

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

MEDITSIINITEADUSTE VALDKOND
BIO- JA SIIRDEMEDITSIINI INSTITUUT

B-raku retseptorite repertuaari uurimine kahel meetodil

Bakalaureusetöö

Bioloogia ja elustiku kaitse õppekava

12 EAP

Liisa Pomerants

Juhendajad PhD, Liina Tserel
 MSc, Ahto Salumets
 PhD, Tõnis Org

B-raku retseptorite repertuaari uurimine kahel meetodil

B-rakud ja nende toodetud antikehad on olulised omandatud immuunsüsteemi osad. Antikehad neutraliseerivad patogeene ning aktiveerivad teisi immuunsüsteemi osasid. Iga antikeha, membraanes vormis B-raku retseptor, on ühele kindlale antigeenile spetsiifiline. Spetsiifika kujuneb erinevate tugevalt reguleeritud protsesside tulemusel, mille käigus moodustub organismi kõikidest B-raku retseptoritest repertuaar. Repertuaari uurimine on oluline uurimaks immuunvahendatud haigusi, nende tekkemehhanisme ning potentsiaalseid ravivõimalusi. Käesoleva bakalaureusetöö eesmärk oli võrrelda kahte erinevat meetodit tervete inimeste B-raku retseptorite repertuaari uurimisel. Lisaks uuriti kahe terve inimese B-raku retseptorite repertuaari klonaalsust, mitmekesisust, geenisegmentide kasutust ning CDR3 järjestuste ülekattet.

Märksõnad: B-rakud, B-raku retseptor, repertuaar, antikehad

CERCS: B500 Immunoloogia, seroloogia, transplantoloogia

Analysis of the B cell receptor repertoire with two methods

B cells and antibodies produced by them are an essential part of the acquired immune system. Antibodies neutralise pathogens by binding them and activate other integral parts of the immune system. Each antibody or its membrane bound form B cell receptor is specific to a particular antigen. The specificity develops through highly regulated processes during which all the B cell receptors in the organism form the B cell receptor repertoire. Analysis of the B cell receptor repertoire is important for researching different immune mediated diseases to improve understanding of their pathogenesis and search for potential treatments. The purpose of this bachelor's thesis was to compare two methods for analysing the B cell receptor repertoire. In addition, the clonality, diversity, gene segment usage and CDR3 sequence overlap of B cell repertoire of the two healthy individuals were studied.

Keywords: B cells, B cell receptor, repertoire, antibodies

CERCS: B500 Immunology, serology, transplantation

SISUKORD

KASUTATUD LÜHENDID	5
SISSEJUHATUS	6
1. KIRJANDUSE ÜLEVAADE.....	7
1.1. Omandatud immuunsüsteem.....	7
1.2. Antikehade ehitus ja spetsiifilisus.....	7
1.3. V(D)J rekombinatsioon.....	9
1.3.1. Immuunglobuliinide lookuse iduliini organisatsioon	9
1.3.2. V(D)J rekombinatsioon	10
1.4. Afiinsusküpsemine.....	12
1.5. B-rakkude retseptorite repertuaari uurimine	13
1.5.1. BCR-i repertuaari uurimine Chudakovi meetodiga.....	14
1.5.2. BCR-i repertuaari uurimine Archer Immunoverse kitiga.....	15
1.5.3. Andmete töötlemise programmid	16
2. EKSPERIMENTAALNE OSA	17
2.1. Töö eesmärgid.....	17
2.2. Materjal ja metoodika	17
2.2.1. Vere võtmine	17
2.2.2. PBMC-de eraldamine	17
2.2.3. RNA eraldamine	18
2.2.4. Chudakovi protokoll.....	18
2.2.5. Archer Immunoverse kit.....	21
2.2.6. Sekvenerimine (NGS).....	22
2.2.7. Bioinformaatiline analüüs	22
2.3. Tulemused.....	23
2.3.1. Sekvenerimise kvaliteet	23
2.3.2. MiXCR tulemused.....	24
2.3.3. Repertuaari klonaalsus.....	25

2.3.4. CDR3 pikkused	28
2.3.5. Repertuaaride CDR3 järjestuste ülekate.....	28
2.3.6. Repertuaaride mitmekesisus	30
2.3.7. Geenisegmentide kasutus	35
2.4. Arutelu	38
KOKKUVÕTE	42
SUMMARY	43
TÄNUSÕNAD	44
KASUTATUD KIRJANDUS	45
LISAD	50
LIHTLITSENTS.....	53

KASUTATUD LÜHENDID

- AID aktivatsioon-indutseeritud tsütidiini deaminaas (*activation-induced cytidine deaminase*)
- AMP *Anchored Multiplex PCR*
- BCR B-raku retseptor (*B cell receptor*)
- cDNA komplementaarne DNA (*complementary DNA*)
- CDR komplementaarsust määrav regioon (*complementarity determining regions*)
- FFPE formaliinis fikseeritud, parafiinis hoitud (*formalin-fixed paraffin-embedded*)
- Ig immuunglobuliin (*immunoglobulin*)
- MBC molekulaarne vötkood (*molecular barcode*)
- PBMC perifeerse vere mononukleaarsed rakud (*peripheral blood mononuclear cells*)
- PCR polümeraasi ahelreaktsioon (*polymerase chain reaction*)
- RACE *rapid amplification of cDNA ends*
- RAG rekombinatsiooni aktiveeriv geen (*recombination activating gene*)
- RIN *RNA integrity number*
- RSS rekombinatsiooni signaaljärjestus (*recombination signal sequence*)
- SHM somaatilised hüpermutatsioonid (*somatic hypermutations*)
- ssDNA üheaahelaline DNA (*single stranded DNA*)
- TCR T-raku retseptor (*T cell receptor*)
- TD tõeline mitmekesisus (*true diversity*)
- TdT terminaalne deoksünukleotidültransferaas (*terminal deoxynucleotidyl transferase*)
- UMI unikaalne molekulaarne identifitseerija (*unique molecular identifier*)

SISSEJUHATUS

B-rakud on oluline omandatud immuunsüsteemi osa. Nad toodavad antikehasid, mis neutraliseerivad patogeene, takistades neil rakku sisenemast ja on olulised teiste immuunrakkude ning kaasasündinud immuunsüsteemi mehhanismide aktiveerijad. Antikehad, membraanseoselise vormina B-raku retseptorid, on spetsiifilised kindlale antigeenile. Spetsiifika tekib tänu erinevatele protsessidele, mis nende afiinsust suurendavad. Nendeks protsessideks on V(D)J rekombinatsioon, isotüübi ümberlülitumine ning afiinsusküpsemine. Kui V(D)J rekombinatsioon toimub ka arenevates T-rakkudes, on isotüübi ümberlülitumine ja afiinsusküpsemine B-rakkudele spetsiifilised protsessid. Antigeeni suhtes kõige suurema afiinsusega antikehad läbivad positiivse ja klonaalse selektsiooni ja prolifereeruvad – toimub klonaalne laienemine. Afiinsust suurendavate protsesside tulemusena tekib organismi kõikide B-raku retseptorite kogumik ehk B-raku retseptorite repertuaar.

B-raku retseptorite repertuaari uurimine on immuunvahendatud haiguste uurimisel väga oluline, sest defektid B-rakkude arengus, selektsioonis ja funktsioonis viivad autoimmuunsuse, halvloomuliste kasvajate, immuunpuudulikkuse või allergia tekkeni. Seega on oluline teada mehhanisme, kuidas B-rakud arenevad ja kuidas kujuneb B-raku retseptorite repertuaar.

Käesoleva bakalaureusetöö eesmärkideks on võrrelda kahte erinevat raamatukogude ettevalmistuse meetodit B-rakkude retseptorite repertuaari uurimiseks ning teiseks analüüsida kahe terve indiviidi BCR repertuaari klonaaalsust, mitmekesisust, geenisegmentide kasutust ja CDR3 järjestuste vahelist ülekattet.

Töö eksperimentaalne osa teostati Tartu Ülikooli meditsiiniteaduste valdkonnas bio- ja siirdemeditsiini instituudis molekulaarpatoloogia uurimisrühmas.

1. KIRJANDUSE ÜLEVAADE

1.1. Omandatud immuunsüsteem

Immuunsüsteem jaguneb laias laastus kaheks: kaasasündinud ning adaptiivne ehk omandatud immuunsus (Parkin ja Cohen, 2010). Inimene kasutab mõlema süsteemi mehhanisme ja nende koosmõju, et patogeene organismist kõrvaldada (Chaplin, 2010; Parkin ja Cohen, 2010). Adaptiivne immuunsus koosneb väiksel hulgal rakkudest, mis on spetsiifilised üksikule patogeenile või allergeenile. Efektiivse immuunvastuse tekkeks peavad spetsiifilised rakud pärast antigeeniga kokkupuudet prolifereeruma (Chaplin, 2010). Erinevalt kaasasündinud immuunsusest, on adaptiivsel immuunsusel ka mälu, mis võimaldab teistkordsel patogeeniga kokkupuutel kiiremat immuunvastust (Marshall *et al.*, 2018). Adaptiivse immuunsüsteemi peakomponendid on B- ja T-lümfotsüüdid, mille pinnal on antigeen-spetsiifilised retseptorid (BCR – B-raku retseptor, *B cell receptor* ja TCR – T-raku retseptor, *T cell receptor*), millest igaüks tunneb ära ühte kindlat antigeeni (Parkin ja Cohen, 2010).

Adaptiivne immuunsus jaotub klassikaliselt rakuliseks ja humoraalseks. Rakulist immuunvastust vahendavad T-rakud ja nende alampopulatsioonid. Nad on olulised immuunvastuse reguleerijad, mis tunnevad ära ja ründavad sihtmärkmolekule ning aktiveerivad teisi immuunrakke. Humoraalne immuunvastus seisneb aga B-rakkudes ja nende toodetud antikehades (Hoffman *et al.*, 2016). B-rakud ja antikehad kaitsevad peremeesorganismi järgenvalt: neutraliseerivad mikroobseid toksine, ennetavad ja takistavad patogeenile seondudes nende levikut, opsoniseerivad patogeeni fagotsütoosiks, aktiveerivad komplemendi süsteemi, märgistavad patogeene T-rakuliseks immuunvastuseks ning ka aktiveerivad enda Fc retseptori kaudu teisi immuunrakke. (Hoffman *et al.*, 2016; Parkin ja Cohen, 2010).

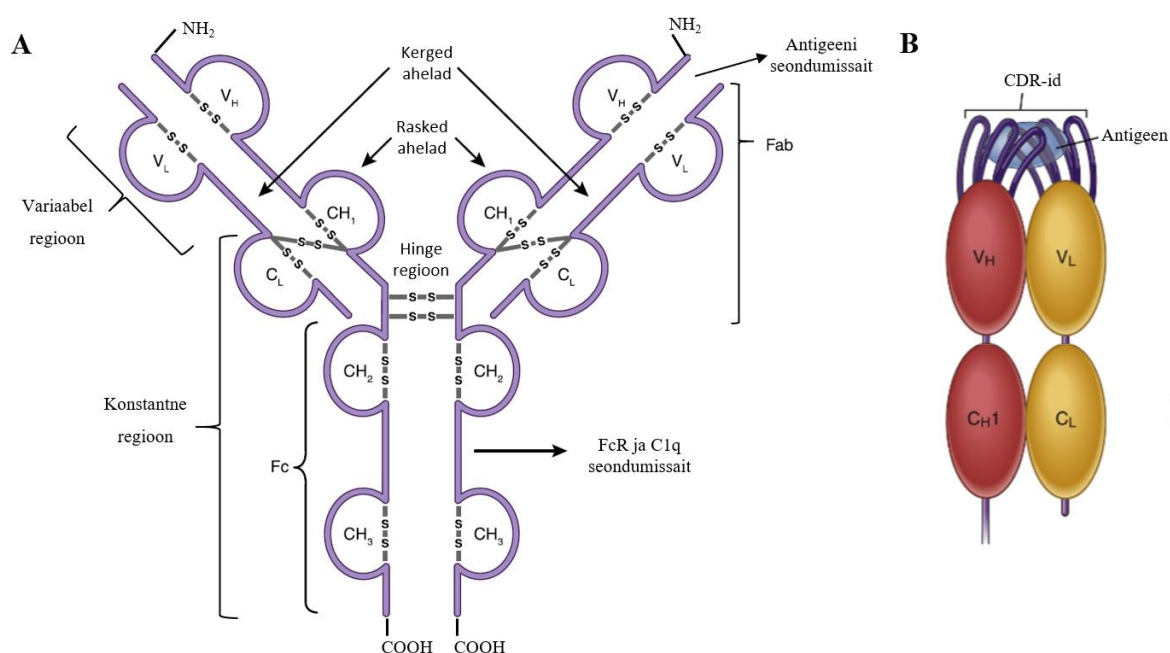
B-rakud ja nende toodetud antikehad on peamised osad patogeeni sissetungi poolt indutseeritud humoraalses immuunvastuses. B-rakuline immuunvastus on seeria järjestikuseid protsesse, mis lõpevad mälu-B-rakkude või antikehasid tootvate plasmarakkude tootmisega (Kurosaki, 2010). BCR-de spetsiifilisuse teke toimub läbi V(D)J rekombinatsiooni, isotüüpide ümberlülitumise ning afiinsusküpsemise. Kõikide organismis leiduvate BCR-ide hulka nimetatakse B-raku retseptorite repertuaariks (Bashford-Rogers *et al.*, 2019). Inimese BCR-de repertuaar on potentsiaalselt võimeline ära tundma 5×10^{13} erinevat antigeeni (Pieper *et al.*, 2013).

1.2. Antikehade ehitus ja spetsiifilisus

Antikehad on Y-kujulised glükoproteiinid, millel on sarnane struktuur (Wang *et al.*, 2019). Antikehad esinevad nii B-raku membraanseoselise vormina ehk B-raku antigeense retseptorina (BCR) kui ka sekreteeritava vormina, mis seonduvad antigeeniga ja neutraliseerivad selle

(Hoffman *et al.*, 2016). Antikeha molekul koosneb kokku neljast valguahelast: kahest identsest raskest ja kahest identsest kergest ahelast, mis on omavahel kovalentselt disulfiidsildadega ühendatud (Chiu *et al.*, 2019; Cohen, 1975; Hoffman *et al.*, 2016, joonis 1A).

Igal antikeha ahelal on üks N-terminaalne variaabel domään ja C-terminaalne konstantne domään. Kergetel ahelatel on ainult 1 C-domään, kuid rasketel ahelatel on neid 3–4, sõltuvalt antikeha isotüübist (Schroeder ja Cavacini, 2010). Ahelate variaabel domäänid on tekitatud V(D)J rekombinatsiooni abil, kus kombineeritakse V (variaabel, *variable*), D (diversiteet, *diversity*, ainult rasketel ahelatel, J (ühendav, *joining*) ja C (konstantne, *constant*) geenisegmente, et tekitada võimalikult suure spetsiifikaga antikeha (Chiu *et al.*, 2019, Stanfield ja Wilson, 2015).



Joonis 1. A: Antikeha struktuur. Joonisel on toodud variaabel ja konstantsed regioonid, Fab ja Fc fragmendid, N- ja C-terminaalsed osad. Näidatud on rasked ja kerged ahelad ning ka hinge region (kohandatud Hoffman *et al.*, 2016). **B: CDR regiooni paiknemine antikehal.** VH ja VL on antikeha variaabel osad ning CH1 ja CL tähistavad konstantseid alasid antikehal vastavalt raskel ja kergel ahelal. (kohandatud Abbas *et al.*, 2018)

Rasketel ahelatel on CH1 ja CH2 domäänide vahel ka hinge region, mis jaotab antikeha Fab ja Fc fragmentideks (Cohen, 1975). Fab regioon paikneb N-terminaalses osas ning on segu kerge ja raske ahela variaabel domäänist (VL ja VH) ja konstantsest domäänist (CL ja CH1) (Stanfield ja Wilson, 2015). Fab regioon sisaldab variaabel domäänide N-terminaalsetes otstes ka kuut komplementaarsust määravat regiooni (CDR, *complementarity determining region*) (Joonis 1B) (Stanfield ja Wilson, 2015). CDR-id osalevad antigeeni ära tundmises moodustades antigeeni seondumissaidi, mille järjestused on antikehas kõige varieeruvamad ja määravad antikeha spetsiifilisuse (Chiu *et al.*, 2019; Hoffman *et al.*, 2016; Schroeder ja Cavacini, 2010).

Kõige varieeruvam CDR-dest on raske ahela CDR3 (CDRH3), kuna see sisaldab nii V, D kui J segmente (Jung *et al.*, 2006). CDRH3 on ühtlasi terve antikeha kõige varieeruvam ala ning suurim antikeha spetsiifilisuse determinant (Arnaout *et al.*, 2011; Stanfield ja Wilson, 2015). Antikeha C-terminaalses osas paikneb Fc fragment, mis koosneb raske ahela konstantse ala dimeeridest (CH₂ ja CH₃ segmendid, Joonis 1A). Fc fragment vahendab antikehade efektorfunktsioone, kui sellele seonduvad Fc retseptori kaudu komplemendi süsteemi valgud (jooniselJoonis 1A toodud C1q) või teised immuunrakud, mis toodavad põletiku mediaatoreid ja reguleerivad immuunvastust (Hoffman *et al.*, 2016; Mkaddem *et al.*, 2019; Stanfield ja Wilson, 2015).

1.3. V(D)J rekombinatsioon

1.3.1. Immuunglobuliinide lookuse iduliini organisatsioon

B-rakkude iduliini lookuses paiknevad kolme tüüpi geenisegmentid, mis ümberkorraldatult moodustavad antikehade ja BCR-i variaabel eksoni. Need on V, D ja J geenisegmentid, millest raske ahelas esinevad kõiki kolme tüüpi segmendid, kuid kerges ahelas, mida on nii κ kui λ tüüpi (Cohen, 1975), esinevad ainult V ja J segmendid (Lewis, 1994). Geenisegmentide arv on toodud tabelis 1.

Tabel 1. Geenisegmentide arv Ig lookuses (kohandatud Abbas *et al.*, 2018)

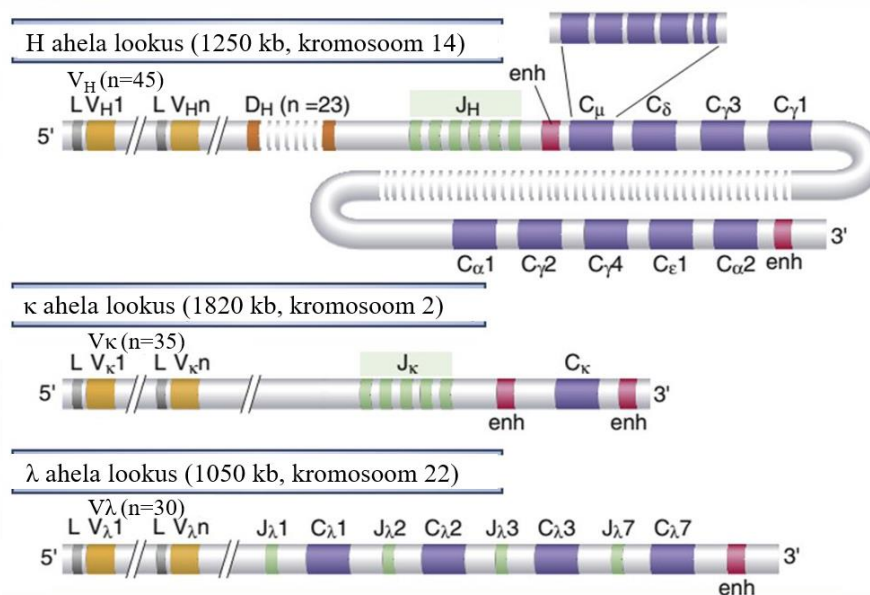
	Ig lookuse geenisegmentide arv		
	Raske ahel	κ	λ
V segmendid	45	35	30
D segmendid	23	0	
J segmendid	6	5	4

Genoomis on antikehade rasket ja kergeid ahelaid kodeerivad IgH, Ig κ , ja Ig λ lookused paigutatud eri kromosoomidesse. Kerge ahela κ tüüpi ahela geenid paiknevad 2. kromosoomis ja λ ahela omad 22. kromosoomis (Chaplin, 2010; Townsend *et al.*, 2016). Raske ahela geenid asuvad 14. kromosoomis (Schroeder ja Cavacini, 2010). V(D)J rekombinatsiooni käigus ühendatud V, D ja J segmendid ühendatakse RNA splaissimise käigus C eksoniga, moodustades terve variaabel geeni (Little *et al.*, 2015). Antikehade erinevate isotüüpide C domääne kodeerivad erinevad C_H eksonite klastrid, mis on paigutatud IgH lookuses kindlasse järjekorda (Chi *et al.*, 2020).

