

TARTU ÜLIKOOL  
LOODUS- JA TÄPPISTEADUSTE VALDKOND  
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT  
EVOLUTSIOONILISE BIOLOOGIA ÕPPETOOL

Laura Laks

**Inimese Y-kromosoomi R1a alamhaplogruppide jaotus Põhja-Euroopas ja Eestis**

Bakalaureusetöö

Geenitehnoloogia (12 EAP)

Juhendajad:

PhD Siiri Rootsi

PhD Kristiina Tambets

TARTU 2019

# Infoleht

## **Inimese Y-kromosoomi R1a alamhaplogruppide jaotus Põhja-Euroopas ja Eestis**

Inimese Y-kromosoomi haplogrupid on fülogeneetilisel lähedased Y-kromosoomide klaadid, mida seovad omavahel ühised unikaalsed mutatsioonid, tavaliselt ühenukleotiidsed polümorfismid (SNPs) või indelid. Haplogrupid on globaalse Y-kromosoomi fülogeneesipuu põhiharudeks ning need on olulised nii populatsioonigeneetikas kui ka kriminalistikas.

Käesoleva töö eesmärgiks on analüüsida Y-kromosoomi hg R1a jaotust Eestis ja Põhja-Euroopas Y-täisjärjestuste põhjal koostatud fülogeneesipuude abil.

Märksõnad: Y-kromosoom, haplogrupp R1a, haplogrupi R1a alamklaadid, populatsioonigeneetika  
CERCS kood: B220 Geneetika, tsütogeneetika

## **Human Y chromosome hg R1a Subclade Distribution in Northern Europe and Estonia**

Human Y chromosomal haplogroups are phylogenetically closely related clades of Y chromosomes that share unique mutations, usually single nucleotide polymorphisms (SNPs) or indels. Haplogroups are part of the Y chromosomal global phylogenetic tree and important both in population genetics as well as in forensic studies.

The aim of this study was to analyze the phylogeographic spread of Y chromosomal haplogroup R1a in Estonia and in Northern Europe using the phylogenetic trees constructed from the data of fully sequenced Y chromosomes.

Keywords: Y chromosome, haplogroup R1a, haplogroup R1a Subclades, population genetics

CERCS code: B220 Genetics, cytogenetics

## Sisukord

Infoleht .....	2
Kasutatud lühendid .....	5
Sissejuhatus .....	6
1. KIRJANDUSE ÜLEVAADE.....	7
1.1. Inimese Y-kromosoom .....	7
1.2. Y-kromosoomi uurimise uuemad meetodid populatsioonigeneetilistes uuringutes .....	9
1.3. Inimese Y-kromosoomi fülogeneesipuu .....	11
1.4. Euroopa asustamine anatoomiliselt kaasaegse inimese poolt ja Y-kromosoomi haplogruppide levik Euroopas .....	12
1.4.1. Geneetilised eelaskomponendid tänaste eurooplaste genoomides .....	12
1.4.2. Y-kromosoomi haplogruppide levik Euroopas .....	14
1.5. Haplogrupi R üldine iseloomustus ja fülogeograafiline levik .....	15
1.5.1. Hg R1 esinemine Eestis ja Põhja-Euroopas .....	19
2. EKSPERIMENTAALOSA .....	20
2.1 Töö eesmärgid.....	20
2.2.Materjal ja meetodika .....	21
2.2.1. Kasutatud DNA proovid ja Y-kromosoomide täisjärjestused.....	21
2.3. Genotüpiseerimine .....	21
2.3.1. Praimerite disainimine.....	21
2.3.2. DNA amplifitseerimine .....	21
2.3.2.1. Polümeraasi ahelreaktsioon (PCR).....	21
2.3.2.2. Geelelektroforees.....	22
2.3.2.3 DNA produkti „puhastamine“ .....	22
2.3.4. Sekvenerimine .....	23
2.3.4.1. DNA sadestamine .....	24
2.3.5. Restriktsioonianalüüs .....	24

2.3.6. Andmeanalüüs .....	25
2.3.6.1. Sekveneeritud järjestuse analüüs ChromasPros .....	25
2.3.7. Genotüüpiseerimise tulemused .....	26
2.3.7.1. Analüüsitud indiviidide Y-kromosoomide varieeruvuse võrdlus EBK andmebaasis olemasolevate eestlaste andmetega .....	27
2.4. Y- kromosoomi täisjärjestuseanalüüs R1a näitel .....	28
2.4.1. Y-kromosoomi täisjärjestuste analüüsi tulemused .....	29
Kokkuvõte .....	37
Summary.....	38
Tänuõnad.....	39
Kasutatud kirjandus .....	40
Kasutatud veebiaadressid .....	48
Lisad .....	49
Lisa 1. Eestlaste valimi (N=49) tüüpiseerimise tulemused.....	49
Lisa 2. Töös kasutatud Y-kromosoomi täisjärjestused (n=556), mida kasutati Y-kromosoom R1a alamharude jaotuse väljaselgitamiseks. ....	50
Lisa 3. Y-kromosoomi täisjärjestuste analüüsi põhjal leitud efektiivse populatsiooni suuruste dünaamika. ....	51
LIHTLITSENTS.....	52
Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks .....	52

## Kasutatud lühendid

AMH – Anatoomiliselt kaasaegne inimene (*anatomically modern human*)

ap - aluspaari

at – aastat tagasi

AZFc – azoospermia faktori regioon c (*azoospermia factor region c*)

BEAST – *Bayesian Evolutionary Analysis Sampling Trees*

EBK – Eesti Biokeskus

EDTA – etüleendiamiintetra-äädikhape (*ethylenediaminetetraacetic acid*)

EtOH – etanool

FastAP™ – temperatuuritundlik aluseline fosfataas (*Thermosensitive Alkaline Phosphatase*)

hg – haplogrupp

ISOGG – Rahvusvaheline Geneetilise Genealoogia Ühing (*International Society of Genetic Genealogy*)

Mb – miljon aluspaari (*mega basepair, Mbp*)

MSY – Y kromosoomi meessoospetsiifiline regioon (*male-specific region of the Y-chromosome*)

mtDNA – mitokondriaalne DNA

Ne – efektiivse populatsiooni suurus

NRY - mitterekombineeruv piirkond (*non-recombining region of the Y-chromosome*)

PAR - pseudoautosomaalne piirkond (*pseudoautosomal region*)

PCR – polümeraasi ahelreaktsioon (*polymerase chain reaction*)

RAxML – *Randomized Accelerated Maximum Likelihood*

SNP - ühenukleotiidne polümorfism (*single nucleotide polymorphism*)

STR – lühike tandemne kordusjärjestus (*short tandem repeat*)

YCC – Y-kromosoomi konsortsium (*The Y Chromosome Consortium*)

## Sissejuhatus

Y-kromosoomi haplogrupid on defineeritud Y-kromosoomi mitterekombineerivas alas (*non-recombining region of the Y-chromosome* – NRY) toimunud ühenukleotiidsete polümorfismide (SNP-id) või insertioonide ja deletsioonide ehk indelidetega. Need mutatsioonid päranduvad isalt-pojale, kuna NRY ei rekombineeru X-kromosoomiga. Tänu sellele on võimalik kasutada Y-kromosoomi populatsioonigeneetilistes uurimustes ning saada infot isaliinide leviku ja mitmekesisuse kohta.

Y kromosoomi haplogrupi (hg) R1 alamklaadide uurimine pakub huvi seetõttu, et hg R1a ja R1b levikus on Euroopas täheldatud vastassuunalist geograafilist gradienti – hg R1a alamklaadid esinevad põhiliselt Ida-Euroopas ning hg R1b esineb põhiliselt Lääne-Euroopas. Hg R1a ja R1b jõudmine Euroopasse on seotud aukhauakultuuri kandjate rändega kõigepealt Ida-Euroopasse ning sealt edasi Lääne-Euroopasse. See tõi endaga kaasa nõndanimetatud stepikomponendi kinnistumise Euroopa populatsioonide genofondis.

Antud töö eesmärgiks on suurendada Tartu Ülikooli Eesti Biokeskuse biopanga genotüpiseeritud proovide valimi suurust ning uurida inimese Y-kromosoomi R1a alamhaplogruppe ja nende jaotust Põhja-Euroopas ja Eestis, kasutades selleks Y-kromosoomi täisjärjestusi, mille põhjal põhjal konstrueeriti fülogeneesipuud, arvutati alamklaadide koalestsentsiajad ja hinnati efektiivse populatsiooni suuruse dünaamikat.

Töö teoreetilises osas antakse ülevaade inimese Y-kromosoomist, Y-kromosoomi uurimismeetodites ning klassifitseerimisel kasutatavatest nomenklatuuridest. Käsitletakse ka Euroopa asustamislugu anatoomiliselt kaasaegse inimese poolt ning sellega seotult Y-kromosoomi haplogruppide levikut Euroopas, põhitähelepanu on hg R kuuluvatel isaliinidel.

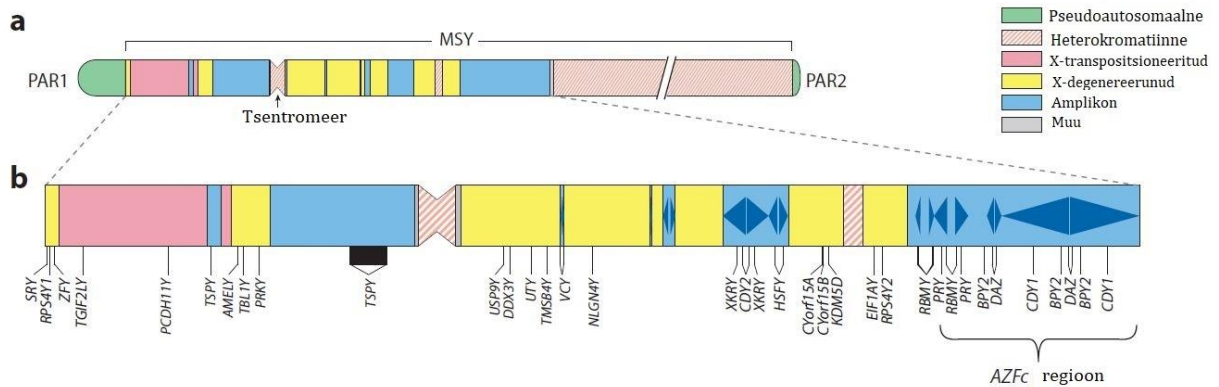
# 1. KIRJANDUSE ÜLEVAADE

## 1.1. Inimese Y-kromosoom

Inimese 23. kromosoomide paar ehk sugukromosoomide paar (XX naistel ja XY meestel) määrab inimese geneetilise soo. Meessoos puhul päritakse isalt üks Y-kromosoom ning emalt üks X-kromosoom. Erandiks on Turneri sündroom (45, X kariotüüp), mida iseloomustab Y-kromosoomi puudumine ning Y-kromosoomi geenide inaktivatsioon või kadumine, mis põhjustab XY naise või hermafrodiidi fenotüübi kujunemist ning viljatust (Sun jt 1999; Jobling and Tyler-Smith 2003; Krausz jt 2008; Hughes and Rozen 2012). Y-kromosoom on oma unikaalse pärandumismustri tõttu oluline informatsiooniallikas nii inimese isaliini evolutsiooniuuringutes, meditsiinis kui ka kriminaalstatistikas (Jobling jt 1997; Jobling and Tyler-Smith 2003 ja selles viidatud artiklid).

Y-kromosoomi suurus on ~57 miljonit aluspaari (Mb; Genoomi Referentsi Konsortsium *Genome Reference Consortium*, avaldatud märts 2019, referentsjärjestus GRCh38.p13, <https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh38.p13>), mis moodustab umbes 2% haploidsest genoomist ning jaguneb kaheks erinevaks osaks: pseudoautosomaalseks (*pseudoautosomal region* – PAR) ja meessoospetsiifiliseks regiooniks (*male-specific region of the Y-chromosome* – MSY) ehk mitterekombineeruvaks alaks (*non-recombining region of the Y-chromosome* – NRY) (joonis 1) (Skaletsky jt 2003; Hughes and Rozen 2012). MSY koosneb heterokromatiinsest, X-transpositsioneeritud, X-degenereerunud ja amplikoonses piirkonnast (Skaletsky jt 2003). Amplikoonses piirkonnas esinevad pikad intrakromosomaalsed kordusjärjestused, mis teevad Y-kromosoomi täieliku sekveneerimise keeruliseks (Hughes and Rozen 2012). Samas on piirkond geenitihe ja spetsialiseerunud, sisaldades testises ekspresseeruvaid gene, ning on vastuvõtlik ulatuslikele deletsioonidele, inversioonidele ja insertioonidele, mis põhjustavad tõsiseid fenotüübilisi muutusi (Skaletsky jt 2003; Hughes and Rozen 2012; Hughes jt 2013). MSY on suurim mitterekombineeruv ala inimese genoomis (Underhill and Kivisild 2007), see on oluline testiste arengus ja spermatogeneesis. Meioosi käigus paardub Y-kromosoom X-kromosoomiga pseudoautosomaalsetes piirkondades, mis paiknevad kromosoomi otstes, kuid PARid moodustavad vähem kui 3 Mb Y-kromosoomi järjestusest. Ainult inimese Y-kromosoomis esineb kaks PARi, teistel liikidel on neid Y-kromosoomis ainult üks (Graves jt 1998; Hughes and Rozen 2012).

Y-kromosoom on lisaks emaliinis päranduvale mitokondrile üks põhilisi populatsioonigeneetika uurimisobjekte. Neil on võrreldes ülejäänud genoomiga suur evolutsioonikiirus (Brown jt 1979; Hallast jt 2015; Karmin jt 2015). Genoomi haploidsete osade kiirem evolutsioon on tingitud kõrge mutatsioonikiirusest – mitokondri DNA (mtDNA) puhul akumuleerub keskmiselt  $1,70 \times 10^{-8}$  ja isaliinis  $0,74 \times 10^{-9}$  mutatsiooni aluspaari (ap) kohta aastas (Ingman jt 2000; Karmin jt 2015) ning neli korda väiksem efektiivse populatsiooni suurus ( $N_e$ ) võrreldes autosoomidega (Thomson jt 2000; Jobling and Tyler-Smith 2003). Neljakordne vahe tuleneb mtDNA ja Y-kromosoomi edasikandumise eripärast – ülekandumisprotsessis järglastele on kaasatud ainult pool populatsiooni (vastavalt naised ja mehed) ning rakkudes on mtDNA ja Y-kromosoom haploidsetel kujul, s.t. nende variatsioonide arv on kaks korda väiksem kui diploidsetel autosoomidel. Mitterekombineeruv Y-kromosoomis (ja mtDNA-s) tekivad muutused ajas kuhjuvate mutatsioonide tõttu.



**Joonis 1.** Inimese Y-kromosoomi skemaatiline ehitus; a) Y-kromosoomi põhilised osad; b) meessoospetsiifiline (MSY) eukromaatiline region. Pseudoautosomaalsed piirkonnad on tähistatud vastavalt PAR1 (pikk ala) ja PAR2-ga (lühike ala). Tumesiniste kolmnurkadega on tähistatud palindroomsed järjestused. Eukromaatilise regiooni all olevad jooned tähistavad valke kodeerivate geenide suhtelist asukohta Y-kromosoomil. AZFc – azoospermia faktori region c. Kohandatud (Hughes and Rozen 2012) järgi.

