomi kodeerivaid järjestusi. Seetõttu on eksoomi sekveneerimisest saanud tänapäeval asendamatu tööriist Mendeliaalsete haiguste uurimisel (Teer and Mullikin, 2010), mida on nüüdseks edukalt rakendatud juba üle 150 haiguse põhjuse väljaselgitamisel (Zhang, 2014). WES-i on edukalt rakendatud monogeensete haiguste nagu Miller'i (Ng *et al.*, 2010) ja Kabuki sündroomi (Ng *et al.*, 2010), aga ka geneetiliselt heterogeensete haiguste, nagu skisofreenia (OMIM #181500) (Girard *et al.*, 2011) ja autismispektri häirete (ASD, *autism spectrum disorder*) (OMIM #209850) (O'Roak *et al.*, 2011) molekulaargeneetiliste põhjuste väljaselgitamisel (joonis 8).



Joonis 8. Ajajoon WES-i rakendamisest Mendeliaalsete haiguste uurimisel (Gilissen *et al.*, 2011). Joonisel on esitatud esimesed publikatsioonid eksoomi sekveneerimise rakendamisest diagnostilise tööriistana.

Eksoomi sekveneerimise rakendamine pole edukas ainult selge kliinilise diagnoosiga patsientide (näiteks Miller'i sündroom) või suunatud grupi uurimisel (näiteks intellektipuue (*ID, intellectual disability*), *ASD*). Yang jt kolleegide poolt läbiviidud uuringus sekveneeriti 250 ebaselge etioloogiaga arenguhäiretega patsientide eksoomid. Uuritud patsientide valimis polnud varasemad diagnostilised uuringud selgitanud haiguse põhjust. WES-i tulemusena identifitseeriti 62 patsiendil haiguspõhjuslik mutatsioon, mis teeb diagnostiliseks saagiseks 25%

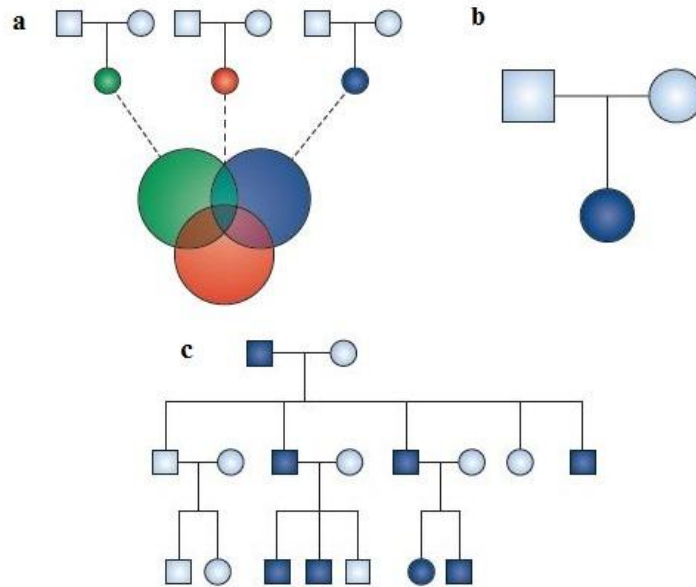
(Yang *et al.*, 2013). See on praeguseks ainuke publitseeritud uuring *WES*-i rakendamisest kliinilises praktikas erinevate neuroloogiliste ja mitteneuroloogiliste haigussümptomitega patsientidel.

1.3.1.1.1 Uuritavate patsientide valiku strateegiad eksoomi sekveneerimisel

Ameerika Meditsiinigeneetika Kolledži (*ACMG, American College of Medical Genetics and Genomics*) poolt on publitseeritud näidustused, millistel juhtudel *WES*-i rakendada (*ACMG Board of Directors, 2012*):

- 1) Patsiendi sümptomite ja suguvõsa anamneesi alusel on kahtlus geneetilise etioloogiaga haigusele, kuid patsiendi kliiniline pilt ei sarnane kindlale teadaolevale haigusele, mille põhjuse väljaselgitamiseks on olemas spetsiifiline geenitest.
- 2) Patsiendi kliiniline pilt vastab geneetiliselt heterogeensele haigusele, mille tõttu pole paljude üksikute kandidaatgeenide Sanger sekveneerimine otstarbekas.
- 3) Ebaselged haigusjuhtumid, mille puhul patsiendil esineb tõenäoliselt geneetiline haigus, kuid eelnevad spetsiifilised geenitestid pole selgitanud häire põhjust.

Haiguspõhjuslike geenivariantide otsingute lihtsustamiseks kasutatakse uuritavate patsientide valikul erinevaid lähenemisi, sõltuvalt haiguse pärandumismustrist ja uuritavast haigusest. *WES* meetodiga on võimalik uurida nii dominantse kui ka retsessiivselt pärandunud või *de novo* mutatsioonide poolt põhjustatud haigusi, rakendades erinevaid valimi valiku strateegiad, nagu mittesuguluses patsientide valim, laps-vanemate triod või perekonnapõhine valim (joonis 9).



Joonis 9. Uuritava valimi valiku strateegiad WES analüüsil (Bamshad *et al.*, 2011) järgi. Joonisel on kujutatud a) kliiniliselt sarnase fenotüübiga mittesuguluses patsientide valim, b) laps-vanemate trio ja c) kliiniliselt sarnase fenotüübiga uuritavad ühes suguvõsas.

Kliiniliselt äratuntavate tunnuste mustriga mittesuguluses patsientide valimi rakendamine on efektiivne, kui haigus pärandub dominantset ja puudub geneetiline heterogeensus. Sellisel juhul otsitakse haiguspõhjuslikke geenivariante, mis esinevad kõigil või suuremal osal patsientidest (Gilissen *et al.*, 2012).

Laps-vanemate trio analüüsil uuritakse terveid vanemaid ja nende haiget last. Laps-vanemate trio lähenemist WES analüüsil rakendatakse eelkõige siis, kui mutatsioonid on tekkinud *de novo*. Sellisel juhul nii lapse kui vanemate eksoomide sekveneerimine lihtsustab andmeanalüüsi, mille tulemusena saab välistada kõik vanematelt päritud muutused (Gilissen *et al.*, 2012). Laps-vanemate triode WES analüüsi on sageli rakendatud näiteks *ID* (Vissers *et al.*, 2010; de Ligt *et al.*, 2012; Rauch *et al.*, 2012) ja *ASD* (O'Roak *et al.*, 2011) põhjuste väljaselgitamisel.

Kliiniliselt sarnase fenotüübiga haigete uurimisel ühes perekonnas/suguvõsas peab haigus päranduma kas dominantset või retsessiivselt. Sellisel juhul peab haiguspõhjusliku mutatsiooni mõju avalduma haiguse fenotüübis (Gilissen *et al.*, 2012).

1.3.2 Autosoom-retsessiivsete haiguste uurimine kogu-eksoomi sekveneerimisega

Retsessiivsete haiguste puhul pärandub haiguspõhjuslik geenivariant vanematelt lapsele, seejuures peavad mõlemad vanemad olema selle haigusvariandi heterosügootsed kandjad. Kui mõlemad vanemad on haigusvariandi kandjad, pärib laps haiguse 25%-lise tõenäosusega. Juhul, kui laps pärib retsessiivse haiguse geenidefekti ainult ühelt vanemalt, on laps haiguse mitesümptoomne kandja⁴.

Esmakordselt rakendati WES-i meetodit retsessiivse pärandumismustriga haiguse uurimiseks Miller'i sündroomi puhul. Miller'i sündroom on väga harvaesinev haigus, mille esinemissageduseks hinnatakse vähem kui 1/1 000 000 elussünni kohta⁵. Miller'i sündroomile on iseloomulik hulgiväärarengute esinemine, näiteks mikrognatia, koanide atreesia, sündaktüülia ja koloboom. Haiguse põhjus oli varasemalt teadmata. Ng ja tema kolleegid sekveneerisid nelja Miller'i sündroomiga (OMIM #263750) patsiendi eksoomid. WES-i andmete analüüsi tulemusena tuvastati ainult üks kandidaatgeen *DHODH* (*dihydroorotate dehydrogenase*), mis sobis retsessiivse pärandumismustriga. Mutatsioonid *DHODH* geenis pärsivad dihüdroorotaadi dehüdrogenaasi aktiivsust, mis omakorda põhjustab mitmesuguseid arenguhäireid (Ng *et al.*, 2010).

WES rakendamine autosoom-retsessiivsete haiguste uurimisel on tänaseks välja selgitanud enam kui 80 retsessiivse haiguse põhjuse (Zhang, 2014).

1.3.3 Autosoom-dominantsete haiguste uurimine kogu-eksoomi sekveneerimisega

Dominantsete haiguste puhul pärandub haiguspõhjuslik heterosügootne geenivariant ühelt vanemalt lapsele ning tavaliselt on haige lapse üks vanematest ka ise haige. Haigus avaldub, kui muteerunud geenikoopia surub maha normaalselt funktsioneeruva geenikoopia. Seetõttu on dominantset tüüpi pärilikkuse korral lapsel 50%-line risk haigus pärida⁶.