Iduliinis on geenisegmentid paigutatud kindlasse järjekorda (Joonis 2). 5' otsas paiknevad V segmendid, seejärel (raske ahela puhul) D segmendid ning J segmendid. C segmendid

paiknevad lookuse lõpus. Raske ahela puhul on funktsionaalseid C_H geenisegmente 9, kuid κ ja λ ahelate puhul vastavalt üks ja neli. Kergete ahelate puhul esineb pärast splaissimist ainult üks C_L ekson, kuid raske ahela puhul moodustab C_H geeni viis kuni kuus eksonit (nagu C_μ joonisel 2). (Abbas *et al.*, 2018; Chaplin, 2010)

Iga V segmendi alguses on 20–30 nt pikkune liiderpeptiid. Lookuses on ka palju mittekodeerivaid alasid (enhanserid, vaigistajad ja promootorid), millel on oluline roll rekombinatsiooni läbiviimisel ja geeniekspressioonil. (Abbas *et al.*, 2018)



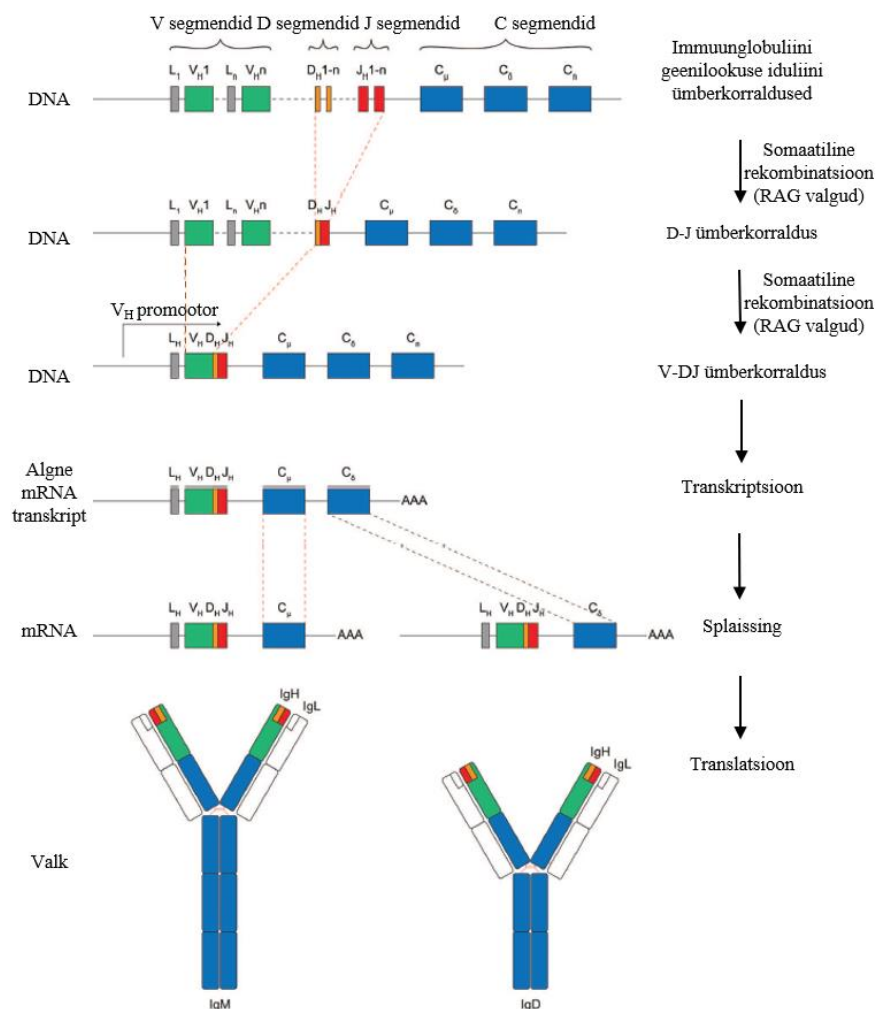
Joonis 2. Immuunglobuliini lookuse iduliini organisatsioon. Joonisel on näidatud inimese raske ahela, κ ja λ kergete ahelate lookused. C_H geenid on joonisel lihtsustamiseks toodud ühe kastina, tegelikkuses koosnevad mitmest eksonist nagu C_μ puhul näidatud. Geenisegmentide tähistused: L – liiderjärjestus, V – variaabel (*variable*), D – diversiteet (*diversity*), J – ühendav (*joining*), C – konstantne (*constant*) ning enh – enhanser (kohandatud Abbas *et al.*, 2018).

1.3.2. V(D)J rekombinatsioon

BCR-e kodeerivad geenid ja antikehade variaabel domääni eksonid tekitatakse varases B-rakkude arengustaadiumis V(D)J rekombinatsiooni käigus, mille jooksul V, D ja J segmendid paigutatakse ümber (Schatz *et al.*, 1989). Selle protsessi käigus tekib antikehade mitmekesisus (Kurosaki, 2010) ning iga B-rakk hakkab tootma kindla spetsiifikaga antikeha (Chaplin, 2010). Protsess algab raske ahela IgH lookuse rekombineerumisega, seejärel rekombineerub kerget ahelat kodeeriv Ig- κ või Ig- λ lookus (Uibo *et al.*, 2015). V(D)J rekombinatsioon on saitspetsiifiline ning toimub ainult arenevates B- ja T-lümfotsüütides ning ainult vastavalt Ig ja TCR geenisegmentide vahel (Bassing *et al.*, 2002).

V(D)J rekombinatsioon on tugevalt reguleeritud protsess, mis toimub kindlas järjekorras. Kõigepealt ühendatakse raskes ahelas D segment J_H segmendiga, seejärel toimub V_H – DJ_H rekombinatsioon, transkriptsioon, splaiising ja translatsioon nagu näidatud joonisel 3 (Oka ja

Kawaichi, 1995). Tulemuseks on raske ahela V_HD_HJ_H ekson, mis kodeerib antikeha variaabel domääni (Kurosaki *et al.*, 2010). Pärast raske ahela V(D)J rekombinatsiooni toimub V–J rekombinatsioon kerges ahelas (Oka ja Kawaichi, 1995).



Joonis 3. V(D)J rekombinatsiooni skeem ja etapid Ig raske ahela lookuses. (kohandatud Backhaus, 2018)

1.3.2.1. V(D)J rekombinatsiooni mehhanism

V(D)J rekombinatsioonis on oluline roll Ig geenisegmentide ümber külgnevatel rekombinatsiooni signaaljärjestustel (edaspidi RSS, *recombination signal sequence*), mis on vajalikud rekombinatsiooni alustamiseks ja suunamiseks (Oettinger *et al.*, 1990; Schatz *et al.*, 1989). V(D)J rekombinatsiooni initsieerib lümfoïdspetsiifiline ja konserveerunud RAG1-RAG2 (rekombinatsiooni aktiveeriv geen, *recombination activating gene*) valkude kompleks (Alt *et al.*, 1992; Oettinger *et al.*, 1990; Schatz *et al.*, 1989). RAG kompleks tunneb ära RSS-ide, seondub nendele ning teeb geenisegmentide ja RSS-ide ühenduskohtadesse kaheahelised katked (Schatz ja Ji, 2011). Tekkinud reaktsiooni tulemusena moodustub kodeerivate geenisegmentide otsi sisaldav juuksenõelastruktuur ning RSS otsi sisaldav tõmbi otsaga struktuur (Little *et al.*, 2015).

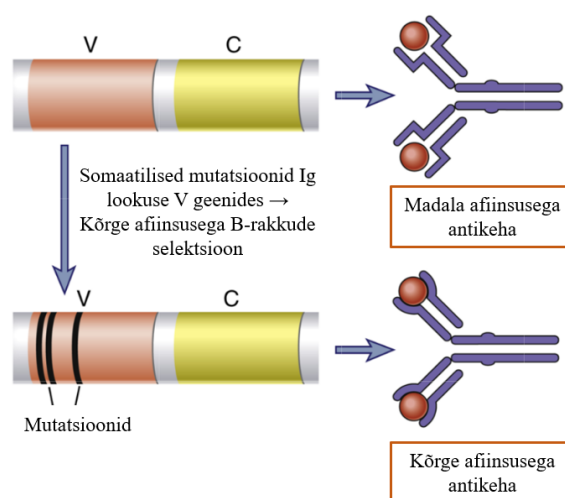
Järgnevalt tehakse juuksenõelastruktuuri katke, mis tekitab 3' üleulatuva otsa, kuhu TdT (terminaalne deoksü nukleotidüül-transferaas, *terminal deoxynucleotidyl transferase*) lisab täiendava mitmekesisuse tekitamiseks lisanukleotiide (Alt *et al.*, 1992; Lieber, 2010; Schroeder ja Cavacini, 2010). Edasi ühendatakse kodeerivad lõpud vigaderohkelt mittehomooloogsete DNA-otste ühendamise süsteemi valkude abil (Schroeder ja Cavacini, 2010; Tonegawa, 1983). Signaaljärjestuste lõpud ühendatakse lisamodifikatsioone lisamata ning protsessi käigus moodustuvad signaalühendused (Oka ja Kawaichi, 1995). Ig lookuses toimuva V(D)J rekombinatsiooni tulemusel transkribeerivad ebaküpsed naiivsed B-rakud vastavaid geene. Geenide transkriptid läbivad alternatiivse splaissingu, ekspresseerides küpsete naiivsete B-rakkudena nii IgD kui IgM tüüpi retseptorit (Chaudhary ja Wesemann, 2018).

1.4. Afiinsusküpsemine

Afiinsusküpsemine toimub siis, kui V(D)J rekombinatsiooni käigus tekkinud funktsionaalsete BCR-i järjestused on moodustunud ning seda BCR-i kandev B-rakk on tundnud ära enda spetsiifilise antigeeni ning aktiveerunud (Little *et al.*, 2015; Chaudhary ja Wesemann, 2018). Aktivatsioonis on toimunud ka isotüübi ümberlülitumine, muutes antikeha efektorfunktsiooni, kuid jättes spetsiifilisuse muutmata (Honjo *et al.*, 2002). Afiinsusküpsemise tulemusel aga tekitatakse antigeeniga kokkupuutel kõrgema afiinsusega antikehasid (Mishra ja Mariuzza, 2018). Protsess ise toimub sekundaarsete lümfoidorganite idukeskuses (Schroeder ja Cavacini, 2010). Afiinsusküpsemine toimub läbi somaatiliste hüpermutatsioonide (SHM, *somatic hypermutation*), mille käigus toimuvad muutused antikehade V-segmentides, mis eristavad neid algsetest IgM või IgD järjestustest (Uibo *et al.*, 2015). Afiinsusküpsemine toimub läbi ensüümi AID (aktivatsioon-indutseeritud tsütidiini deaminaas, *activation-induced cytidine deaminase*), mis deamineerib ssDNA (üheahelaline DNA, *single strand DNA*) tsütidiini jäägid, muutes selle uridiiniks (Schroeder ja Cavacini, 2010). Tsütidiini mutatsioonid tekivad tihti nn *hotspot* motiivide asukohta, kus mutatsioonisagedus on suurem (Schroeder ja Cavacini, 2010). Seejärel töötlevad neid kahjustatud kohti vearohked DNA parandussüsteemid, mis tekitavad laialdaselt punktmutatsioone, suurendades seeläbi veelgi mitmekesisust (Schroeder ja Cavacini, 2010). Eriti palju punktmutatsioone tekib CDR3 kodeerivates alades, mis tõstavad antikeha afiinsust antigeeni suhtes (Schroeder ja Cavacini, 2010).

SHM esinemissagedus on umbes 1 mutatsioon 1000 nukleotiidi kohta (Uibo *et al.*, 2015). Samas on see protsess juhuslik, nii et antikeha afiinsus võib nii tõusta kui langeda. BCR-i antigeeniga seondumisel töödeldakse antigeen ning esitletakse T-helper rakkudele. Kõige afiinsema BCR-ga B-rakud läbivad idukeskuses positiivse selektsiooni ning diferentseeruvad plasma- ja mälorakkudeks (Chaplin, 2010). Kõige suuremate afiinsustega küpsed B-rakud

hakkavad prolifereruma ja läbivad klonaalset laienemist, tootes tuhandeid sama CDR3 järjestuse ja ühe spetsiifikaga antikeha kloonide. Afiinsusküpsemise skeem on toodud joonisel 4.



Joonis 4. Afiinsusküpsemise üldskeem (kohandatud Abbas *et al.*, 2018).

1.5. B-rakkude retseptorite repertuaari uurimine

Immuunrepertuaari uuringud on üha arenev uurimisvaldkond, kuna *high throughput* sekveneerimise meetodid on suurendanud võimalusi nii B- kui T-rakkude mitmekesisuse uurimiseks (Chaudhary ja Wesemann, 2018). Kuid sekveneerimise odavnemine ja selle kättesaadavuse suurenemine ei muuda sekveneerimisandmete, eriti immuunrakkude retseptorite (BCR ja TCR) repertuaari järjestuste, analüüsi lihtsaks (Yaari ja Kleinstein, 2015). Pigem teeb andmemahu kasv keeruliseks nende analüüsiks õige meetodika leidmise ning nõuab efektiivset ja spetsialiseeritud analüüsi (Yaari ja Kleinstein, 2015). Immuunrakkude järjestuste sekveneerimisele eelnevad olulised protsessid on ka proovide ettevalmistamine ja järjestuste analüüsimeetodid (Chaudhary ja Wesemann, 2018). Küll aga pole veel leitud standardset meetodit, kuidas immuunretseptorite repertuaari uurida, kuna võimalusi on palju ja iga teadlane usub, et tema analüüsimeetod on kõige õigem.

B-raku retseptorite repertuaari analüüs on oluline haiguste uurimisel, kuna immuunhaiguste käigus muutub BCR-i repertuaar ning see võib omakorda haiguste kulgu mõjutada (Ghraichy *et al.*, 2018). Muutused võivad toimuda näiteks läbi V(D)J rekombinatsioonis osalevate geenide defektide, autoreaktiivsete B-rakkude tekke või mõne immuunretseptori vähese hulga tõttu (Ghraichy *et al.*, 2018; Rawlings *et al.*, 2017). BCR-ide repertuaari on uuritud erinevate immuunvahendatud haiguste, näiteks HIV-1, primaarse immuunpuudulikkuse, süsteemse erütematoosse luupuse ja Crohni tõve puhul (Bashford-Rogers *et al.*, 2019; Ghraichy *et al.*, 2018; Roskin *et al.*, 2020; Rovin ja Erickson, 2020). Lisaks mõjutab BCR-i repertuaari vaktsineerimine (Khavrutskii *et al.*, 2017), seega on vaktsiinide välja töötamisel ning nende

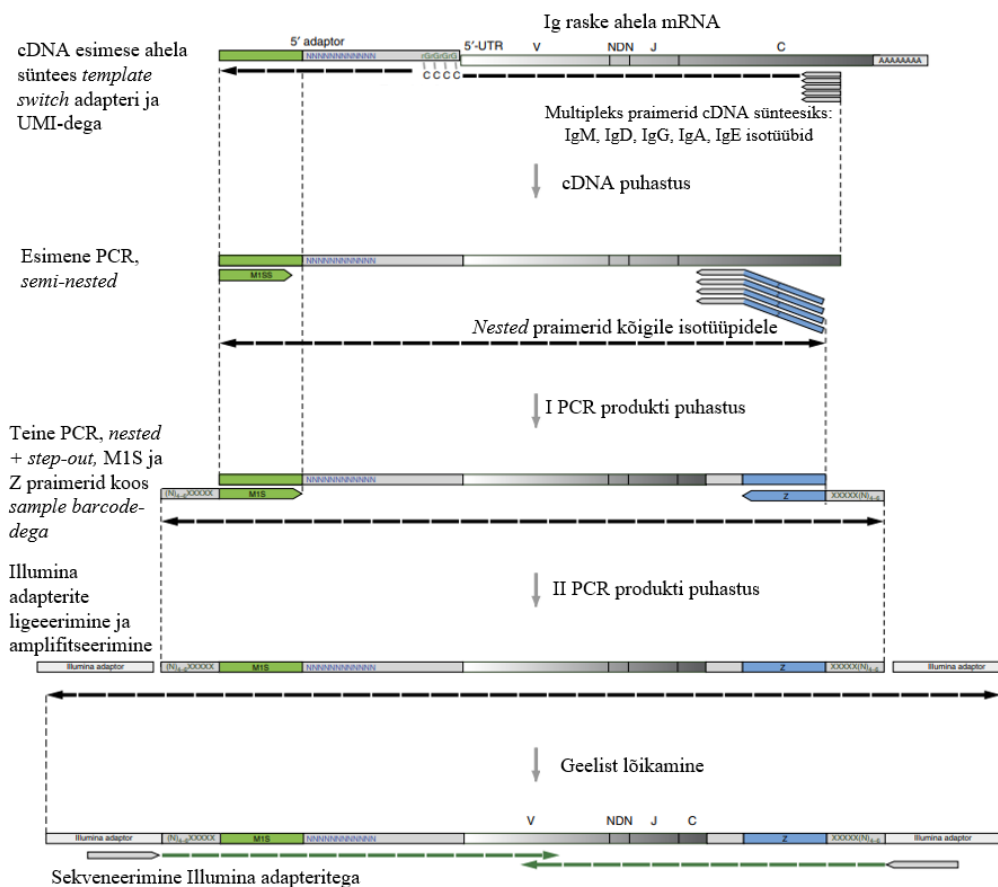
mõju uurimisel võimalik kasutada BCR-i repertuaaride analüüsimist. BCR-i repertuaari uurimine on oluline ka selleks, et leida uusi haiguste diagnostilisi markereid (Medina et al., 2020). Kokkuvõttes võimaldavad BCR-i repertuaari uuringud saada rohkem teadmisi haiguste kulu, nende tekkemehhanismide ning potentsiaalsete ravivõimaluste kohta.

1.5.1. BCR-i repertuaari uurimine Chudakovi meetodiga

Turchaninova *et al.* (2016) avaldasid artiklis meetodi (edaspidi Chudakovi protokoll), kuidas täispikkuses nii inimese kui hiire immuunglobuliinide repertuaari kirjeldada. Meetod põhineb 5' RACE (*Rapid Amplification of cDNA Ends*) lähenemisel, mis kasutab antikehade kirjeldamisel ja nende repertuaari analüüsil unikaalseid molekulaarseid identifitseerijaid (*unique molecular identifiers*, edaspidi UMI-d), mis lisatakse juba cDNA (komplementaarne DNA, *complementary DNA*) sünteesi käigus. Chudakovi protokoll eelisteks, mis autorid artiklis (2016) välja tõid, on esiteks see, et tänu sekveneerimislugemite UMI-de alla grupeerimisele on võimalik iga algse cDNA molekuli täpne konsensusjärjestus kindlaks teha. Meetodi eeliseks on ka PCR (polümeraasi ahelreaktsioon) amplifikatsioonide käigus kuhjunud ning ka sekveneerimisel tekkinud vigade eemaldamine edasise analüüsi käigus. See on võimalik, kuna UMI-d lisatakse proovidele cDNA sünteesi käigus, kui amplifikatsiooni pole veel toimunud. See vähendab PCR-ist tingitud vigade hulka lõppandmestikus ja seega võimaldab saada uuritavast proovist tõelähedasema ülevaate. Chudakovi protokoll käigus amplifitseeritud järjestus katab BCR-i kogu variaabel eksoni koos isotüübi konstantse alaga.

Chudakovi protokoll käigus sünteesitakse kõigepealt RNA-st cDNA multipleks praimeritega kõikide Ig isotüüpide jaoks ning lisatakse 5' otsa ka UMI-d (Joonis 5 NNNNNNNNNNNN), märgistades sellega iga cDNA molekuli. Esimese PCR käigus amplifitseeritakse regiooni 3' *semi-nested* praimeritega, mis lähenevad J segmendile, kuid katavad ka C segmendi paarikümne nt kauguselt. Siin lisatakse ka universaalne 3' adapter (Joonis 5 M1SS), mis võimaldab järgnevat PCR-i läbi viia ainult kahe praimeriga. Teine PCR viiakse läbi *sample-barcoded* praimeritega (Joonis 5 M1S ja Z), mis loob Ig raamatukogud, kus igal molekulil on nii 5' kui 3' otsas unikaalne vöötkood ((N)₄₋₆XXXXX). See väldib edasises analüüsis proovide vahelist saastumist. Viimases etapis lisatakse igale proovile sekveneerimise jaoks vajalikud Illumina adapterid ning raamatukogud amplifitseeritakse veelkord (Turchaninova *et al.*, 2016).

Meie laboris modifitseeritud Chudakovi protokoll kasutab kättesaadavaid reagente ja ensüüme, kui artiklis välja toodud. Samuti oleme modifitseerinud võrreldes artiklis tooduga indekseid lisamist, tuues selle jaoks sisse eraldi reaktsiooni, ehk teises PCR-is kasutatud M1S ja Z praimeritel pole eraldi vöötkoode, vaid iga proov märgistatakse Illumina indekseid.