## 1.2. Y-kromosoomi uurimise uuemad meetodid populatsioonigeneetilistes uuringutes

Y-kromosoomi uurimine on tähtis nii meditsiinis, kohtuekspertiisis isaduse testimisel ja kuritegude uurimisel (Jobling jt 1997; Batini and Jobling 2017), antropoloogias kui ka evolutsiooniliste protsesside ja inimpopulatsioonide vaheliste geneetiliste seoste väljaselgitamisel (Jobling and Tyler-Smith 2003). Y-kromosoomi mutatsioonikiiruste väljaselgitamisel on evolutsiooniliste protsesside ja geneetiliste haiguste mõistmisel oluline roll. Võrreldes teiste kromosoomidega, mis alluvad rekombinatsioonile, esinevad Y-kromosoomis ainult nn. intra-alleelsed mutatsioonid (Jobling and Tyler-Smith 2003). Meditsiiniliste uuringutega on näidatud, et Y-kromosoomi MSY regioonis toimunud mutatsioonid ja struktuursed muutused, eriti deletsioonid, on seotud mitmete haiguslike fenotüüpidega, nt. viljatuse ning mõnede vähivormidega, mis võivad olla nii testikulaarsed kui ka somaatilised (Hughes and Rozen 2012; Kido jt 2015).

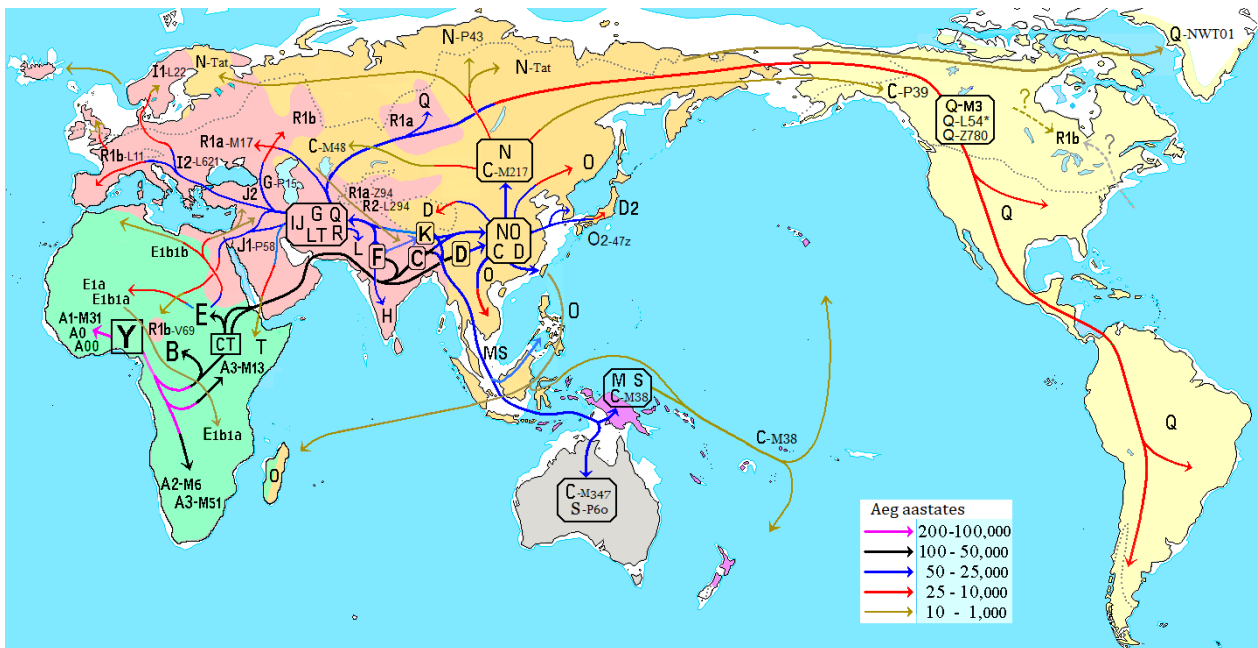
Tänu Y-kromosoomi peaaegu täielikule sekveneerimisele on evolutsiooniliste protsesside uurimiseks leitud uusi lähenemisviise (Poznik jt 2013; Batini jt 2015; Hallast jt 2015; Karmin jt 2015), mis võimaldavad väga täpselt isaliinide fülogeneetilisi suhteid rekonstrueerida, nende võimalikku päritolu ning geograafilist jaotust ja migratsioonisuundi uurida (joonis 2). Y-kromosoomide (ja mtDNA) variantide eristamiseks jagatakse need haplogruppideks (hg) - teatud kindlate mutatsioonide kombinatsiooniga defineeritud monofüleetilisteks klaadideks (vt. 1.3), mille tekkeloo kohta saab järeldusi teha isaliinide ajalise-ruumilise fülogeograafilise analüüsi alusel (YCC 2002; Jobling and Tyler-Smith 2003).

Kuigi Y-kromosoom ning mtDNA võimaldavad meeste ja naistega seotuid rändeid ja päritolu eraldi vaadelda, on neil ka omad puudused: mtDNA ja Y-kromosoom moodustavad vaid väga väikese osa kogu genoomist (vastavalt 0,0005% ja 2%) ning neid mõjutab geneetiline triiv ja soospetsiifilised käitumismustrid (Jobling and Tyler-Smith 2017). Soospetsiifilisteks erinevusteks on näiteks meeste ja naiste erinev panus laste sünni ning abiellumistraditsioonidega seotud eripärad, mis soodustavad naiste rännet (Jobling and Tyler-Smith 2017). Eeldatakse, et meeste  $N_e$  on madalam kui naistel (Brown jt 2009) ning seega ka rohkem mõjutatud geneetilisest triivist. Meeste ebavõrdsus sigimisel on suurem populatsioonides, kus väiksel grupil meestel on ühiskonnas kõrge staatus (Heyer jt 2012; Jobling and Tyler-Smith 2017). Ebatavaliselt kõrge sagedusega esinevaid Y-kromosoomi variante, nn. lühikeste kordusjärjestuste (STR-ide, *short tandem repeats*) klastreid, seostatakse sageli patriarhaalsete dünastiatega (Batini and Jobling 2017)

- näiteks Tšingis-khaani loodud dünastia (Zerjal jt 2003), Hiina Qing'i dünastia (Xue jt 2005) ning Iiri varase keskaja Uí Néill'i dünastiaga (Moore jt 2006).

Inimpopulatsioonidest umbes 70% on patrilokaalsed, ehk eri regioonidest pärit naise ja mehe abiellumise korral kolib uude elukohta elama naine (Burton jt 1996). See praktika suurendab Y-kromosoomide hg-de lokaalset diferentseerumist. MtDNA hg-de puhul on täheldud vastupidist toimet (Oota jt 2001). Matrilokaalsuse puhul – mis esineb umbes 20% populatsioonidest – kolib mees uude kohta ning suureneb mtDNA lokaalne diferentseeritus (Burton jt 1996; Oota jt 2001; Heyer jt 2012).

Tänapäeval on evolutsiooniliste ja populatsioonigeneetiliste protsesside uurimisel ja inimese rändeteede ja demograafilise ajaloo rekonstrueerimisel mtDNA ja Y-kromosoomi genotüüpiseerimise kõrval üha suurem roll Y täisjärjestuste (Hallast jt 2015; Karmin jt 2015), samuti täisgenoomide sekveneerimistulemuste analüüsil (Lachance jt 2012; Pagani jt 2016; Sherman jt 2019) ning nn. „vana“ DNA andmestiku kaasamisel (Haak jt 2005; Willerslev and Cooper 2005; Sarkissian jt 2013; Lazaridis jt 2014; Haak jt 2015; Fu jt 2016; Saag jt 2019).

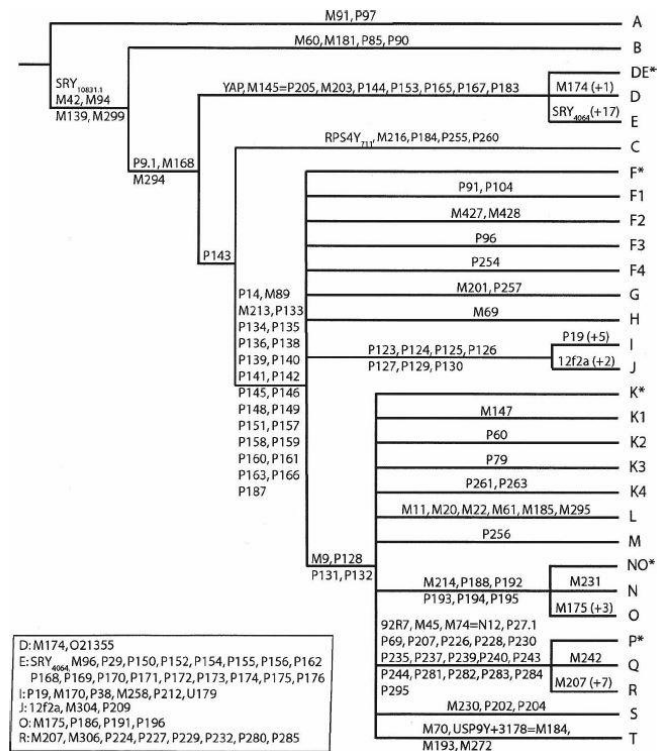


**Joonis 2.** Põhiliste Y-kromosoomi haplogrupid (hg) globaalne jaotus ning võimalikud migratsiooniteed. Kaardil on erinevate värvidega tähistatud erinevate Y-kromosoomide hg-de levilad ja nende nomenklatuuris kasutatavad tähised (allikas: [https://en.wikipedia.org/wiki/Human\\_Y-](https://en.wikipedia.org/wiki/Human_Y-)

chromosome\_DNA\_haplogroup#/media/File:Migraciones\_humanas\_en\_haplogrupos\_de\_ADN-Y.PNG).

### 1.3. Inimese Y-kromosoomi fülogeneesipuu

Y-kromosoomi populatsioonigeneetilistes uuringutes kasutatakse paralleelset mitut nomenklatuuri. Nendest tuntumad on Y-kromosoomi konsortsiumi (YCC) (YCC 2000, Jobling 2003; Karafet jt. (2008) ; van Oven jt. (2014) ), ISOGG (Rahvusvaheline Geneetilise Genealoogia Ühing (*International Society of Genetic Genealogy* <https://isogg.org/>), Yfull (<https://www.yfull.com/tree/>), ning Karmini jt. (2015) alusel koostatud nomenklatuurid, viimast kasutatakse ka siinses töös. Nomenklatuure täiendatakse pidevalt uute markerite ja fülogeneesipuu harude defineerimisega. Fülogeneetilisel lähedased Y-kromosoomide variandid jaotatakse haplogruppidesse. Haplogrupp on monofüleetiline rühm, mida seovad omavahel ühised unikaalsed mutatsioonid, tavaliselt ühenukleotiidsed polümorfismid (SNP-id) või insertioonid ja deletsioonid ehk indelid, mis on tõenäoliselt evolutsioonis tekkinud ainult ühel korral ning siis edasi kandunud kõiki seda järjestusvarianti kandvate indiviidide meessoost järglastele. Haplogrupid on globaalse Y-kromosoomi fülogeneesipuu põhiharudeks (Joonis 3). Samadel põhimõtetel on koostatud ka ühtne mtDNA nomenklatuur (van Oven and Kayser 2009).



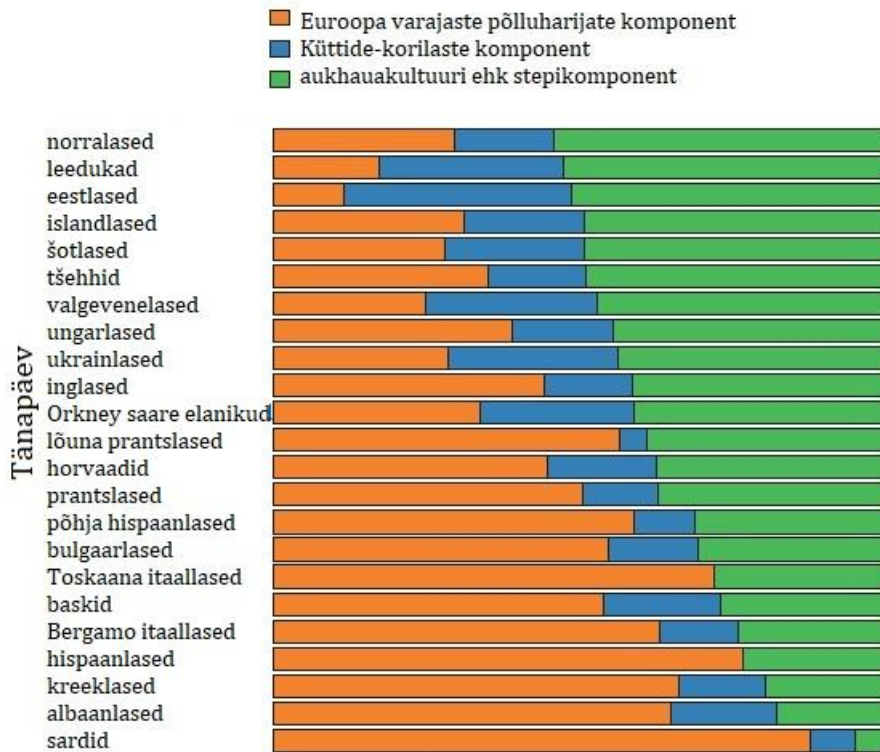
**Joonis 3.** Lihtsustatud Y-kromosoomi fülogeneesi puu. Skeemil ei ole harude pikkused proportsionaalsed mutatsioonide arvuga. Haplogruppe defineerivate markerite nimetused on välja toodud puu harudel. Haplogruppide nimetused on esitatud harude tipus. Kohandatud (Karafet jt 2008) järgi.

## 1.4. Euroopa asustamine anatoomiliselt kaasaegse inimese poolt ja Y-kromosoomi haplogruppide levik Euroopas

### 1.4.1. Geneetilised elaskomponendid tänaste eurooplaste genoomides

Arvatakse, et kaasaegsete eurooplaste esivanemad pärinevad asustuslainest, milles osalenud anatoomiliselt kaasaegsed inimesed (AMH – *anatomically modern human*) lahkusid Aafrikast umbes 50 000 – 60 000 aastat tagasi ning liikusid Lähis-Ida kaudu Euroopasse (Mellars 1992; Benazzi jt 2011). Esmased jäljed ülempaleoliitikumile omastest AMH tööriistadest on väljaspool Aafrikat leitud Levantist (Klein 2000; Rebollo jt 2011; Seguin-Orlando jt 2014). Euroopa geneetilise mitmekesisuse kujunemisel peetakse oluliseks kolme põhilist lainet, mis on jätnud oma jälje tänaste eurooplaste geenipärandisse ja mõjutanud hg-de levikut (Lazaridis jt 2014; Haak jt