⁴ Vaata <http://www.nlm.nih.gov/medlineplus/ency/article/002052.htm>

⁵ Vaata <http://ghr.nlm.nih.gov/condition/miller-syndrome>

⁶ Vaata http://genome.wellcome.ac.uk/doc_WTD020849.html

Esmakordselt rakendati *WES*-i dominantse haiguse põhjuse väljaselgitamiseks Kabuki sündroomi (OMIM #147920) puhul. Selleks sekveneeriti kümne omavahel mittesuguluses Kabuki sündroomiga patsientide eksoomid. Kabuki sündroom esineb 1/32 000 elussünni kohta ning haigusele on iseloomulik *ID*, iseloomulikud näojooned, mikrotsefaalia, hüpotoonia ja krambihood⁷. Just haiguse vähene esinemissagedus ja kliiniliselt heterogeene taust on raskendanud traditsiooniliste uurimismeetodite rakendamist Kabuki sündroomi põhjuste väljaselgitamisel. *WES* andmete analüüsi tulemusena tuvastati üheksal patsiendil kümnest mutatsioon *MLL2* (*myeloid/lymphoid or mixed-lineage leukemia 2*) geenis. *MLL2* geenilt kodeeritakse metüültransferaasi, mis metüleerib histoonid ja osaleb kromatiini epigeneetilisel kontrollil (Ng *et al.*, 2010). Mutatsioonid *MLL2* geenis vähendavad metüültransferaasi ensüümi aktiivsust, mille tulemusena avalduvad Kabuki sündroomile iseloomulikud tunnused⁸.

1.3.3.1 Eksoomi sekveneerimine ja *de novo* mutatsioonid

Eksoomi sekveneerimise kasutusele võtuga on saanud võimalikuks *de novo* mutatsioonide uurimine kõikide geenide ulatuses, mis pole traditsiooniliste meetoditega olnud võimalik (de Ligt *et al.*, 2013). *De novo* mutatsioon on sugurakkudes juhuslikult tekkinud muutused (Ng *et al.*, 2010), mis võivad avaldada mõju tervisele ja põhjustada arenguhäireid, kui sellised muutused kahjustavad funktsionaalselt olulisi järjestusi genoomis (Vissers *et al.*, 2010).

Esmakordselt rakendati *WES*-i *de novo* mutatsioonide poolt põhjustatud haiguste uurimisel, kui sekveneeriti nelja Schinzel-Giedion'i sündroomiga (OMIM #269150) patsiendi eksoomid. Schinzel-Giedion'i sündroomi esinemissagedus populatsioonis on teadmata. Sündroomile on iseloomulik raske *ID*, iseloomulikud näojooned ja hulgiväärarengute, nagu laiad roided, lühike rinnaluu ja genitaalide alaareng, esinemine⁹. *WES* analüüsi tulemusena tuvastati kõigil patsientidel heterosügootsed *de novo* mutatsioonid *SETBP1* (*SET binding protein 1*) geenis (Hoischen *et al.*, 2010).

⁷ Vaata <http://ghr.nlm.nih.gov/condition/kabuki-syndrome>

⁸ Vaata <http://ghr.nlm.nih.gov/gene/KMT2D>

⁹ Vaata <http://ghr.nlm.nih.gov/condition/schinzel-giedion-syndrome>

Eksoomide uuringud kliinilises praktikas on näidanud, et umbes 1/3 juhtudest tekivad mutatsioonid *de novo* ning selliste mutatsioonide rolli on väga ulatuslikult uuritud eelkõige *ID*, *ASD* ja skisofreenia puhul (de Ligt *et al.*, 2013).

1.3.4 Eksoomi sekveneerimise puudused

Kuigi eksoomi sekveneerimine on väga tõhus ja usaldusväärne geenidiagnostika tööriist, esineb sellel analüüsimeetodil ka teatud puudusi:

- 1) *WES*-ga on võimalik uurida ainult teadaolevate geenide eksonite ja nendega külgnevate reguleerivate piirkondade geneetilist varieeruvust (sünonüümsed/mittesünonüümsed asendused ning väikesed deletsioonid või insertioonid ehk indelid). Muutused mittekodeerivates alades jäävad *WES*-i rakendades tuvastamata (Ng *et al.*, 2010).
- 2) Tavaliselt kaetakse eksoomi sekveneerimisega umbes 95% kodeerivast järjestusest, mistõttu läheb väike osa geneetilisest infost analüüsi käigus „kaduma“. See on põhjustatud osade eksoomi piirkondade halvast katvusest, mistõttu ei läbi need alad kvaliteedikontrolli, või väljapüüdmata jäänud eksoomi fragmentidest, mis ei jõuagi sekveneerimisele (Bamshad *et al.*, 2011).
- 3) *WES*-i rakendades pole võimalik tuvastada tandeemseid kordusjärjestusi (Shashi *et al.*, 2014). See on tingitud sellest, et eksoomi sekveneerimise etapile eelnevalt amplifitseeritakse DNA proovid PCR-i meetodit kasutades, mis on aga kordusjärjestustega regioonide tuvastamisel madala efektiivsusega. Samuti on bioinformaatilise töötlemise etapis raskendatud kordusjärjestustega *read*'ide joondamine referentsgenoomi vastu. Seetõttu jäävad tandeemse kordusjärjestusega regioonid tavaliselt edasisest analüüsist välja, sest ei vasta joondamise kvaliteediparameetritele. Sellest tulenevalt pole eksoomi sekveneerimise meetodit võimalik usaldusväärsetl rakendada näiteks selliste haiguste vastu nagu Huntington'i tõve ja fragiilse X sündroomi diagnoosimiseks (Behjati and Tarpey, 2013). Mõlema haiguse põhjuseks on vastavalt trinukleotiidsete CAG¹⁰ ja CGG¹¹ korduste arvu suurenemine üle teatud kriitilise piiri.

¹⁰ Vaata <http://ghr.nlm.nih.gov/condition/huntington-disease>

- 4) Samuti ei võimalda *WES* leida mitokondriaalse DNA ega epigeneetiliselt reguleeritud muutusi (Shashi *et al.*, 2014).
- 5) *WES*-iga on võimalik tuvastada ühenukleotiidseid asendusi ja väikseid indeleid (Ng *et al.*, 2010), kuid struktuursete muutuste tuvastamine ainult *NGS* platvorme kasutades on keeruline (Gonzaga-Jauregui *et al.*, 2012), mistõttu võivad suuremad muutused jääda märkamata. Koopiaarvu varieeruvuste (*CNV*, *copy number variation*) määramiseks *NGS* sekveneerimisandmetest on loodud mitmesuguseid algoritme, näiteks CoNIFER (Krumm *et al.*, 2012) ja Convex (Amarasinghe *et al.*, 2013), kuid ükski neist pole täiesti usaldusväärne.

¹¹ Vaata <http://ghr.nlm.nih.gov/condition/fragile-x-syndrome>

2 EKSPERIMENTAALOSA

2.1 Töö eesmärgid

Käesoleva bakalaureusetöö eesmärgid olid:

- 1) analüüsida eksoomi sekveneerimise andmeid, et tuvastada haiguspõhjuslikke geenivariante kaasasündinud arenguhäiretega patsientide valimis;
- 2) valideerida eksoomi sekveneerimise tõenäoliselt leiud Sanger sekveneerimisega;
- 3) valideerida *Children's Mercy Hospital and Clinics* meditsiinikeskuse diagnostikalaboris saadud eksoomi sekveneerimise tulemused uuritud patsientide valimis, kasutades TÜ Eesti Geenivaramu bioinformaatilist andmetötlust ja –analüüsi.

2.2 Materjal ja metoodika

2.2.1 Valimi kirjeldamine

Koostöös SA TÜK Lastekliinikumiga koguti 11 patsiendist koosnev valim. Uurimisgruppi valiti patsiendid, kellel on vastündinu- või lapseas avaldunud ebaselge etioloogiaga päriliku komponendiga haigused ja sündroomid, näiteks füüsilise või vaimse arengu mahajäämus, kaasasündinud arenguanomaaliad, ainevahetushäired jms. Eelnevad diagnostilised uuringud haiguste põhjuse väljaselgitamiseks, nagu kromosoomvöödistus, submikroskoopiline kromosoomianalüüs, ainevahetusuuringud, üksikute geenide sekveneerimine jne, ei olnud selgitanud häire tekkepõhjust.

Eksoomi sekveneerimise tulemuste valideerimiseks koguti ka patsientide tervete vanemate genoomse DNA proovid.

Antud uurimistöö toimus Tartu Ülikooli inimuuringu eetikakomitee loal. Uuritavad on andnud oma kirjaliku nõusoleku uuringus osalemiseks.

2.2.2 Genoomse DNA eraldamine

Genoomse DNA eraldamine perifeerses verest viidi läbi klassikalise väljasoolamise meetodil standardprotokollil alusel. Eraldatud DNA kontsentratsiooni ja kvaliteeti hinnati geelelektroforeesil ja spektrofotomeetriga NanoDrop ND-1000 (Thermo Fisher Scientific Inc., Wilmington, De, USA). Genoomse DNA eraldamine ja kvaliteedi hindamine teostati SA TÜK Ühendlabori Geneetikakeskuses.