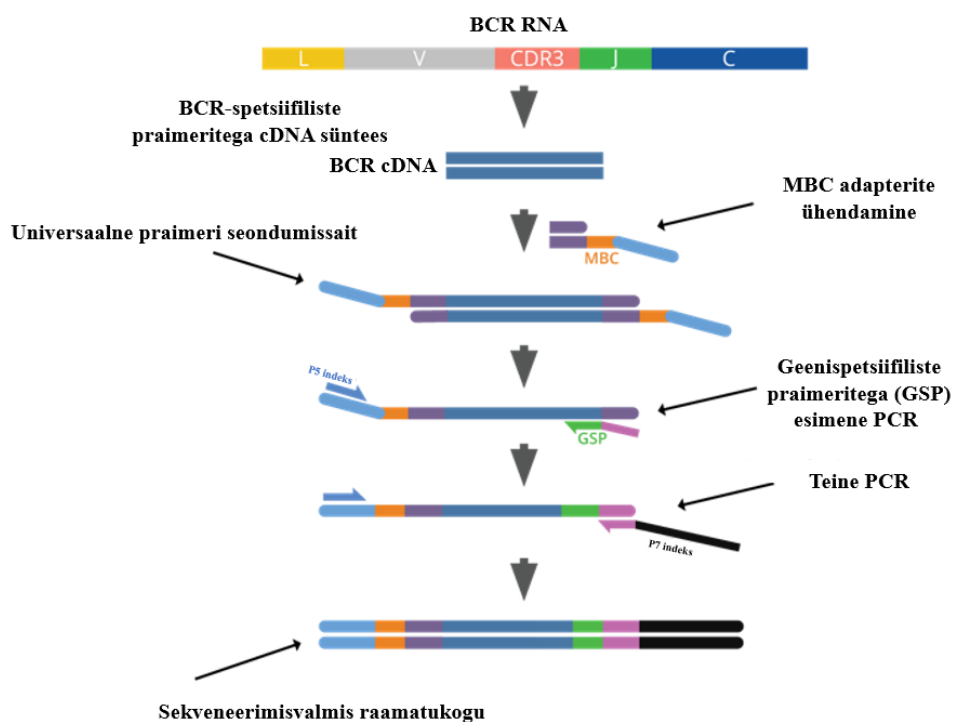


Joonis 5. Chudakovi protokoll järgi cDNA raamatukogu ettevalmistuse ja sekveneerimise skeem. Joonisel toodud järjestuste ülesehitus 5' suunast: UMI, 5' mittetransleeritav region (UTR), V segment, V ja J segmentide ühenduskohtades nukleotiididega ümbritsetud D segment (NDN), J segment, konstantne region C ja polü(A) saba (kohandatud Turchaninova *et al.*, 2016).

1.5.2. BCR-i repertuaari uurimine Archer Immunoverse kitiga

Archer Immunoverse firma (ArcherDX, 2018) on välja töötanud kommertsiaalse protokoll (Archer® Immunoverse™ - HS BCR IGH Protocol for Illumina®), millega amplifitseeritakse sekveneerimiseks ja analüüsiks üles BCR-de rasked ahelad. Protokollis sisendmaterjal on RNA, mis võib olla pärit nii PBMC-dest, FFPE (*formalin-fixed paraffin-embedded*) kui ka värskest koest. Reagendid on lüofiliseeritud kujul 8-katsutiga ribades, mis on iga reaktsiooni jaoks eraldi pakendatud. See teeb raamatukogu ettevalmistuse protsessi lihtsaks ning väldib kasutaja vigu ja saastamist. Töövoog kuni sekveneerimiseni nõuab kokku 8,5 h, millest <3,5 on praktiline aeg. Archeri kiti meetoodika põhineb Anchored Multiplex PCR-il (AMP) (Zheng *et al.*, 2014), mis kasutab sarnaselt Chudakovi protokolliga (punkt 1.5.1) molekulaarseid märgiseid (MBC, *molecular barcodes*), et analüüsida vigu eemaldada. Protokoll lubab ühe reaktsiooni saagiseks rohkem kui 100 000 klonotüüpi. (ArcherDX, 2019)

Protokolli käigus sünteesitakse RNA-st cDNA (Joonis 6), seejärel ühendatakse MBC adapterid koos praimerite seondumiskohtadega. Sellele järgneb geenispetsiifiliste praimeritega PCR ning Illumina P7 adapterite ligeerimine.



Joonis 6. Archer Immunoverse kiti järgi raamatukogu ettevalmistamise skeem (kohandatud tootja protokoll järgi).

1.5.3. Andmete töötlemise programmid

Nii T- kui B-rakkude uurimiseks on kirjutatud kümneid arvutiprogramme, mis suuresti automatiseerivad järjestuste analüüsi (Dmitriy A Bolotin et al., 2017). Antud töös kasutatakse MiNNN (Dmitry A. Bolotin et al., 2020) ja MiXCR (Dmitriy A. Bolotin et al., 2015; Smakaj et al., 2020) programme, et fastq toorfaile edasiseks analüüsiks ette valmistada. MiNNN kasutab sekveneerimisest saadud fastq toorfaile ning töötleb need ümber, kasutades molekulaarseid vöötkoode (UMI-d või MBC-d). Lõppväljundiks on fastq formaadis iga vöötkoodi konsensusjärjestused (Shugay *et al.*, 2019). Bolotin *et al.* välja töötatud MiXCR (2015) joondab MiNNN töötuse läbinud paarislugemite failid ning saadud nukleotiidsed järjestused põhjal paneb kokku BCR-i V, D ja J geenisegmentide järjestused vastavalt GenBank andmebaasis olevatele referentsjärjestustele. Selles etapis parandab MiXCR PCR-il tekkinud vigu, eemaldades vigased järjestused, mis muidu tekitaks repertuaari kunstlikku mitmekesisust. Lisaks paneb MiXCR kokku klonotüüpide järjestused kasutades eelmises sammus assambleeritud geenijärjestusi, et CDR3 järjestused saada. Viimasena ekspordib MiXCR joondused või kloonid vastavalt programmile antud parameetritele tekstifaili, mis on kergesti loetav ning edasiseks analüüsiks sisend R-i statistikatarkvarasse.

2. EKSPERIMENTAALNE OSA

2.1. Töö eesmärgid

Käesoleva bakalaureusetöö eesmärgid on:

- 1) võrrelda Archer Immunoverse kiti ja Chudakovi protokollide inimese B-raku retseptorite repertuaari uurimiseks;
- 2) analüüsida kahe terve indiviidi B-raku retseptorite repertuaari klonalsust, mitmekesisust, geenisegmentide kasutust ning ülekatet CDR3 järjestuste vahel.

2.2. Materjal ja meetodika

2.2.1. Vere võtmine

Vere võtmine teostati Tartu Ülikooli Kliinikumi Ühendlaboris teise töögrupi liikme poolt. Verd võeti kahes ajapunktis kuuekuulise vahega kahelt naissoost indiviidilt (töös tähistatud P ja T). Indiviid P oli vere võtmise ajal 20 a vana, indiviid T oli 35. Mõlemas ajapunktis võeti veeniverd 5 ml vaakumkatsutisse (K2E-EDTA, BD Vacutainer) nii vasakust kui paremast käest.

2.2.2. PBMC-de eraldamine

Perifeerse vere mononukleaarsete rakkude (PBMC, *peripheral blood mononuclear cells*) eraldamiseks kasutati Ficolli gradient-tsentrifugimist. Selleks valati veri 50 ml katsutitesse ning lisati 10 ml 1x DPBS-i (Corning), millega loputati ka verevõtu katsutid. Teistesse 50 ml katsutitesse pipeteeriti 15 ml Ficolli lahust (Ficoll-Paque PLUS, GE Healthcare). Vere ja PBS-i segu pipeteeriti fraktsioone segi ajamata Ficolli lahusele. Katsuteid tsentrifugiti toatemperatuuril 30 min 400 rcf (Eppendorf 5810R, rotor A-4-62) minimaalse piduri ja kiirendusega. Pärast tsentrifugimist eemaldati plasma, jättes alles u 1 cm. Pasteuri pipetiga tõsteti eraldunud mononuklearsed rakud uude 50 ml katsutisse.

Pesemiseks lisati PBMC-dele PBS-i 50 ml-ni, tsentrifugiti toatemperatuuril 10 min 300 rcf (kiirendus ja pidur peal) ja eemaldati supernatant. Rakusadet pesti 20 ml PBS-iga, 18 µl rakususpensiooni võeti rakkude lugemiseks. Ülejäänud rakususpensiooniga katsuteid tsentrifugiti 10 min 300 rcf toatemperatuuril ning eemaldati supernatant. Tsentrifugimise ajal loeti rakud rakulugejaga Luna-FL Dual Fluorescence Cell Counter (Logos Biosystems), lisades 18 µl rakususpensioonile 2 µl akridiinoranž-propiidiumjodiid värvi (katalooginumber F23001).

Rakud võeti üles 5 ml-s PBS-is ning suspendeeriti. PBMC-d tõsteti kahte 1.5 ml katsutisse 2.5×10^6 raku kaupa ning tsentrifugiti 10 min 300 rcf 4 °C juures (Centrifuge 5415R, Eppendorf). PBS eemaldati ja rakusade säilitati -150 °C juures.

2.2.3. RNA eraldamine

Külmutatud rakusade suspenderiti üles 1 ml-s Qiazolis (QIAzol Lysis Reagent, Qiagen), inkubeeriti toatemperatuuril 5 min ning lisati 0.2 ml kloroformi (ThermoFischer Chemical). Suspensiooni segati vortexil ning inkubeeriti toatemperatuuril 3 minutit. Proove tsentrifuugiti 15 min 12000 rcf 4 °C (Centrifuge 5415 R, Eppendorf). Seejärel pipeteeriti vesifaas uude 1.5 ml katsutisse. Edasi eraldati RNA kasutades RNeasy Mini Kit-i (Qiagen) vastavalt tootja protokollile. Pärast RW1 puhvriga pesemise sammu tehti proovidele vastavalt tootja protokollile ka DNaa I töötlus. RNA elueerimiseks kasutati 30 µl RNAasi vaba vett (Qiagen). RNA kontsentratsiooni mõõdeti NanoDropiga (ThermoFischer Scientific), nullides eelnevalt 1 µl RNAasi vaba vee (RNeasy Mini Kit, Qiagen) vastu ning kvaliteeti mõõdeti Tartu Ülikooli Kliinilise Genoomika Tuumiklaboris Agilent 2100 Bioanalyzeriga. Edasistes katsetes kasutati RNA-d RIN-ga (*RNA integrity number*) vahemikus 5.3–10 (keskmine 7.54 ja mediaan 8.2). RNA säilitati -80 °C juures. Edasisteks Chudakovi ja Archeri meetoditeks tehti iga RNA proov pooleks, jättes kummagi meetodi jaoks piisavas koguses RNA-d.

2.2.4. Chudakovi protokoll

Chudakovi protokoll koosneb neljast etapist, mille vahel on produkti puhastamise etapid. Etapid on järgmised: cDNA süntees, I PCR, II PCR ning raamatukogu ettevalmistus. Antud töös kasutati praimereid ainult raskete ahelate amplifikatsiooniks. Turchaninova *et al.*, 2016 avaldatud protokollis lisatakse proovide eristamiseks vajalikud indeksid teise PCR-i käigus koos praimeritega, kuid protokollis modifitseeriti, lisades need eraldi etapina raamatukogu ettevalmistuse etapis (punktis 2.2.4.5).

2.2.4.1. cDNA süntees

cDNA sünteesi reaktsioon viidi läbi 20,5 µl mahus ning koosnes kahest erinevast etapist. Esimese etapi reaktsioon viidi läbi 8 µl mahus ning koosnes tabelis 2 toodud komponentidest. Antud reaktsioonis kasutati hIGG_r1, hIGM_r1, hIGA_r1, hIGD_r1 ja hIGE_r1 praimeritest koosnevat segu. Praimerite järjestused on toodud lisas 1.

Tabel 2. cDNA sünteesi esimese etapi komponendid

Komponent	Kogus (µl)	Lõppkogus/kontsentratsioon
Punktis 2.2.3 eraldatud RNA	5–6	81–415 ng
cDNA sünteesi praimerite segu (10µM igat)	2	1 µM iga praimeri jaoks
Nukleaasivaba vesi (Ambion, Thermo Scientific)	0–1	–

Proove inkubeeriti termotsükleris (Mastercycler, Eppendorf) 2 min 70 °C juures ja 3 min 42 °C juures. Seejärel lisati 12,5 µl tabelis toodud reaktsioonisegu, suspendeeriti õrnalt pipetiga ning inkubeeriti 60 min 42 °C juures.

Tabel 3. cDNA sünteesi teise etapi komponendid

Komponent	Kogus (µl)	Lõppkogus/kontsentratsioon
<i>First-strand</i> puhver (5x, Takara)	4	1x
DTT (20mM, Takara)	2	2 mM
<i>5'-template switch</i> adapter (10µM) SMARTNNNa)	2	1 µM
dNTP segu (10 mM igat)	2	1 mM igat
SMARTScribe pöördtranskriptaas (10×, TaKaRa)	2	10 U/µl
Ribolock (40 U/µl, Thermo Scientific)	0,5	1,6 U/µl

Pärast inkubatsiooni lisati proovidele 1 µl uratsiil DNA glükosülaasi (5 U/µl, New England Biolabs) ning inkubeeriti uuesti 40 min 37 °C juures.

2.2.4.2. *Produkti puhastamine*

Proovide puhastamiseks kasutati Zymo DNA Clean and Concentrator kiti (Zymo Research) vastavalt tootja protokollile. Produkt elueeriti 15 µl DNA Elution puhvril (Zymo Research). Proove säilitati -20 °C juures või üleöö 4 °C juures.

2.2.4.3. *Esimene PCR reaktsioon*

Esimene PCR amplifikatsiooni reaktsioon viidi läbi 50 µl mahus ning koosnes tabelis 4 toodud komponentidest ning viidi läbi parameetritel, mis on toodud tabelis 5. Praimerite segu hIGX_r2 koosneb praimeritest hIGGE_r2, hIGM_r2, hIGA_r2 ja hIGD_r2. Praimerite järjestused on toodud lisas 1.

Tabel 4. I PCR-il kasutatud reagensid ning nende lõppkontsentratsioonid

Komponent	Kogus (µl)	Lõppkogus/kontsentratsioon
cDNA	2,5	–
Phusion HF polümeraasi puhver (5x, Thermo Scientific)	10	1x
dNTP segu (igat 10 mM)	1	0,2 mM igat
Rep_M1SS praimer (10 µM)	1	0,2 µM
hIGX_r2 praimerite segu (10µM)	1	0,2 µM igat praimerit

Phusion Hot Start II DNA polümeraas (Thermo Scientific) (2 U/μL)	0,5	0,02 U/μl
Nukleasivaba vesi (Ambion, Thermo Scientific)	34	—

Tabel 5. PCR programm ning parameetrid

Tsüklite arv	Denaturatsioon	Praimerite seondumine	Produkti pikenemine
1	95 °C 1 min 30 s	—	—
18	95 °C 10 s	60 °C 20 s	72 °C 40 s
1	—	—	72 °C 4 min

Proovide puhastamiseks kasutati Zymo DNA Clean and Concentrator kiti (Zymo Research) vastavalt alapunktis 2.2.4.2. toodud juhistele. DNA elueerimiseks kasutati 30 μl DNA Elution puhvrit.

2.2.4.4. Teine PCR reaktsioon

Teise PCR amplifikatsiooni segu valmistati tabelis 6 toodud komponentidest ning programm teostati vastavalt tabelis 7 toodud parameetritele. Praimerite järjestused on toodud lisa 1.

Tabel 6. II PCR-il kasutatud reagentid ja nende lõppkontsentratsioonid

Komponent	Kogus (μl)	Lõppkontsentratsioon
Puhastatud esimese PCR-i produkt	2	—
Phusion HF polümeraasi puhver (5×, ThermoScientific)	10	1×
dNTP segu (10 mM igat)	1	0,2 mM igat
FwdOH_M1S praimer (10 μM)	1	0,2 μM
RevOH_Z2 praimer (10 μM)	1	0,2 μM
Phusion Hot Start II DNA polümeraas (ThermoScientific) (2 U/μL)	0,5	0,02 U/μl
Nukleasivaba vesi (Ambion, Thermo Scientific)	34,5	—

Tabel 7. PCR programm ning parameetrid

Tsükkel	Denatureerimine	Ahelate seondumine	Ahela pikenemine
1	95 °C 1 min 30 s	—	—
12	95 °C 10 s	60 °C 20 s	72 °C 40 s
1	—	—	72 °C 4 min

Saadud produkt puhastati Zymo DNA Clean and Concentrator kitiga vastavalt punktis 2.2.4.2 toodud juhistele. DNA elueeriti 10 μl elueerimispuhvrts.

2.2.4.5. Raamatukogude ettevalmistus

PCR produktide kontsentratsioone mõõdeti Qubit 2.0 fluoromeetriga kasutades Qubit dsDNA HS Assay Kiti (Thermo Scientific) vastavalt tootja protokollile. Mõõtmistulemuste põhjal tehti proovidest 10 ng/μl lahjendused. Sekvenerimiseks vajalike adapterite lisamiseks segati 15 μl mahus reaktsioon vastavalt tabelis 8 toodud juhistele. PCR-i programm teostati vastavalt tabelis 9 toodud parameetritele. Sekvenerimiseks vajalike indeksite järjestused on toodud lisas 2.

Tabel 8. Raamatukogude märgistamise reaktsiooni komponendid ja lõppkontsentratsioonid

Komponent	Kogus (μl)
DNA (10 ng/ul)	2,5
Nextera XT indeks 1 praimer (i7) (Illumina)	2,5
Nextera XT indeks 2 praimer (i5) (Illumina)	2,5
2x KAPA HiFi HotStart PCR segu (Kapa Biosystems)	7,5

Tabel 9. PCR reaktsiooni programm ja parameetrid

Tsükkel	Denaturatsioon	Praimerite seondumine	Ahela pikenemine
1	95 °C 3 min	—	—
7	95 °C 30 s	55 °C 30 s	72 °C 30 s
1	—	—	72 °C 5 min

2.2.4.6. Produkti puhastamine Ampure keradega

PCR produkt puhastati Ampure AmpXP keradega (Beckman Coulter). Selleks lisati igale proovile 27 μl AMPure XP suspensioonis olevaid kerasid (1,8 μl kerasid 1 μl proovi kohta). Proove ja kerasid segati 10 korda pipetiga ning inkubeeriti 5 minutit toatemperatuuril. Proovid asetati 2 minutiks magnetplaadile, et kerad lahusest eraldada. Magnetil olevast segust eemaldati selginenud suspensioon. Keradele lisati 200 μl 70% etanooli ning inkubeeriti toatemperatuuril 30 sekundit, etanool eemaldati ning pesu korrati. Kui kerad olid kuivad, eemaldati proovid magnetilt ning elueeriti 27,5 μl Tris-elueerimispuhvrts (pH 8,0, 10 mM, ArcherDX) ning segati pipetiga 10 korda. Seejärel proove inkubeeriti toatemperatuuril 2 minutit. Proovid asetati 1 minutiks magnetile ning 25 μl eluaati viidi uutesse 0,2 ml katsutitesse. Proove säilitati -20 °C juures või üleöö 4 °C juures.

2.2.5. Archer Immunoverse kit

Archer Immunoverse HS BCR Protocol for Illumina kitiga (ArcherDX) tehtud reaktsioonid viidi läbi vastavalt tootja protokollile, sisendiks oli RNA vahemikus 202–361 ng. Protokoll koosneb üheksast etapist, mille hulgas on neli produkti puhastuse etappi (Joonis 6 peatükis

1.5.2). Kõigepealt toimus BCR spetsiifiline praimimine, seejärel esimese cDNA ahela süntees, siis teise ahela süntees. Sellele järgnes lõppude parandus (*end repair*). Pärast seda toimus esimene ühenduse (*ligation*) etapp, MBC adapterite (Illumina P5 indeks, Lisa 2 indeks 2) lisamine ning teine ühendamise etapp. Nendele etappidele järgnesid kaks PCR reaktsiooni, millest viimasega lisati Illumina P7/i7 indeksid (lisa 2. sekveneerimise indeksid indeks 1). P7 indeksid on igas Archeri kitis samad, kuid MBC adapterid tuleb iga kiti juurde eraldi tellida.

2.2.6. Sekveneerimine (NGS)

Proovide sekveneerimine teostati Tartu Ülikooli genoomika instituudi tuumiklaboris. Proovide kvaliteedi kontrollimiseks kasutati TapeStation D1000 ScreenTape (TapeStation Analysis Software, Agilent Technologies) süsteemi ja Qubit 2.0 fluoromeetrit dsDNA HS Assay Kit-i (Thermo Scientific) vastavalt tootjate protokollidele. Seejärel kasutati raamatukogude kvantifitseerimiseks KAPA Library Quantification Kit-i (KAPA Biosystems) vastavalt tootja protokollile. Sekveneerimiseks kasutati MiSeq v3 sekvenaatorit (Illumina) ning sekveneeriti 2x300 bp formaadis.

2.2.7. Bioinformaatiline analüüs

2.2.7.1. Kvaliteedikontroll

Sekveneerimistulemuste kvaliteedi kontrollimiseks kasutati programmi FastQC (versioon 0.11.2; Andrews, 2014) ja tulemused sisestati paremaks visualiseerimiseks MultiQC programmi (Ewels *et al.*, 2016).