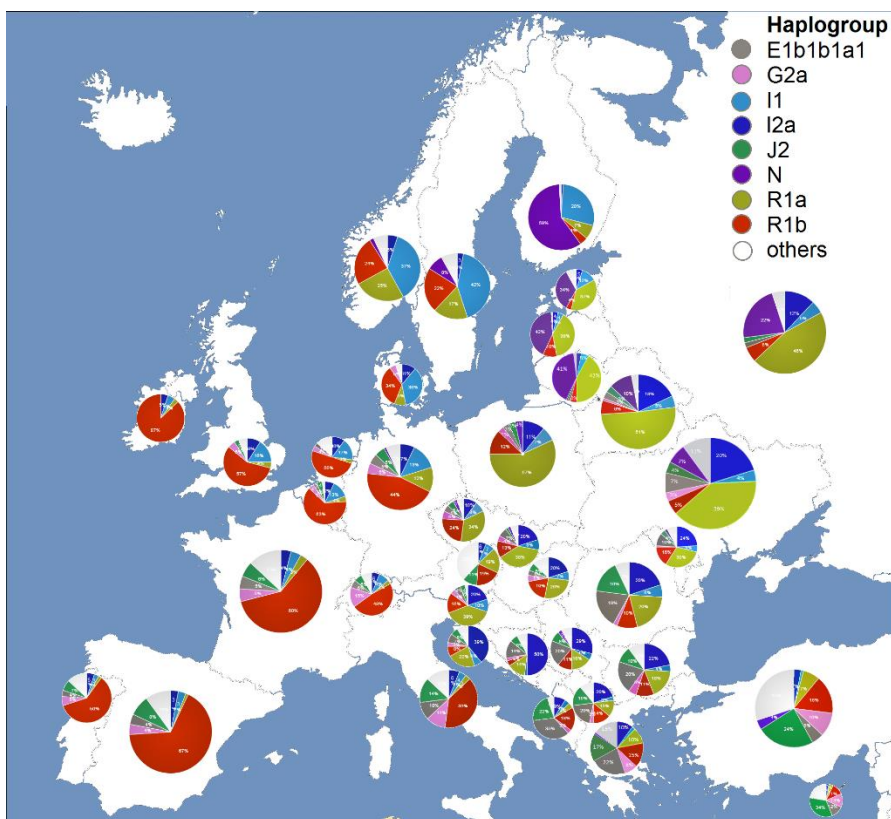
2015; Mathieson jt 2015) (Joonis 4). Varaseim asustuslaine jõudis Euroopasse vanemal kiviajal ehk paleoliitikumis, ca 45 000 at, sellest ajast pärinevad ka vanimad Euroopast leitud AMH jäänused (Benazzi jt 2011). Keskmisel kiviajal ehk mesoliitikumis, pärast viimase jääaja maksimumi, toimus Euroopa põhjapoolsemate alade järk-järguline taasasustamine lõunas asunud jääajajärgsetest refuugiumidest (Housley jt 1997). Tolleaegsed asukad, nagu ka Euroopa esmaasustajad, tegelesid küttimise ja korilusega. Nooremal kiviajal ehk neoliitikumis toimus küttide-korilaste segunemine varajaste Euroopa põlluharijatega, kes pärinesid Lähis-Idast ja tõid Euroopasse liikudes kaasa põlluharimiskultuuri, mis levis aegamööda lõunast põhjasuunas. Pronksiajal, umbes 5000 at, toimus järgmine oluline rahvastikumuutus, mida seostatakse Musta ja Kaspia merega piirnevate stepialadelt pärineva aukhauakultuuri (ingl. k. *Yamnaya culture*) kandjate rändega Ida-Euroopasse ja sealt edasi läände. See ränne tõi kaasa nn. stepikomponendi, mis meie aladele levis koos nöörikeramika kultuuri kandjatega (Saag jt 2017). Aukhauakultuuriga seostatakse ka indoeuroopa keelte levikut (Lazaridis jt 2014; Haak jt 2015). Kõikide nende rännete tulemusena saab tänapäevase eurooplase geneetiliselt „kokku panna“ kolmest põhilisest eellaskomponendist, mis vastavad esmaasukatest küttide-korilaste, nooremal kiviajal lisandunud põlluharijate ja hilisemate stepirahvaste geenipärandile (joonis 4). Hiljuti näidati, et Kirde-Euroopa uurali keeli rääkivate populatsioonide (sh. eestlaste puhul) lisandub neile põhilistele komponentidele veel nn. Siberi päritolu komponent, mille osakaal eestlaste genoomis on umbes 5% (Tambets jt 2018).



**Joonis 4.** Geneetiliste eellaskomponentide osakaalud (%) tänapäevastes Euroopa populatsioonides. Kohandatud (Haak jt 2015) järgi.

#### 1.4.2. Y-kromosoomi haplogruppide levik Euroopas

Y-kromosoomi hg-dest esineb tänapäeval Euroopas kõige sagedamini hg R1b (43,1%), mis on levinud põhiliselt Lääne- ja Loode-Euroopas, sellele järgnevad I1 (13,8%), I2 (9,0%) ja J2 (7,5%) (Batini jt 2015; joonis 5). Põhja- ja Kirde-Euroopas on väga levinud hg N (15-60%), mis Lääne-Euroopas praktiliselt puudub (Rootsi jt 2007; Ilumäe jt 2016) ning Ida-Euroopas on sagedasemaks haplogrupiks R1a (Underhill jt 2015). Y-kromosoomi hg-de levik on tihti geograafiliselt struktureeritud, antud töös käsitletav hg R1a ning tema alamharud on levinud Kesk- ja Ida-Euroopas (Underhill jt 2010; Underhill jt 2015). Seevastu hg E, G ja J on levinud põhiliselt Vahemere ümbruses (joonis 5) (Batini jt 2015). Hg R1a sagedus on suurem slaavi keelt kõnelevates populatsioonides (Balanovsky jt 2008; Underhill jt 2015).



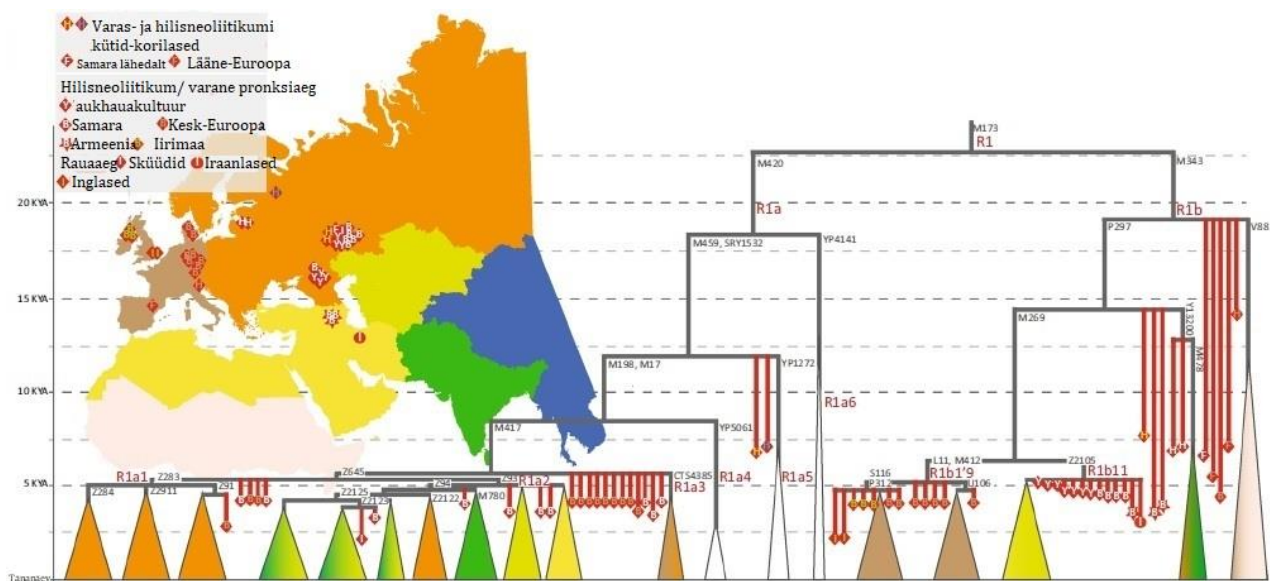
**Joonis 5.** Sektordiagrammid Y-kromosoomi haplogruppide levikust Euroopas, andmed on antud protsentides (allikas: [https://commons.wikimedia.org/wiki/File:Percentage\\_of\\_major\\_Y-DNA\\_haplogroups\\_in\\_Europe.png](https://commons.wikimedia.org/wiki/File:Percentage_of_major_Y-DNA_haplogroups_in_Europe.png)).

### 1.5. Haplogrupi R üldine iseloomustus ja fülogeograafiline levik

Hg R on välja kasvanud haplogrupist P-M45, mille kiire mitmekesisustumine toimus Kesk-Aasias (Karafet jt 2015) 48 000 aastat tagasi (Karmin jt 2015) ning millel on kaks tütarklaadi R-M207 (Underhill jt 2015) ja Q-M242 (joonis 3) (Karmin jt 2015). Hg R arvatakse olevat tekkinud umbes 29 000 aastat tagasi (Karmin jt 2015). Seda toetab ka Mal'ta-Buret' ülempaleoliitilise kultuuri kontekstist leitud poisi jäänused (edaspidi MA-1), mis leiti Kesk-Siberi lõunaosast Mal'ta küla lähedalt, ning mille vanuseks on radioaktiivse süsiniku dateeringu järgi 24,423–23,891 aastat. MA-1 Y-kromosoom kuulus tänaste Lääne-Euraasia asukate Y-kromosoomidega võrreldes veel lahknemata basaalsesse hg R harusse (Raghavan jt 2014).

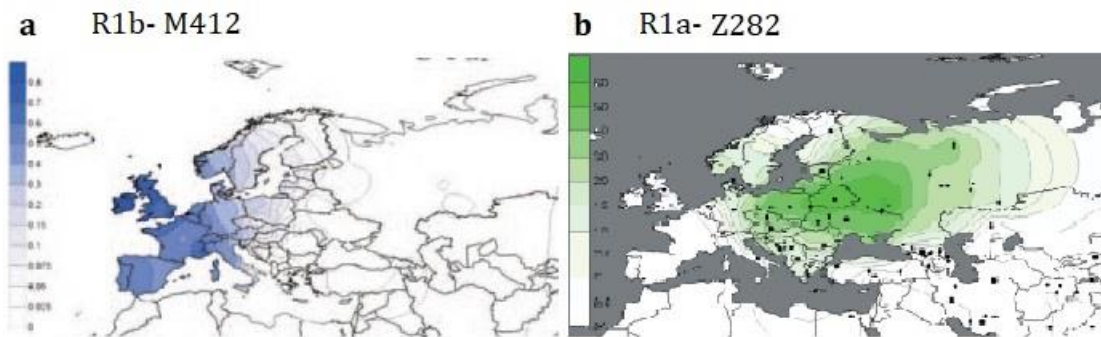
Hg R jaguneb kaheks haruks: R1-M173 ja R2-M479 (joonis 6), millest R1 klaadi kuulub suurem osa Euroopa hg R variatsioonidest ning R2 esineb Kesk- ja Lõuna-Aasias.





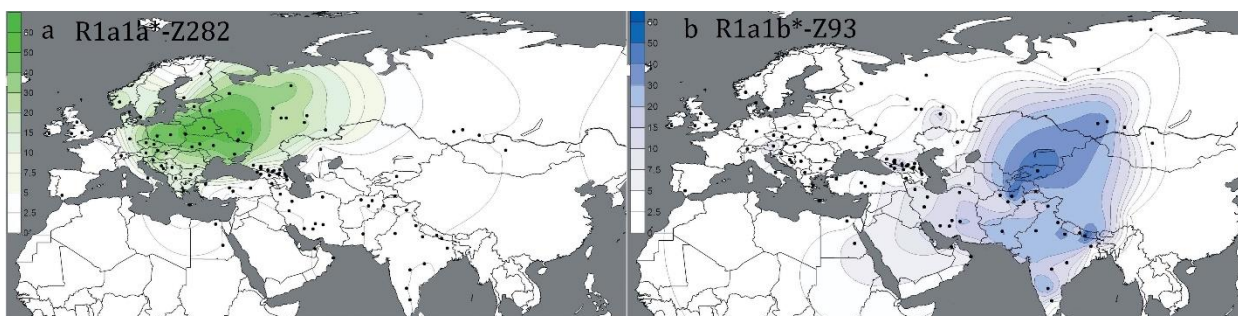
**Joonis 7.** Haplogrupid R1a ja R1b alamharude geograafiline jaotus tänapäevaste ja vana DNA leidude leviku kohaselt. Punasete sümbolitega on tähistatud vana DNA proovid. Kohandatud (Kivisild 2017) järgi.

Euroopas on selgelt näha haplogrupid R1a-M420 ja R1b-M343 vastupidine geograafiline gradient (joonis 8), mida täheldati juba esimestes üleeuroopalistes Y-kromosoomi levikut käsitlevates töodes (Rosser jt 2000; Semino jt 2000). Hg R1a on levinud põhiliselt Euroopa idaosas, seevastu R1b esineb suurema sagedusega Euroopa lääneosas (joonis 8a ja b). Euroopas kõige levinuma R1b alamklaadi R1b-M269 koalestseerumine toimus hiljuti, 5 000-7 000 aastat tagasi (Batini jt 2015; Hallast jt 2015; Karmin jt 2015; Poznik jt 2016; Kivisild 2017). Eriti haruldased hg R1a alamklaadid R1a4-YP5061, R1a5-YP1272 ja R1a6-YP4141 (Yfull nomenklatuuri järgi) võivad viidata sellele, et nende levik on olnud pikaajaliselt järjepidev. Seevastu enamus tänapäeval levinud R1a ja R1b alamharudest on alguse saanud väikesest hulgast hilisneoliitikumi/varase pronksiaja meessoost asutajatest (Kivisild 2017).



**Joonis 8.** Haplogrupi R1 põhiliste Euroopa-spetsiifiliste tütarharude: **a)** R1b-M412 ning **b)** R1a-Z282 levikumuster. Kohandatud (Myres jt 2011) ja (Underhill jt 2015) järgi.

Hg R1a tervikuna on suhteliselt laialt levinud, ulatudes Lõuna-Aasiast Kesk- ja Ida-Euroopani ja Lõuna-Siberini ning esineb umbes 10% meestel. Underhill jt. (2015) andmete järgi on hg R1a alguse saanud Lääne-Euraasiast, tõenäoliselt tänapäeva Iraani ja Türgi territooriumil, kuna R1a haruldased basaalsed harud on leitud põhiliselt sealt. Hg R1a kaks põhilist alamklaadi on R1a1a\*-Z282 (R1a1-Z282 Karmin jt 2015 nomenklatuuri järgi) ning R1a1b\*-Z93 (R1a2-Z93 Karmin jt. 2015 nomenklatuuri järgi), millel on selgelt erinev geograafiline levikumuster (joonis 9a ja b) (Underhill jt 2015). R1a1a\* (vastavalt R1a1-Z282 Karmin jt. 2015 nomenklatuuri järgi) ning tema alamharud esinevad põhiliselt Euroopas, Ida- ja Kesk-Euroopast Venemaa keskosani, seevastu R1a1b\* (vastavalt R1a2-Z93 Karmini jt. 2015 nomenklatuuri järgi) on levinud peamiselt Lähis-Idas, Lõuna-Aasias, Kesk-Aasias ja Siberis. R1a1a\*-Z282 ja R1a1b\*-Z93 liinide ekspansioon toimus 5 000 aastat tagasi (Karmin jt 2015; Kivisild 2017) (joonis 7). R1a-Z282 alamklaadidest on kõige laiema levikuga R1a-M558, mis esineb Kesk- ja Ida-Euroopas sh. Volga-Uurali populatsioonides ning on kõige sagedasem R1a-Z282 alamklaad (Underhill jt 2015). M558-le sarnase, kuid kitsama jaotusega on R1a-M458, mille esinemissagedus on kõrgeim Kesk- ja Ida-Euroopas, eriti Tšehhis, Slovakkias ning Poolas. R1a-Z284 on levikult piiratud, esinedes ainult Loode-Euroopas, kus selle sagedus on kõrgeim Norras (20%). Tänu täisjärjestuste analüüsile, mis ei olnud tehniliselt teostatav Underhill jt. (2015) uuringu teostamise ajal, on täna võimalik määrata ka R1a-Z282\* alamgrupi kuuluvate isaliinide levikusagedus. Need esinevad põhiliselt Põhja-Ukrainas, Valgevenes ja Venemaal spetsiifiliste alamklaadidena (Karmin jt. 2015).



**Joonis 9.** Haplogrupi (hg) R1a levinumate tütarharude levikukaardid: **a)** R1a1a\*-Z282; **b)** R1a1b\*-Z93 levikualad. Kohandatud (Underhill jt 2015) järgi.