2.2.3 Eksoomi sekveneerimine ja andmete analüüsimine

2.2.3.1 Eksoomi sekveneerimine

Eksoomi sekveneerimise katse ja andmete analüüs teostati koostööprojekti raames *Children's Mercy Hospital and Clinics* meditsiinikeskuse diagnostikalaboris (Kansas, Missouri, USA). *Children's Mercy Hospital and Clinics* on üks maailma juhtivamaid keskusi, kus teostatakse eksoomi sekveneerimise analüüse diagnostilisel eesmärgil. Eksoomi rikastamiseks kasutati *TruSeq Exome Enrichment Kit*'i (Illumina Inc., San Diego, CA, USA) ja sekveneerimiseks HiSeq2500 platvormi (Illumina Inc., San Diego, CA, USA). Patsientide haigusloo põhjal otsustati üheksal patsiendil sekveneerida kogu eksoom ja kahel patsiendil (patsiendil EST3 ja

EST8) viidi läbi *TaGSCAN (Targeted Gene Sequencing and Custom Analysis*¹²) geenipaneeli test. *TaGSCAN* geenipaneel võimaldab uurida 514 geeni kodeerivaid järjestusi, mille mutatsioone seostatakse sagedamini esinevate vastsündinu- ja lapseas esinevate haigustega, mis põhjustavad näiteks füüsilise ja vaimse arengu mahajäämuse ja kaasasündinud arenguanomaaliaid. *TaGSCAN* geenipaneeli uuringuga on võimalik tuvastada enam kui 750 haiguse põhjus, näiteks Joubert'i sündroom 3, tsüstiline fibroos ja fenüülketonuuria.

2.2.3.2 Eksoomi sekveneerimisandmete analüüs TÜ Eesti Geenivaramus

Children's Mercy Hospital and Clinics meditsiinikeskuse diagnostikalaboris läbi viidud eksoomi sekveneerimise analüüsi tulemused ja toorandmed (fastq ja bam failid) saadeti koostööprojekti raames TÜ EGV-sse, et valideerida tulemused uuritud patsientide valimis kasutades TÜ Eesti Geenivaramu bioinformaatilist andmetöötlust ja –analüüsi. Eksoomi sekveneerimisandmete bioinformaatilise analüüsi TÜ EGV-s teostas spetsialist Mart Kals. 100 bp *paired-end* fragmendid joondati inimese referentsgenoomi (hg19, GRCh37) vastu, kasutades programmi Burrows-Wheeler Aligner (BWA, version 0.6.1) (Li and Durbin, 2009). Variantide määramiseks kasutati tarkvara *Genome Analysis Toolkit (GATK, Version 1.6-7)* (McKenna *et al.*, 2010). Ühenukleotiidsed variandid ja indelid annoteeriti Ensembl andmebaasi vastu, kasutades majasiseseid (*in-house*) skripte. Annoteeritud eksoomiandmete analüüsimise ja interpretatsiooni viis läbi töö autor TÜMRI doktorandi ja TÜ EGV spetsialisti juhendamisel (M. Nõukas).

2.2.3.3 Andmete interpretatsioon

Eksoomi sekveneerimise andmetest võimalike haiguspõhjuslike geenivariantide tuvastamiseks kasutati veebipõhist programmi SimulConsult¹³ (Chestnut Hill, MA, USA) ja variantide manuaalset filtreerimist ja prioritseerimist, kuhu on integreeritud ka tarkvara Genome Trax^{TM14} (Beverly, MA, USA) andmebaaside andmed.

Eksoomi sekveneerimise andmete manuaalsel filtreerimisel ja prioritseerimisel keskenduti eksonites ja kanoonilistes splaissingu saitides paiknevatele valgu järjestust muutvatele

¹² Vaata http://www.childrensmercy.org/Health_Care_Professionals/Research/Pediatric_Genomic_Medicine/TaGSCAN_Diagnostic_Test/

¹³ Vaata <http://www.simulconsult.com/>

¹⁴ Vaata <http://www.biobase-international.com/product/genome-trax>

variantidele (mittesünonüümsed asendused ja raaminihke mutatsioonid), mis esinevad sagedusega 1% või vähem vastavalt 1000 Genoomi projekti (*1000G, 1000 Genome Project*)¹⁵ andmebaasile. Kandidaatgeenides tuvastatud asendusi võrreldi erinevate avalike andmebaasidega, näiteks NHLBI GO Exome Sequencing Project (ESP), NCBI dbSNP Build 139 ja TÜ EGV majasisese referentsandmebaasiga, mis sisaldab 122 sekveneeritud eksoomide andmeid. Kandidaatgeenide puhul tutvuti erialase kirjandusega OMIM ja PubMed andmebaasides. Ühenukleotiidsete asenduse mõju valku kodeerivale järjestusele hinnati kasutades PolyPhen-2, SIFT ja Condel skooride. Näiteks PolyPhen-2 skoor varieerub 0-1 ja ühele lähenevad väärtused on tõenäolisemalt haiguspõhjuslikud võrreldes nullile lähenevate väärtustega. Evolutsioonilist konserveerumist hinnati PhyloP skoori järgi.

Andmete manuaalse filtreerimise ja prioritseerimise etappi on integreeritud ka tarkvara Genome TraxTM andmed. Genome TraxTM on tarkvara, mis hõlmab informatsiooni erinevatest andmebaasidest, näiteks *HGMD (Human Gene Mutation Database)*, ClinVar, Orphanet ja *PGMD (PharmacoGenomic Mutation Database)*. Nendes erinevates andmebaasides sisalduvat infot on võimalik integreerida eksoomi sekveneerimisandmete analüüsi. Võimalike haiguspõhjuslike geenivariantide tuvastamiseks uuritud patsientidel kasutati eelkõige selliste andmebaaside infot nagu *HGMD*, ClinVar ja Orphanet. Vastavad andmebaasid sisaldavad informatsiooni teadaolevate haigusseoselistest geenivariantide kohta, mis on publitseeritud erinevate töögruppide poolt üle maailma.

SimulConsult on veebipõhine programm, mille tööpõhimõte seisneb võimalike haiguspõhjuslike geenivariantide tuvastamises *WES* andmetest. SimulConsult tarkvara võimaldab integreerida omavahel patsiendi haigustunnused, eelnevad diagnostiliste uuringute tulemused ja eksoomi sekveneerimise andmed.

SimulConsult tarkvara kasutamiseks koostati iga uuritava patsiendi kohta kaks sisendfaili, milleks olid fenotüübi andmete põhjal koostatud fail (html-failivorming) ja annoteeritud variantide fail (txt-failivorming). Mõlema faili sisestamisel SimulConsult'i pakub tarkvara annoteeritud variantide faili põhjal tõenäoliselt haiguspõhjuslikud geenivariandid. Leid määrati kui „haiguspõhjuslik“ juhul, kui seda kinnitasid nii manuaalse filtreerimise ja SimulConsult tulemused.

¹⁵ Vaata <http://www.1000genomes.org/>

2.2.4 Eksoomi sekveneerimise tulemuste valideerimine Sangeri sekveneerimisega

2.2.4.1 PCR reaktsiooni läbiviimine

Eksoomi sekveneerimisandmete analüüsi tulemusena tuvastatud võimalikud haiguspõhjuslikud leiud valideeriti kasutades Sangeri sekveneerimise meetodit. Selleks amplifitseeriti uuritavad DNA piirkonnad PCR-i meetodiga. Sobivad praimerid disainiti kasutades veebipõhist programmi Primer3 Input 4.0.0¹⁶ (Untergasser *et al.*, 2012). Disainitud praimerijärjestused on esitatud tabelis 1.

Tabel 1. PCR reaktsiooni läbiviimiseks kasutatud praimerid.

Praimerite paar	Praimeri nimetus	Järjestus 5'-3'	Amplikoni pikkus
RUNX2	RUNX2_F	GTCTTTGTTTCATTGCCTCCTT	350 bp
	RUNX2_R	ATGCAGATAGCAAAGTCCACAA	
CHD7	CHD7_F	GGAATCGGACAGAGGAATATGG	400 bp
	CHD7_R	TCAGACAAGTTAGCCATGTTGAA	

PCR-i reaktsioonid viidi läbi mahus 20 µl. Reaktsioonisegu sisaldas 2 µl reaktsioonipuhvrit B (0,8 M Tris-HCl, 0,2 M (NH₄)₂SO₄, 0,2% w/v Tween 20; Solis BioDyne, Eesti), 3 µl dNTP-d (2 mM dATP, dCTP, dTTP, dGTP; Thermo Scientific, USA), 1,2 µl MgCl₂ (25 mM; Thermo Scientific, USA), 1 µl pärisuunalist ja 1 µl vastassuunalist praimerit (10 mM; Metabion, Saksamaa), 0,3 µl Hot FirePol Taq DNA polümeraasi (5U/µl Solis BioDyne, Eesti), 2 µl genoomset DNA-d (50 ng/µl) ja lõppmahuni ddH₂O-d. PCR-i reaktsioonid viidi läbi programmeeritavas termotsükleris (GeneAmp® PCR System 2700, Applied Biosystems, USA). DNA fragmentide amplifitseerimiseks kasutati PCR programmi, mis on esitatud tabelis 2.