2.2.7.2. Sekveneerimise toorandmete töötlemine ja andmete analüüs

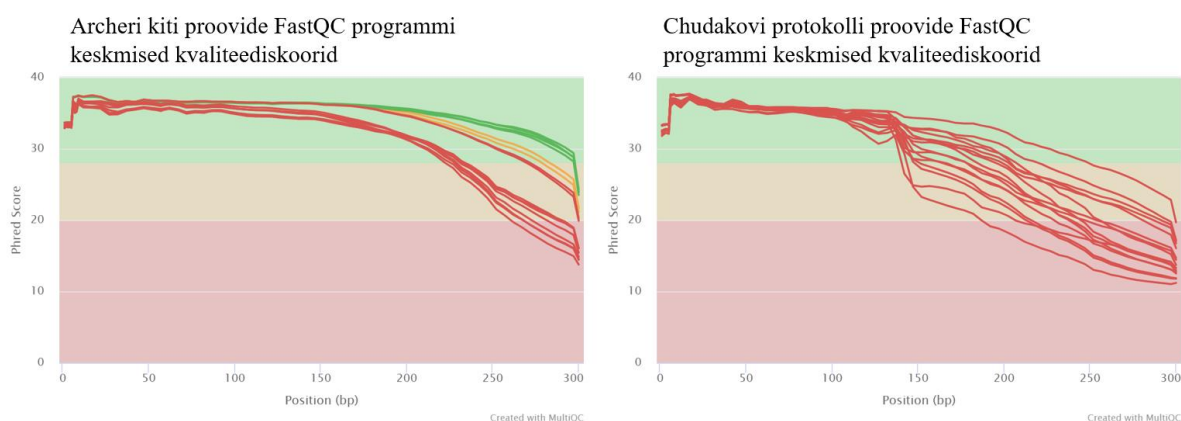
Sekveneerimise toorandmete töötlemiseks kasutati tööriistu MiNNN (versioon 10.1; Bolotin *et al.*, 2020; Shugay *et al.*, 2019) ja MiXCR (versioon 3.0.13; Bolotin *et al.*, 2015). Programme ja nende ülesandeid on pikemalt kirjutatud punktis 1.5.3. Andmeanalüüsiks kasutati R programmeerimiskeelt (versioon 4.0.3; R Core Team, 2020) tarkvaras RStudio (versioon 1.3.1093; RStudio Team, 2020). Analüüs viidi läbi R-i paketiga Immunarch (versioon 0.6.5; ImmunoMind Team, 2019). Tulemuste visualiseerimiseks kasutati R-i pakette ggplot2 (versioon 3.3.3; Wickham, 2016), VennDiagram (versioon 1.6.20; Chen ja Boutros, 2011), ggpubr (versioon 0.4.0; Kassambara, 2020), pheatmap (versioon 1.0.12; Kolde, 2019), cowplot (versioon 1.1.1; Wilke, 2020), dplyr (versioon 1.0.6; Wickham *et al.*, 2021) ning ggplotify (versioon 0.0.5; Guangchuang, 2020).

2.3. Tulemused

Joonistel olevate proovide tähistused on toodud lisas 3. Proovid on tähistatud meetodite kaupa, millega proov valmistati, vastavalt kas Archer või Chudakov ning indiviidi järgi, milleks on kas individ P või T. Lisaks tähistati proovid bioloogiliste replikaatide järgi, mis näitab kumma käe verest proov on valmistatud ja viimaks ajapunktide I ja II järgi, mis olid kuuekuulise vahega. Näiteks „C-TVII“ on proov, mis on tehtud Chudakovi meetodil ning võetud individ T vasakust käest II ajapunktil.

2.3.1. Sekveneerimise kvaliteet

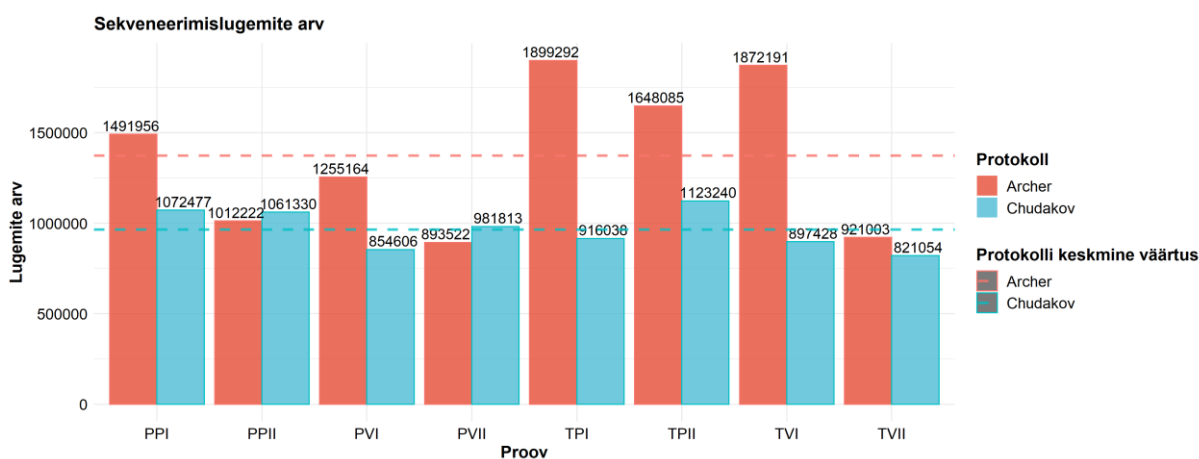
BCR-i repertuaari uurimisel on väga oluline sekveneerimise kvaliteet. Kuna antud töö eesmärk on võrrelda kahte erinevat meetodit, on ka sekveneerimise kvaliteet võrdluse oluline aspekt, kuna sellel on otsene mõju töö tulemustele. Lugemite kvaliteedi mõõtmiseks kasutati programmi FastQC ning kahe protokolliga proovide tulemused visualiseeriti MultiQC programmiga. Joonisel 7 on toodud vastavalt Archeri ja Chudakovi meetodi iga proovi vastava lugemi positsiooni keskmine kvaliteediskoor. Graafikutelt on näha, et Chudakovi protokolliga tehtud proovide keskmised kvaliteediskoorid on väga madalad ning langus üle kriitilise piiri (graafikul kvaliteediskoor alla 20) toimub juba umbes 150 bp juures. Samas Archeri kitiga tehtud proovide kvaliteet langeb alla kriitilise piiri alles lugemite lõpus u 250 bp juures, mis on Illumina sekveneerimise puhul tavaline. Chudakovi proovide puhul oli märgata ka duplikaatlugemite vähesus võrreldes Archeri proovidega, st enamikes lugemites oli esimeses lugemis palju nukleotiidide vigu, mistõttu ei kattunud kaks lugemit omavahel. Seetõttu vähenes ka duplikaat lugemite arv (lisas 4).



Joonis 7. FastQC/MultiQC tulemused Archeri ja Chudakovi protokolliga. Graafikul on toodud kvaliteediskooride keskmine väärtus iga vastava nukleotiidipositsiooni kohta. X-teljel on toodud lugemi bp positsioon 1-300 ning y-teljel *phred* kvaliteediskoor. Iga joon graafikul vastab ühele proovile.

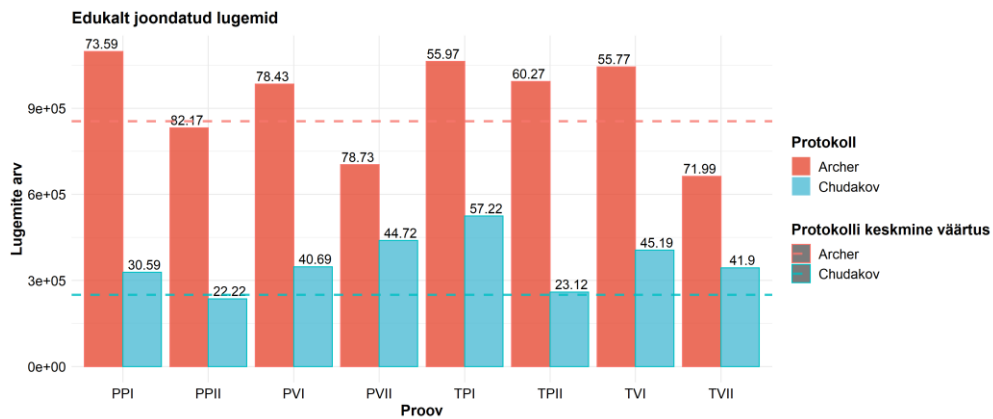
2.3.2. MiXCR tulemused

MiXCR programmi väljundfail sisaldab mitmesuguseid andmeid repertuaari kohta, millest kõige olulisem info on joondamise edukuse ning lõpliku klonotüüpide arvu kohta. Lõplikud klonotüübid on kõik klonotüübid, mis jäävad pärast kõiki vaeemalduse etappe alles. MiXCR-i tulemused sõltuvad suuresti sekveneerimise kvaliteedist, kuna programm eemaldab kõik järjestused, mille kvaliteet on halb. Kõigepealt leiti esimeses analüüsi etapis järjestused, kus on nukleotiide, mille kvaliteediskoor on alla 20. Kui edasises analüüsis ilmneb, et järjestuses on 70% või rohkem alla 20 kvaliteediskooriga nukleotiide, eemaldatakse see järjestus analüüsist. Joonisel 8 on toodud sekveneerimislugemite arv koos kummagi protokolliga keskmise väärtusega. Keskmiselt analüüsiti 1 374 179 Archeri lugemite ning 965 998 Chudakovi lugemite (protokollide keskmiste väärtuste standardhälve 288 267). MiXCR-i töötuse läbinud sekveneerimislugemite arv on Chudakovi proovidel madalam halvema sekveneerimise kvaliteedi tõttu. Kui Archeri proovidel oli individil T keskmiselt kaks korda rohkem lugemiteid kui individil P, siis Chudakovi proovidel individide lugemite arvul olulist vahet polnud.



Joonis 8. Sekveneerimislugemite arv koos protokollide keskmistega. X-teljel on toodud analüüsitud proovi tähistus ning y-teljel lugemite arv, mida MiXCR programm analüüsib. Graafik on värvikoodi järgi grupeeritud protokollide kaupa ning joontega on toodud ka mõlema protokolliga keskmise väärtuse joon.

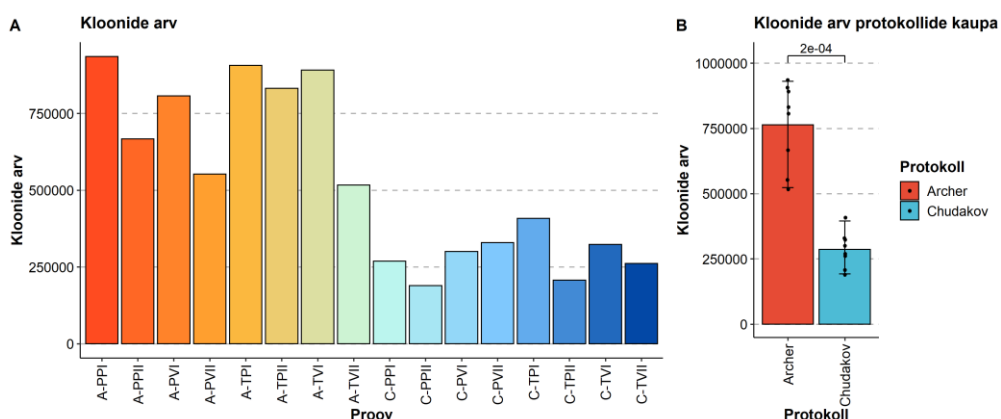
Joonisel 9 on toodud MiXCR-i väljundfaili edukalt joondatud järjestuste graafik koos mõlema meetodi keskmiste väärtustega. Joonisel on toodud, mitu protsenti kõikidest paaris-lugemitest MiXCR suutis Ig referentsjärjestustele (V-, D- ja J-geenidele) edukalt joondada. Jooniselt on näha, et Archeri proovide puhul on edukalt joondatud lugemite arv keskmiselt 1,8 korda suurem kui Chudakovi proovidel (Archeril keskmiselt 922 592 ja Chudakovil 360 499 lugemiteid). Archeri proovidel oli edukalt joondatud järjestusi kõikidest järjestustest keskmiselt 69% ja Chudakovi proovidel keskmiselt 38% järjestustest.



Joonis 9. Edukalt joondatud lugemite arv protsentides. X-teljel on toodud 8 proovi tähistused, mis on grupeeritud vastavalt protokollile ning y-teljel on toodud MiXCR programmi edukalt joondatud lugemite arv. Graafik on värvikoodi järgi grupeeritud protokollide kaupa ning joontega on toodud ka mõlema protokollide keskmise väärtuse joon.

2.3.3. Repertuaari klonaalsus

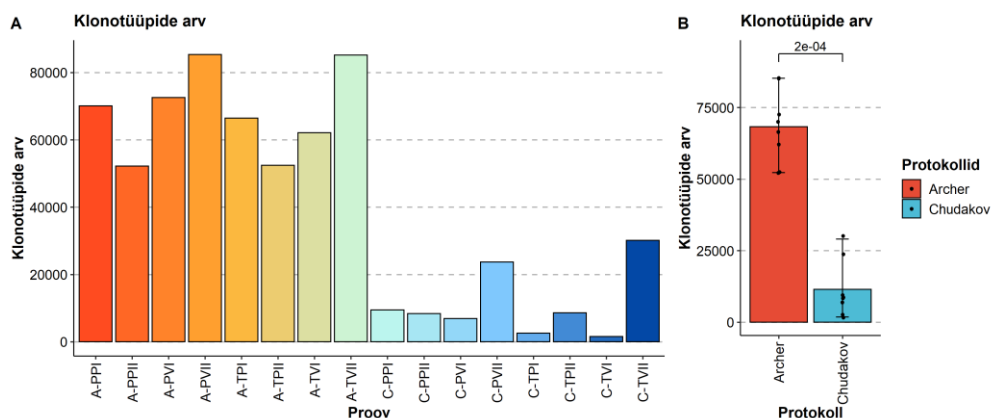
Järgnevalt uuriti BCR-i repertuaari kloonide arvu ning klonaalsust R-i Immunarch paketiga. Joonistel 10 A ja B on toodud kloonide arv pärast Immunarch paketiga töötlemist vastavalt iga proovi kaupa ja ka meetodite kaupa eraldi (joonis B). Klooniks loetakse BCR-e, millel on sama CDR3 järjestus. Jooniselt on näha, et Archeri meetodi abil tuvastati keskmiselt 2,5 korda rohkem kloonide kui Chudakovi meetodi abil (Archeri meetodiga keskmiselt 763 362, Chudakoviga keskmiselt 285 952 kloonid, p-väärtus 0,0002). Lisaks oli kloonide arv peaaegu kõikides proovides teises ajapunktis väiksem kui esimeses, kuigi RNA kvaliteet (RIN) oli esimeses ajapunktis keskmiselt 8,83 ning teises 7,57, mis ei ole väga suur erinevus. Chudakovi proovide madalam kloonide arv tuleneb tõenäoliselt ka madalamast sekveneerimise kvaliteedist, mis omakorda väljendub madalamas joondamise efektiivsuses.



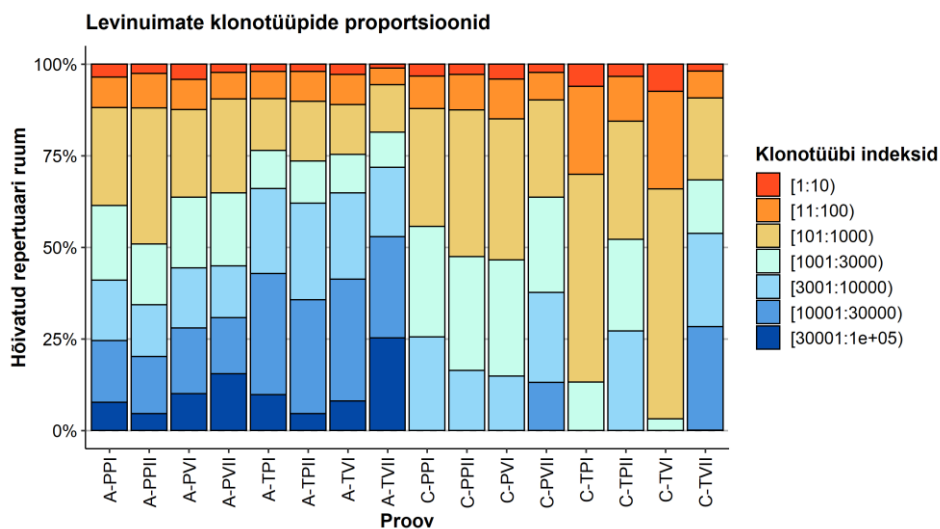
Joonis 10. A: Tuvastatud kloonide arv. B: Tuvastatud kloonide arv protokollide kaupa. X-teljel on toodud proovi tähistus ning y-teljel tuvastatud kloonide arv. P-väärtus on arvatud Wilcoxon testiga.

Joonisel 11A on näha, et Archeri meetodi abil tuvastati rohkem klonotüpe kui Chudakovi meetodi abil. Klonotüübiks loetakse kõiki kloonid, millel on sama järjestusega CDR3.

Kõikumised meetodite vahel on suured, kuna kahe meetodi keskmiste kloonide arvude standardhälve on 11 470 ja p-väärtus 0,0002. Archeri klonotüüpe oli keskmiselt 68 293 ning Chudakovi meetodiga tuvastati keskmiselt 11 415 klonotüüpi. Kõige väiksem klonotüüpide arv oli proovil C-TVI, milles oli 1568 klonotüüpi, kõige rohkem oli klonotüüpe A-PVII proovis (85 302). Klonotüüpide arv protokollide vahel on summeeritult toodud ka joonisel 11B, kus on näha 6-kordne vahe Archeri ja Chudakovi meetoditel.



Joonis 11. A: on toodud klonotüüpide arv iga proovi kohta. **B:** klonotüüpide keskmine arv kahe protokollide võrdluses. X-telgedel on tähistatud proov ning protokoll, y-teljel klonotüüpide arv. P-väärtus on arvutatud Wilcoxon testiga.

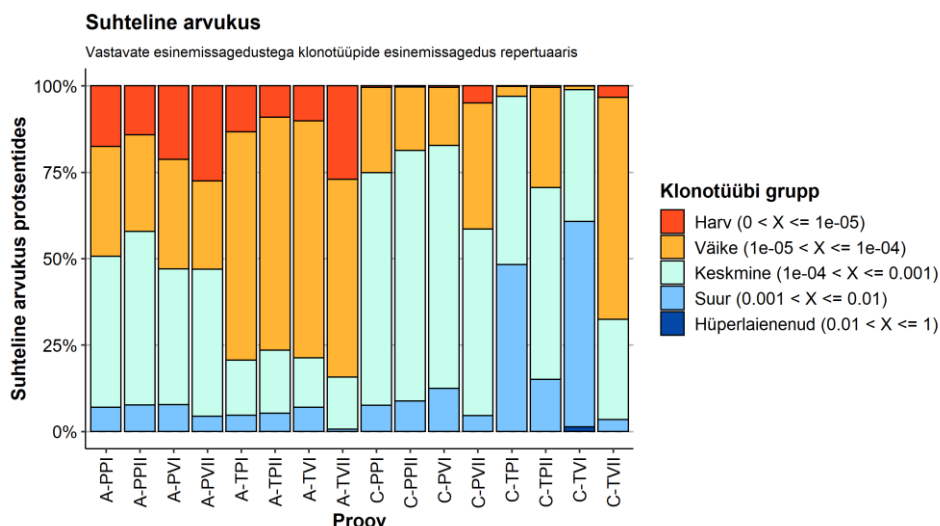


Joonis 12. Levinuimate klonotüüpide proportsioonid. X-teljel on toodud proovi tähistus ja y-teljel kogu repertuaarist hõivatud osa protsentides. Legendis toodud klonotüüpide indeksid näitavad vastavalt klonotüübi levikule (indeks 1 on antud repertuaaris suurima esinemissagedusega) ning indeksid vähenevad vastavalt klonotüübi leviku vähenemisele.

Järgnevas on joonisel 12 ära toodud enim esindatud klonotüüpide proportsioonid. Joonisel on kujutatud, kui palju repertuaarist võtavad enda alla vastavate esinemissagedustega klonotüübid. Antud funktsiooni sisendiks on sorteeritud klonotüüpide tabel. See on sorteeritud iga klonotüübi arvukuse järgi nii, et suurima arvukusega klonotüübid on tabeli eesotsas. Joonisel

toodud indeksid vastavad tabeli ridadele, näiteks indeks [1:10] jaotuses on kümme kõige levinumat klonotüüpi. Jooniselt on näha, et Archeri meetodiga tuvastati rohkem harva esinevaid klonotüüpe (alates indeksitest 10 001), mida esines Archeri proovide repertuaarides proportsionaalselt rohkem, samas Chudakovi proovidel oli repertuaaris suhteliselt suurem osakaal klonotüüpidel, mille indeksid jäid vahemikku 101–30 000. Seda ka sellepärast, et Chudakovi proovide maksimaalne klonotüüpide arv oli võrreldes Archeri proovidega oluliselt madalam. Seepärast ongi joonisel kahe meetodi vahel ebavõrdne jaotus, mis tuleneb kõigepealt kloonide arvust, mis sõltub veel omakorda sekveneerimise kvaliteedist.