### 1.5.1. Hg R1 esinemine Eestis ja Põhja-Euroopas

Eestis esineb hg R1a 33-34,9% meestel ning R1b 3,7-8,1% meestel (Tambets jt 2004; Underhill jt 2015; Myres jt 2011; Tambets jt 2018). Hg R1 kandvate meeste saabumine Eesti aladele pidi tõenäoliselt toimuma enne nöörikeramika levikut Eestis. Nelja 4 800–4 000 aastat vana Eesti esimestest põlluharijatest nöörikeramikute säilmete DNA uurimisel selgus, et kõik nad kuulusid R1a-Z645 alamklaadi, mis on ainult kauges suguluses R1a5 liiniga, millesse kuulusid 5 900–3 800 aastat vanad küttidest-korilastest kammkeramikute säilmed (Saag jt 2017). Läti aladelt on leitud mesoliitikumi küttide-korilaste jäänused, mis kuuluvad hg-sse R1b ning kõige tõenäolisemalt selle alamhaplogruppi R1b1b (Pinhasi jt 2017). Hg R1b esinemissagedus Lääne-Euroopas hakkas tõusma alles hilises neoliitikumis, Venemaal oli tema esinemissagedus juba enne seda väga kõrge, mistõttu võib oletada, et R1b levik Lääne-Euroopasse algas idast (Pinhasi jt 2017). R1b levikut Euroopasse seostatakse aukhauakultuuri komponenti kandvate indiviidide rändega Ida-Euroopasse.

Soomlastel esineb hg R1 vähe (7,28%), selle esinemissagedus on enam-vähem võrdne nii Lääne-kui ka Ida-Soomes. R1b seevastu esineb 3,73% Soome meestel ning on selgelt näha, et suurem esinemissagedus on läänes elavatel soomlastel (Lappalainen jt 2006; Lappalainen jt 2008).

Rootslastel on näha märgatav Lääne-Euroopa mõjude suurenemine. Hg R1a esineb ainult 13,4% meestel, seevastu R1b 20,9% meestel (Karlsson jt 2006; Lappalainen jt 2008). Norras on hg R1 esinemissagedus Soome, Rootsi ja Eestiga võrreldes suurem: hg R1a esineb 22-26% ja R1b 32% Norra meestel (Dupuy jt 2006; Underhill jt 2015).

## 2. EKSPERIMENTAALOSA

### 2.1 Töö eesmärgid

Käesoleval bakalaureusetöö peamised eesmärgid:

1. Suurendada Tartu Ülikooli Eesti Biokeskuse DNA biopanga genotüpiseeritud proovide hulka ja võrrelda töö autori genotüpiseeritud proovide Y-kromosoomi hg-de jaotust seni analüüsitud eestlaste üldvalimi jaotusega;
2. Analüüsida Y-kromosoomi hg R1a jaotust Eestis ja Põhja-Euroopas Y-täisjärjestuste põhjal koostatud fülogeneesipuude kaudu;

Selleks tehti töö erinevates etappides järgnevat:

1. Uuriti eestlaste valimi (N=49) Y-kromosoomi haplogrupilist kuuluvust genotüpiseerimise alusel, kasutades markereid M9, M253, TAT, L550, M423, Z282, Z1936, M78, M438, M412, M173, VL29, VL39, M198, Z93, M558, M458, M223, M269 (Underhill jt 1997; Zerjal jt 1997). Kuna selgus, et üks analüüsitud proov kuulunuks tõenäoliselt hg-sse R1a3, mis defineeriti Karmin jt. (2015) avaldatud nomenklatuuri alusel, disainiti seda defineeriva markeri CTS4385 määramiseks Karmin jt. (2015) avaldatud info alusel uued praimerid, mida kasutati proovi genotüpiseerimisel;
2. konstrueeriti Y-kromosoomi täisjärjestuste alusel (N=565) tarkvaraprogrammi BEAST abil fülogeneetilised puud ning määrati nende põhiharude koalestsentsiajad ning Ne-d.

## 2.2.Materjal ja metoodika

### 2.2.1. Kasutatud DNA proovid ja Y-kromosoomide täisjärjestused

Töös kasutati:

1. Genotüpiseerimiseks ja laborimeetodite omandamiseks kasutatud DNA proovid (N=49) pärinevad Tartu Ülikooli EBK biopangast (Lisa 1. )
2. Y-kromosoomi hg R1a alamhaplogruppide jaotuse väljaselgitamiseks kasutatud Y-kromosoomi järjestused (N=565) pärinevad Tartu Ülikooli EBK andmekogust (Lisa 2.)

### 2.3. Genotüpiseerimine

Proovide haplogrupilise kuuluvuse kindlakstegemiseks genotüpiseerisin Y-kromosoomi markerid M9, M253, TAT, L550, M423, Z282, Z1936, M78, M438, M412, M173, VL29, VL39, M198, Z93, M558, M458, M223, M269.

#### 2.3.1. Praimerite disainimine

Genotüpiseerimise käigus selgus, et üks proov kuulub tõenäoliselt hg-sse R1a3, mis defineeriti Karmin jt. (2015) uurimistöös. Selle kinnitamiseks disainisin hg R1a3 markerit CTS4385 defineerivad uued praimerid (Karmin jt 2015). Selleks võtsin Karmin jt 2015.aastal avaldatud hg R1a3 defineeriva SNP-i positsiooni. Seejärel otsisin positsiooni üles Ensembl'is (Ensembl GRCh37; <http://www.ensembl.org>), valisin sekveneeritud ala vahemiku nii, et enne ja pärast SNP-i jääks umbes 300 bp ning kopeerisin selle tekstiredigeerimisprogrammi Notepad++, et ära märkida mind huvitav SNP. Kasutades programmi Primer3 (<http://primer3.ut.ee/>) (Koressaar and Remm 2007; Untergasser jt 2012; Koressaar jt 2018) leidsin erinevad võimalikud praimerid, mille sobilikkust kontrollisin programmidega GenomeTester 1.3 (<http://bioinfo.ut.ee/genometester/>) (Andreson jt 2006) ja PrimerBlast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

#### 2.3.2. DNA amplifitseerimine

##### 2.3.2.1. Polümeraasi ahelreaktsioon (PCR)

Polümeraasi ahelreaktsiooni (*polymerase chain reaction* – PCR) reaktsioonisegu tegin ühe proovi kohta 13µl. Ahelate sünteesiks vajalikke primereid (10pmol/µl kohta Technology A/S) võtsin 0,3µl, reaktsioonipuhvrit (5x FIREpol® Master Mix, mis sisaldab 12,5 mM MgCl<sub>2</sub>, Solis BioDyne) 3µl ja milli-Q vett (Millipore) 9,7µl iga DNA proovi kohta. Segu kokku segamine toimus 1,5ml katsetuubis, mille järel kandsin segu 13µl kaupa 200µl-se mahuga PCR-i tuubiribadesse. Igasse tuubi lisasin 2µl uuritavat DNA-d.

PCR-i läbiviimiseks kasutasin termotsüklerit UNOII (Biometra). Genotüpiseerimiseks kasutatud termotsükleri programm on toodud tabelis 1.

**Tabel 1.** Termotsükleri programm

Aeg	Temperatuur	Etapp
1. 3 min	94°C	Eelkuumutus
2. 15 – 25 sek	94°C	Kaheaahelalise DNA denaturatsioon
3. 15 – 30 sek	55-60°C	Praimerite seundumine DNAlle
4. 15 – 40 sek	72°	DNA ahela elongatsioon
5. 2 – 3 min	72°	Lõplik elongatsioon, tagamaks kõikide DNA ahelate sünteesi
		30-40 tsüklit (etapid 2.-4.)
∞	4°C või 10°C	DNA lagunemise ära hoidmine

Denaturatsiooni, praimerite seundumise ja elongatsiooni etappide pikkus sõltub amplifitseeritud DNA lõigu pikkusest.

### 2.3.2.2. Geelelektroforees

PCR õnnestumist kontrollisin horisontaalsel geelelektroforeesil 2% agarosgeelil 0,5x TBE puhvris (45mM Trisboraat, 0,05 M EDTA, pH=8,3). Agarosgeeli valmistamiseks kasutasin 0,5x TBE puhvrit, millele lisasin 100µl etiidiumbromiidi iga 100ml geeli kohta. Forees toimus 170 V juures, elektroodide vahe oli 32 cm, 25-30 min. DNA-produkti kvaliteedi ja koguse hindamiseks pildistasin geeli UV-valguses (UVIPro).

### 2.3.2.3 DNA produkti „puhastamine“

Madala kontsentratsiooniga produktide esinemise korral (hägusad jooned foresipildil) tuli mul PCR-i segu jäägid eemaldada. Selleks lisasin produktidele kahe ensüümi segu. Kogus on toodud 1µl proovi kohta. Praimerijääkide lagundamiseks 0,05µl eksonukleas I (*ExoI*, Thermo Scientific, 20 U/µl) ja üleliigsete nukleotiidide lagundamiseks 0,1µl FastAP™ (*Thermosensitive Alkaline Phosphatas* temperatuuritundlik aluseline fosfataas, Thermo Scientific, 1U/µl). Tuube inkubeerisin termotsükleris 15 minutit 37°C juures üleliigsete jääkide lagundamiseks, seejärel toimus 15 minutit 85°C juures ensüümide inaktiveerimine.

### 2.3.4. Sekvenerimine

DNA sekvenerimine toimus Sangeri meetodil. Sekvenerimise ettevalmistuseks vajalik PCR ja sadestamine teostati minu poolt, sekvenerimisprodukti kapillaarne geelelektroforees teostati Tartu Ülikooli genoomika instituudi tuumiklabori sekvenaatoril (3730xl DNA Analyzer, Applied Biosystems). Sekvenerimise eesmärgiks oli tuvastada genotüpiseerimata proovide hg-d. Selleks uurisin, kas uuritavaal proovil on toimunud alamharusid määravas nukleotiidses positsioonis referentsjärjestusega võrreldes asendusreaktsioon või mitte. Uurisin selleks hg N3 (L550, Z1936, VL39), R1a (Z282, CTS4385), R1b (M412) alamharusid defineerivaid markereid.

Sekvenerimisreaktsiooni läbiviimiseks segasin 1,5 ml tuubis kokku 1 proovi kohta tabelis 2 märgitud reagentid.

**Tabel 2.a)** Sekvenerimisreaktsiooni segu ja **b)** termotsükleri programm

a) Sekvenerimissegu	µl	b) Termotsükleri programm	Kestvus
<b>BigDye Terminator v3.1 reagent premix (BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems)</b>	0,75	1. Denaturatsioon 95°C	15 sek
<b>BigDye 5x sekvenerimispuhvrit (BigDye Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems)</b>	2	2. Praimeri seondumine 50°-55°C	10 sek
<b>milli-Q H2O</b>	6,1	3. DNA süntees 60°C	60 sek
<b>Sekvenerimispraimer (1,6pmol/ul)</b>	0,16	30 tsükli (etapid 1.-3.)	
		<b>4°C</b>	∞

Seejärel kandsin segu 200µl PCR-i tuubidesse, millesse lisasin 1µl PCR-i produkti. Proovi kohta on reaktsioonisegu kogumaht 10µl. Reaktsiooni viisin läbi termotsükleris tabelis 2 toodud programmi järgi.

### 2.3.4.1. DNA sadestamine

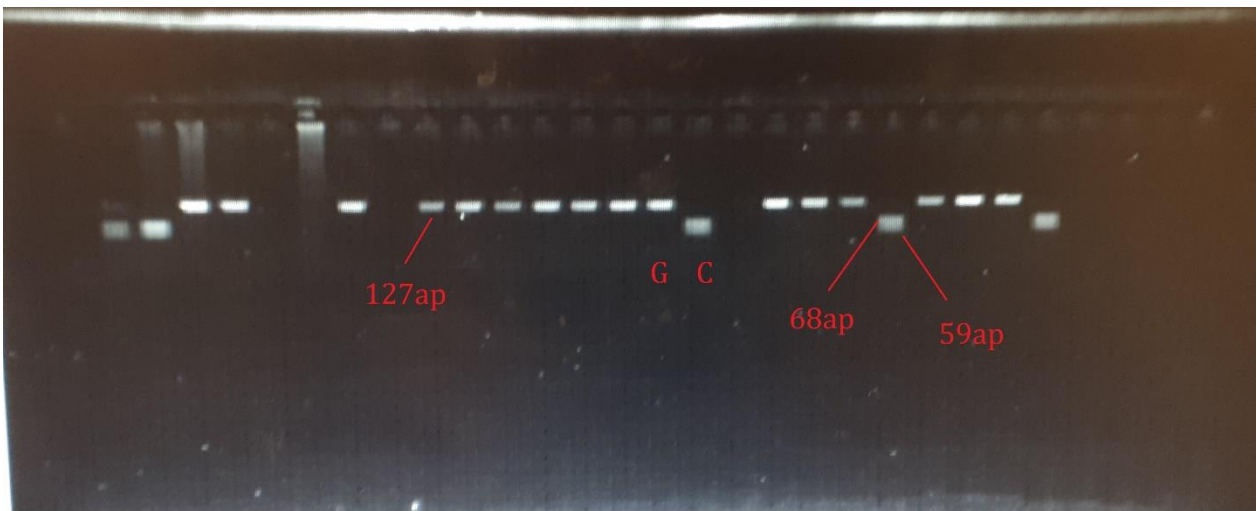
Pärast sekveneerimisreaktsiooni lõppu teostasin DNA sadestamise. Selleks lisasin alguses igasse tuubi 30µl 96% EtOH-d, mis oli -20°C juures ning 2µl ammooniumatsetaadi ja dekstraani segu [3,5M CH<sub>3</sub>COONH<sub>4</sub> [pH>8]/EDTA (250 mM) ja 20 mg/ml punane dekstraan suhtes 1:1]

Ammooniumatsetaat ja dekstraani segu on vajalik teostamiseks DNA puhastamiseks vajalikud etapid. Ammooniumatsetaat sadestab DNA ja dekstraan aitab selle seondada tahkele kandjale. DNA väljasadenemiseks hoidsin tuube 10 minutit sügavkülmas -20°C juures, seejärel tsentrifugeisin tuube 10 minuti jooksul 13 000 rpm juures ning eemaldasid supernatandi veejoapumbaga nii, et DNA sade ja dekstraan jäid tuubi põhja. Seejärel lisasin igasse tuubi 200µl 70% EtOH-d, mis oli -20°C juures, et DNAd pesta. Tsentrifugeisin tuube veel 5 minutit 13000 rpm juures, ning eemaldasid supernatandi. Järgmiseks kandsin igasse tuubi 200µl -20°C 70% EtOH-d teistkordseks pesemiseks. Supernatandi eemaldamine toimus ilma eelneva tsentrifugeimiseta. Seejärel lasin proovidel 15 minutit 37°C termostaadis kuivada, et EtOH ära aurustuks. Lõpus lisasin igasse tuubi 10µl 70% formamiidi ning viisin proovid edasiseks sekveneerimiseks Tartu Ülikooli genoomika instituudi tuumiklaborisse.

### 2.3.5. Restriksioonianalüüs

Restriksioon on DNA molekuli fragmentideks lõikamine spetsiifilise nukleotiidses järjestuses põhjal. Restriksiooni viivad läbi restriksioonilised endonukleaasid ehk restriктаasid, mis lõhuvad ära nukleotiididevahelised fosfodiesterid. Restriksioonisaadid võivad tekkida või kaduda DNA järjestuses tekkinud mutatsioonide tõttu. Viisin läbi restriksioonianalüüsi, et teada saada, kas restriктаasid lõikavad amplifitseeritud DNAd või mitte.

Uurimiseks valisin 49 eestlase proovi, kelle isaliini haplogrupiline kuuluvus ei olnud teada. Esmaseks uuringuks amplifitseerisin neil hg K defineeriva markeri M9 (C→G) järjestust sisaldava DNA lõigu, mille pikkus on 127 ap. Restriктаasina kasutasin *HinfI* ensüümi (5 U/µl). Mutatsiooni kohalt lõigatud proovid on lühemad ja liiguvad elektroforeesil kiiremini, jagunedes 68 ap ja 59 ap pikkusteks fragmentideks (joonis 10).