¹⁶ Vaata <http://primer3.ut.ee/>

Tabel 2. PCR programm.

	Temperatuur (°C)	Aeg	Korduste arv
Algne denaturatsioon	95 °C	15 min	1x
Denaturatsioon	95 °C	30 s	40x
Praimeri seondumine	62 °C		
Ekstensioon	72 °C		
Ekstensioon	72 °C	10 min	1x
Produktide hoidmine	4 °C		

Produktid lahutati 2% agarosgeelis geelelektroforeesil 0,5-kordses TBE puhvril (0,045 M Trisboraat; 0,001 M EDTA). Visualiseerimiseks kasutati 6x laadimispuhvrit (6x Loading Dye; Thermo Scientific, USA) ning pikkusmarkeriks oli pUC Mix Marker, 8 (Thermo Scientific, USA).

2.2.4.2 Sekvenerimisreaktsioon

Pärast PCR-i reaktsiooni puhastati produktid ensüümtötlusega kasutamata jäänud praimeritest ja desoksüribonukleotiididest ning sadestati sekvenerimisproduktid. Ensüümtötlus ja sekvenerimisproduktide sadestamine ja järjestuse määramine viidi läbi Eesti Biokeskuse tuumiklaboris, kasutades Applied Biosystems 3730 xl DNA Analyzer masinat (Applied Biosystems, USA).

2.3 Tulemused

Käesoleva töö raames uuriti haiguspõhjuslike geenivariantide esinemist ebaselge etioloogiaga arenguhäiretega patsientide valimis. Eksoomi sekveneerimise analüüs teostati kokku 11 patsiendile. Kogu-eksoomi sekveneerimise keskmine katvus referentsgenoomi suhtes on 73x (kogu eksoom sekveneeriti kokku 9 patsiendil), *TaGSCAN* geenipaneeli uuringu keskmine katvus sihtmärk-regioonide suhtes on vähemalt 500x (*TaGSCAN* geenipaneeli uuring teostati 2 patsiendil).

Eksoomi sekveneerimisandmete analüüsimise tulemusena tuvastati haigust põhjustav mutatsioon kahel patsiendil. Patsiendil EST5 identifitseeriti heterosügootne mittesünonüümne c.674G>A, p.(Arg225Gln) mutatsioon geenis *RUNX2*. Patsiendil EST10 tuvastati samuti heterosügootne mittesünonüümne c.1480C>T, p.(Arg494X) mutatsioon geenis *CHD7*. Vastavad leiud valideeriti Sanger sekveneerimise meetodiga nii patsiendil kui ka tema vanematel, et selgitada välja, kas mutatsioonid c.674G>A *RUNX2* geenis ja c.1480C>T *CHD7* geenis on lapsel tekkinud *de novo*.

Uuritud ebaselge etioloogiaga arenguhäiretega patsientide valimis ei tuvastanud üheksal patsiendil (EST1, EST2, EST3, EST4, EST6, EST7, EST8, EST9, EST11) eksoomi sekveneerimise andmete analüüsi tulemusena selliseid geenivariante, mille puhul oleks olnud põhjust kahtlustada mõju kliinilise fenotüübi väljakujunemisele.

2.3.1 Patsient EST5

WES analüüsi tulemusena tuvastati patsiendil EST5 heterosügootne mittesünonüümne c.674G>A, p.(Arg225Gln) (NM_001024630.3) mutatsioon *RUNX2* (*runt-related transcription factor 2*) geenis (tabel 3). Vastav leid valideeriti Sanger sekveneerimise meetodiga patsiendil ja tema tervetel vanematel. Sanger sekveneerimise tulemused kinnitavad *WES* analüüsi tulemusi, mille järgi esineb lapsel *RUNX2* c.674G>A *de novo* mutatsioon (joonised 11 ja 12). Joonisel 11 on esitatud programmi IGV (*Integrative Genomics Viewer*, versioon 2.3) väljavõtte joondatud järjestusest, kus on kujutatud *RUNX2* 5. eksonis tuvastatud heterosügootne c.674G>A mutatsioon, mille suhtes on patsient EST5 heterosügootse A/G genotüübiga. Joonisel 12 on esitatud Sanger sekveneerimise tulemused, mis kinnitavad, et patsiendil esineb *RUNX2* c.674G>A heterosügootne mutatsioon ning vanematel mitte (tähistatud musta raamiga).

Patsiendil EST5 tuvastatud c.674G>A mutatsiooni *RUNX2* geenis on varasemalt teadaoleva kliinilise tähtsusega (Baumert *et al.*, 2005; Hansen *et al.*, 2011; Lee *et al.*, 2013), mis põhjustab kleidokraniaalset düsplaasiat. Vastav mutatsioon c.674G>A esineb ka *HGMD* andmebaasis kui teadaolev haiguspõhjuslik variant. Kleidokraniaalne düsplaasia esineb umbes 1/1 000 000 elussünni kohta. *RUNX2* geeni mutatsioonid põhjustavad *RUNX2* valgu tootmise katkemist või vähendavad geenilt kodeeritud valgu aktiivsust¹⁷, mis mõjutab peamiselt luude ja kõhre normaalset arengut. Kleidokraniaalsele düsplaasiale on iseloomulik lühike kasv, alaarenenud või puudulik rangluu, mis põhjustab kitsaid ja längus õlgu, hilinenud lõgede (pehme luustumata koht imiku koljul) sulgumine, selgroo anomaaliad (skolioos ja kүүrus selg), ülehammastus (hüperdontia) ja teised luude ning hammastega seotud anomaaliad¹⁸. Kleidokraniaalse düsplaasia sobib patsiendi EST5 kliinilise pildiga, sest patsiendil esinevad hilinenud lõgede (*fontanel*) sulgumine, pisilõugsus (*micrognathia*), hüpertelorism, kitsas rindkere ja kүүrus selg (*kyphosis*), mida on kirjeldatud kui haigusele iseloomulikke tunnuseid.

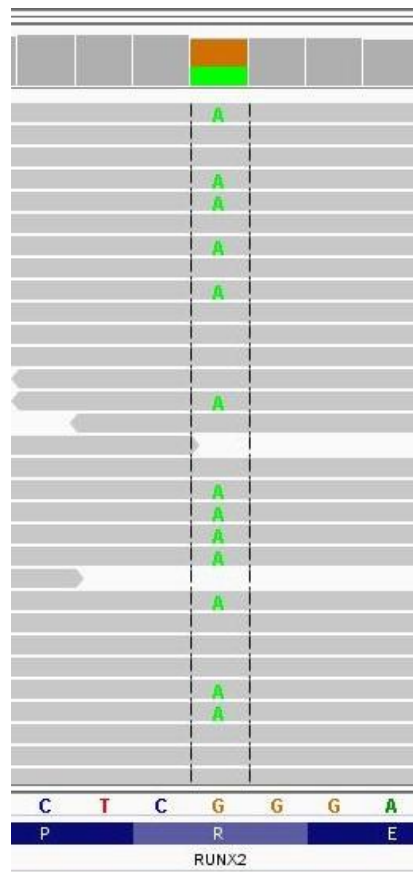
Tabel 3. *WES* analüüsi tulemusena patsiendil EST5 tuvastatud haiguspõhjuslik mutatsioon geenis *RUNX2*.

Kromosoom/ positsioon	Geen	Asendus nukleotiidses järjestuses	Asendus valgu- järjestuses	Asukoht	PolyPhen tulemus
6: 45 405 777*	<i>RUNX2</i>	c.674G>A	p.Arg225Gln	Eksonis 5/9, mitte- sünonüümne asendus	haigus- põhjuslik

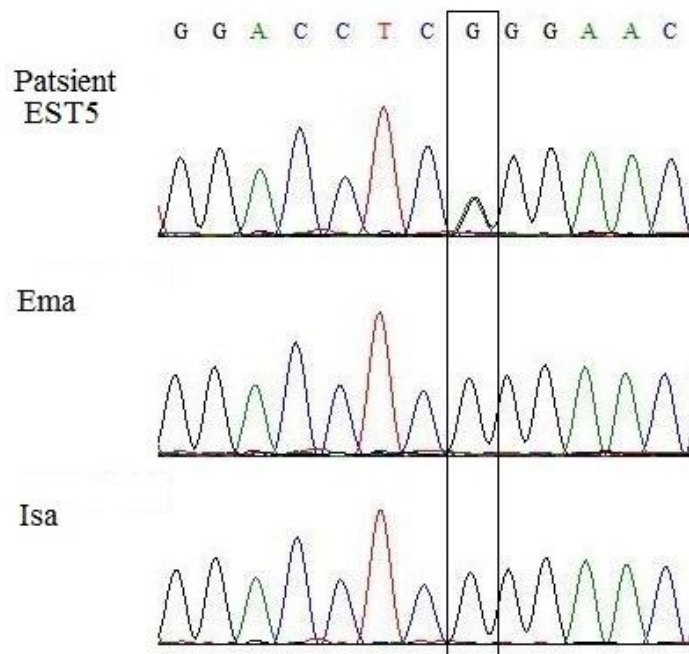
(*tähistab positsiooni vastavalt hg19 inimese genoomi assambleeringule)

¹⁷ Vaata <http://ghr.nlm.nih.gov/gene/RUNX2>

¹⁸ Vaata <http://ghr.nlm.nih.gov/condition/cleidocranial-dysplasia>



Joonis 11. Väljavõte programmist IGV. Musta punktiirjoonega on tähistatud positsioon c.674, mille suhtes on patsient EST5 heterosügootse A/G genotüübiga. Nimetatud positsioon on 72-kordse katvusega, millest G alleeli esineb 42x (59%) ja A alleeli 30x (41%).



