

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

Eesti keskaegse populatsiooni struktuur vana DNA vaatevinklist

Magistritöö

30 EAP

Kadri Irdt

Juhendajad

PhD Lehti Saag

PhD Kristiina Tambets

TARTU 2020

## **Eesti keskaegse populatsiooni struktuur vana DNA vaatevinklist**

**Lühikokkuvõte:** Rändamine on inimestele läbi aegade omane olnud. Üheks näiteks sellest on sakslaste saabumine Eestisse keskajal ristiustamise käigus. Eesti keskaegset rahvastikku on uurinud muuhulgas arheoloogid ja ajaloolased, kuid geneetiliselt on see periood seni kirjeldamata. Käesolevas töös uuriti keskaegse Eesti geneetilist struktuuri. Selleks eraldati DNA 30 keskaegse indiviidi hammastest. Tulemustest selgus, et keskaegsed eestlased sarnanesid enim tänapäeva uurali keeli kõnelevate populatsioonide ning Eesti pronksi- ja rauaaja populatsioonidega. Lisaks nähti, et Tartu toomkirikusse ja osad Tartu Jaani kirikusse maetud individid olid kõige sarnasemad hoopis tänapäeva lääneeurooplaste, sh sakslastega, mis on kooskõlas arheoloogide ja ajaloolaste teadmistega neisse surnuaedadesse maetutest. Lähemalt vaadati ka nn. „Siberi“ geneetilise komponendi olemasolu keskaegses Eestis – siinsetel nn. sakslastel see praktiliselt puudus, samas kui eestlastel oli see komponent olemas juba rauaajal ning on leitav ka tänapäeva eestlastel. Saadud tulemused annavad esmase ülevaate keskaegse Eesti elanikkonna geneetilisest struktuurist.

**Märksõnad:** Eesti keskaeg, vana DNA, arheogeneetika, populatsioonigeneetika

**CERCS:** B220 Geneetika, tsütogeneetika, H220 Keskaja ajalugu, H340 Arheoloogia

## **Population structure of Medieval Estonia through aDNA perspective**

**Abstract:** People have been prone to migrate throughout time. An example of this is the arrival of Germans in medieval Estonia. The medieval population of Estonia has been studied by archaeologists and historians among others but genetic studies have not been conducted so far. This thesis focuses on the genetic structure of medieval Estonia. DNA was extracted from the teeth of 30 individuals. The results showed that medieval Estonians were most similar to modern-day Uralic speaking populations and Estonian Bronze and Iron Age people. People buried in Tartu Cathedral and some in St. John's church were, however, most similar to modern Germans, which is in concordance with archaeological and historical knowledge. Furthermore, a “Siberian” genetic component was almost absent in medieval Estonian Germans but has been present in Estonia from the Iron Age. This thesis gives a first glance of the genetical structure of medieval Estonians.

**Keywords:** Medieval Estonia, ancient DNA, archaeogenetics, population genetics

**CERCS:** B220 Genetics, cytogenetics, H220 Medieval history, H340 Archaeology

## Sisukord

Kasutatud lühendid .....	6
Sissejuhatus .....	8
1.1 Populatsioonigeneetilised uuringud inimese demograafilise ajaloo rekonstrueerimisel .....	9
1.1.1 Mitokondriaalne DNA ja Y kromosoom.....	10
1.1.2 Autosomaalne DNA .....	11
1.2 Eesti demograafiline ajalugu Euroopa kontekstis.....	11
1.2.1 Kütt-korilased.....	11
1.2.2 Põlluharijad .....	12
1.3 Eesti keskaeg.....	16
1.3.1 Eesti keskaegne kiriku- ja matmiskultuur .....	16
1.3.1.1 Matmiskultuur külades .....	17
1.3.1.2 Matmiskultuur linnades .....	18
1.3.1.2.1 Tartu kirikud ja matused.....	19
2. Eksperimentaalosa .....	21
2.1 Töö eesmärgid.....	21
2.2 Materjal ja meetodika .....	21
2.2.1 Bioarheoloogiline materjal vana DNA analüüsiks.....	21
2.2.2 Materjali kataloogimine .....	23
2.2.3 DNA eraldamine .....	23
2.2.3.1 Saastuse vältimine .....	23
2.2.3.2 Proovi võtmine DNA analüüsiks.....	24
2.2.3.3 Pinnasaastuse eemaldamine ja DNA eraldamine hambajuurtest.....	25
2.2.3.4 DNA ekstrakti puhastamine.....	25
2.2.4 DNA sekveneerimisraamatukogude koostamine .....	26
2.2.4.1 DNA fragmentide üheaheelaliste otste täitmine ja reaktsioonisegust puhastamine .	26