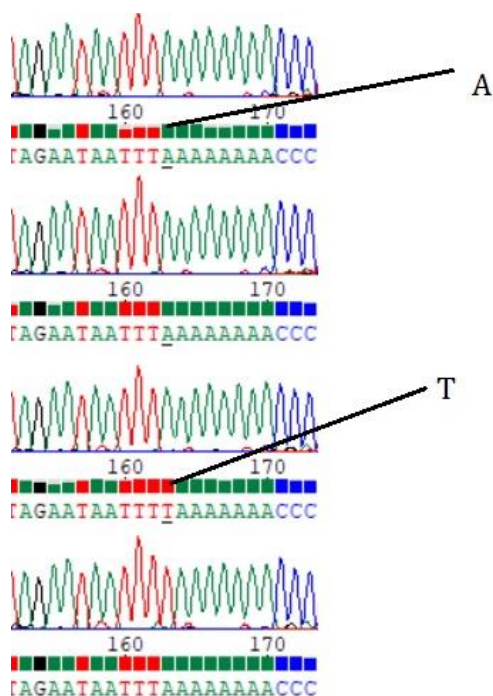
**Joonis 10.** Markeri M9 (C→G) allelide tuvastamine *Hinf*I restriktsoonanalüüsil. Restriktaas lõikab markeri algse (ingl. k. *ancestral allele*) alleelivariandi 68 ap ja 59 ap pikkusteks fragmentideks. Restriktaas ei lõika markeri derivaatalleeli (ingl. k. *derived allele*) G, mis jääb 127 ap pikkuseks fragmendiks.

Restriktsooni läbiviimiseks lisasin eelnevalt amplifitseeritud PCR-i produktidele 0,5 U ehk 0,1 µl *Hin*III ensüümi ja 5 U ehk 1 µl puhvrit ühe proovi kohta. Seejärel hoidsin proove üleöö 37°C juures termostaadis. Reaktsiooni tulemust kontrollisin geelelektroforeesil (vt. 2.3.2.2).

## 2.3.6. Andmeanalüüs

### 2.3.6.1. Sekveneeritud järjestuse analüüs ChromasPros

Sekveneerimistulemuste analüüsiks kasutasin programmi ChromasPro 2.1.8 (Technelysium, <https://technelysium.com.au/wp/chromaspro/>). Programmi abiga on võimalik tuvastada, millise alleeliga on vastava markeriga uuritavas proovis mutatsiooni määravas positsioonis tegu (joonis 11).



**Joonis 11.** Sekveneeritud DNA lõigu järjestus markeri CTS4385(A→T) näitel, kus on näha mõlemad alleelivariandid; algne alleel A ja derivaatalleel T. PCR-il kasutati markerit CTS4385. (Chromas Pro 2.1.8 Technelysium, ekraanitõmmis)

### 2.3.7. Genotüpiseerimise tulemused

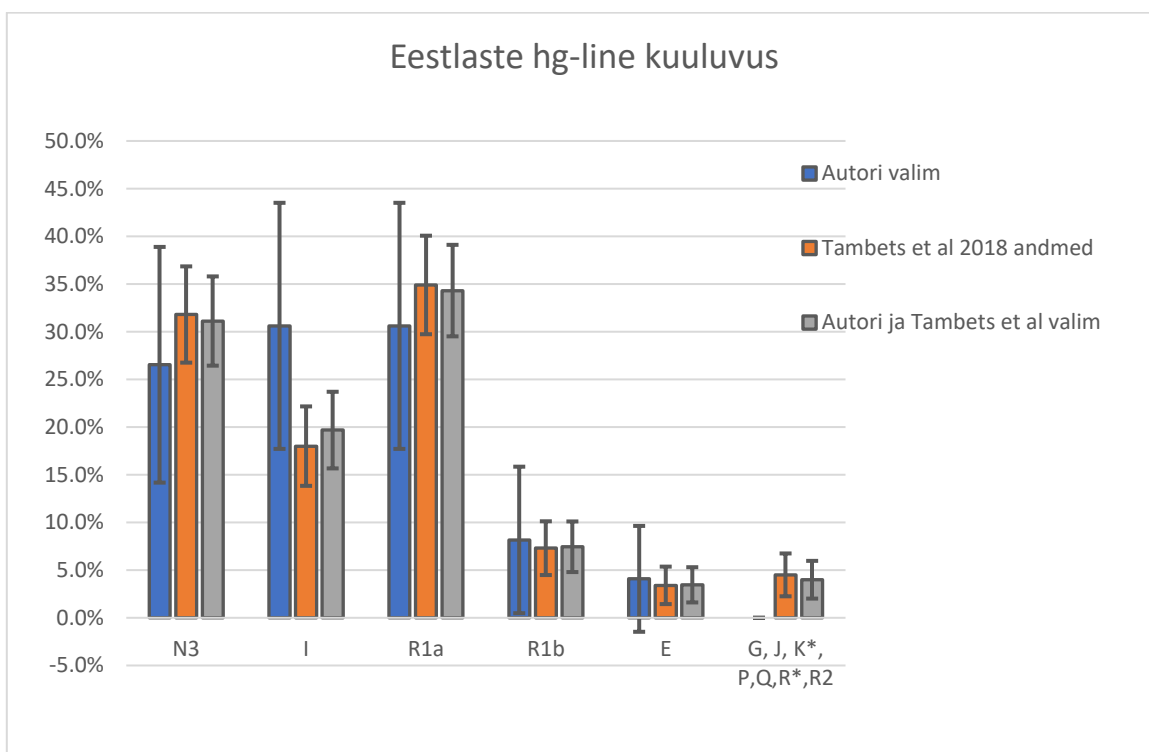
Y-kromosoomi haplogrupid määrati kõikide uuritud proovide jaoks (Lisa 1.). Tabelis 3 on näha uuritud proovide jaotumise sagedus suurematesse haplogruppidesse, detailne jaotus on toodud Lisis 1. Analüüsi tulemustest selgus, et uuritud indiviidest enamus kuulus kolme haplogruppi – R1a, N3 ja I1 (kokku 81,6%).

**Tabel 3.** Uuritud indiviidide Y-kromosoomi hg-de sagedused ning hg-de esinemise protsentuaalne sagedus.

Haplogrupid	I1	I2	N3	R1a	R1b	E	Kokku
Indiviidide arv	12	3	13	15	4	2	49
	(24,5%)	(6,1%)	(26,5%)	(30,6%)	(8,2%)	(4,1%)	(100%)

### 2.3.7.1. Analüüsitud indiviidide Y-kromosoomide varieeruvuse võrdlus EBK andmebaasis olemasolevate eestlaste andmetega

Analüüsitud indiviidide Y-kromosoomide haplogruppe kõrvutati olemasolevate eestlaste andmetega (joonis 12). Selleks kasutati kahte allikat: 2018. aasta uuringut (Tambets jt 2018) ning EBK Y-kromosoomi hg andmebaasi lisandunud minu poolt genotüpiseeritud eestlaste proove. Uuring näitas, et hg R1a esineb 34,9% eesti meestel, N3 - 31,8%, levinud on ka I (18%) ja R1b (7,3%) ning esineb ka E3b (3,4%), J (1,8%), G (1,2%), F(0,3%). Nagu graafikult on näha, muutub veapiir seda väiksemaks, mida suurem on uuritud valim, olles siinse väikse valimi puhul suhteliselt suur. Summeerides minu ja Tambets jt 2018 andmed muutub leitud veapiir veidi väiksemaks võrreldes ainult Tambets jt andmetega ning tunduvalt väiksemaks kui minu valimi puhul. Saadud tulemus kinnitab veelkord suuremate proovivalimite vajalikkust populatsiooniuringutes.



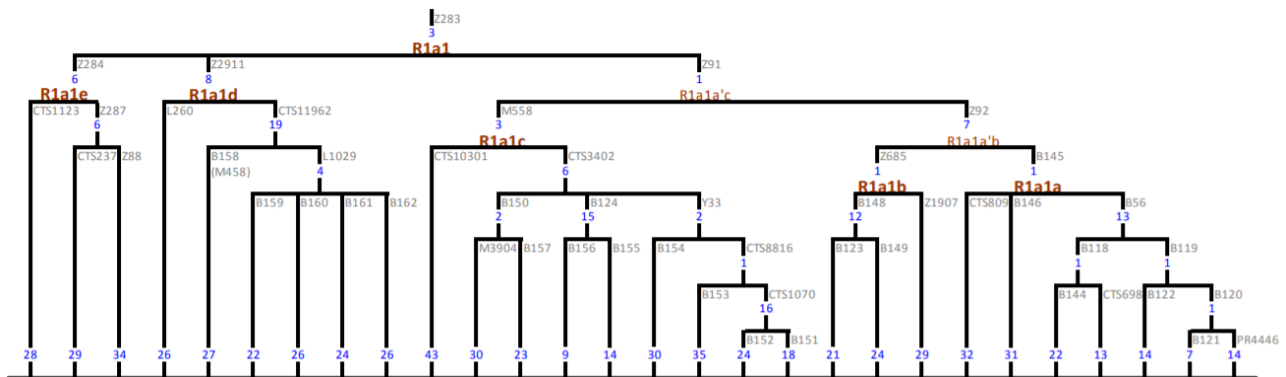
**Joonis 12.** Tulpdiaagramm Y-kromosoomi genotüpiseerimise tulemuste kohta erineva suurusega valimites. Eesti meeste Y-kromosoom hg-de protsentuaalne jaotus (Tambets jt 2018) andmete järgi (N=327), autori valimis olnud eestlaste hg-line jaotus (N=49) ning summeeritud valimi haplogrupiline jaotus (N=365) ning 95% veapiirid.

## 2.4. Y- kromosoomi täisjärjestuseanalüüs R1a näitel

Töös kasutatud Y-kromosoomi täisjärjestused (N=565) pärinevad EBK Y-kromosoomi täisjärjestuste kogust ning on saadud Illumina platvormil sekveneeritud täisgenoomidest Y-kromosoomi täisjärjestuste väljaeraldamisel ja Y-kromosoomi väljapüügi meetodit (ingl. k. *Y-capture* BigY, Gene by Gene Ltd.) kasutades. Järjestused on järeltöötuse käigus joondatud ja määratud on nende aluspaarid. Antud töös kasutasin Y-kromosoomi fülogeneetiliste puude rekonstrueerimiseks Tartu Ülikooli genoomika instituudi EBK vanemteadur Lauri Saagi väljatöötatud Y-kromosoomi täisjärjestuste andmeanalüüsi programmide töövoogu ja BEAST v 1.7.5 tarkvarapaketti (Drummond jt 2002; Drummond and Rambaut 2007). Fülogeneesi puud on konstrueeritud ainult X-degenereerunud piirkonna ja varieeruvate nukleotiidipositsioonide põhjal. Kladide vanuste määramisel kasutasin mutatsioonikiirust, mis on saadud, teisendades  $0.74 \times 10^{-9}$  asendust nukleotiidi kohta aastas (Karmin jt 2015) ümber varieeruvatele positsioonidele. Puude visualiseerimiseks kasutasin programmi FigTree v. 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>). Tracer v.1.7.1 programmiga (<https://github.com/beast-dev/tracer/releases>) (Rambaut jt 2018) oli võimalik hinnata Ne dünaamikat. Analüüsi lihtsustamiseks jaotasin valimi (N= 565), millesse kuulus 395 eestlast, 38 soomlast ja 123 rootslast, alamhaplogruppide alusel kuueks grupiks – R1a1, R1a1c\_esimene ja R1a1c\_teine, R1a1d, R1a3\_R1a2\_R1a5 ning R1a3\_R1a1e, ning kasutasin Y-kromosoomi töövoogu, et saada edasiseks analüüsiks vajalikud väljundfailid. Kõigepealt rekonstrueerisin fülogeneesipuu programmiga BEAST kasutades välisgrupina haplogrupi R1b järjestust, et saada teada hg R1a5 eeldatav koalestsentsaeg, milleks sain 14577 aastat. Järgnevalt kasutasin fülogeneesipuude tegemisel välisgrupina R1a5 järjestust ning leitud aega välisgrupi ehk fülogeneesi puu juure koalestsentsiajana. Kasutasin Y-kromosoomi töövooväljundeid, et leida harudel esinenud mutatsioonid ning saada RAxML (*Randomized Axelerated Maximum Likelihood*) puud, millel on olemas haru pikkused, ja Nexus failid edasiseks BEAST'i analüüsiks. Esmalt tegin Nexus failide põhjal BEAST programmi sisendfailid (laiendiga .xml), kasutades samasse tarkvarapaketti kuuluvat programmi BEAUTi. Sisendfaile kasutasin BEAST'i analüüsis, mille väljundiks on parim fülogeneesipuu, ning .log-fail ja trees-fail, milles on parameetrite vaatamiseks ja Ne muutuse visualiseerimiseks vajalik info.

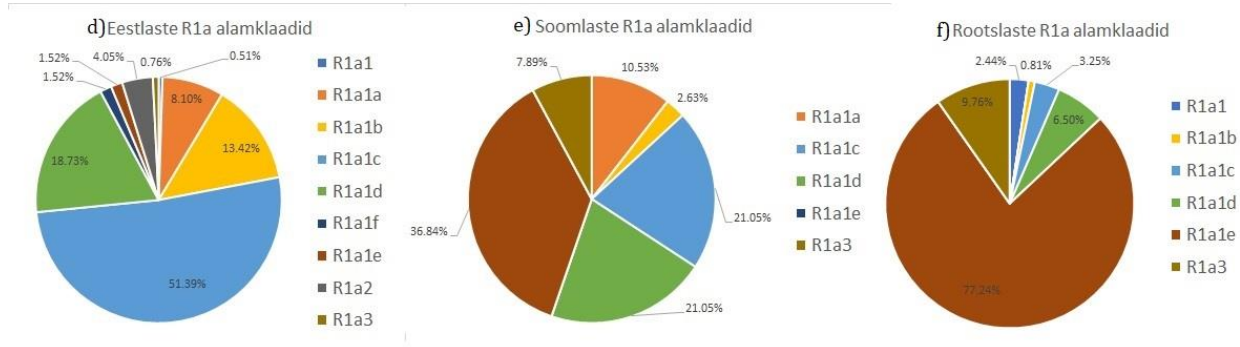
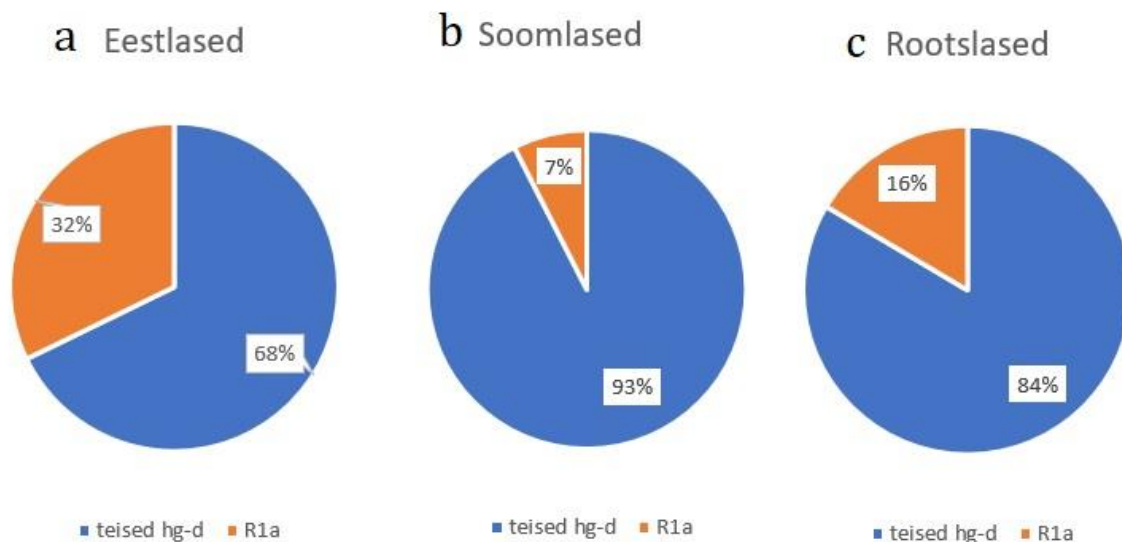
### 2.4.1. Y-kromosoomi täisjärjestuste analüüsi tulemused

Töös kasutasin Karmin jt 2015 R1a1 topoloogiat (joonis 13). Uuritud populatsioonid jagunesid 9 hg R1a alamklaadi vahel (Lisa 2), millest 7 olid R1a1 alamklaadid. Tööd läbi viies selgus, et suurema proovide hulga tõttu on suurenenud varem defineerimata markerite arv ning valimis olid esindatud ka Karmin jt 2015 defineerimata uued alamklaadid, mis vajavad täiendavat tööd harude defineerimisel, et olemasolevat nomenklatuuri edasi arendada (R1a1\* ja R1a1f - töövariant nomenklatuuri täiendamisest, Siiri Rootsi).