Joonis 12. Sanger sekveneerimise graafik. Musta raamiga on tähistatud uuritav positsioon c.674 geenil *RUNX2*. Patsiendil EST5 esineb vastavas positsioonis heterosügootne A/G genotüüp, patsiendi vanematel esineb referentsjärjestusega sama homosügootne G/G genotüüp.

2.3.2 Patsient EST10

WES analüüsi tulemusena tuvastati patsiendi EST10 heterosügootne mittesünonüümne c.1480C>T, p.(Arg494X) (NM_017780.3) mutatsioon *CHD7* (*chromodomain helicase DNA-binding protein 7*) geenis (tabel 4). Vastav leid valideeriti Sanger sekveneerimisega patsiendil ja tema tervetel vanematel. Sanger sekveneerimise tulemused kinnitavad *WES* analüüsi tulemusi, mille järgi esineb lapsel *CHD7* geenis c.1480C>T *de novo* mutatsioon (joonised 13 ja 14). Joonisel 13 on esitatud programmi IGV väljavõtte joondatud järjestusest, kus on kujutatud *CHD7* eksonis tuvastatud heterosügootse c.1480C>T mutatsiooniga, mille suhtes on patsient EST10 heterosügootse C/T genotüübiga. Joonisel 14 on esitatud Sanger sekveneerimise tulemused, mis kinnitavad, et patsiendil esineb *CHD7* c.1480CC>T heterosügootne mutatsioon ning vanematel mitte (tähistatud musta raamiga).

Patsiendil EST10 tuvastatud c.1480C>T mutatsioon *CHD7* geenis, mis viib enneaegse stoppkoodoni tekkele, on varasemalt kirjeldatud kui teadaoleva kliinilise tähtsusega (Wincent *et al.*, 2008; Janssen *et al.*, 2012), mis põhjustab CHARGE'i sündroomi. Nimetatud c.1480C>T variant esineb ka *HGMD* andmebaasis kui teadaolev haiguspõhjuslik mutatsioon. Mutatsioonid *CHD7* geenis põhjustavad häireid kromantiini remodelleerumisprotsessides ja geeniekspressiooni regulatsioonis¹⁹. CHARGE'i sündroomi esineb keskmiselt 1/8500 elussünni kohta. Sündroomi nimi tuleneb haigusele iseloomulikest tunnustest, milleks on koloboom (Coloboma), südamehäired (Hearth defect), koaanide atreesia (Atresia choanae), alaarenenud kasv ja areng (Retarded growth and development), genitaalide anomaaliad (Genital abnormality) ja kõrva anomaaliad (Ear abnormality)²⁰. Patsiendil EST10 olid juba vastasündinueas avaldunud koloboom ja koaanide atreesia, mistõttu kahtlustati CHARGE'i sündroomi. *WES* analüüsil tuvastatud c.1480C>T mutatsioon kinnitas patsiendi EST10 kliinilist diagnoosi.

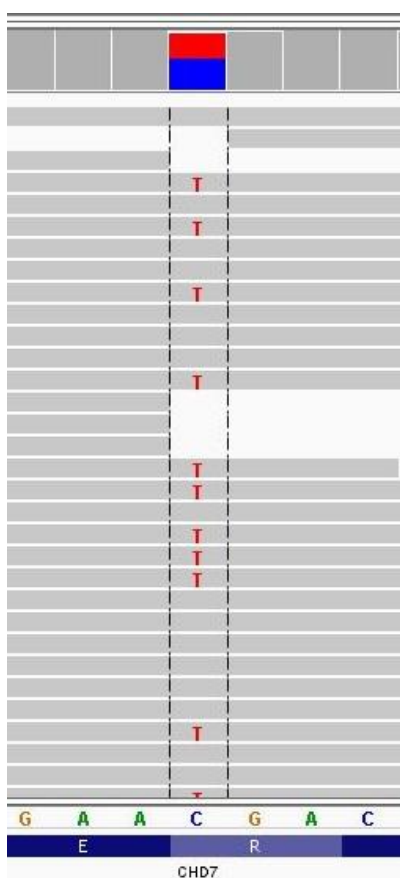
¹⁹ Vaata <http://ghr.nlm.nih.gov/gene/CHD7>

²⁰ Vaata <http://ghr.nlm.nih.gov/condition/charge-syndrome>

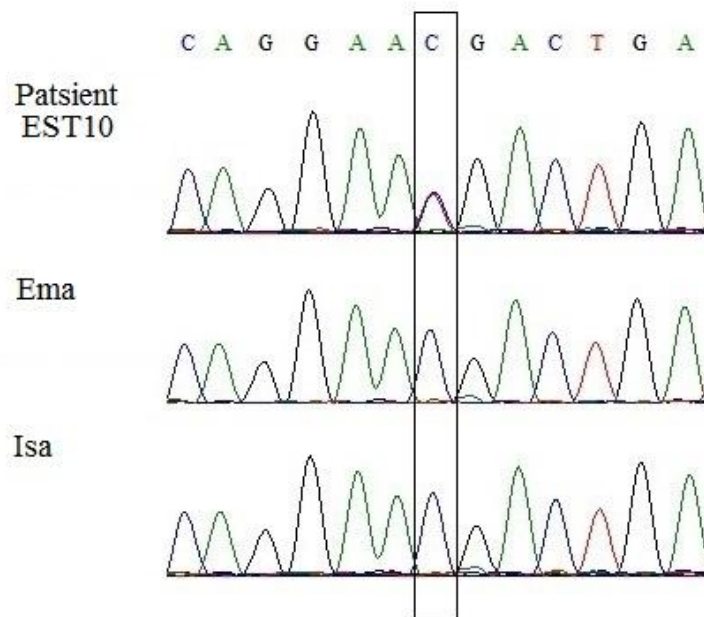
Tabel 4. WES analüüsi tulemusena patsiendil EST10 tuvastatud haiguspõhjulik mutatsioon geenis *CHD7*.

Kromosoom/ positsioon	Geen	Asendus nukleotiidses järjestuses	Asendus valgu- järjestuses	Asukoht	PolyPhen tulemus
8: 61 655 471*	CHD7	c.1480C>T	p. Arg494X	Eksonis 2/38, stoppkoodon	haigus- põhjuslik

(*tähistab positsiooni vastavalt hg19 inimese genoomi assambleeringule)



Joonis 13. Väljavõte programmist IGV. Musta punktiirjoonega on tähistatud positsioon c.1480, mille suhtes on patsient EST10 heterosügootse C/T genotüübiga. Nimetatud positsioon on 38-kordse katvusega, millest C alleeli esineb 21x (55%) ja T alleeli 17x (45%).



Joonis 14. Sanger sekveneerimise graafik. Musta raamiga on tähistatud uuritav positsioon c.1480 geenis *CHD7*. Patsiendil EST10 esineb vastavalt positsioonis heterosügootne C/T genotüüp, patsiendi vanematel esineb referentsjärjestusega sama homosügootne C/C genotüüp.

2.4 Arutelu

Käesoleva töö eesmärgiks oli uurida võimalike haiguspõhjuslike geenivariantide esinemist ebaselge etioloogiaga arenguhäiretega patsientide valimis kasutades eksoomi sekveneerimise meetodit.

Kuna enamik haigusi on põhjustatud mutatsioonidest just valke kodeerivates alades (Choi *et al.*, 2009; Lalonde *et al.*, 2010), on *WES*-ist saanud kliinilises praktikas väga tõhus ja usaldusväärne tööriist. Tänapäeval on tänu *WES*-ile võimalik lahendada keerulisi haigusjuhtumeid, mida seniste diagnostiliste võimalustega on olnud võimatu diagnoosida. Praeguseks on *WES* analüüsi edukalt rakendatud juba rohkem kui 150 haiguse põhjuse väljaselgitamisel (Zhang, 2014).