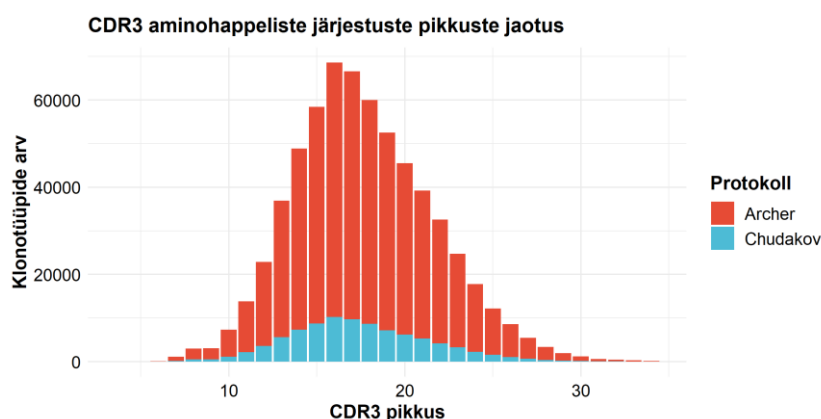
Joonisel 13 on toodud klonotüüpide suhteline arvukus (ka klonaalse ruumi homöostaas) ehk kui suure osa vastavas grupis (joonisel „harv“, „väike“, „keskmise“, suur“ ning „hüperlainenud“) olevad klonotüübid repertuaarist moodustavad. Gruppidesse jaotumine on tehtud klonotüübi esinemissageduse alusel ehk kui näiteks võtta „väikesest“ grupist üks suvaline klonotüüp, esineb ta repertuaaris tõenäosusega 0,0001 kuni 0,001. Mõlema meetodi puhul domineerivad väikse ja keskmise esinemissagedusega klonotüübid. Archeri proovide puhul on harvade klonotüüpide osakaal enamikel proovidel peaaegu 25%, samas Chudakovi proovidel harvad klonotüübid peaaegu puuduvad. See näitab, et kuigi Archeri proovide repertuaarid olid kloonide hulgalt suuremad, on meetod paremini võimeline harva esinevaid klonotüüpide järjestusi tuvastama. Chudakovi proovidel esineb ka proportsionaalselt rohkem suuri kloone ning ühel proovil (C-TVI) ka hüperlainenud klonotüüpe, mida Archeri meetodi proovidel ei esinenud peaaegu üldse. See on ilmselt sellepärast, kuna Chudakovi proovide repertuaarid olid võrreldes Archeri omadega väiksemad ning on suurem võimalus leida suure levimisega klonotüüpe.



Joonis 13. Klonotüüpide suhteline arvukus. X-teljel on tähistatud proovi nimi ning y-teljel klonotüübi gruppide suhteline arvukus protsentides. Legendis toodud klonotüübi grupp näitab vastavat vahemikku, millise esinemistõenäosusega seal grupis olev klonotüüp repertuaaris esineb.

2.3.4. CDR3 pikkused

Järgnevalt uuriti CDR3 aminohappeliste järjestuste pikkuseid Archeri ja Chudakovi meetodi kasutamisel, graafik on toodud joonisel 14. Mõlema meetodi puhul oli CDR3 järjestus keskmiselt 20 aminohapet pikk ning kõik järjestuste pikkused moodustavad histogrammil normaaljaotuse kurvi, ka esimese ja kolmanda kvantiili väärtused olid mõlema protokolliga puhul sama pikad või maksimaalselt ühe aminohappe erinevusega. Antud jooniselt tuleb väga selgelt välja ka Archeri kitiga tehtud proovide oluliselt suurem repertuaar.



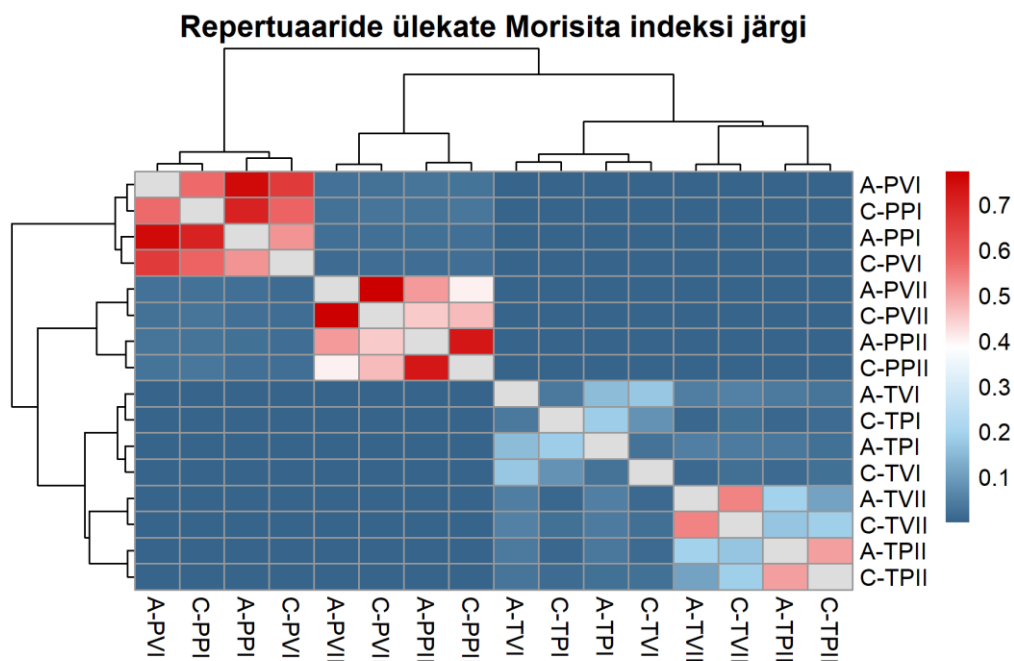
Joonis 14. CDR3 aminohappeline järjestuse pikkus. X-teljel on toodud CDR3 ala järjestuse pikkus aminohapetes ning y-teljel klonotüüpide arv, milles vastava pikkusega järjestus esines. Värvikoodiga on eristatud kaks meetodit.

2.3.5. Repertuaaride CDR3 järjestuste ülekate

Seejärel uuriti proovide CDR3 aminohappeliste järjestuste ülekatet Morisita statistilise indeksi abil. Morisita indeks mõõdab valimi kahe isendi ülekatet eeldusel, et valimi suurenedes suureneb ka mitmekesisus. Kuna varasemad tervete inimeste BCR-i repertuaaride uuringud on näidanud, et CDR3 järjestuste ülekate on indiviidide vahel madal (Soto et al., 2019), eeldati seda ka antud töös. Samas eeldati, et sama proovi erinevate raamatukogu ettevalmistusmeetodiga analüüsides on ülekatte suurem, kuna proovid läbisid sama informaatilise töötuse.

Morisita indeksi põhjal tehtud soojakaardilt (Joonis 15) on näha, et omavahel klasterduvad indiviidid P ja T ajapunktide kaupa nelja suuremasse gruppi, ehk kogu 16 proovi repertuaaris on kõige suuremad ülekatted sama indiviidi sama ajapunkti proovide vahel. Arvesse tuleb võtta ka seda, et proove ei *downsampletud* ning seetõttu on Archeri proovide omavaheline ülekatte suurem kui Chudakovi proovide vahel, kuna Archeri proovidel oli rohkem kloone. Ka on suured sarnasused erineva protokolliga tehtud sama proovi vahel (A-PPI ja C-PPI; A-PVII ja C-PVII) ning indiviidi P sama proovi replikaatide ehk paremast ja vasakust käest võetud proovide vahel (A-PVI ja A-PPI), ka erinevate protokollide replikaatide vahel (A-PVI ja C-PPI, A-PPI ja PVI; A-PPII ja C-TVII). Indiviidi T puhul olid ülekatted väiksemad kui indiviid P puhul, kuid tulid ka välja ülekatted sama proovi erinevate protokollide vahel. T indiviidi puhul olid ülekatted

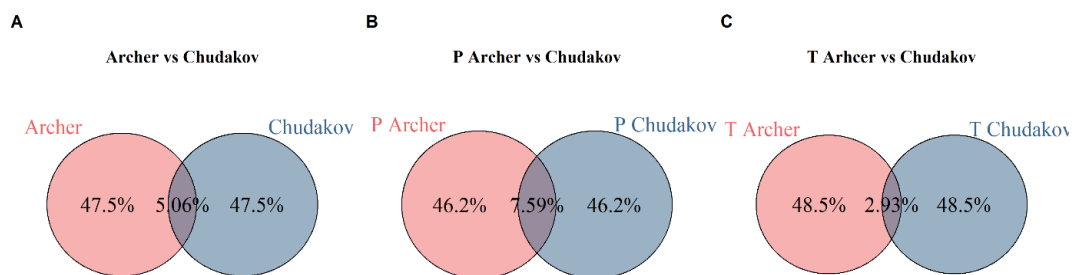
ühtlaselt väiksemad erinevate replikaatide vahel (helesinise värviga proovid A-TVI ja A-TPI), ning erinevate protokollide replikaatide vahel (A-TPII ja C-TVII), kuid ülekatted ei ole eriti suured võrreldes indiviidiga P. Indiviid T puhul olid ülekatted teises ajapunktis suuremad kui esimeses ajapunktis, indiviid P puhul just esimeses ajapunktis. Väiksem ülekate indiviidi kahe ajapunkti vahel oli märgata mõlema protokolliga puhul, olles veidi suurem indiviid T puhul (proovid A-TVI ja A-TVII).



Joonis 15. Repertuaaride CDR3 aminohappeliste järjestuste ülekate Morisita indeksiga. Värvikoodiga on tähistatud Morisita indeksi väärtus, kõrgemad väärtused on toodud punase ja madalamad väärtused sinise värviga.

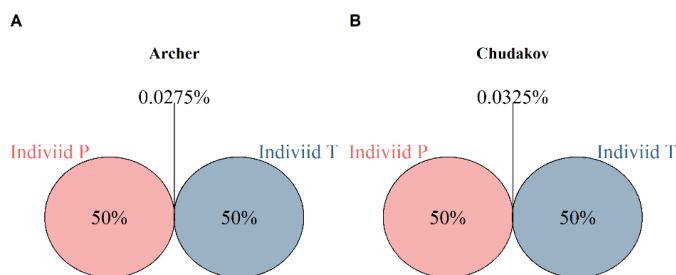
CDR3 aminohappe järjestuste ülekate visualiseeriti ka Venni diagrammidega. Selleks võeti igast proovist 10 000 kloni suurune juhuvalim, et võrdluses olevaid proove arvestada võrdselt ja kõikide võrdluste puhul võeti arvesse ka kloonide arvu ehk levikut. Esmalt vaadati kloone, mis olid olemas nii Archeri ja Chudakovi proovides (Joonis 16A). Ülekate kahe protokolliga kõikide kloonide vahel oli 5,06%. See protsent on väga väike, arvestades, et kõikide proovide peale oli kloone kokku 8 394 519, sellest Archeri omi 6 106 898 ning Chudakovi 2 287 621.

Järgnevalt vaadati Venni diagrammide abil indiviidide kaupa protokollide vahelist ühisosa, et teada saada, kui palju sama indiviidi proovidest saadi mõlema protokolliga kätte. Joonis 16B on näha, et P indiviidi puhul oli Archeri ja Chudakovi protokolliga CDR3 aminohappeliste järjestuste ühisosa 7,59%, samas indiviidi T (Joonis 16C) puhul oli see ainult 2,93%. Seda näitas ka soojakaart (Joonis 15), kus oli näha, et T indiviidi ülekatted olid väiksemad kui indiviid P puhul. Indiviidide vaheline erinevus võib tuleneda sellest, et indiviid P puhul oli rohkem suurema levikuga klonotüüpe, mis olid suurema arvuliselt repertuaaris esindatud.



Joonis 16. A: Archer ja Chudakov protokollide jagatud kloonid. B, C: Vastavalt individ P ja T kloonide ühisosa Archeri ja Chudakovi protokollidel. Joonistel on toodud proovide CDR3 aminohappeliste järjestuste ühisosa protsent.

Vaadati ka ühisosa indiviidide ja kahe protokollide vahel, võrreldes Archeri protokolliga kõiki indiviid P ja T proove ning samamoodi Chudakovi protokollide proovidega. Indiviidide vaheline ülekate Archeri protokolliga oli 0,0275% ja Chudakovi protokolliga 0,0325% (Joonis 17). Seda näitas ka soojakaart (Joonis 15), kus indiviidide vahelisi ülekatteid visuaalselt näha ei olnud.



Joonis 17. Indiviidide klonotüüpide ühisosa Archeri (A) ja Chudakovi (B) proovidel. Joonistel on toodud proovide CDR3 aminohappeliste järjestuste ühisosa protsent.

Järgnevalt võrreldi CDR3 järjestuste ülekate protokollide vahel ning indiviidide ja ajapunktide kaupa. Lisa 5 A ja B joonistel on toodud individ P võrdlused vastavalt Archeri ja Chudakovi proovide kohta ning joonistel C ja D samad võrdlused individ T kohta. Individ P puhul oli ülekate Archeri ja Chudakovi protokollidel esimese ja teise ajapunkti vahel vastavalt 0,48% ja 0,52%. See on väga väike vahe (0,04%) ning näitab üldiselt, et kahes ajapunktis võetud verel on B-rakkude poolest väike, peaaegu olematu sarnasus. Individ T puhul oli meetodite vahel aga 3,4 kordne erinevus, Archeri proovide ülekate oli 0,982%, Chudakovi proovidel aga 0,286%: Ilmselt oli erinevus nii suur, sest individ T-l oli kahe meetodi klonotüüpide hulga erinevus suurem, kui individ P puhul.

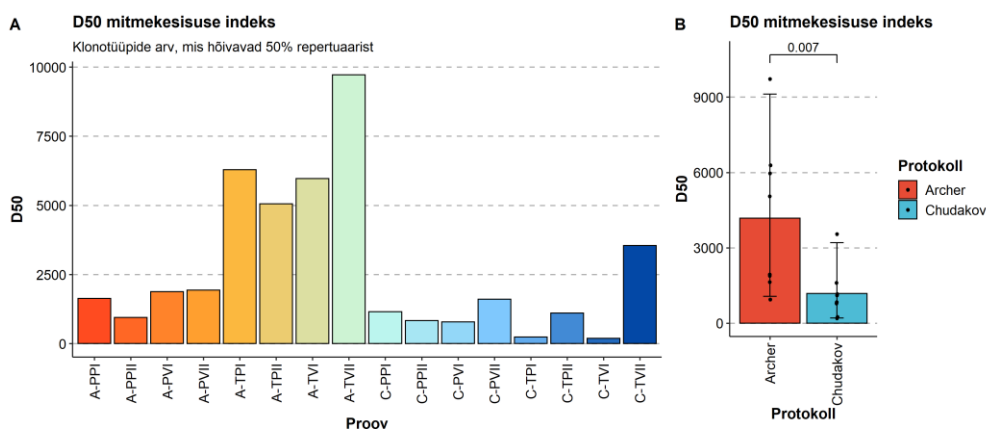
2.3.6. Repertuaaride mitmekesisus

Repertuaari mitmekesisuse indeksite arvutamiseks *downsampleti* kõik proovid kõige väiksema kloonide hulgaga repertuaari tasemele. Selleks oli proov C-TVI, kus oli 189 456 klooni. Seda tehti, kuna Archeri protokollide klonotüüpide hulk oli ka pärast *downsamplimist* 2,5 korda Chudakovi omast suurem. Mitmekesisuse arvutamise valemid sõltuvad suurel määral just

repertuaari suurusel ning ebavõrdsete repertuaari suurusel võib mitmekesisuse arvutamine olla kallutatud. Repertuaari uurimisel kasutati D50 ja *true diversity* indekseid. Kasutati just neid indekseid, kuna need annavad hea ülevaate repertuaari erinevate omaduste kohta.

2.3.6.1. D50 indeks

Esmalt vaadati repertuaari D50 indeksit, mis näitab klonotüüpide arvu, mis hõivavad 50% repertuaarist. Indeks sõltub ainult dominantsete CDR3 järjestuste arvust ning ei ole mõjutatud harvade liikide arvukusest (Chaudhary ja Wesemann, 2018). Mida kõrgem on D50 indeks, seda mitmekesisem on antud repertuaar. Joonisel 18 on toodud Archeri ja Chudakovi protokollide vaheline D50 indeksi vahe, mis on Archeri meetodi puhul kolm korda suurem kui Chudakovi proovidel (Archeri proovidel keskmiselt 3496, Chudakovil keskmiselt 1183, p-väärtus 0,007).

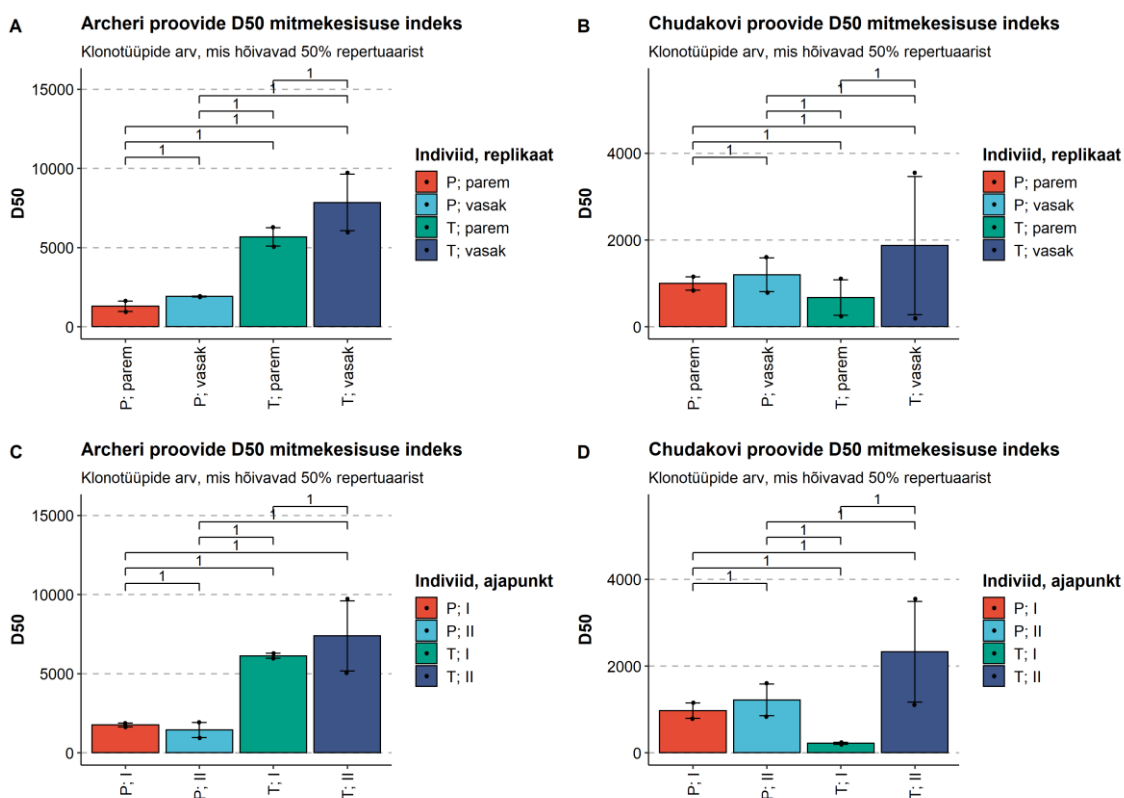


Joonis 18. D50 indeks proovide (A) ning protokollide (B) kaupa. X-teljel on toodud värvikoodiga eristatud protokoll, ning y-teljel klonotüüpide arv, mis hõivavad 50% repertuaarist. P-väärtus on arvatatud Wilcoxon testiga.

D50 indeksit võrreldi ka indiviidide, replikaatide ja ajapunktide kaupa. Joonisel toodud p-väärtused on väikese valimi suuruse tõttu kõik suurusega 1, seega kirjeldatakse edaspidi joonisel näha olevaid trende meetodi keskmiste väärtuste abil. Joonistel 19 A ja B on toodud indiviidide ja parema-vasaku käe replikaate, grupeerides D50 indeksi võrdlus eri meetodite puhul. Archeri (Joonis 19A) puhul on näha, et T indiviid on P-st mitmekesisem keskmiselt 4,2 korda (indiviid P D50 indeks 1600, indiviid T oma 6756) ning vasaku ja parema käe mitmekesisusel ei ole olulist erinevust (standardhälve P indiviidi replikaatide vahel 440, T indiviidi replikaatide vahel 1535). Joonisel 19B Chudakovi proovide puhul on kohe märgata, et väiksema arvu klonotüüpide pärast on ka mitmekesisuse indeks keskmiselt 4 korda väiksem kui Archeri proovide puhul (Archeri proovidel 4177, Chudakovi proovidel 1138). Ka pole Chudakovi meetodi puhul indiviidide vahel nii suurt erinevust, kui Archeri proovidel (Chudakovi proovidel oli T indiviidi D50 indeks keskmiselt 1,16 suurem kui indiviid P oma (995,5 vs 674)). Märgata on ka T indiviidi vasaku käe proovide suur erinevus, I ajapunkti D50

oli 195 ja II 3546. Erinevus tulebki nendes proovides levinuimate klonotüüpide suuruse erinevustest (vt Joonis 13).