**Joonis 13.** Hg R1a1 topoloogia. Kohandatud (Karmin jt 2015).

Alamklaadide puhul on näha selged erinevused populatsioonide vahel nii R1a esinemissageduse (joonis 14 a, b, c), klaadide hulga kui ka sageduste osas (joonis 14 d, e, f.). Eestlastel on R1a sagedus 32% (95% usalduspiirid 29,7-34,9%) ning neil esineb rohkem hg R1a alamklaade kui soomlastel ja rootslastel, mille põhjuseks võivad olla erinevused demograafilises ajaloos ja migratsioonivoogude hulgas ning suunas (joonis 14 d, e, f). Eestlastel on põhiliselt esindatud haplogrupid R1a1c ja R1a1d. Soomlastel leidub hg R1a suhteliselt vähe (7%; 95% usalduspiirid 5,5-10,1%), põhilised esindatud alamklaadid on R1a1e-Z284 ja R1a1d. Rootslaste hg R1a (16%; 95% usalduspiirid 14-19,3%) alamklaadide esinemise muster erineb soomlaste ja eestlaste omast. Peamiselt leiduvad rootslastel alamklaad R1a1e (76% kogu R1a-st) ning R1a3, mida rootslastel on 9,8% (95% usalduspiirid 5,7-16,3%) ja mis on eestlaste seas väga haruldane (0,8%; 95% usalduspiirid 0,3-2,2%). Eestlastel ja rootslastel esinevad alamhaplogrupid on väga erinevad – rootslastel on põhiliseks alamhaplogrupiks R1a1e, mis on eestlastel esindatud vaid marginaalselt, seevastu soomlaste alamhaplogrupid sarnanevad mõlema populatsiooni hg-dega.

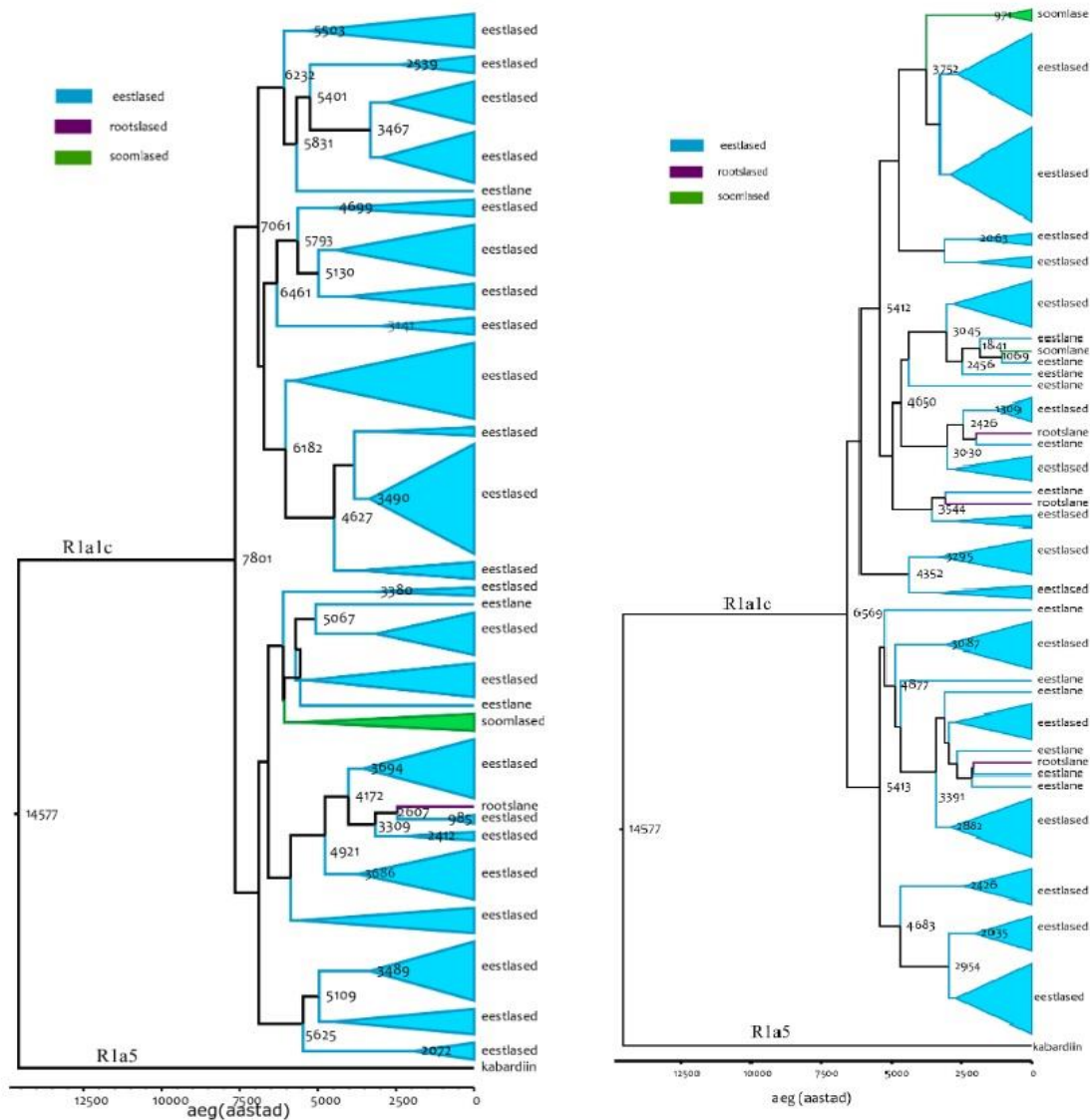


**Joonis 14.** Hg R1a esinemissagedus **a)** eestlastel **b)** soomlastel ja **c)** rootslastel. Hg R1a alamklaadide jaotus R1a-d kandvatel indiviididel **d)** eestlastel, **e)** soomlastel ja **f)** rootslastel

Eestlastel esinevad põhiliselt R1a1 alamklaadid, R1a1a ja R1a1b ja lisaks ka Karmin jt 2015 järjestuste puudumise tõttu defineerimata alamklaadid R1a1\* ning R1a1f, mida defineerib Yfulli andmebaasis olev marker S24902 (joonis 15, Lisa 2.). Fülogeneesipuul on näha ka uus rootslaste R1a1 haru, mida ei ole varem Karmin jt 2015 kirjeldatud ja mille koalestsentsiaeg on umbes 2 000 (95% usaldusvahemik: 1 007 – 2 432 aastat) at ning seda, et R1a1 alamklaadid esinevad põhiliselt eestlastel. R1a1a’b haru koondab Karmini jt järgi Underhill jt R1a-Z282 (xM458, M558) harud. Keskmiselt toimus R1a1 alamklaadide ekspansioon 2 500 aastat tagasi (lisa 3). Uute, seni defineerimata klaadide esinemine viitab vajadusele kaasata populatsiooniuuringutes suuremaid valimeid ja täiendada fülogeneesipuid uute markerite ja harudega.

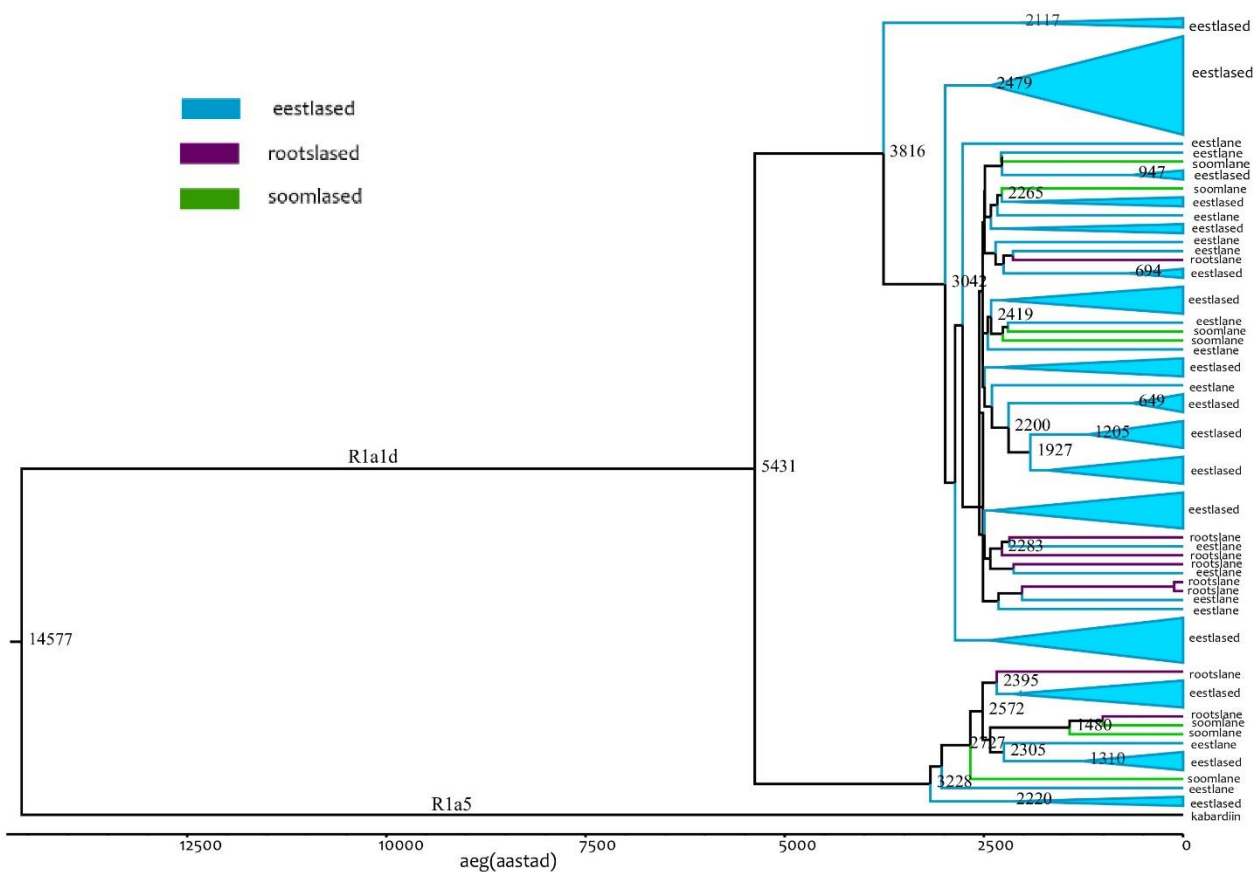


esineb R1a1-M558 kõige rohkem Kesk- ja Ida-Euroopas sh. Volga-Uurali populatsioonides ning on kõige sagedasem R1a-Z282 alamklaad Euroopas.



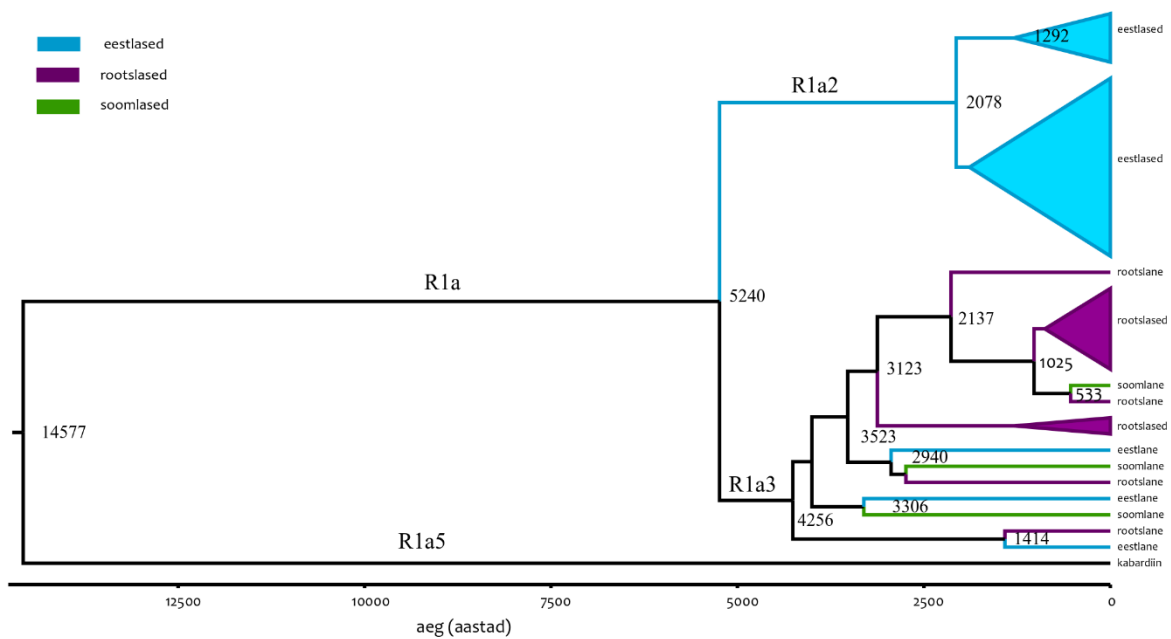
**Joonis 16.** Haplogrupi R1a1c fülogeneetilised puud, millel on ära märgitud ekspansioonijad harudel, mille aposterioorne väärtus on üle 0,85..

Hg R1a1d – Z2911, mis on uemate andmete põhjal ka defineeritav kui hg R1a1d-M458, esineb põhiliselt soomlastel ja eestlastel ning ka osadel rootslastel (joonis 17). Seda klaadi iseloomustab suhteliselt ühtlane jaotus kõigis kolmes uuritud populatsioonis ja Underhill jt 2015 järgi on see klaad kõige suurema sagedusega esindatud Kesk- ja Ida Euroopas. Haplogrupi erinevate alamklaatri ekspansioon toimus umbes 2 828 – 4 786 aastat tagasi.



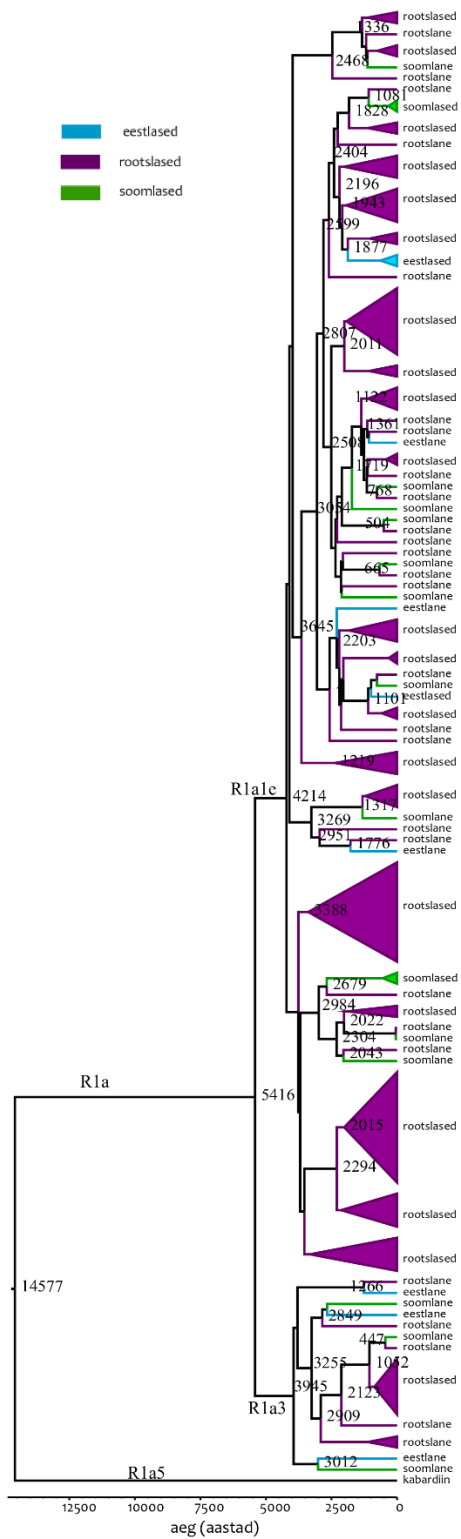
**Joonis 17.** Haplogrupi R1a1d fülogeneesipuu, millel on ära märgitud ekspansiooniajad harudel, mille aposterioorne väärtus on üle 0,85.