Võrreldes traditsiooniliste lookusspetsiifiliste uurimismeetoditega, võimaldab *WES* uurida ühe eksperimendi tulemusena hüpoteesivabalt kõikide teadaolevate geenide järjestusi. Traditsiooniliste uurimismeetodite rakendamisel peab raviarst või meditsiinigeneetik teadma, millise geeni defekti patsiendil uurida. Selline lähenemine võib selge kliinilise diagnoosiga patsiendi puhul olla edukas (Katsanis and Katsanis, 2013). Näiteks *CHARGE* (Vissers *et al.*, 2004) või *Apert*'i sündroomiga (Bochukova *et al.*, 2009) patsientidel, mille puhul on haigus põhjustatud mutatsioonidest väga vähestes geenides. Sanger sekveneerimise meetodit rakendatakse tänapäeval molekulaardiagnostilise analüüsina rutiinses kliinilises praktikas pärilike haiguste ja vähkkasvajate diagnoosimiseks. Sanger sekveneerimist kasutatakse ka sõltumatu analüüsimeetodina *NGS* tulemuste valideerimisel. Traditsioonilistest uurimismeetodidest rakendatakse geenipõhiste uurimismeetoditena veel lisaks aheldatuse analüüsi koos positsioonilise kloonimisega, mis siiski tänapäeval enam nii laialdast kasutust ei leia. Traditsiooniliste uurimismeetodite rakendamise tulemusena on tänaseks teada umbes 3000 Mendeliaalse haiguse põhjus (Bamshad *et al.*, 2011). Esimene edukas aheldatuse analüüsi rakendamine haiguspõhjusliku geeni leidmiseks oli 1989. aastal, kui tuvastati tsüstilist fibroosi põhjustav geen - *CFTR* (*cystic fibrosis transmembrane conductance regulator*) (Riordan *et al.*, 1989).

Käesoleval ajal puudub siiski täpne ettekujutus, kui paljudel juhtudel aitab eksoomi sekveneerimine selgitada välja ebaselge etioloogiaga geneetilise haiguse põhjuse. Erinevate töögruppide poolt on publitseeritud arvukaid artikleid *WES* rakendamisest *ID*, epilepsia, *ASD*, skisofreenia jt haiguste geneetiliste põhjuste uurimisel. Nendes uurimistöodes raporteeritud diagnostiline saagis varieerub, mis on tingitud haiguse geneetilisest ja kliinilisest

heterogeensusest. Näiteks de Ligt jt kolleegid uurisid 100 idiopaatilise *ID*-ga patsiendi *WES* andmeid, tuvastades haiguspõhjusliku mutatsiooni 16 patsiendil, mis teeb diagnostiliseks saagiseks 16% (de Ligt *et al.*, 2012). *WES* rakendamine haiguse põhjuse väljaselgitamisel võib olla märksa edukam, kui uurida selge kliinilise diagnoosiga, näiteks Miller'i sündroomi (Ng *et al.*, 2010) või Schinzel-Giedioni sündroomiga (Hoischen *et al.*, 2010) patsiente. Käesoleval ajal ainus *WES* diagnostiline uuring, mille puhul patsiendid ei olnud valitud teatud kindla haigusgrupi alusel, on publitseeritud Yangi jt kolleegide poolt. Selle uuringu raames uuriti 250 teadmata põhjusega neuroloogiliste ja mitteneuroloogiliste sümptomitega patsiente. Haiguspõhjuslikud geenivariandid tuvastati 62 patsiendil, mis teeb diagnostiliseks saagiseks 25% (Yang *et al.*, 2013). Arvatakse, et *WES*-i rakendamisel kliinilises praktikas ebaselge etioloogiaga patsientide valimis võikski diagnostiline saagis olla umbes 25%.

Käesolevas töös tuvastati 11 patsiendi eksoomi sekveneerimise analüüsi tulemusena haiguspõhjuslik leid kahel patsiendil, mis teeb diagnostiliseks saagiseks ~18%. Arvestades, et eelnevad diagnostilised uuringud ei olnud selgitanud haiguse põhjust, võimaldas eksoomi sekveneerimine lahendada kaks keerulist haigusjuhtumit (patsiendid EST5 ja EST10). Samas ei selgita c.674G>A mutatsioon *RUNX2* geenis kõiki patsiendi kliinilisi tunnuseid, eelkõige siseorganite arenguanomaaliad (maksafibroos, maksa ja põrna suurenemine), hüperelastset nahka ja liigeste hüpermobiilsust, mis annab alust arvata, et patsiendi EST5 kliinilise pildiga on seotud mitu haiguspõhjuslikku muutust. Patsiendi EST10 kliinilise pildi järgi kahtlustati *CHARGE* sündroomi, mis võib olla põhjustatud mutatsioonidest kas *CHD7* või *SEMA3E* geenides. Seega arvestades patsiendi selget kliinilist diagnoosi, oleks selle juhtumi puhul õigustatud ka nende geenide Sanger sekveneerimine. Seejuures on teada, et mutatsioonid *CHD7* geenis vastutavad rohkem kui poolte *CHARGE* sündroomi juhtude eest²¹.

Kuigi üheksal patsiendil ei tuvastatud eksoomi sekveneerimise uuringu tulemusena haiguspõhjuslikku leidu, võivad selle põhjused olla mitmesugused. Näiteks võivad haiguspõhjuslikud mutatsioonid asuda mittekodeerivates alades (Ng *et al.*, 2010) või on haigus põhjustatud tandemsetest kordusjärjestustest, mida *WES*-i rakendades pole võimalik tuvastada (Shashi *et al.*, 2014). Lisaks pole 100% usaldusväärsusega võimalik kõiki kodeerivaid järjestusi uurida. Kuna eksoomi sekveneerimist on analüüsimeetodina rakendatud vaid paar aastat, pole teadmised Mendeliaalsete haigustega seotud geenidest ka täielikud. *WES*-iga saadakse informatsiooni kõikide teadaolevate geenide kodeerivate järjestuste kohta, mistõttu saab neid andmeid ka aastate möödudes uuesti analüüsida ja võrrelda erinevate

²¹ Vaata <http://ghr.nlm.nih.gov/condition/charge-syndrome>

andmebaaside vastu. Näiteks polnud veel mõned aastad tagasi teada üks Coffin-Sirise sündroomi põhjustavatest kandidaatgeenidest- *ARID1B* (*AT rich interactive domain 1B*), mis identifitseeriti alles paari aasta eest (Tsurusaki *et al.*, 2012). Koostöös SA TÜK Lastekliinikumiga tuvastati alles hiljaaegu ühel uuritud patsiendist c.1584delG heterosügootne raaminihke mutatsioon *ARID1B* geenis, mis põhjustab Coffin-Sirise sündroomi (Vals *et al.*, 2014). Tegemist oli komplekse kliinilise pildiga patsiendiga, kes oli uurimise hetkel juba 16-aastane. Eelnevad diagnostilised uuringud ei olnud selgitanud haiguse põhjust. Meditsiinigeneetikud kahtlustasid ühe võimalusena Coffin-Sirise sündroomi, kuid uuringuid raskendas ka mitmete kandidaatgeenide testimise vajadus. WES lahendas ühe eksperimendi tulemusena pikalt uuritud haiguse põhjuse. Tuvastatud c.1584delG *ARID1B* geenis polnud varasemalt haiguspõhjuslikuna kirjeldanud.

Kuigi praegu on WGS analüüs kliinilises praktikas veel liiga kallis ja andmete analüüsimine keeruline (Majewski *et al.*, 2011), siis tehnoloogia arenedes võib juba mõne aasta pärast leida laiemalt rakendust ka WGS diagnostilise tööriistana. Näiteid WGS-i diagnostilisest rakendamisest on teada juba praeguseks, näiteks tuvastati neli aastat tagasi WGS-ga haiguspõhjuslik geenivariant *SH3TC2* (*SH3 domain and tetratricopeptide repeat domain 2*) geenis patsiendil, kellel esines Charcot-Marie-Toothi haigus (Lupski *et al.*, 2010).

Eksoomi sekveneerimise meditsiiniliseks näidustuseks diagnostilise analüüsina on eeskätt vastündinu- ja lapseas avalduvad ebaselge etioloogiaga haigused (näiteks intellektipuue, ainevahetushäired, epilepsia jt), sest just selles vanuses on õige diagnoosi ja võimaliku ravi mõju edasisele tervisele suurim. Täpne diagnoos võimaldab paremini hinnata ka haiguse kordusriske perekonnas²².

Antud töös saadud tulemused näitavad, et eksoomi sekveneerimine on väga tõhus ja usaldusväärne analüüsimeetod, mis aitab haiguse põhjuse välja selgitada väga kiiresti. Käesoleva uurimistöö raames tuvastati haiguspõhjuslik leid kahel patsiendil 11-st. Siiski on alust arvata, et tulevikus võivad ka praegu lahendamata jäänud juhtumid jõuda molekulaarse diagnoosini seoses uute haiguspõhjuslike geenivariantide tuvastamisega.

²² Vaata <http://www.mu.ee/?PublicationId=de4ed9b5-3713-4e97-a33c-76270cf264e4>

KOKKUVÕTE

WES võimaldab ühe eksperimendi tulemusena uurida kõikide teadaolevate geenide kodeerivaid järjestusi. Kuna enamik haigusi on põhjustatud mutatsioonidest just valke kodeerivates alades (Choi *et al.*, 2009), on *WES*-ist saanud kliinilises praktikas väga tõhus ja usaldusväärne geenidiagnostika tööriist.