Joonistel 19 C ja D on võrreldud Archeri ja Chudakovi meetodi D50 indeksit indiviidide ja ajapunktide kaupa. C joonisel on endiselt on mitmekesisem individ T, kuid mõlema indiviidi puhul pole ajapunktide mitmekesisusel statistilist olulist vahet (standardhälve P indiviidi ajapunktide D50 indeksi väärtuse vahel 223, T indiviidi replikaatide vahel 890). Joonise 19D Chudakovi proovide puhul on D50 indeks madalam kui Archeri proovidel ja P indiviidi mõlemad ajapunktid on väga sarnase mitmekesisusega. Kuid individ T II ajapunktis on kahe proovi D50 indeks niivõrd erinev, sest proovis C-TVII olid enamlevinud kloonid väikeste arvukustega, kuid C-TPII proovis oli rohkem „keskmise“ esinemissagedusega klonotüüpe (vt Joonis 13), mis mõjutab suuresti D50 indeksit. Samas indiviidi T I ajapunkti D50 väärtused on mõlemad väga sarnased ning madalad.



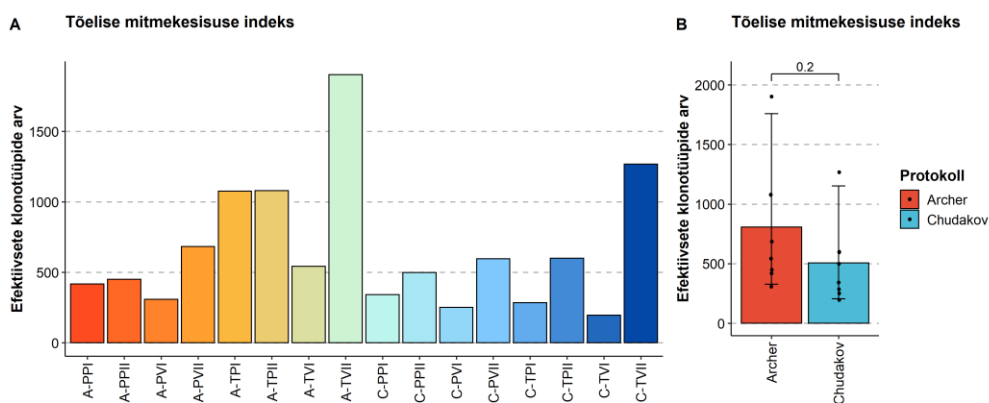
Joonis 19. A ja B: D50 indeks Archeri ja Chudakovi protokollidel grupeeritud indiviidide ja replikaatide kaupa. C ja D: D50 indeks Archeri ja Chudakovi proovidel grupeeritud indiviidide ja ajapunktide kaupa. X-teljel on vastav värvikoodiga eristatud grupeering ning y-teljel 50% repertuaarist hõivavate klonotüüpide arv. P-väärtused on arvatatud Kruskal-Wallise testiga.

2.3.6.2. Tõelise mitmekesisuse indeks

Teise mitmekesisuse indeksina arvutati repertuaaride nn „tõeline mitmekesisus“ (TD, *true diversity*), mis näitab „efektiivset“ klonotüüpide arvu. TD indeksi arvutamisel leitakse kõigepealt repertuaari Shannon-Wieneri indeks, seejärel leitakse sellele indeksile vastav

klonotüüpide arv tingimusel, et kõik klonotüübid on võrdse arvukusega. Viimane nn võrdse arvukusega klonotüüpide arv ongi „efektiivne“ klonotüüpide arv ehk tõelise mitmekesisuse indeks. See indeks annab lihtsasti interpreteeritava ja standardiseeritud mõõdiku, mis arvestab nii repertuaari kloonide arvu kui ka nende kloonide proportsioone (Rosenfeld et al., 2018).

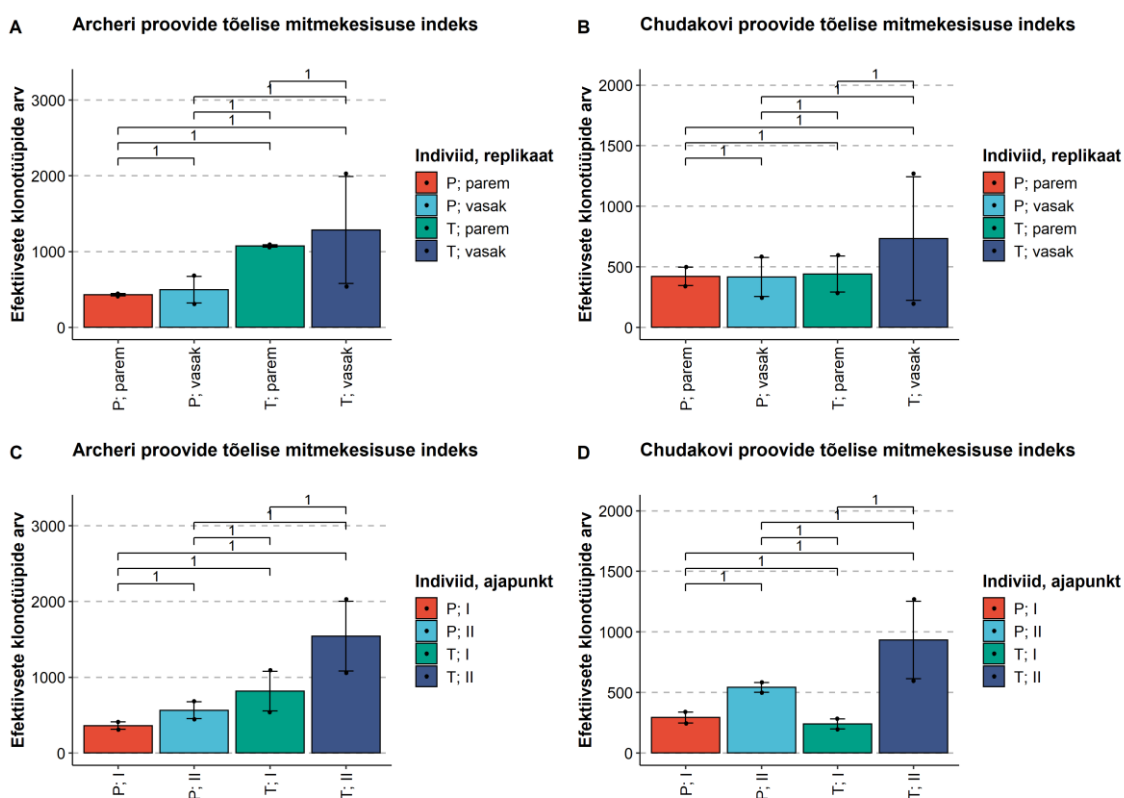
Joonisel 20A on toodud TD indeks iga proovi kohta ning sellel graafikul on hästi näha, et kahel proovil (A-TVII ja C-TVII) on võrreldes sama indiviidi teiste proovidega suurema klonotüüpide arvu tõttu kõikide proovide keskmisest 2,5 korda suurem TD indeks. Joonisel 20B on toodud Archeri ja Chudakovi protokollide võrdlus tõelise mitmekesisuse indeksi kohta. Vahe kahe protokollil vahel pole enam nii suur kui D50 puhul (D50 indeksi puhul oli p-väärtus 0,007, TD p-väärtus 0,2). Archeri TD on 1,6 korda kõrgem kui Chudakovi protokollil (Archeril keskmiselt 806, Chudakovil 502). Mõlema protokollil puhul on efektiivsete klonotüüpide arv suhteliselt madal, arvestades ka *downsampletud* andmestiku klonotüüpide koguhulka (Archeri protokollil kokku 546 344 klonotüüpi ja Chudakovil 91 317 klonotüüpi).



Joonis 20. A: Tõelise mitmekesisuse indeks proovide vahel. B: Tõelise mitmekesisuse indeks värvikoodiga eristatud protokollide kaupa. X-teljel on toodud protokoll, ning y-teljel efektiivsete klonotüüpide arv. P-väärtus on arvutatud Wilcoxon testiga.

Edasi võrreldi TD indeksit Archeri ja Chudakovi protokollide puhul grupeerides indiviide ja ajapunkte. Joonistel 20 A ja B on toodud TD indeksi indiviidide ja parema-vasaku käe replikaatide vaheline võrdlus mõlema meetodi puhul. Sarnaselt D50 indeksiga on joonistel toodud p-väärtused väikese valimi suuruse tõttu kõik suurusega 1, seega kirjeldatakse edaspidi jooniselt näha olevaid trende keskmiste väärtuste abil. Jooniselt A võib välja lugeda, et Archeri proovide puhul on mitmekesisus indiviidis T jällegi suurem ning replikaatide vahel suuri erinevusi ei olnud (indiviid P replikaatide vaheline standardhälve 47, T puhul 147). Chudakovi protokollil puhul joonisel B on TD indeksid kolmel proovil neljast väga sarnased, kuid T indiviidi vasaku käe proov suure klonotüüpide erinevuse tõttu eristub suurema mitmekesisuse poolest (indiviid P replikaatide vaheline standardhälve 3, T puhul 206).

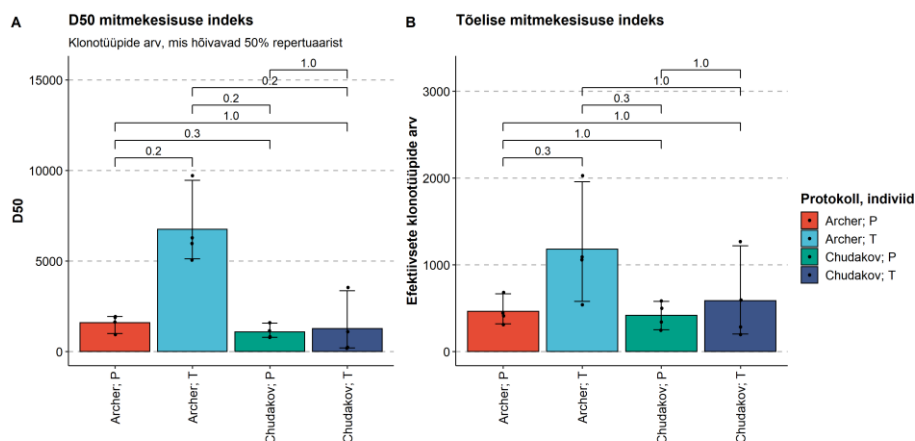
Joonistel 21 C ja D on võrreldud indiviidide ajapunktide mitmekesisust Archeri ja Chudakovi meetodi võrdluses. Archeri proovidel ei olnud indiviid P-I nii suurt erinevust (standardhälve 143), kuid indiviid T proovide TD indeks erines kahe ajapunkti vahel suuresti (standardhälve 514). Joonisel D on Chudakovi proovide indiviidi P kahe ajapunkti proovid omavahel suhteliselt sarnase mitmekesisusega, kuid T indiviidi II ajapunkti proovid oli keskmisest 1,8 korda kõrgema TD indeksiga (932 vs 502), kuna jällegi vasaku käe proov C-TVII oli väga suure klonotüübi hulgaga ja seega repertuaari proportsioonid eristusid teistest Chudakovi proovidest (vt Joonis 13). Nii joonisel C kui D on huvitaval kombel näha, et mõlema indiviidi puhul on II ajapunkti TD indeks kõrgem kui esimese ajapunkti puhul, kuigi teises ajapunktis ei olnud kummalgi otseselt rohkem klonotüüpe.



Joonis 21A ja B: Tõelise mitmekesisuse indeks Archeri ja Chudakovi protokollidel grupeeritud indiviidide ja replikaatide kaupa. C ja D: Tõelise mitmekesisuse indeks Archeri ja Chudakovi proovidel grupeeritud indiviidide ja ajapunktide kaupa. X-teljel on toodud vastav värvikoodiga eristatud grupeerimise ning y-teljel efektiivsete klonotüüpide arv. P-väärtused on arvatud Kruskal-Wallise testiga.

Kokkuvõttes võib mitmekesisuse kohta öelda, et Archeri meetodiga tehtud proovid olid suuremate repertuaaride tõttu mitmekesisemad, ka pärast *downsampling*. Mõlema antud töös kasutatud indeksi puhul olid kõige mitmekesisemad indiviid T Archeri meetodiga tehtud proovid, kuna nende repertuaar oli kõige suurem (p-väärtus teiste gruppidega võrreldes <0,3). Chudakovi meetodiga ei olnud kummagi indeksi puhul indiviidide vahel erinevust (p-väärtus

1). Samas oli erinevus D50 indeksi ja tõelise mitmekesisuse indeksi vahel seal, et TD indeksi puhul ei erinenud indiviid P kummastki Chudakovi meetodiga tehtud indiviidide TD indeksi väärtusest (p-väärtus mõlemal juhul 1), kuid D50 indeksi puhul oli indiviid P Archeri meetodiga tehtud proovide väärtus kõrgem kui Chudakovil (p-väärtus 0,3) (Joonis 22).



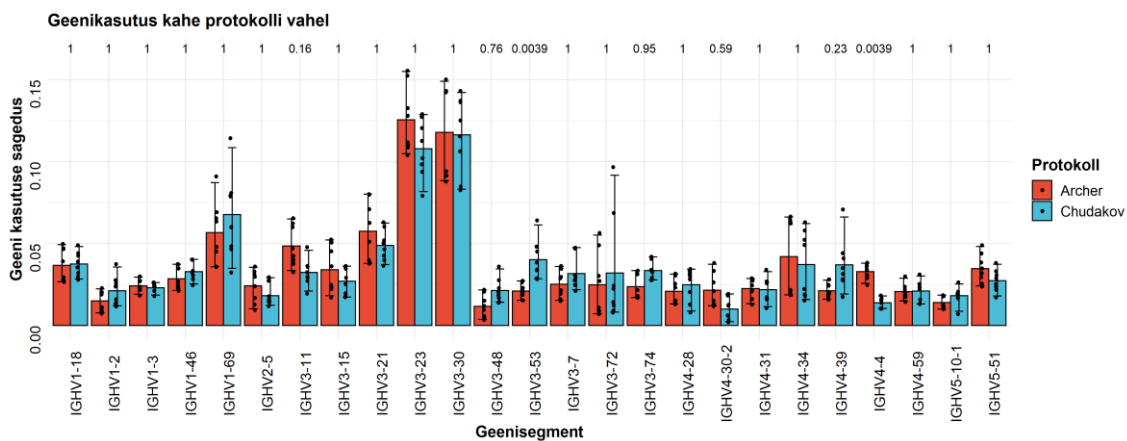
Joonis 22. A ja B: D50 indeksi ja TD indeks kahe meetodi ja indiviidide võrdluses. X-teljel on toodud meetodi nimi ja indiviidi tähistus, y-teljel vastav mitmekesisuse indeks. P-väärtused on arvutatud Kruskal-Wallise testiga.

2.3.7. Geenisegmentide kasutus

Viimaks vaadati repertuaari geenikasutust, et teada saada, kas Archeri ja Chudakovi meetodid tuvastasid samade geenisegmentide kasutust ning kas indiviidide vahel on segmentide kasutuses leida sarnasusi. Kõikide jooniste puhul vaadati ainult enim kasutatavate geenisegmentide kasutust, kuna paljude segmentide või nende kombinatsioonide kasutuse sagedus jäi vahemikku, mis erines kõikide proovide vahel tuhandeid kordi. Lisaks ei tuvastatud paljude geenisegmentide kasutust üldse, seega jäeti need graafikutelt välja. Joonistel on toodud geenisegmentid, mille kasutuse sagedus jäi üle 0,025 piiri. Lisaks võeti arvesse ka vastava klonotüübi arvukust. Joonistel on näha ainult V-geenide kasutus, kuna D- ning J-geenide kasutuse sagedus oli võrreldes V-segmentidega liiga madal ja üleüldiselt oli V-geenide kasutus domineeriv.

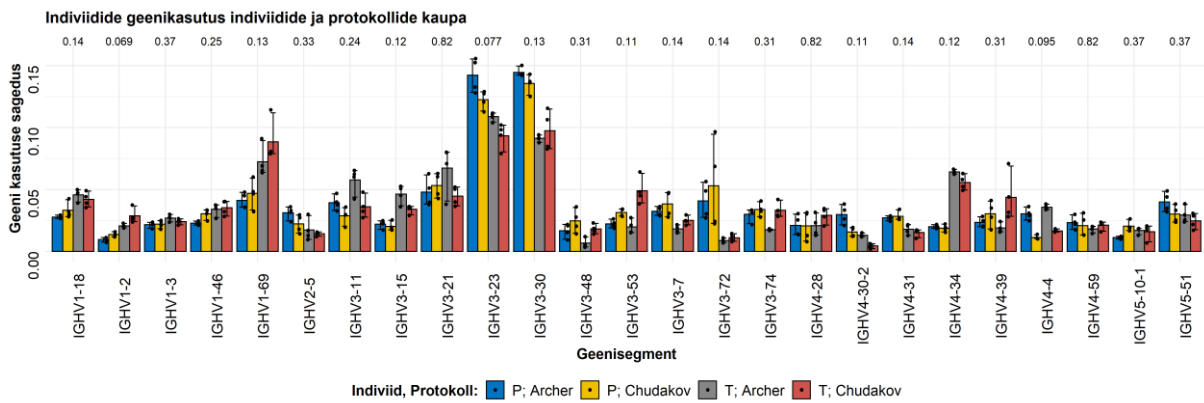
Joonisel 23 on kujutatud geenide kasutus kahe meetodi võrdluses. Joonisel on toodud 25 kõige enim kasutatud segmenti, mille kasutuse sagedus jäi üle 0,025. Üldiselt on näha, et mõlemad protokollid on hästi tuvastanud samu enim kasutatavaid geenisegmente. P-väärtused on väiksemad kui 1 seitsmel segmendil ning kõige suuremad erinevused meetodite vahel on segmentidel IGHV3-53 ning IGHV4-4. Kõige rohkem on kasutatud IGHV3 perekonna gene IGVH3-23 ning IGHV3-30 ning mõlema sagedus jääb üle 0,1 piiri. Mõlemad protokollid on seda tuvastanud võrdsel määral. Enim on protokollid tuvastanud segmente perekonnast IGHV3. IGHV3 perekonnast kasutati suurema sagedusega ka segmente IGHV3-11, IGHV3-15, IGHV3-

21, IGHV3-48, IGHV3-49, IGHV3-7, IGHV3-72 ja IGHV3-74. Lisaks kasutati palju perekonna IGHV1 geene, sealt kõige rohkem segmenti IGHV1-69. Kasutuse sagedus oli suur ka perekonnas IGHV4, kus on mõlema meetodi abil tuvastatud segmentide IGHV4-28, IGHV4-30-2, IGHV4-31, IGHV4-34, IGHV4-39, IGHV4-44, IGHV4-59 kasutust. Selle perekonna kasutuse sagedus on segmentide lõikes palju ühtlasem kui IGHV3 perekonnas, kus oli 2 väga domineeriva sagedusega segmenti ning palju madalate sagedustega segmente. Väiksema sagedusega tuvastati segmente perekondadest IGHV2 ja IGHV5.



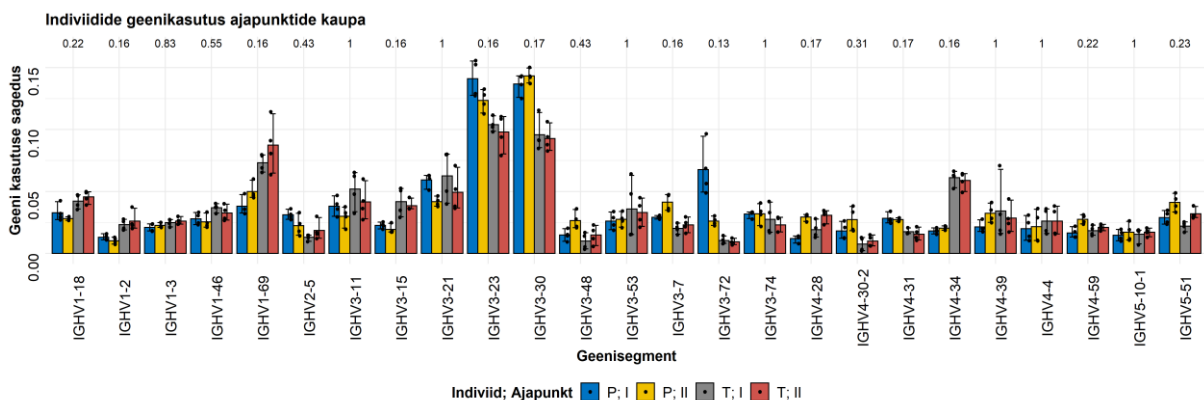
Joonis 23. Protokollide kaupa grupeeritud geenisegmentide kasutus. X-teljel on välja toodud vastava geenisegmenti nimi ning y-teljel on välja toodud geenisegmenti kasutuse sagedus. Värvikoodiga on eristatud protokoll. P-väärtused on arvatud Wilcoxon testiga.