2.2.4.2 Adapterite lisamine ja reaktsioonisegust puhastamine .....	27
2.2.4.3 Adapterite otste täitmine.....	27
2.2.4.4 Indeksite lisamine, DNA-fragmentide amplifitseerimine ja reaktsioonisegust puhastamine .....	27
2.2.5 DNA sekveneerimine .....	28
2.2.6 Sekveneeritud järjestuse joondamine referentsjärjestusele .....	28
2.2.6.1 Järjestuste ettevalmistamine joondamiseks .....	28
2.2.6.2 Järjestuste referentsile joondamine.....	28
2.2.7 Järjestuste <i>in silico</i> kvaliteedikontroll .....	29
2.2.8 Genotüüpide määramine .....	30
2.2.9 MtDNA ja Y kromosoomi haplogrupid määramine.....	31
2.2.10 Võrdlusandmetega liitmine .....	31
2.2.11 Populatsioonigeneetilised analüüsid .....	32
2.2.11.1 Peakomponentanalüüs .....	32
2.2.11.2 Sugulusanalüüs .....	33
2.2.11.3 f3 ja D statistikud.....	34
2.2.11.4 Admixture.....	34
2.2.11.5 qpAdm ja CP/NNLS .....	35
2.3 Tulemused ja arutelu.....	36
2.3.1 Uuritavate proovide kvaliteet ja sekveneerimise tulemused .....	36
2.3.1.1 Soo määramine ning mtDNA ja Y kromosoomi haplogrupid.....	37
2.3.2 Populatsioonigeneetiline analüüs .....	38
2.3.2.1 Peakomponentanalüüs .....	38
2.3.2.2 Sugulusanalüüs .....	41
2.3.2.3 f3 ja D statistikud.....	43
2.3.2.4 Admixture analüüs.....	46

2.3.2.5 qpADM ja ChromoPainter/NNLS .....	48
Kokkuvõte .....	50
Summary.....	51
Tänuõnad.....	53
Kasutatud kirjanduse loetelu .....	54
Kasutatud veebiaadressid .....	63
Lihtlitsents .....	90

## **Kasutatud lühendid**

AMH – anatoomiliselt moodsad inimesed (ingl. k. *anatomically modern humans*)

ap – aluspaar

BA – pronksiaeg (ingl. k. *Bronze Age*)

CCC – kammkeraamika kultuur (ingl. k. *Comb Ceramic culture*)

CHG – Kaukaasia kütt-korilased (ingl. k. *Caucasus hunter-gatherers*)

ChL – eneoliitikum e. vase-kiviaeg (ingl. k. *Chalcolithic*)

CP/NNLS – ingl. k. *ChromoPainter/non-negative least squares*

CWC – nöörikeramika kultuur (ingl. k. *Corded Ware culture*)

EBA – vanem pronksiaeg (ingl. k. *Early Bronze Age*)

EF – varajased põlluharijad (ingl. k. *early farmers*)

EHG – ida kütt-korilased (ingl. k. *eastern hunter-gatherers*)