Käesolevas töös analüüsiti 11 ebaselge etioloogiaga arenguhäiretega patsiendi eksoomi sekveneerimisandmeid, et tuvastada võimalikke haiguspõhjuslikke geenivariante. Analüüsi tulemusena tuvastati haiguspõhjuslik leid kahel patsiendil (patsiendid EST5 ja EST10). Patsiendil EST5 tuvastati heterosügootne mittesünonüümne c.674G>A, p.(Arg225Gln) (NM_001024630.3) mutatsioon *RUNX2* geenis. Nimetatud mutatsioon on teadaoleva kliinilise tähtsusega ja põhjustab kleidokraniaalset düsplaasiat, mis sobib ka patsiendi kliiniliste haigustunnustega. Patsiendil EST10 tuvastati heterosügootne mittesünonüümne c.1480C>T, p.(Arg494X) (NM_017780.3) mutatsioon *CHD7* geenis, mis viib enneaegse stoppkoodoni tekkeni. Nimetatud c.1480C>T on teadaoleva kliinilise tähtsusega ja põhjustab CHARGE sündroomi, mis sobib samuti patsiendi kliinilise fenotüübiga. Patsientidel EST5 ja EST10 tuvastatud mutatsioon valideeriti ka Sanger sekveneerimisega patsientidel ja tervetel vanematel. Sanger sekveneerimise tulemused kinnitavad *WES* tulemusi- patsientidel on EST5 ja EST10 tuvastatud haiguspõhjuslikud mutatsioonid on tekkinud *de novo*.

WES-i rakendamine diagnostilise analüüsimeetodina kliinilises praktikas aitab lahendada keerulisi haigusjuhtumeid, mida traditsiooniliste uurimismeetoditega on olnud võimatu diagnoosida. Haiguse molekulaarsete aluste mõistmine ja väljaselgitamine on vajalik teave haiguste diagnoosimiseks ja adekvaatse geneetilise nõustamise pakkumisel peredele.

SUMMARY

Diagnostic exome sequencing in 11 patients with neurodevelopmental disorders

Kadri Onemar

In recent years, with the advent of technology and application of the new sequencing methods it is now possible to sequence whole human genome or only coding-regions – exons – which constitute 1-2% of the human genome. Exome sequencing has become very effective and reliable diagnostic tool because most Mendelian disorders are caused by mutations in the protein coding regions. Establishing a molecular diagnosis gives information about inheritance pattern, allows to offer adequate genetic counseling, estimate risks of recurrence in the family and apply possible treatment.

The aim of this study was to use exome sequencing to search for possible disease-causing variants in 11 patients with unknown etiology of neurodevelopmental disorders. The application of exome sequencing revealed disease-causing mutations in two patients (patients EST5 and EST10). In patient EST5 exome sequencing identified heterozygous nonsynonymous c.674G>A, p.(Arg225Gln) (NM_001024630.3) mutation in *RUNX2* gene. The mutation has previously been described as clinically significant and causes cleidocranial dysplasia. In patient EST10 exome sequencing identified heterozygous nonsynonymous c.1480C>T, p.(Arg494X) (NM_017780.3) mutation in *CHD7* gene which causes premature stop codon. The identified mutation in patient EST10 has previously also been described as clinically significant and causes CHARGE syndrome. These findings were validated in patients and their healthy parents with Sanger sequencing. Sanger sequencing confirmed the mutations c.674G>A and c.1480C>T as *de novo* in patients EST5 and EST10, accordingly.

Exome sequencing as a diagnostic tool is an effective approach for disease gene discovery where traditional methods have failed.

TÄNUAVALDUSED

Eelkõige sooviksin tänada oma juhendajat Margit Nõukast väsimatu juhendamise ja abi eest ning kaasjuhendajat prof Ants Kurge.

Samuti tänan ma TÜ EGV spetsialisti, Mart Kalsi, eksoomi sekveneerimisandmete bioinformaatilise analüüsi eest.

Tänan Eesti Biokeskuse tuumiklabori töötajaid, kes viisid läbi Sanger sekveneerimise.

Veel soovin ma tänada oma pereliikmeid ja sõpru toetuse ja innustuse eest.

KASUTATUD KIRJANDUS

- ACMG Board of Directors (2012). Points to consider in the clinical application of genomic sequencing. *Genet Med* 14(8): 759-761.
- Albert, T. J., Molla, M. N., Muzny, D. M., Nazareth, L., Wheeler, D., Song, X., Richmond, T. A., Middle, C. M., Rodesch, M. J., Packard, C. J., Weinstock, G. M. and Gibbs, R. A. (2007). Direct selection of human genomic loci by microarray hybridization. *Nat Methods* 4(11): 903-905.
- Amarasinghe, K. C., Li, J. and Halgamuge, S. K. (2013). CoNVEX: copy number variation estimation in exome sequencing data using HMM. *BMC Bioinformatics* 14 Suppl 2: S2.
- Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A. and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 12(11): 745-755.
- Baumert, U., Golan, I., Redlich, M., Akin, J. J. and Muessig, D. (2005). Cleidocranial dysplasia: molecular genetic analysis and phenotypic-based description of a Middle European patient group. *Am J Med Genet A* 139A(2): 78-85.
- Behjati, S. and Tarpey, P. S. (2013). What is next generation sequencing? *Arch Dis Child Educ Pract Ed* 98(6): 236-238.
- Bochukova, E. G., Roscioli, T., Hedges, D. J., Taylor, I. B., Johnson, D., David, D. J., Deininger, P. L. and Wilkie, A. O. (2009). Rare mutations of FGFR2 causing apert syndrome: identification of the first partial gene deletion, and an Alu element insertion from a new subfamily. *Hum Mutat* 30(2): 204-211.
- Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32(3): 314-331.
- Choi, M., Scholl, U. I., Ji, W., Liu, T., Tikhonova, I. R., Zumbo, P., Nayir, A., Bakkaloglu, A., Ozen, S., Sanjad, S., Nelson-Williams, C., Farhi, A., Mane, S. and Lifton, R. P. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 106(45): 19096-19101.
- de Ligt, J., Veltman, J. A. and Vissers, L. E. (2013). Point mutations as a source of de novo genetic disease. *Curr Opin Genet Dev* 23(3): 257-263.
- de Ligt, J., Willemsen, M. H., van Bon, B. W., Kleefstra, T., Yntema, H. G., Kroes, T., Vulto-van Silfhout, A. T., Koolen, D. A., de Vries, P., Gilissen, C., del Rosario, M., Hoischen, A., Scheffer, H., de Vries, B. B., Brunner, H. G., Veltman, J. A. and Vissers, L. E. (2012).

Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 367(20): 1921-1929.

Gilissen, C., Hoischen, A., Brunner, H. G. and Veltman, J. A. (2011). Unlocking Mendelian disease using exome sequencing. *Genome Biol* 12(9): 228.

Gilissen, C., Hoischen, A., Brunner, H. G. and Veltman, J. A. (2012). Disease gene identification strategies for exome sequencing. *Eur J Hum Genet* 20(5): 490-497.

Girard, S. L., Gauthier, J., Noreau, A., Xiong, L., Zhou, S., Jouan, L., Dionne-Laporte, A., Spiegelman, D., Henrion, E., Diallo, O., Thibodeau, P., Bachand, I., Bao, J. Y., Tong, A. H., Lin, C. H., Millet, B., Jaafari, N., Joobar, R., Dion, P. A., Lok, S., Krebs, M. O. and Rouleau, G. A. (2011). Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet* 43(9): 860-863.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D. B., Lander, E. S. and Nusbaum, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27(2): 182-189.

Gonzaga-Jauregui, C., Lupski, J. R. and Gibbs, R. A. (2012). Human genome sequencing in health and disease. *Annu Rev Med* 63: 35-61.

Hansen, L., Riis, A. K., Silaharoglu, A., Hove, H., Lauridsen, E., Eiberg, H. and Kreiborg, S. (2011). RUNX2 analysis of Danish cleidocranial dysplasia families. *Clin Genet* 79(3): 254-263.

Hoischen, A., van Bon, B. W., Gilissen, C., Arts, P., van Lier, B., Steehouwer, M., de Vries, P., de Reuver, R., Wieskamp, N., Mortier, G., Devriendt, K., Amorim, M. Z., Revencu, N., Kidd, A., Barbosa, M., Turner, A., Smith, J., Oley, C., Henderson, A., Hayes, I. M., Thompson, E. M., Brunner, H. G., de Vries, B. B. and Veltman, J. A. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* 42(6): 483-485.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431(7011): 931-945.

Janssen, N., Bergman, J. E., Swertz, M. A., Tranebjaerg, L., Lodahl, M., Schoots, J., Hofstra, R. M., van Ravenswaaij-Arts, C. M. and Hoefsloot, L. H. (2012). Mutation update on the CHD7 gene involved in CHARGE syndrome. *Hum Mutat* 33(8): 1149-1160.

Katsanis, S. H. and Katsanis, N. (2013). Molecular genetic testing and the future of clinical genomics. *Nat Rev Genet* 14(6): 415-426.