Vaadati geenikasutust ka indiviidide ja protokollide vahel (Joonis 24). Joonistel on näha, et mõlemad indiviidid on kasutanud samu geenisegmente, kuid erineva sagedusega (p-väärtus pole ühelgi juhul 1). Kõige ühtlasematel tasemetel indiviidide ja protokollide vahel tuvastati segmentidel IGHV3-21, IGHV4-28 ja IGHV4-69 puhul, kus kõikidel oli p-väärtus 0,82. IGHV3-23 ja IGHV3-30 segmente olid meetodid (Joonis 23) tuvastanud samadel tasemetel, kuid siit jooniselt on näha, et mõlema protokolliga puhul on mõlema segmenti puhul individ P sagedused umbes 0.05 võrra kõrgemad (p-väärtused vastavalt 0,077 ja 0,13). Mõlema indiviidi puhul on suured erinevused ka segmentide IGHV1-69, IGHV3-11 ja IGHV4-34 juures, kus indiviidi T geeni kasutuse sagedus on palju kõrgem (p-väärtused vastavalt 0,13, 0,12 ja 0,12). Lisaks olid indiviidide ja meetodite vahelised erinevused näha ka segmenti IGHV3-53 juures, kus on näha, et individ T puhul on Chudakovi meetod segmenti kasutust oluliselt rohkem tuvastanud. Individ P on individ T-st kasutanud rohkem segmente IGHV3-7 ja IGHV3-72. Võrdsetel tasemetel on indiviidid kasutanud segmente IGHV1-3, IGHV3-48, IGHV4-59, IGHV5-10-1 ja IGHV4-51.



Joonis 24. Geenikasutus protokollide ja indiviidide kaupa. X-teljel on välja toodud geenisegmentide nimed ning y-teljel on välja toodud geenisegmendi kasutuse sagedus. Värvikoodiga on eristatud indiviidide P ja T kahe meetodiga valmistatud proovid. P-väärtus on arvutatud Kruskal-Wallis testiga.

Vaadati ka geenikasutust indiviidide ja ajapunktide vahel, kombineerides mõlema meetodiga tehtud proovid (Joonis 25), kuid olulisi erinevusi ajapunktide vahel ei nähtud. Ainuke segment, mille sagedusel oli oluline erinevus ajapunkti vahel oli IGHV3-72 indiviidi P puhul, kuid segmendil pole olulist funktsiooni mõne haiguse esinemise puhul.



Joonis 25. Geenikasutus indiviidide ja ajapunktide vahel. X-teljel on välja toodud geenisegmentide nimed ning y-teljel on välja toodud geenisegmendi kasutuse sagedus. Värvikoodiga on eristatud indiviidide P ja T kahe ajapunkti proovid. P-väärtus on arvutatud Kruskal-Wallis testiga.

Üldiselt võib geenikasutuse kohta öelda, et meetodid on samal tasemel geenisegmentide kasutust tuvastanud, sest indiviidide erinevate protokollidega tehtud proovide sageduste erinevus ei ole enamusel proovidel suur. Kõige rohkem kasutati mõlema inimese puhul segmente IGHV3-23 ja IGHV3-30, mille esinemissagedus oli mitu korda kõrgem teiste segmentide omast. Teiste enim kasutatud segmentide esinemise sagedus jäi enamikul juhtudel alla 0,05 piiri.

2.4. Arutelu

B-raku repertuaari uurimine on oluline, sest paljude immuunvahendatud haiguste käigus muutub B-rakkude repertuaar ning selle uurimine aitab paremini mõista nende haiguste tekkemehhanisme ning potentsiaalseid ravivõimalusi. Antud bakalaureusetöös uuriti kahe terve inimese B-raku retseptorite repertuaari kahe erineva meetodiga. Esimene töös kasutatud meetod, Chudakovi protokoll (Turchaninova et al., 2016), põhineb 5' RACE meetodil. Teine meetod, Archeri kit (ArcherDX, 2018) põhineb aga Anchored Multiplex PCR-il. Mõlematel meetoditel märgistatakse proovid UMI-dega ning amplifitseeritakse multipleks praimeritega inimese BCR-i variaabel osa V-, D- ja J geenisegmendid.

Kahe meetodi erinevuste võrdlemiseks vaadati kõigepealt sekveneerimise kvaliteeti. Nähti, et kõik Chudakovi meetodiga tehtud proovid olid madalama kvaliteediga kui Archeri omad, mis omakorda mõjutas ka edasisi tulemusi. Chudakovi meetodi puhul hakkas sekveneerimise kvaliteet alla kriitilise piiri langema keskmiselt 150 bp juures, samas Archeri kiti puhul toimus see 250 bp juures, mis on antud sekveneerimise meetodi eripära. Kuna MiXCR programm eemaldas analüüsi käigus halvema kvaliteediga järjestused, jäi selle tulemusena Chudakovi proovidele alles vähem antikeha järjestusi ja seega olid nendel proovidel repertuaarid väiksemad kui Archeri proovidel. Chudakovi proovide väiksem repertuaar väljendus madalamas kloonide ja klonotüüpide arvus ja seeläbi vähenes ka proovide mitmekesisus. Käesolevas töös vaadati ka CDR3 järjestuste aminohappelist pikkust ning leiti, et mõlema protokolliga puhul oli see keskmiselt 20 aminohapet pikk. CDR3 pikkus on mõjutatud erinevate immuunhaiguste poolt, muutes seda kas lühemaks või pikemaks (Bashford-Rogers et al., 2019). Sama uuring näitas, et keskmiselt on tervetel inimestel CDR3 pikkus 20 aminohapet, mis on kooskõlas antud tööga.

Nagu eelnevalt mainitud, oli Chudakovi meetodiga tehtud proovidel väiksemate repertuaari suuruste tõttu väiksem mitmekesisus ning see väljendus omakorda madalamas D50 ja TD indeksi väärtuses. D50 indeksi puhul oli meetodite vaheline erinevus statistiliselt olulisem (p-väärtus 0,007), mis näitab, et meetodid tuvastasid enim levinud klonotüüpe väga erinevatel tasemetel. TD indeksi puhul oli meetodite vahel väiksem erinevus (p-väärtus 0,2). Sellest võib järeldada, et tegelik mitmekesisus, mis hõlmab terves repertuaaris esinevate klonotüüpide arvukusi, on kahel meetodil sarnasem kui kloonide arv kõige levinumates klonotüüpides.

Antud töös leiti ka, et indiviidide vahel esineb eri meetodite puhul samade klonotüüpide vahel väga väike ühisosa (Joonis 17). Seda toetab ka Soto *et al.* (2019) poolt läbi viidud uuring, kus näidati, et kolme terve indiviidi vahel jagati ainult 0,2% klonotüüpidest. Kahe terve indiviidi vahel jagati keskmiselt 2,6% repertuaarist, kuid seal uuringus ei võetud igast indiviidist võrdsel

arvul klonotüüpe nagu tehti käesolevas töös. Kuna soojakaardil (Joonis 15) klasterdusid kõik proovid indiviidide ja ajapunktide kaupa, näitab see seda, et ka sama indiviidi ajapunktide vahel oli sarnasus väga väike. Kõige suurem kattuvus oli erinevate meetoditega tehtud sama proovi vahel Käesolevas töös nähti, et mõnel üksikul juhul esines ülekate vasaku ja parema käe replikaadi vahel. Oodatav oleks, et nende vahel ei esine väga suurt ülekate, kuna see oleks umbes nagu „basseini kahest erinevast nurgast lusikaga võtmine ja lootmine, et saad sama tulemuse“. Siin töös loodeti näha, et sama indiviidi erineva meetodiga tehtud proovi vahel on suur ülekate. See pidas paika, sest Venni diagrammidega visualiseeritud ülekate näitas, et kahe meetodi vahel oli P indiviidi puhul ülekate 7,59% ja T indiviidi puhul 2,93% (Joonis 16B, C). Seda oli näha ka soojakaardil (Joonis 15), kus oli näha erineva meetodiga tehtud sama proovi ülekate. Meetodite vahelised ülekatted näitavad, et mõlemad protokollid on efektiivsed tuvastamiseks samu CDR3 järjestusi.

Käesolevas töös nähti, et mõlemad meetodid suutsid tuvastada samade geenisegmentide kasutust. Antud bakalaureusetöös tuvastati enim kasutatavaid geenisegmente, mille sagedus repertuaaris oli suurem kui 0,025. D- ja J segmentide kasutust ei tuvastatud peaaegu üldse. Enim kasutati segmente IGHV3-23 ja IGHV3-30, millest mõlema sagedus ületas 0,1 piiri. Mõlemat segmenti kasutas rohkem indiviid P, kuid erinevus polnud statistiliselt oluline. Kusjuures nii Chudakovi kui Archeri meetod tuvastasid samu segmente statistiliselt väikese erinevusega. Kõige suuremad erinevused protokollide vahel olid segmentide IGHV3-53 (Archeri proovide sagedus 0,019, Chudakovil 0,04) ning IGHV4-4 juures (Archeri sagedus 0,033, Chudakovil 0,014), kus p-väärtus oli mõlema puhul 0,0039 (Joonis 23). IGHV3-53 segmenti puhul on välja toodud, et selle poolt kodeeritud antikehade tase tõuseb pärast COVID-19 infektsiooni (Wu et al., 2020). Antud töös vaadati põgusalt ka indiviidide geenikasutuse erinevusi ajapunktide vahel (Joonis 25) ning nähti, et antud segmenti puhul polnud kummalgi indiviidil sagedus teises ajapunktis suurem (II ajapunkt võeti peale COVID-19 pandeemia algust). Arnaout *et al.* (2011) poolt läbi viidud terveid indiviide kaasanud uuring tuvastas indiviididel enim kasutatud geenisegmente ning enamik neist tuvastati ka käesolevas töös. 10-st Arnaout *et al.* (2011) töös tuvastatud geenisegmendist tuvastati käesolevas töös 8, mis viitab nende segmentide kasutuse universaalsusele. Nendeks segmentideks olid IGHV3-23, IGHV3-48, IGHV1-3, IGHV1-18, IGHV4-34 ja IGHV3-11.

Kuna antud bakalaureusetöö eesmärk oli võrrelda kahe meetodi efektiivsust B-raku retseptorite repertuaari uurimisel, on olulised ka praktilised aspektid. Meie laboris on varasemalt repertuaari uurimisel kasutatud Archeri kiti, kuid tekkis vajadus protokollile järele, mida saaks modifitseerida vastavalt katsele (erinevate immunoglobuliinide kombineerimine või lahutamine, sobilikkus

hiirele ja rotile) ning vähemtähtis pole ka kogu katse hind. Sellest tulenevalt võeti kasutusele modifitseeritud Chudakovi meetod, mis võimaldab raamatukogu ettevalmistusel Archeri kitile sarnast ja efektiivset meetodit, aga palju odavamalt ja paindlikumalt.

Antud töös modifitseeriti Chudakovi protokoll, et eksperimendi kulgu lihtsustada ning ka hinda vähendada. Originaalprotokollis on II PCR-i praimerite küljes nn „vöotkoodid“, et sekveneerimisel proove üksteisest eristada. See eeldab aga seda, et iga proovi jaoks disainitakse eraldi praimer. Antud töös oli kokku ainult 8 proovi, kuid meie laboris on tihti katseid, kus proovide arv võib ulatuda mõnesajani. Sellisel hulgal praimerite disainimine on väga töömahukas ning kulukas. Selle vältimiseks lisati II PCR praimeritele Illumina üleulatuvad otsad (Lisa 1, praimerid FwdOH_M1S ja RevOH_Z2), ning proovid märgistati eraldi etapis Nextera indeksitega, mis võimaldas teha ka suuremahulisi katseid.

Chudakovi protokoll suureks eeliseks on ka see, et seda saab lihtsasti modifitseerida, kasutades erinevaid primereid, näiteks isotüübi vahetuse täpsemaks uurimiseks, või erinevate organismide uurimiseks. Lisaks võimaldab Chudakovi meetod kasutada originaalprotokollist erinevate tootjate poolt tehtud, kuid sama funktsiooniga reagente või proovide puhastamise kitte. Archeri kommertsiaalset kiti ei saa aga üldse modifitseerida, kuna kõikide reaktsioonide jaoks vajalikud reagentid on juba tuubides lüofiliseeritud kujul. See omakorda vähendab inimlike eksimuste vigu, mis Chudakovi protokollil on tõenäolisemad, kuna kõik reaktsioonid tuleb ise kokku segada.

Archeri kiti puuduseks võib kindlasti lugeda ka kõrget hinda, mis 8 proovi puhul on umbes 1200€. Lisaks on Archeri kitiga võimalik koos sekveneerida vaid maksimaalselt 48 proovi, kuna see on maksimaalne MBC adapterite hulk. Chudakovi protokoll kombinimine Nextera indeksitega võimaldab korraga sekveneerida kuni 384 proovi, mis teeb sellest suuremate eksperimentide puhul palju käepärasema meetodi.

Käesolevas töös saadud tulemuste põhjal võib oletada, et Chudakovi meetodi halvemad tulemused võivad tuleneda sellest, et praimerite multipleksimine ei tööta väga hästi ning praimerid hakkavad üksteist nii-öelda segama. Selle kindlakstegemine nõuab suuremat proovide arvu ning lisakatseid üksikute praimerite kaupa. Võimalik ka, et II PCR käigus kasutatud Illumina *overhangiga* praimeritel on kahjulik mõju ülejäänud reaktsioonile, kuna need on praimeriga suhteliselt pikad (vastavalt 57 ja 51 nt) ning selle tulemusel võib matriitsile seondumise efektiivsus olla madalam. Kuigi kahe meetodi sarnased tulemused nii CDR3 järjestuste pikkuste, geenisegmentide kasutuse kui ka CDR3 järjestuste ülekatte osas näitavad, et tegelikult Chudakovi meetod töötab ja näitab sarnaseid tulemusi nagu Archeri kit, kuid mitmed repertuaari olulised näitajad, näiteks antud töös kasutatud D50 ja TD indeks, on

mõjutatud sekveneerimise kehvemast kvaliteedist ning antud töös jäid Chudakovi meetodi head tulemused halva sekveneerimise kvaliteedi varju. Tulevaseks eesmärgiks on optimeerida Chudakovi meetodit, kuna see meetod on laboritöökäes väga vajalik tänu selle mugavale kasutusele, modifitseerimisvõimalusele ja odavamale hinnale. Kindlasti oleks vaja laiaulatuslike järelduste tegemiseks suurendada proovide arvu.

Antud bakalaureusetöö peamine eesmärk oli võrrelda Archer Immunoverse kiti ja Chudakovi protokolliga B-raku retseptorite repertuaari analüüsimisel ning teada saada, kumb meetod on parem ja efektiivsem. Archeri kiti abil saadi kvaliteetsemad andmed ja sellest tulenevalt tuvastati ka rohkem kloonide, klonotüüpide ning repertuaarid olid mitmekesisemad. Samuti tuvastati Archeri kiti abil rohkem harva esinevaid klonotüüpe, mida Chudakovi protokolliga puhul ei suudetud teha. Samas enamlevinud klonotüüpide tuvastamisega said mõlemad meetodid hästi hakkama. Kokkuvõttes oli Archer Immunoverse kit küll hea meetod tuvastamiseks eri levimissagedusega klonotüüpe, andes mitmekesised repertuaarid, kuid ka Chudakovi meetod pole otseselt halb, kuna CDR3 järjestuste ülekatted kahe meetodi vahel näitasid, et Chudakovi meetod suudab sama hästi kloonide tuvastada. Käesolevas töös mõjutas Chudakovi protokolliga tulemusi negatiivses suunas halb sekveneerimise kvaliteet. Tulevikus loodetakse optimeerimise käigus siiski Chudakovi protokolliga edasi kasutada.

KOKKUVÕTE

Käesoleva töö peamine eesmärk oli võrrelda Archer Immunoverse Kiti ja Chudakovi protokolliga B-raku retseptorite repertuaari uurimiseks. Selleks uuriti kahte indiviidi vanustes 20 ja 35, võttes vereproovi vasakust ja paremast käest ning kahes ajapunktis, mis olid kuuekuulise vahega. Töös leiti, et Archeri meetod oli BCR-i repertuaari analüüsil efektiivsem kui Chudakovi protokoll, sest selle meetodi sekveneerimise kvaliteet oli parem ning seetõttu tuvastati rohkem kloone ja klonotüüpe. Chudakovi meetod ei olnud harva esinevate klonotüüpide tuvastamisel nii efektiivne, ilmselt madalama sekveneerimise kvaliteedi tõttu. Samas enamlevinud klonotüüpide tuvastamisega said mõlemad meetodid hästi hakkama. Töös arvatud mitmekesisuse indeksi väärtused näitasid, et Archeri kiti abil tehtud proovid olid mitmekesisemad, seda ka suurema repertuaari tõttu. Geenisegmentide kasutuse poolest olid kaks protokolliga samaväärsed, tuvastades samu geenisegmente samadel sagedustel. Kuigi antud töö tulemustes tuli välja, et Archeri meetod on repertuaari uurimisel efektiivsem, oli siiski näha, et Chudakovi meetod tegelikult töötab, kuid head tulemused jäävad kehva sekveneerimise kvaliteedi varju.

Töö teine eesmärk oli analüüsida BCR-i repertuaari klonaalsust, mitmekesisust, geenide kasutust, ja CDR3 järjestuste ülekate kahe indiviidi ning ajapunktide võrdluses. Töös leiti, et kahe indiviidi BCR repertuaarid ei ole kattuvad ning sarnaste järjestuste klasterdumine toimub indiviidide ja ajapunktide kaupa. Mõningane ülekate oli sama indiviidi erinevatest käest võetud proovide vahel. Geenisegmentide kasutus oli indiviidide vahel sarnane, sest kõrgete sagedustega kasutati samu segmente.

Kokkuvõttes oli Archeri kit hea meetod tuvastamiseks mitmekesisuid ja eri levikuga klonotüüpe, kuid ka Chudakovi protokoll polnud otseselt halb, kuna kahe meetodi vaheline CDR3 järjestuste ülekate näitas, et Chudakovi meetod suudab sama hästi kloone tuvastada.

Analysis of the B cell receptor repertoire with two methods

Liisa Pomerants

SUMMARY

B cells and antibodies are an important part of the acquired immune system. They neutralize invading pathogens by binding them and prohibiting them from entering the cell. In addition, they activate other components of the immune system, for example the complement system and other immune cells. Each antibody or its membrane form BCR is specific to one specific antigen. The specificity is created through highly regulated processes and as a result, all the BCRs in the organism form the B cell repertoire. Studying of the BCR repertoire is important to better understand immune mediated diseases, their pathogenesis and to find potential treatments or vaccines. In recent years, the development of high throughput sequencing has allowed for better quality obtaining of immune data.

In this thesis, the library for BCR repertoire of two healthy individuals was prepared with two different methods, the commercial Archer Immunoverse Kit and the “Chudakov protocol”. The goal of this thesis was to compare which of the two is more effective for BCR repertoire analysis. The results showed that the Archer kit was more efficient. It was better already in sequencing quality and therefore detected more clones and clonotypes than the Chudakov protocol. Also, the Chudakov protocol was not as good in detecting rarely occurring clonotypes as the Archer kit. Due to the smaller repertoire sizes, the diversity was also lower for the Chudakov protocol samples. The CDR3 sequence overlap was the biggest for the same sample analysed with different methods. There was little to no similarity between the two individuals or the two time points of the same individual. Gene segment usage was also analysed. It was very similar in both methods and between the two individuals. The segments that were used the most were IGHV3-23 and IGHV3-30.

Since many of the repertoire characteristics depend on sequencing quality and repertoire size, Archer kit’s results were more efficient in BCR repertoire analysis, even though the Chudakov protocols’ results were not that bad, just affected by the poor sequencing quality. The goal for the future is to optimise the Chudakov protocol to obtain better sequencing quality and therefore hopefully get better results.

TÄNUSÕNAD

Kõige suuremad tänusõnad lähevad juhendajatele Liina Tserelile ja Ahto Salumetsale, kes aitasid töö valmimisele kaasa, abistasid kõigega ja andsid töö kirjutamiseks häid soovitusi.

Veel tahan tänada Priit Adlerit, kes aitas serveris keeruliste andmete töötlemisel ning Maksym Zarodniuki heade nõuannete eest andmete analüüsil. Lisaks tänan meie uurimisgrupi juhti professor Pärt Petersoni võimaluse eest teha lõputööd nii lähedas laboris tohutult tarkade inimeste keskel ja usaldades mulle keerulise teema. Viimaks tänan kogu uurimisgrupi inimesi abivalmiduse ja lõbusa töömeeleolu loomise eest.