EMBA – vanem/keskmise pronksiaeg (ingl. k. *Early/Middle Bronze Age*)

EN – varaneoliitikum (ingl. k. *Early Neolithic*)

Est – eestlased (ingl. k. *Estonians*)

Ger – sakslased (ingl. k. *Germans*)

HG – kütt-korilased (ingl. k. *hunter-gatherers*)

IA – rauaaeg (ingl. k. *Iron Age*)

IMA – raua-/keskaeg (ingl. k. *Iron/Middle Ages*)

LBK – paelkeramika kultuur (saksa k. *Linearbandkeramik*)

LN – hilisneoliitikum (ingl. k. *Late Neolithic*)

LNBA – hilisneoliitikum/pronksiaeg (ingl. k. *Late Neolithic/Bronze Age*)

MA – keskaeg (ingl. k. *Middle Ages*)

Med – keskaeg (ingl. k. *medieval*)

MLBA – keskmine/noorem pronksiaeg (ingl. k. *Middle/Late Bronze Age*)

MN – keskneoliitikum (ingl. k. *Middle Neolithic*)

mtDNA – mitokondriaalne DNA

N – neoliitikum (ingl. k. *Neolithic*)

NA – pole saadaval (ingl. k. *not available*)

NaOCl – naatrium hüpoklorit

NCA – neoliitikum/vase-kiviaeg (ingl. k. *Neolithic/Copper Age*)

PCA – peakomponentanalüüs (ingl. k. *principal component analysis*)

PCR – polümeraasi ahelreaktsioon (ingl. k. *polymerase chain reaction*)

rpm – pööret minutis (ingl. k. *revolutions per minute*)

SE – standardviga (ingl. k. *standard error*)

SHG – Skandinaavia kütt-korilased (ingl. k. *Scandinavian hunter-gatherers*)

SNP – üksiknukleotiidne polümorfism (ingl. k. *single nucleotide polymorphism*)

T – Tartu

TJ – Tartu Jaani kirik

TM – Tartu Maarja kirik

TRB – lehterpeekrite kultuur (saksa k. *Trichter-rand-becherkultur*)

TT – Tartu toomkirik

TÜ GI – Tartu Ülikooli genoomika instituut

WHG – lääne kütt-korilased (ingl. k. *western hunter-gatherers*)

## Sissejuhatus

Läbi aegade on inimestele omane olnud pidev liikumine ja uute maade avastamine. Rahvaste rändamine on toimunud sadu tuhandeid aastaid ning see on jätnud oma jälje nii kultuuri kui ka geenidesse. Erinevate rahvaste geneetilised komponendid on põlvkonniti edasi pärandunud ning osad nendest jälgedest on leitavad ka tänapäeva inimestel. Need komponendid, mis tänapäevasesse geenipärandisse ei ole jõudnud, kuid olid iseloomulikud vanade kultuuride kandjatele, on võimalik uuesti üles leida vana DNA-d uurides.

Eesti keskaega on uuritud muuhulgas nii ajaloolaste kui arheoloogide poolt, kuid geneetiliselt on keskaegne rahvastik veel uurimata. Viimastel aastatel on tehtud teadustöid, kus on lähemalt vaadeldud Eesti muinasaegseid kütt-korilasi ja põlluharijaid ning ka pronksi- ja rauaaegseid inimesi. See on andnud olulist informatsiooni rahvaste rännete ning geneetiliste iseärasuste kohta.