Krumm, N., Sudmant, P. H., Ko, A., O'Roak, B. J., Malig, M., Coe, B. P., Project, N. E. S., Quinlan, A. R., Nickerson, D. A. and Eichler, E. E. (2012). Copy number variation detection and genotyping from exome sequence data. *Genome Res* 22(8): 1525-1532.

Ku, C. S., Cooper, D. N., Polychronakos, C., Naidoo, N., Wu, M. and Soong, R. (2012). Exome sequencing: dual role as a discovery and diagnostic tool. *Ann Neurol* 71(1): 5-14.

Lalonde, E., Albrecht, S., Ha, K. C., Jacob, K., Bolduc, N., Polychronakos, C., Dechelotte, P., Majewski, J. and Jabado, N. (2010). Unexpected allelic heterogeneity and spectrum of mutations in Fowler syndrome revealed by next-generation exome sequencing. *Hum Mutat* 31(8): 918-923.

Lee, K. E., Seymen, F., Ko, J., Yildirim, M., Tuna, E. B., Gencay, K. and Kim, J. W. (2013). RUNX2 mutations in cleidocranial dysplasia. *Genet Mol Res* 12(4): 4567-4574.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14): 1754-1760.

Luckey, J. A., Drossman, H., Kostichka, A. J., Mead, D. A., D'Cunha, J., Norris, T. B. and Smith, L. M. (1990). High speed DNA sequencing by capillary electrophoresis. *Nucleic Acids Res* 18(15): 4417-4421.

Lupski, J. R., Reid, J. G., Gonzaga-Jauregui, C., Rio Deiros, D., Chen, D. C., Nazareth, L., Bainbridge, M., Dinh, H., Jing, C., Wheeler, D. A., McGuire, A. L., Zhang, F., Stankiewicz, P., Halperin, J. J., Yang, C., Gehman, C., Guo, D., Irikat, R. K., Tom, W., Fantin, N. J., Muzny, D. M. and Gibbs, R. A. (2010). Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* 362(13): 1181-1191.

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A. and Jabado, N. (2011). What can exome sequencing do for you? *J Med Genet* 48(9): 580-589.

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., Howard, E., Shendure, J. and Turner, D. J. (2010). Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7(2): 111-118.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9: 387-402.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. and Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057): 376-380.

- Maxam, A. M. and Gilbert, W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A* 74(2): 560-564.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9): 1297-1303.
- Metzker, M. L. (2005). Emerging technologies in DNA sequencing. *Genome Res* 15(12): 1767-1776.
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet* 11(1): 31-46.
- Morey, M., Fernandez-Marmiesse, A., Castineiras, D., Fraga, J. M., Couce, M. L. and Cocho, J. A. (2013). A glimpse into past, present, and future DNA sequencing. *Mol Genet Metab* 110(1-2): 3-24.
- Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., Beck, A. E., Tabor, H. K., Cooper, G. M., Mefford, H. C., Lee, C., Turner, E. H., Smith, J. D., Rieder, M. J., Yoshiura, K., Matsumoto, N., Ohta, T., Niikawa, N., Nickerson, D. A., Bamshad, M. J. and Shendure, J. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 42(9): 790-793.
- Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M., Huff, C. D., Shannon, P. T., Jabs, E. W., Nickerson, D. A., Shendure, J. and Bamshad, M. J. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 42(1): 30-35.
- Ng, S. B., Nickerson, D. A., Bamshad, M. J. and Shendure, J. (2010). Massively parallel sequencing and rare disease. *Hum Mol Genet* 19(R2): R119-124.
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A. and Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461(7261): 272-276.
- O'Roak, B. J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J. J., Girirajan, S., Karakoc, E., Mackenzie, A. P., Ng, S. B., Baker, C., Rieder, M. J., Nickerson, D. A., Bernier, R., Fisher, S. E., Shendure, J. and Eichler, E. E. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 43(6): 585-589.
- Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endeke, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., Dufke, A., Cremer, K., Hempel, M., Horn, D., Hoyer, J., Joset, P., Ropke, A., Moog, U., Riess, A., Thiel, C. T., Tzschach, A., Wiesener, A., Wohlleber, E., Zweier, C., Ekici, A. B., Zink, A. M., Rump, A., Meisinger, C., Grallert, H.,

Sticht, H., Schenck, A., Engels, H., Rappold, G., Schrock, E., Wieacker, P., Riess, O., Meitinger, T., Reis, A. and Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380(9854): 1674-1682.

Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J. L. and et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245(4922): 1066-1073.

Royer-Pokora, B., Kunkel, L. M., Monaco, A. P., Goff, S. C., Newburger, P. E., Baehner, R. L., Cole, F. S., Curnutte, J. T. and Orkin, S. H. (1986). Cloning the gene for an inherited human disorder--chronic granulomatous disease--on the basis of its chromosomal location. *Nature* 322(6074): 32-38.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74(12): 5463-5467.

Shashi, V., McConkie-Rosell, A., Rosell, B., Schoch, K., Vellore, K., McDonald, M., Jiang, Y. H., Xie, P., Need, A. and Goldstein, D. B. (2014). The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genet Med* 16(2): 176-182.

Shendure, J. and Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol* 26(10): 1135-1145.

Shendure, J., Mitra, R. D., Varma, C. and Church, G. M. (2004). Advanced sequencing technologies: methods and goals. *Nat Rev Genet* 5(5): 335-344.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. and Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature* 321(6071): 674-679.

Zhang, X. (2014). Exome sequencing greatly expedites the progressive research of Mendelian diseases. *Front Med* 8(1): 42-57.

Teer, J. K. and Mullikin, J. C. (2010). Exome sequencing: the sweet spot before whole genomes. *Hum Mol Genet* 19(R2): R145-151.

Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., Fukushima, Y., Homma, T., Kato, M., Hiraki, Y., Yamagata, T., Yano, S., Mizuno, S., Sakazume, S., Ishii, T., Nagai, T., Shiina, M., Ogata, K., Ohta, T., Niikawa, N., Miyatake, S., Okada, I., Mizuguchi, T., Doi, H., Saitsu, H., Miyake, N. and Matsumoto, N. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* 44(4): 376-378.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res* 40(15): e115.

Vals, M. A., Oiglane-Shlik, E., Noukas, M., Shor, R., Peet, A., Kals, M., Kivistik, P. A., Metspalu, A. and Ounap, K. (2014). Coffin-Siris Syndrome with obesity, macrocephaly, hepatomegaly and hyperinsulinism caused by a mutation in the ARID1B gene. *Eur J Hum Genet*.

Wheeler, D. A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y. J., Makhijani, V., Roth, G. T., Gomes, X., Tartaro, K., Niazi, F., Turcotte, C. L., Irzyk, G. P., Lupski, J. R., Chinault, C., Song, X. Z., Liu, Y., Yuan, Y., Nazareth, L., Qin, X., Muzny, D. M., Margulies, M., Weinstock, G. M., Gibbs, R. A. and Rothberg, J. M. (2008). The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452(7189): 872-876.

Wincent, J., Holmberg, E., Stromland, K., Soller, M., Mirzaei, L., Djureinovic, T., Robinson, K., Anderlid, B. and Schoumans, J. (2008). CHD7 mutation spectrum in 28 Swedish patients diagnosed with CHARGE syndrome. *Clin Genet* 74(1): 31-38.

Winnick, E. (2004). DNA sequencing industry sets its sights on the future. *Scientist* 18(18): 44-47.

Vissers, L. E., de Ligt, J., Gilissen, C., Janssen, I., Stehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., van Bon, B. W., Hoischen, A., de Vries, B. B., Brunner, H. G. and Veltman, J. A. (2010). A de novo paradigm for mental retardation. *Nat Genet* 42(12): 1109-1112.

Vissers, L. E., van Ravenswaaij, C. M., Admiraal, R., Hurst, J. A., de Vries, B. B., Janssen, I. M., van der Vliet, W. A., Huys, E. H., de Jong, P. J., Hamel, B. C., Schoenmakers, E. F., Brunner, H. G., Veltman, J. A. and van Kessel, A. G. (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 36(9): 955-957.

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M. R., Leduc, M. S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., Plon, S. E., Lupski, J. R., Beaudet, A. L., Gibbs, R. A. and Eng, C. M. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 369(16): 1502-1511.

KASUTATUD VEEBILEHED

1000 Genomes	http://www.1000genomes.org
Biobase	http://www.biobase-international.com
Children's Mercy Hospital and Clinics	http://www.childrensmercy.org
Genetics Home Reference	http://ghr.nlm.nih.gov
March of Dimes	http://www.marchofdimes.com
Meditisiiniudised	http://www.mu.ee
Primer3	http://primer3.ut.ee
SimulConsult	http://www.simulconsult.com
U.S. National Library of Medicine	http://www.nlm.nih.gov
Wellcome Trust	http://genome.wellcome.ac.uk

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina Kadri Onemar (sünnikuupäev: 20.04.1992)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

Eksoomi sekveneerimise rakendamine kliinilises praktikas 11 haigusjuhtumi näitel

mille juhendajad on Margit Nõukas, M.Sc.; Prof. Ants Kurg, Ph.D.

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates **31.12.2014** kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 26.05.2014