Hg R1a3 on põhiliselt Rootsis esinev R1a alamklaad, mille ekspansioon toimus hilises mesoliitikumis 2164 – 6177 aastat tagasi. Eestlastel esineva haruldase alamklaadi R1a2 ekspansioon toimus hg R1a3-st hiljem 709 - 2091 aastat tagasi.



**Joonis 18.** Haplogrupi R1a3 ja R1a2 fülogeneesipuu, millel on ära märgitud ekspansiooniajad harudel, mille aposterioorne väärtus on üle 0,85.

Haplogrupid R1a3 ja R1a1e on iseloomulikud rootslastele (joonised 18 ja 19). Ligikaudu 77% rootslaste R1a kandjatest kuuluvad hg-sse R1a1e. Hg R1a1e leidub ka soomlastel ja vähesel määral ka eestlastel. Hg R1a1e ekspansioon toimus 3435 – 5476 at.



**Joonis 19.** Haplogrupi R1a3 ja R1a1e fülogeneesipuu, millel on ära märgitud ekspansiooniajad harudel, mille aposterioorne väärtus on üle 0,85.

Vaadeldes R1a alamhaplogruppide Ne dünaamikat (lisa 3.) saab hinnata nende ekspansioonieaegu. Nende järgi toimus R1a1, R1a1d ning R1a3 (millel on näha kõige madalamat Ne muutust) alamhaplogruppide ekspansioon umbes 2 500 aastat tagasi, hg R1a1c ja R1a1e ekspansioon 3 500 at.

Töö tulemusena saab öelda, et eestlastel, soomlastel ja rootslastel on väga erinev R1a esinemissagedus populatsioonis ning R1a alamklaadide esinemismuster. Eestlasi ning nende R1a alamklaade tuleks edaspidistes uuringutes kindlasti võrrelda teiste naaberpopulatsioonidega (venelased, lätlased, leedukad), kuna ka nendel on täheldatud kõrget R1a1c-M458 ja R1a1d-M558 esinemissagedust.

## Kokkuvõte

Antud töös käsitletakse hg R1a alamklaadide sagedust ja jaotust eestlaste ja Põhja-Euroopa populatsioonide – rootslased ning soomlased – hulgas. Hg R1a on huvipakkuv, kuna ta esineb põhiliselt Kesk- ja Ida-Euroopas sh. ka Volga-Uurali populatsioonides ning on eestlaste seas väga sage.

Töö eksperimentaalosa vaatleb kõigepealt eestlaste juhuvalimi (n=49) genotüpiseerimise tulemusi. Indiviidid jagunesid 13 haplogrupi vahel, millest kõige sagedamini esinesid haplogrupid I1 (24,5%), N3 (26,5%) ja R1a (31%). Tulemustest selgub, et autori valimi haplogrupiline jaotus sarnaneb Tartu Ülikooli EBK andmebaasis olevate andmetega. Väiksed erinevused olid seotud autori valimi suurusega, mis ei võimaldanud leida Eestis haruldasemaid haplogruppe. Analüüsi tulemustest selgus, et populatsiooniuuringutes on vajalik suuremate proovivalimite kogumine, kuna see võimaldab hoida haplogruppide sageduse veapiiri väiksena ja tuvastada ka populatsioonis vähemlevinud haplogruppe.

Y-kromosoomi täisjärjestuste analüüsi tulemusena selgus, et autori valimis esinesid R1a alamhaplogrupid (R1a1\*, R1a1f), mis ei olnud veel Karmin jt 2015 avaldatud. See tõestab veelkord, et populatsiooniuuringutes on vajalikud suured valimid, et tuvastada varem defineerimata alamklaade. Rootslastel esineb põhiliselt alamhaplogrup R1a1e ja lisaks leidub madala sagedusega hg R1a3, mille ekspansioon toimus 2200-6200 aastat tagasi hilismesoliitikumis. Eestlaste (ja soomlaste) seas on kõige sagedasemad alamklaadid R1a1c-M458 ja R1a1d-M558, mida rootslaste hulgas leidub vähe. Sellest saab järeldada, et tulevikus on vaja võrrelda eestlaste ja nende teiste naaberpopulatsioonide hg R1a jaotust.

## **Y- chromosome hg R1a Subclade Distribution in Estonia and Northern Europe**

Laura Laks

### **Summary**

The aim of this study was to analyse the phylogeographic spread and frequency of Y chromosomal haplogroup R1a in Estonia and in Northern Europe (among Finns and Swedish). Hg R1a stands out among other European paternal lineages as it occurs mostly in Central and Eastern Europe, including in Volga-Uralic populations and is one of the most common haplogroups in Estonia.

The first part of the study focuses on the genotyping of an Estonian randomized sample (n=49). Haplogroups were determined for all samples and it was concluded that the individuals were divided into 13 haplogroups. The most often occurring haplogroups were I1 (24,5%), N3 (26,5%) and R1a (31%). The results showed that the haplogroup distribution of randomized samples was similar to the data in the University of Tartu EBC database. It was concluded that the small changes that did exist were due to the small sample size.

The analysis of whole Y-chromosome sequences revealed two haplogroup R1a subclades (R1a1\* and R1a1f\*) that were not described in previous population genetic studies. This shows that there is a need for larger sample sizes to determine new subclades. The main subclades found among Swedes are R1a1e and R1a3. The expansion of Haplogroup R1a3 happened around 2 200-6 200 years ago in the late Mesolithic. The most common subclades found in Estonians (and Finns) are R1a1c-M458 and R1a1d-M558, both of which are rare in Swedes. In the future, comparison of hg R1a distribution among Estonians and their neighbouring populations is needed.

## **Tänuõnad**

Soovin eelkõige tänada oma juhendajaid Siiri Rootsit ja Kristiina Tambetsit. Samuti tänan Helen Posti, kes aitas mul läbi viia Y-kromosoomi täisjärjestusanalüüsi ja Anne-Mai Ilumäed konstruktiivse kriitika eest ning kõiki evolutsioonibioloogia töögrupi liikmeid, kes mind alati aitasid, nõustasid ja õpetasid.

## Kasutatud kirjandus

- Andreson R, Reppo E, Kaplinski L, Remm M. 2006. GENOMEMASKER package for designing unique genomic PCR primers. *BMC Bioinformatics*. 7(172):1–11. doi:10.1186/1471-2105-7-172.
- Balanovsky O, Rootsi S, Pshenichnov A, Kivisild T, Churnosov M, Evseeva I, Pocheshkhova E, Boldyreva M, Yankovsky N, Balanovska E, jt 2008. Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context. *Am J Hum Genet*. 82(1):236–250. doi:10.1016/j.ajhg.2007.09.019.
- Batini C, Hallast P, Zadik D, Delser PM, Benazzo A, Ghirotto S, Arroyo-pardo E, Cavalleri GL, Knijff P De, Dupuy BM, jt 2015. Large-scale recent expansion of European patrilineages shown by population resequencing. *Nat Commun*. 6(7152):1–8. doi:10.1038/ncomms8152.
- Batini C, Jobling MA. 2017. Detecting past male - mediated expansions using the Y chromosome. *Hum Genet*. 136(5):547–557. doi:10.1007/s00439-017-1781-z.
- Benazzi S, Douka K, Fornai C, Bauer CC, Kullmer O, Svoboda J, Pap I, Mallegni F, Bayle P, Coquerelle M, jt 2011. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature*. 479(7374):525–528. doi:10.1038/nature10617.
- Brown GR, Laland KN, Mulder MB. 2009. Bateman ' s principles and human sex roles. *Cell Press*. 24(6):297–304. doi:10.1016/j.tree.2009.02.005.
- Brown WM, George MJ, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *J Raman Spectrosc*. 76(4):1967–1971. doi:10.1002/(sici)1097-4555(199706)28:6<433::aid-jrs125>3.3.co;2-5.
- Burton ML, Moore CC, Romney AK, Aberle DF, Barcelo JA, Dow MM, Guyer JI, Kronenfeld DB, Levy JE, Linnekin J. 1996. Regions Based on Social Structure. *Curr Anthropol*. 37(1):87–123. doi:10.1086/204474.
- Drummond AJ, Rambaut A. 2007. BEAST : Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 8:1–8. doi:10.1186/1471-2148-7-214.
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2002. Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Mol Biol Evol*. 22(5):1185–1192.

doi:10.1093/molbev/msi103.

Dupuy BM, Stenersen M, Lu TT, Olaisen B. 2006. Geographical heterogeneity of Y-chromosomal lineages in Norway. *Forensic Sci Int.* 164(1):10–19.

doi:10.1016/j.forsciint.2005.11.009.

Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, et al. 2016. The genetic history of Ice Age Europe. *Nature.* 534(7606):200–205.

doi:10.1038/nature17993.

Graves JAM, Wakefield MJ, Toder R. 1998. The origin and evolution of the pseudoautosomal regions of human sex chromosomes. *Hum Mol Genet.* 7(13):1991–1996.

doi:10.1093/hmg/7.13.1991.

Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tänzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, et al. 2005. Evolution: Ancient DNA from the first European farmers in 7500-year-old neolithic sites. *Science* (80- ). 310(5750):1016–1018.

doi:10.1126/science.1118725.

Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature.* 522:207–211.

doi:10.1038/nature14317.

Hallast P, Batini C, Zadik D, Delser PM, Wetton JH, Arroyo-pardo E, Cavalleri GL, Knijff P De, Bisol GD, Dupuy BM, et al. 2015. The Y-Chromosome Tree Bursts into Leaf: 13,000 High-Confidence SNPs Covering the Majority of Known Clades. *Mol Biol Evol.* 32(3):661–673.

doi:10.1093/molbev/msu327.

Heyer E, Chaix R, Pavard S, Austerlitz F. 2012. Sex-specific demographic behaviours that shape human genomic variation. *Mol Ecol.* 21(3):597–612. doi:10.1111/j.1365-294X.2011.05406.x.

Housley R, Holloway R, Gamble C, Street M, Pettitt PB. 1997. Housley, R. A., Gamble, C. S., Street, M. & Pettitt, P. 1997: Radiocarbon evidence for the Lateglacial human recolonisation of Northern Europe. *Proceedings of the Prehistoric Socie... Proc Prehist Soc.* 63:25–54.

doi:10.1017/S0079497X0000236X.

Hughes JF, Rozen S. 2012. Genomics and Genetics of Human and Primate Y Chromosomes. *Annu Rev Genomics Hum Genet.* 13:83–108. doi:10.1146/annurev-genom-090711-163855.

- Hughes JF, Skaletsky H, Page DC. 2013. Sequencing of rhesus macaque Y chromosome clarifies origins and evolution of the DAZ (Deleted in AZoospermia) genes. *BioEssays*. 34(12):1035–1044. doi:10.1002/bies.201200066.Sequencing.
- Illumäe AM, Reidla M, Chukhryaeva M, Järve M, Post H, Karmin M, Saag L, Agdzhoyan A, Kushniarevich A, Litvinov S, jt 2016. Human Y Chromosome Haplogroup N: A Non-trivial Time-Resolved Phylogeography that Cuts across Language Families. *Am J Hum Genet*. 99(1):163–173. doi:10.1016/j.ajhg.2016.05.025.
- Ingman M, Kaessmann H, Pääbo S, Gyllensten U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*. 410(6828):708–713. doi:10.1038/35069127.
- Jobling MA, Pandya A, Tyler-Smith C. 1997. The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med*. 110(3):118–124. doi:10.1007/s004140050050.
- Jobling MA, Tyler-Smith C. 2003. THE HUMAN Y CHROMOSOME : AN EVOLUTIONARY MARKER COMES OF AGE. *Nature*. 4(August):598–612. doi:10.1038/nrg1124.
- Jobling MA, Tyler-Smith C. 2017. Human Y-chromosome variation in the genome-sequencing era. *Nat Publ Gr*. 18(8):485–497. doi:10.1038/nrg.2017.36.
- Karafet TM, Mendez FL, Sudoyo H, Lansing JS, Hammer MF. 2015. Improved phylogenetic resolution and rapid diversification of Y-chromosome haplogroup K-M526 in Southeast Asia. *Eur J Hum Genet*. 23(3):369–373. doi:10.1038/ejhg.2014.106.
- Karlsson AO, Wallerström T, Götherström A, Holmlund G. 2006. Y-chromosome diversity in Sweden - A long-time perspective. *Eur J Hum Genet*. 14(8):963–970. doi:10.1038/sj.ejhg.5201651.
- Karmin M, Saag L, Vicente M, Sayres MAW, Järve M, Talas UG, Rootsi S, Illumäe A-M, Mägi R, Mitt M, jt 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res*. 25:459–466. doi:10.1101/gr.186684.114.67.
- Kido T, Fai Y, Lau C. 2015. Roles of the Y chromosome genes in human cancers. *Asian J Androl*. 17:373–380. doi:10.4103/1008-682X.150842.
- Kivisild T. 2017. The study of human Y chromosome variation through ancient. *Hum Genet*. 136(5):529–546. doi:10.1007/s00439-017-1773-z.

- Klein RG. 2000. Archeology and the evolution of human behavior. *Evol Anthropol.* 9(1):17–36. doi:10.1002/(SICI)1520-6505(2000)9:1<17::AID-EVAN3>3.0.CO;2-A.
- Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinforma Appl NOTE.* 23(10):1289–1291. doi:10.1093/bioinformatics/btm091.
- Krausz C, Forti G, Mcelreavey K. 2008. The Y chromosome and male infertility. *Curr Opin Urol.* 18(6):628–632. doi:10.1097/MOU.0b013e32831367dc.
- Lachance J, Vernot B, Elbers CC, Ferwerda B, Froment A, Bodo JM, Lema G, Fu W, Nyambo TB, Rebbeck TR, jt 2012. Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. *Cell.* 150(3):457–469. doi:10.1016/j.cell.2012.07.009.
- Lappalainen T, Koivumäki S, Salmela E, Huoponen K, Sistonen P, Savontaus ML, Lahermo P. 2006. Regional differences among the finns: A Y-chromosomal perspective. *Gene.* 376(2):207–215. doi:10.1016/j.gene.2006.03.004.
- Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus ML, Lahermo P. 2008. Migration waves to the baltic sea region. *Ann Hum Genet.* 72(3):337–348. doi:10.1111/j.1469-1809.2007.00429.x.
- Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, jt 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature.* 513(7518):409–413. doi:10.1038/nature13673.
- Malyarchuk B, Derenko M, Denisova G, Maksimov A, Wozniak M, Grzybowski T, Dambueva I, Zakharov I. 2011. Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a. *J Hum Genet.* 56(8):583–588. doi:10.1038/jhgm.2011.64.
- Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, jt 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature.* 528(7583):499–503. doi:10.1038/nature16152.
- Mellars PA. 1992. Archaeology and the population-dispersal hypothesis of modern human origins in Europe. *Philos Trans R Soc London Ser B Biol Sci.* 337:225–234. doi:10.1098/rstb.1992.0100.