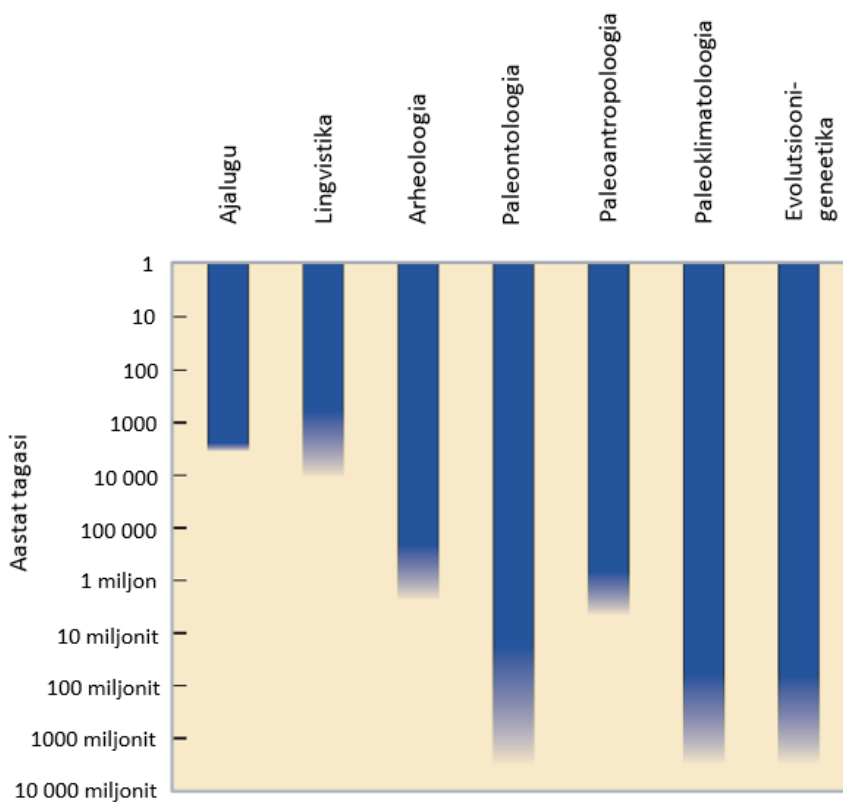
Keskaegses Eestis elas erineva etnilise taustaga inimesi. Suurema osa elanikkonnast moodustasid alamklassi kuuluvad eestlastest talupojad ning ülemklassi moodustasid sakslased. Rahva kihistumist on näha ka matusepaikades, kus ülemklassi inimesed said endale lubada matusekohta isegi piiskopkonna toomkirikus, samas kui lihtrahva jaoks olid peamisteks matusepaikadeks kihelkonna kirikuaiad ja külakalmistud. Sellist rahvaste jagunemist võiks näha ka geneetiliselt.

Käesolev töö on osa käimasolevast Eesti keskaega uurivast projektist. Antud töös vaadatakse väikest osa Eesti keskaegsest elanikkonnast. Töö eesmärgiks oli lähemalt uurida ning iseloomustada keskaegsete Eesti elanike geneetilist struktuuri. Selleks eraldati ja sekveneeriti DNA keskaegsete Eesti ala indiviidide hammastest. Saadud andmeid võrreldi teiste ammuste ja tänapäevaste populatsioonide andmetega. Lisaks uuriti lähemalt erineva sotsiaalse taustaga linnaelanike geneetilist päritolu keskaegse Tartu näitel ning keskaegsete inimeste sarnasust muinasaja ja tänapäevaste populatsioonidega. Lähemalt vaadeldi ka „Siberi“ komponendi olemasolu ja jaotust Eesti keskaegses populatsioonis, varasemalt tuvastati see tänaastege eestlaste seas levinud komponent meie aladel esmakordselt Eesti rauaaegsetes tarandkalmetes. Samuti uuriti võimalikke sugulussidemeid uuringusse kaasatud indiviidide vahel.