KASUTATUD KIRJANDUS

- Abbas, A., Lichtmann, A., and Pillai, S. (2018). *Cellular and Molecular Immunology* (9th ed.). 103, 125–126, 184–185. Elsevier Inc., Philadelphia.
- Alt, F. W., Oltz, E. M., Young, F., Gorman, J., Taccioli, G., and Chen, J. (1992). VDJ recombination. *Immunology Today*, 13(8), 306–314.
- Andrews, S. (2014). *FastQC A Quality Control tool for High Throughput Sequence Data* (0.11.2). Babraham Bioinformatics.
<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- ArcherDX. (2018). *Archer® Immunoverse™- HS BCR Protocol for Illumina®*. 1–24.
- ArcherDX. (2019). *Repertoire analysis on your own terms*. <https://dlongwood.com/wp-content/uploads/2019/11/archer-Immunoverse.pdf>
- Arnaut, R., Lee, W., Cahill, P., Honan, T., Sparrow, T., Weiland, M., Nusbaum, C., Rajewsky, K., and Korralov, S. B. (2011). High-resolution description of antibody heavy-chain repertoires in humans. *PLoS ONE*, 6(8).
- Backhaus, O. (2018). Generation of Antibody Diversity. In *Antibody Engineering*. InTech.
- Bashford-Rogers, R. J. M., Bergamaschi, L., McKinney, E. F., ... and Smith, K. G. C. (2019). Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature*, 574(7776), 122–126.
- Bassing, C. H., Swat, W., and Alt, F. W. (2002). The mechanism and regulation of chromosomal V(D)J recombination. *Cell*, 109(2 SUPPL. 1), 45–55.
- Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V., and Chudakov, D. M. (2015). MiXCR: Software for comprehensive adaptive immunity profiling. *Nature Methods*, 12(5), 380–381.
- Bolotin, Dmitriy A, Stanislav Poslavsky, Alexey N Davydov, ... Chudakov, D. M. (2017). Antigen receptor repertoire profiling from RNA-seq data. *Nature Biotechnology*, 35(10), 908–911.
- Bolotin, Dmitry A., Poslavsky, S., and Shugay, M. (2020). *MiNNN: Universal tool for processing molecular-barcoded Next Generation Sequencing data*.
<https://github.com/milaboratory/minnn>
- Chaplin, D. D. (2010). Overview of the immune response. *Journal of Allergy and Clinical Immunology*, 125(2 SUPPL. 2), S3–S23.

- Chaudhary, N., and Wesemann, D. R. (2018). Analyzing immunoglobulin repertoires. *Frontiers in Immunology*, 9(MAR), 1–18.
- Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, 12(35).
- Chi, X., Li, Y., and Qiu, X. (2020). V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. *Immunology*, 160(3), 233–247.
- Chiu, M. L., Goulet, D. R., Teplyakov, A., and Gilliland, G. L. (2019). Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies*, 8(4), 55.
- Cohen, S. (1975). Antibody structure. *Journal of Clinical Pathology*, 28(6), 1–7.
- Ewels, P., Magnusson, M., Lundin, S., and Källér, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.
- Ghraichy, M., Galson, J. D., Kelly, D. F., and Trück, J. (2018). B-cell receptor repertoire sequencing in patients with primary immunodeficiency: a review. *Immunology*, 153(2), 145–160.
- Guangchuan, Y. (2020). *ggplotify: plot everything in ggplot2* (0.0.5). School of Basic Medical Sciences, Southern Medical University. <https://cran.r-project.org/web/packages/ggplotify/index.html>
- Hoffman, W., Lakkis, F. G., and Chalasani, G. (2016). B cells, antibodies, and more. *Clinical Journal of the American Society of Nephrology*, 11(1), 137–154.
- Honjo, T., Kinoshita, K., and Muramatsu, M. (2002). Molecular mechanism of class switch recombination: Linkage with somatic hypermutation. *Annual Review of Immunology*, 20, 165–196.
- ImmunoMind Team. (2019). *immunarch: An R Package for Painless Bioinformatics Analysis of T-Cell and B-Cell Immune Repertoires* (0.6.5). Zenodo.
- Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F. W. (2006). Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annual Review of Immunology*, 24(D), 541–570.
- Kassambara, A. (2020). *ggpubr: “ggplot2” Based Publication Ready Plots*. <https://rpkgs.datanovia.com/ggpubr/>

- Khavrutskii, I. V., Chaudhury, S., Stronsky, S. M., Lee, D. W., Benko, J. G., Wallqvist, A., Bavari, S., and Cooper, C. L. (2017). Quantitative analysis of repertoire-scale immunoglobulin properties in vaccine-induced B-cell responses. *Frontiers in Immunology*, 8(AUG).
- Kolde, R. (2019). *pheatmap: Pretty Heatmaps*. <https://cran.r-project.org/web/packages/pheatmap/index.html>
- Kurosaki, T. (2010). B-lymphocyte biology. *Immunological Reviews*, 237, 5–9.
- Kurosaki, T., Shinohara, H., and Baba, Y. (2010). B cell signaling and fate decision. *Annual Review of Immunology*, 28, 21–55.
- Lewis, S. M. (1994). The mechanism of V(D)J joining: Lessons from molecular, immunological, and comparative analyses. *Advances in Immunology*, 56(D), 27–150.
- Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry*, 79(D), 181–211.
- Little, A. J., Matthews, A., Oettinger, M., Roth, D. B., and Schatz, D. G. (2015). The Mechanism of V(D)J Recombination. In *Molecular Biology of B Cells: Second Edition* (Second Edi, Issue D, pp. 13–34). Elsevier Ltd.
- Marshall, J. S., Warrington, R., Watson, W., and Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy, Asthma and Clinical Immunology*, 14(s2), 1–10.
- Medina, A., Jiménez, C., Sarasquete, M. E., ... García-Sanz, R. (2020). Molecular profiling of immunoglobulin heavy-chain gene rearrangements unveils new potential prognostic markers for multiple myeloma patients. *Blood Cancer Journal*, 10(2).
- Mishra, A. K., and Mariuzza, R. A. (2018). Insights into the structural basis of antibody affinity maturation from next-generation sequencing. *Frontiers in Immunology*, 9(FEB).
- Mkaddem, S. Ben, Benhamou, M., and Monteiro, R. C. (2019). Understanding Fc receptor involvement in inflammatory diseases: From mechanisms to new therapeutic tools. *Frontiers in Immunology*, 10(APR).
- Oettinger, M. A., Schatz, D. G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, Adjacent Genes That Synergistically Activate V(D)J Recombination. *Science*, 248(4962), 1517–1523.
- Oka, C., and Kawaichi, M. (1995). V(D)J recombination of immunoglobulin genes. *Advances*

- in *Biophysics*, 31(C), 163–180.
- Parkin, J., and Cohen, B. (2010). An Overview of the Immune System. *Veterinary Immunology*, 357, 9–18.
- Pieper, K., Grimbacher, B., and Eibel, H. (2013). B-cell biology and development. *Journal of Allergy and Clinical Immunology*, 131(4), 959–971.
- R Core Team. (2020). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. <https://www.r-project.org/>
- Rawlings, D. J., Metzler, G., Wray-Dutra, M., and Jackson, S. W. (2017). Altered B cell signalling in autoimmunity. *Nature Reviews Immunology*, 17(7), 421–436.
- Rosenfeld, A. M., Meng, W., Chen, D. Y., Zhang, B., Granot, T., Farber, D. L., Hershberg, U., and Prak, E. T. L. (2018). Computational evaluation of B-cell clone sizes in bulk populations. *Frontiers in Immunology*, 9(JUN), 1–18.
- Roskin, K. M., Jackson, K. J. L., Lee, J. Y., ... Boyd, S. D. (2020). Aberrant B cell repertoire selection associated with HIV neutralizing antibody breadth. *Nature Immunology*, 21(2), 199–209.
- Rovin, B. H., and Erickson, L. D. (2020). Variability in the B cell–receptor repertoire across immune-mediated diseases. *Kidney International*, 98(3), 536–538.
- Schatz, D. G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews Immunology*, 11(4), 251–263.
- Schatz, D. G., Oettinger, M. A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. *Cell*, 59(6), 1035–1048.
- Schroeder, H. W., and Cavacini, L. (2010). Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(2 SUPPL. 2), S41–S52.
- Shugay, M., Bolotin, D. A., Chudakov, D. M., and Popov, A. (2019). *MiNNN 10.1 documentation*. Milaboratory.Com. <https://minnn.readthedocs.io/index.html>
- Smakaj, E., Babrak, L., Ohlin, M., ... Miho, E. (2020). Benchmarking immunoinformatic tools for the analysis of antibody repertoire sequences. *Bioinformatics (Oxford, England)*, 36(6), 1731–1739.
- Soto, C., Bombardi, R. G., Branchizio, A., Kose, N., Matta, P., Sevy, A. M., Sinkovits, R. S., Gilchuk, P., Finn, J. A., and Crowe, J. E. (2019). High frequency of shared clonotypes in human B cell receptor repertoires. *Nature*, 566(7744), 398–402.

- Stanfield, R. L., and Wilson, I. A. (2015). Antibody structure. *Antibodies for Infectious Diseases*, 49–62.
- Team, Rs. (2020). *RStudio: Integrated Development for R*. RStudio, PBC.
<http://www.rstudio.com>
- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, 302(5909), 575–581.
- Townsend, C. L., Laffy, J. M. J., Wu, Y. C. B., O’Hare, J. S., Martin, V., Kipling, D., Fraternali, F., and Dunn-Walters, D. K. (2016). Significant differences in physicochemical properties of human immunoglobulin kappa and lambda CDR3 regions. *Frontiers in Immunology*, 7(SEP), 1–12.
- Turchaninova, M. A., Davydov, A., Britanova, O. V., ... Chudakov, D. M. (2016). High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nature Protocols*, 11(9), 1599–1616.
- Uibo, R., Kisand, K., Peterson, P., and Reimand, K. (2015). *Immunoloogia. Õpik kõrgkoolidele*. 125–126, 132–133. Tartu Ülikooli Kirjastus, Tartu.
- Wang, L. X., Tong, X., Li, C., Giddens, J. P., and Li, T. (2019). Glycoengineering of antibodies for modulating functions. *Annual Review of Biochemistry*, 88, 433–459.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (3.3.3). Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Wickham, H., François, R., Henry, L., and Müller, K. (2021). *dplyr: A Grammar of Data Manipulation* (1.0.6). Comprehensive R Archive Network (CRAN). <https://cran.r-project.org/package=dplyr>
- Wilke, C. O. (2020). *cowplot - Streamlined Plot Theme and Plot Annotations for ggplot2*. <https://wilkelab.org/cowplot/index.html>
- Wu, N. C., Yuan, M., Liu, H., ... Wilson, I. A. (2020). An alternative binding mode of IGHV3-53 antibodies to the SARS-CoV-2 receptor binding domain. *BioRxiv: The Preprint Server for Biology*.
- Yaari, G., and Kleinstein, S. H. (2015). Practical guidelines for B-cell receptor repertoire sequencing analysis. *Genome Medicine*, 7(1), 1–14.
- Zheng, Z., Liebers, M., Zhelyazkova, B., ... Le, L. P. (2014). Anchored multiplex PCR for targeted next-generation sequencing. *Nature Medicine*, 20(12), 1479–1484.

LISAD

Lisa 1. Chudakovi protokollis kasutatud praimerite järjestused.

Praimer	Kasutus	Järjestus
Matriitsahela cDNA süntees		
SmartNNNa	5' – <i>template-switch</i> adapter. U = dU	AAGCAGUGGTAUCAACGCAGAG UNNNNUNNNNUNNNNUCTT(rG) ₄
Inimese IGH cDNA sünteesi praimerite segu		
hIGG_r1	Praimer cDNA sünteesiks, inimese IgG raske ahela mRNA	GAAGTAGTCCTTGACCAGGCA
hIGM_r1	Praimer cDNA sünteesiks, inimese IgM raske ahela mRNA	GTGATGGAGTCGGGAAGGAAG
hIGA_r1	Praimer cDNA sünteesiks, inimese IgA raske ahela mRNA	GCGACGACCACGTTCCCATCT
hIGD_r1	Praimer cDNA sünteesiks, inimese IgD raske ahela mRNA	GGACCACAGGGCTGTTATC
hIGE_r1	Praimer cDNA sünteesiks, inimese IgE raske ahela mRNA	AGTCACGGAGGTGGCATTG
Esimene PCR amplifikatsioon		
Rep_M1SS	<i>Step-out</i> praimer 1, seondub <i>template switch</i> adapterile	AAGCAGTGGTATCAACGCA
Inimese IGH reverse praimerite segu		
hIGGE_r2	<i>Nested</i> praimer Z adapteriga, inimese IgG/IgE raske ahela cDNA	ATTGGGCAGCCCTGATTARGGGG AAGACSGATG
hIGA_r2	<i>Nested</i> praimer Z adapteriga, inimese IgA raske ahela cDNA	ATTGGGCAGCCCTGATTCAGCGG GAAGACCTTG
hIGM_r2	<i>Nested</i> praimer Z adapteriga, inimese IgM raske ahela cDNA	ATTGGGCAGCCCTGATTAGGGG GAAAAGGGTTG
hIGD_r2	<i>Nested</i> praimer Z adapteriga, inimese IgD raske ahela cDNA	ATTGGGCAGCCCTGATTATATGA TGGGGAACAC
Teine PCR amplifikatsioon		
FwdOH_M1S	<i>Nested</i> praimer, seondub M1SS-le	TCGTCGGCAGCGTCAGATGTGTATAAAGAG ACAG NNNNN CAGTGGTATCAACGCAGAG
RevOH_Z2	<i>Step-out</i> praimer	GTCTCGTGGGCTCGGAGATGTGTATAAAGAG ACAG ATTGGGCAGCCCTGATT

Lisa 2. Sekveneermise indeksid

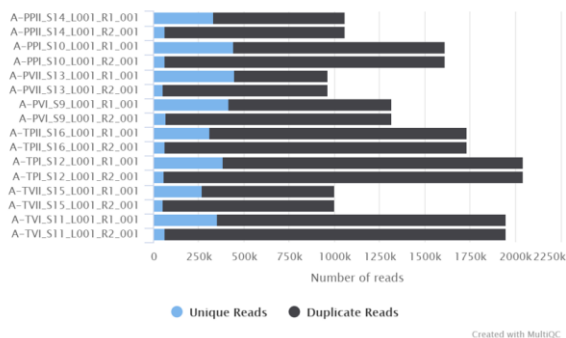
Proovi nimi	Indeks 1	Järjestus	Indeks 2	Järjestus
C_PVI	S503	AGAGGATA	N710	CGAGGCTG
C_PPI	S506	TATGCAGT	N710	CGAGGCTG
C_TVI	S508	AGGCTTAG	N710	CGAGGCTG
C_TPI	S510	ATTAGACG	N710	CGAGGCTG
C_PVII	S503	AGAGGATA	N711	AAGAGGCA
C_PPII	S506	TATGCAGT	N711	AAGAGGCA
C_TVII	S508	AGGCTTAG	N711	AAGAGGCA
C_TPII	S510	ATTAGACG	N711	AAGAGGCA
A_PVI	1	TAAGGCGA	A25	AGAGACCT
A_PPI	2	CGTACTAG	A26	TGTGCATG
A_TVI	3	AGGCAGAA	A27	ATAAGCGG
A_TPI	4	TCCTGAGC	A28	CATTCCAC
A_PVII	5	GGACTCCT	A29	TCCATGTC
A_PPII	6	TAGGCATG	A30	CATGTGAC
A_TVI	7	CTCTCTAC	A31	AGGAGATC
A_TPII	8	CAGAGAGG	A32	GTCAGTTG

Lisa 3. Proovide tähistused

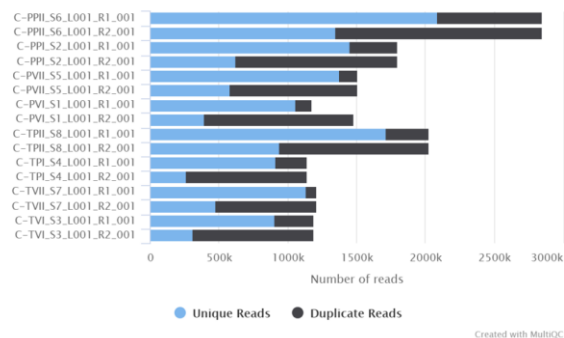
Proov	Protokoll	Indiviid	Replikaat	Ajapunkt
A-PPI	Archer	P	parem	I
A-PPII	Archer	P	parem	II
A-PVI	Archer	P	vasak	I
A-PVII	Archer	P	vasak	II
A-TPI	Archer	T	parem	I
A-TPII	Archer	T	parem	II
A-TVI	Archer	T	vasak	I
A-TVII	Archer	T	vasak	II
C-PPI	Chudakov	P	parem	I
C-PPII	Chudakov	P	parem	II
C-PVI	Chudakov	P	vasak	I
C-PVII	Chudakov	P	vasak	II
C-TPI	Chudakov	T	parem	I
C-TPII	Chudakov	T	parem	II
C-TVI	Chudakov	T	vasak	I
C-TVII	Chudakov	T	vasak	II

Lisa 4. Duplikaatlugemite osakaal Archeri ja Chudakovi proovides. Joonistel on toodud unikaalsete- ja duplikaatlugemite osakaalud FastQC kvaliteedikontrolli läbinud failides.

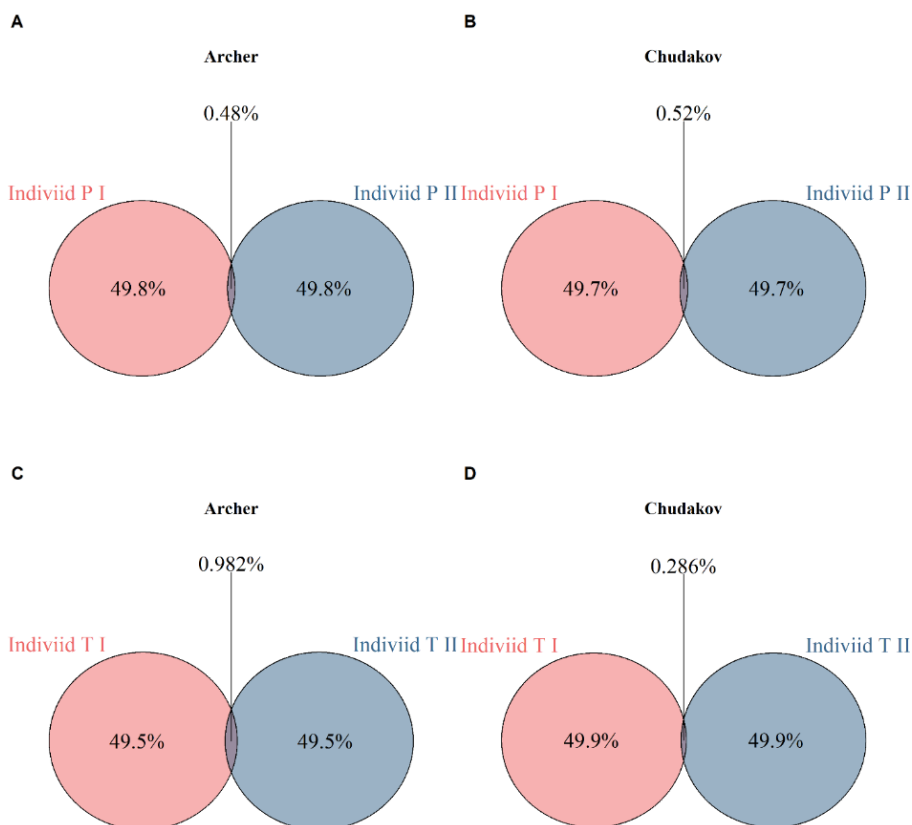
Archeri kiti proovide FastQC programmi unikaalsete- ja duplikaatlugemite jaotused



Chudakovi protokollide proovide FastQC programmi unikaalsete- ja duplikaatlugemite jaotused



Lisa 5. A ja B: indiviid P CDR3 järjestuste ülekate esimese ja teise ajapunkti proovide vahel vastavalt Archeri ja Chudakovi protokollide kaupa. C ja D: indiviid T CDR3 järjestuste ülekate esimese ja teise ajapunkti proovide vahel vastavalt Archeri ja Chudakovi protokollide kaupa. Joonistel on toodud proovide CDR3 aminohappeliste järjestuste ühisosa protsent.



LIHTLITSENTS

Lihlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Liisa Pomerants,

annan Tartu Ülikoolile tasuta loa (lihlitsentsi) minu loodud teose „B-raku retseptorite repertuaari uurimine kahel meetodil“,

mille juhendajad on Liina Tserel, Ahto Salumets ja Tõnis Org,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

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, kuni autoriõiguse kehtivuse lõppemiseni.

Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

Kinnitan, et lihlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Liisa Pomerants
31.05.2021