- Moore LT, McEvoy B, Cape E, Simms K, Bradley DG. 2006. A Y-Chromosome Signature of Hegemony in Gaelic Ireland. *Am J Hum Genet.* 78(2):334–338. doi:10.1086/500055.
- Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, jt 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur J Hum Genet.* 19(1):95–101. doi:10.1038/ejhg.2010.146.
- Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat Genet.* 29(1):20–21. doi:10.1038/ng711.
- van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat.* 30(2):386–394. doi:10.1002/humu.20921.
- Pagani L, Lawson DJ, Jagoda E, Mörseburg A, Eriksson A, Mitt M, Clemente F, Hudjashov G, Degiorgio M, Saag L, jt 2016. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature.* 538(7624):238–242. doi:10.1038/nature19792.
- Pinhasi R, Zarina G, Moiseyev V, Lightfoot E, Nigst PR, Manica A, Pinhasi R, Bradley DG. 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Curr Biol.* 27(4):576–582. doi:10.1016/j.cub.2016.12.060.
- Poznik GD, Henn BM, Yee M-C, Sliwerska E, Euskirchen GM, Lin AA, Snyder M, Quintana-Murci L, Kidd JM, Underhill PA, jt 2013. Sequencing Y Chromosomes Resolves Discrepancy in Time to Common Ancestor of Males versus Females. 341(6145):562–565. doi:10.1007/s10803-012-1725-3.Brief.
- Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, jt 2016. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet.* 48(6):593–599. doi:10.1038/ng.3559.
- Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford TW, Orlando L, Metspalu E, jt 2014. Upper palaeolithic Siberian genome reveals dual ancestry of native Americans. *Nature.* 505(7481):87–91. doi:10.1038/nature12736.
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in

Bayesian phylogenetics using Tracer 1.7. *Syst Biol.* 67(5):901–904. doi:10.1093/sysbio/syy032.

Rebollo NR, Weiner S, Brock F, Meignen L, Goldberg P, Belfer-Cohen A, Bar-Yosef O, Boaretto E. 2011. New radiocarbon dating of the transition from the Middle to the Upper Paleolithic in Kebara Cave, Israel. *J Archaeol Sci.* 38(9):2424–2433. doi:10.1016/j.jas.2011.05.010.

Rootsi S, Zhivotovsky LA, Baldovič M, Kayser M, Kutuev IA, Khusainova R, Bermisheva MA, Gubina M, Fedorova SA, Ilumäe AM, jt 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet.* 15(2):204–211. doi:10.1038/sj.ejhg.5201748.

Rosser H, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, jt 2000. Ychromosomal diversity in europe Rosser 2000.pdf. *Am J Hum Genet.* 67:1526–1543.

Saag Lehti, Laneman M, Varul L, Malve M, Valk H, Razzak MA, Shirobokov IG, Khartanovich VI, Mikhaylova ER, Kushniarevich A, jt 2019. The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers further East. *Curr Biol.* 29:1701–1711. doi:10.1016/j.cub.2019.04.026.

Saag Lehti, Varul L, Scheib CL, Stenderup J, Allentoft ME, Saag Lauri, Pagani L, Reidla M, Tambets K, Metspalu E, jt 2017. Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Curr Biol.* 27(14):2185-2193.e6. doi:10.1016/j.cub.2017.06.022.

Sarkissian C Der, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, Koshel S, Zaporozhchenko V, Gronenborn D, Moiseyev V, Kolpakov E, jt 2013. Ancient DNA Reveals Prehistoric Gene-Flow from Siberia in the Complex Human Population History of North East Europe. *PLOS Genet.* 9(2):1–17. doi:10.1371/journal.pgen.1003296.

Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas A-S, Manica A, Moltke I, Albrechtsen A, Ko A, Margaryan A, Moiseyev V, jt 2014. Genomic structure in Europeans dating back at least 36,200 years Andaine. *Science (80- ).* 346(6213):1113–1118.

Semino O, Skrabanek L, Wolfe KH, Opini C, Dev G, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, jt 2000. The Genetic Legacy of Paleolithic Homo sapiens sapiens in Extant Europeans : A Y Chromosome Perspective. *Science (80- ).* 290(November):1155–1159.

doi:10.1126/science.290.5494.1155.

Sherman RM, Forman J, Antonescu V, Puiu D, Daya M, Rafaels N, Boorgula MP, Chavan S, Vergara C, Ortega VE, jt 2019. Assembly of a pan-genome from deep sequencing of 910 humans of African descent. *Nat Genet.* 51(1):30–35. doi:10.1038/s41588-018-0273-y.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, jt 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature.* 423(6942):825–837. doi:10.1038/nature01722.

Sun C, Skaletsky H, Birren B, Devon K, Tang Z, Silber S, Oates R, Page DC. 1999. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat Genet.* 23(4):429–432. doi:10.1038/70539.

Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogväli E-L, Tolk H-V, Reidla M, Metspalu E, Pliss L, jt 2004. The Western and Eastern Roots of the Saami—the Story of Genetic “Outliers” Told by Mitochondrial DNA and Y Chromosomes. *Am J Hum Genet.* 74(4):661–682. doi:10.1086/383203.

Tambets K, Yunusbayev B, Hudjashov G, Ilumäe AM, Rootsi S, Honkola T, Vesakoski O, Atkinson Q, Skoglund P, Kushniarevich A, jt 2018. Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biol.* 19(1):1–20. doi:10.1186/s13059-018-1522-1.

Thomson R, Pritchard JK, Shen P, Oefner PJ, Feldman MW. 2000. Recent common ancestry of human Y chromosomes: Evidence from DNA sequence data. *Proc Natl Acad Sci.* 97(13):7360–7365. doi:10.1073/pnas.97.13.7360.

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-sforza LL, Oefner PJ. 1997. Detection of Numerous Y Chromosome Biallelic Polymorphisms by Denaturing High-Performance Liquid Chromatography. *Genome Res.* 7:996–1005.

Underhill PA, Kivisild T. 2007. Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annu Rev Genet.* 41(1):539–564. doi:10.1146/annurev.genet.41.110306.130407.

Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CET, Semino O, Battaglia V, jt 2010. Separating the post-Glacial coancestry of European and

Asian y chromosomes within haplogroup R1a. *Eur J Hum Genet.* 18(4):479–484.  
doi:10.1038/ejhg.2009.194.

Underhill PA, Poznik GD, Rootsi S, Järve M, Lin AA, Wang J, Passarelli B, Kanbar J, Myres NM, King RJ, jt 2015. The phylogenetic and geographic structure of Y-chromosome haplogroup R1a. *Eur J Hum Genet.* 23(1):124–131. doi:10.1038/ejhg.2014.50.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3 — new capabilities and interfaces. *40(15):1–12.* doi:10.1093/nar/gks596.

Willerslev E, Cooper A. 2005. Review Paper. Ancient DNA. *Proc R Soc B Biol Sci.* 272(1558):3–16. doi:10.1098/rspb.2004.2813.

Xue Y, Zerjal T, Bao W, Zhu S, Lim S-K, Shu Q, Xu J, Du R, Fu S, Li P, jt 2005. Recent Spread of a Y-Chromosomal Lineage in Northern China and Mongolia. *Am J Hum Genet.* 77(6):1112–1116. doi:10.1086/498583.

YCC. 2002. A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups. *Genome Res.* 12(2):339–348. doi:10.1101/gr.217602.

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhover W, Fretwell N, Jobling MA, Harihara S, jt 1997. Genetic Relationships of Asians and Northern Europeans, Revealed by Y-Chromosomal DNA Analysis. *Am J Hum Genet.* 60:1174–1183.

Zerjal T, Xue Y, Bertorelle G, Wells RS, Bao W, Zhu S, Qamar R, Ayub Q, Mohyuddin A, Fu S, jt 2003. The Genetic Legacy of the Mongols. *Am J Hum Genet.* 72(3):717–721.  
doi:10.1086/367774.

## Kasutatud veebiaadressid

- [1] <https://www.thermofisher.com/order/catalog/product/EF0651> (viimati kasutatud 25.4.2019)
- [2] <https://technelysium.com.au/wp/chromaspro/> (viimati kasutatud 25.4.2019)
- [3] <http://www.ensembl.org> (viimati kasutatud 05.09.2018)
- [4] <http://primer3.ut.ee/> (viimati kasutatud 05.09.2018)
- [5] <http://bioinfo.ut.ee/genometester/> (viimati kasutatud 05.09.2018)
- [6] <https://www.ncbi.nlm.nih.gov/tools/primer-blast/> (viimati kasutatud 05.09.2018)
- [7] <https://isogg.org/> (viimati kasutatud 1.05.2019)
- [8] <https://www.yfull.com/tree/> (viimati kasutatud 1.05.2019)
- [9] <https://github.com/beast-dev/tracer/releases> (viimati kasutatud 1.05.2019)
- [10] <http://tree.bio.ed.ac.uk/software/figtree/> (viimati kasutatud 1.05.2019)

## Lisad

### Lisa 1. Eestlaste valimi (N=49) tüpiseerimise tulemused.

Hg Viimane defineeriv marker	E- M78	I1- M253	I2- M423	I2a1b- M423	N3a3a- L550,	N3a3b- VL39	N3a4- Z1936	R1a1d -M458	R1a1c- M558	R1a1- Z282	R1a- M198(x- Z282,Z93)	R1a3- CTS4385	R1b- M412	Kokku
Indiviidide arv	2	12	1	2	4	7	2	1	8	4	1 *	1	4	49
Sagedus (%)	(4,1)	(24,5)	(2,0)	(4,1)	(8,2)	(14,3)	(4,1)	(2,0)	(16,3)	(8,2)	(2,0)	(2,0)	(8,2)	(100)

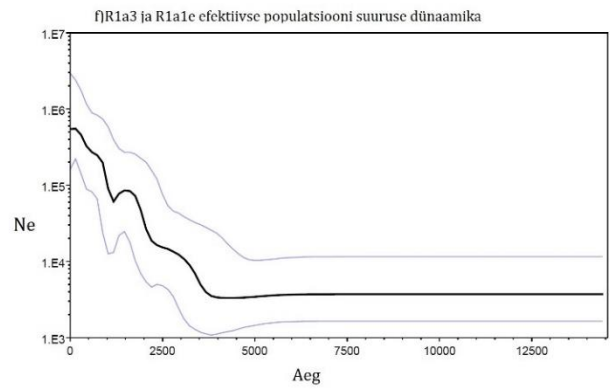
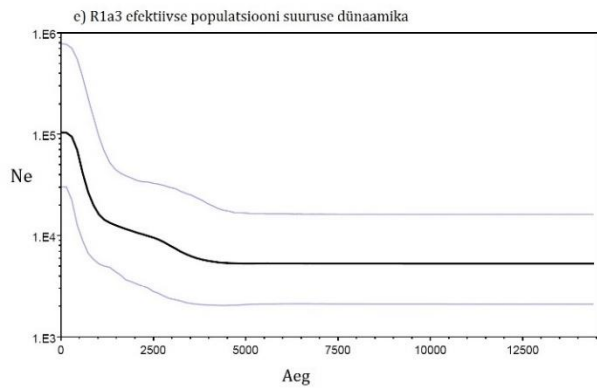
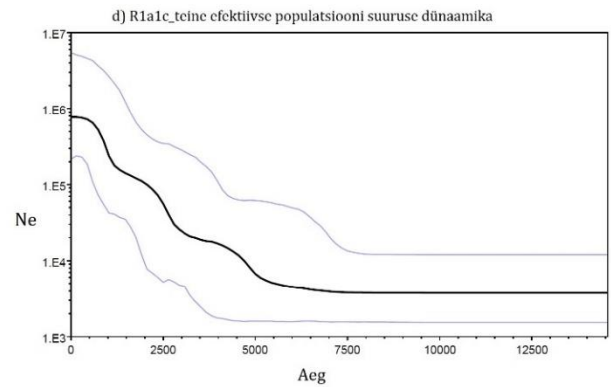
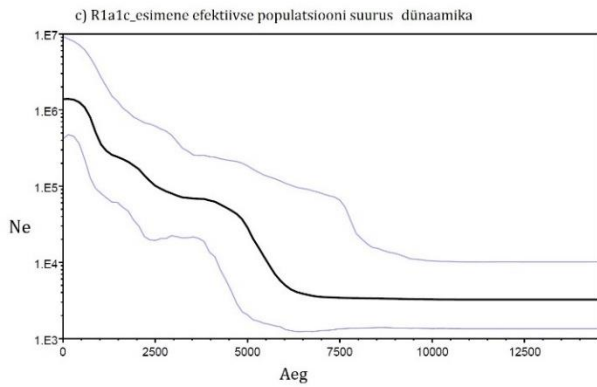
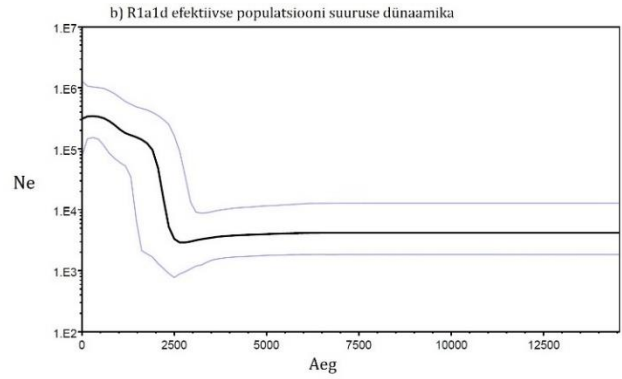
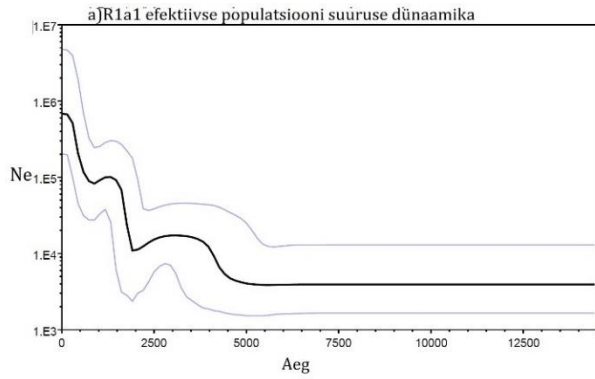
\*DNA vähesuse tõttu jäi see proov lõpuni tüpiseerimata.

**Lisa 2. Töös kasutatud Y-kromosoomi täisjärjestused (n=556), mida kasutati Y-kromosoom R1a alamharude jaotuse väljaselgitamiseks.**

Populatsioon	N	n	R1a1*- Z283	R1a1a – B145	R1a1b – Z685	R1a1c – M558	R1a1d – Z2911/ M458	R1a1f*- S24902	R1a1e – Z284	R1a2 – Z93	R1a3 – CTS4385	Allikas
Eestlased	1226	395	2	32	53	203	74	6	6	16	3	BigY kinnipüügimeetodil sekveneeritud proovid ja Illumina täisjärjestused
Soomlased	507	38	0	4	2	8	8	0	14	0	3	BigY kinnipüügimeetodil sekveneeritud proovid
Rootslased	746	123	3	0	1	4	8	0	95	0	12	BigY kinnipüügimeetodil sekveneeritud proovid

\* Karmin jt. 2015 poolt defineerimata alamklaad

### Lisa 3. Y-kromosoomi täisjärjestuste analüüsi põhjal leitud efektiivse populatsiooni suuruste dünaamika.



# LIHTLITSENTS

## Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Laura Laks,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Inimese Y-kromosoomi R1a alamhaplogruppide levik Põhja-Euroopas ja Eestis“, mille juhendajad on Siiri Rootsi ja Kristiina Tambets,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace alates 31.12.2022 kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

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

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

*Laura Laks*

**27.05.2019**