# 1. Kirjanduse ülevaade

## 1.1 Populatsioonigeneetilised uuringud inimese demograafilise ajaloo rekonstrueerimisel

Aja jooksul toimunud geneetilised muutused on salvestunud inimese genoomi ning nende kaudu on võimalik uurida inimese evolutsioonilist ajalugu (Jobling et al., 2014). Ühendades omavahel arheoloogia, geneetika ja keeleteaduse, on võimalik uurida inimkonna ajaloo erinevaid tahke, näiteks inimeste migratsioone ning kultuuride ja keelte levikut (Haak et al., 2015). Üks selline interdistsiplinaarne teadusharu on ka arheogeneetika, mis uurib inimpopulatsioonide minevikku, kasutades selleks erinevaid geneetilisi meetodeid (Soares et al., 2010). Ainult ühele valdkonnale keskendudes ei saa me piisavalt informatsiooni tervikpildi kokkupanekuks (joonis 1). Erinevad teadusharud täiendavad üksteist ning annavad infot erinevas ajaloolises sügavuses toimunud sündmuste kohta.



**Joonis 1. Erinevad teadusharud minevikusündmuste uurimiseks ning nende ajas tagasivaatamise võimekus (Jobling et al., 2014).**

Täna aitab kaasaegse geneetilise mitmekesisuse kõrval inimpopulatsioonide arengut ja migratsioone uurida vana DNA, mille abil saab vaadelda populatsioonidega ajas toimunud geneetilisi muutusi (Jobling et al., 2014; Mathieson et al., 2015). Vana DNA-ks nimetatakse surnud

organismist eraldatud DNA-d, milles on hakanud toimuma erinevad lagunemise protsessid, mistõttu on geneetilise info hankimine arheoloogilisest materjalist tehniliselt märksa keerukam (Herrmann ja Hummel, 1994). Esimesed vana DNA uuringud tehti mitokondriaalse DNA (mtDNA) põhjal (Gilbert et al., 2008; Poinar et al., 2006). Kuna mtDNA-d on rakus palju (~500 mtDNA molekuli inimese raku kohta), siis on see molekul heaks sihtmärgiks vana DNA uuringute tegemiseks (Handt et al., 1994). Suur koopiate arv tagab suurema tõenäosuse, et vaatamata surmajärgselt tekkivatele DNA kahjustustele (Handt et al., 1994), millest üheks levinuimaks on tsütosiini deamineerimine uratsiiliks (Briggs et al., 2007; Hofreiter, 2001), säilib piisavalt geneetiliseks analüüsiks sobivat materjali.

### **1.1.1 Mitokondriaalne DNA ja Y kromosoom**

MtDNA on väike rõngaskromosoom energia tootmise eest vastutavas rakuorganellis mitokondris, mis pärandub järglastele edasi ainult emaliinis (Wallace et al., 1999). Inimese mtDNA sisaldab umbes 16 500 aluspaari (ap) (Anderson et al., 1981). Uute mutatsioonide tekkimine mtDNA-s on muude genoomi osadega võrreldes üsnagi kiire:  $\sim 1.8899 \times 10^{-7}$  mutatsiooni ap kohta aastas (Soares et al., 2009). Võrdluseks, autosomaalse genoomi keskmine mutatsioonikiirus on  $\sim 10^{-8}$  ap generatsiooni kohta (M. Jobling et al., 2014) ehk  $\sim 3,3 \times 10^{-10}$  mutatsiooni ap kohta aastas. Kuna mtDNA pärandub põlvkondade lõikes edasi ainult emaliinis, siis on selle abil võimalik uurida täpsemalt naisliini demograafilist ajalugu (Torrioni et al., 2006). Isalt pojale kanduva Y kromosoomi (mutatsioonikiirus:  $\sim 0,74 \times 10^{-9}$  mutatsiooni ap kohta aastas (Karmin et al., 2015)) abil on omakorda võimalik uurida meesliini ajalugu. MtDNA ja Y kromosoom on mõlemad haploidsed ning ei rekombineeru, mis teeb nende genoomi osade uurimise autosoomidega võrreldes lihtsamaks – ajas tekkivad uued variandid on võimalik paigutada liinidevahelisi sugulussuhteid peegeldavale fülogeneesipuule (Jobling ja Tyler-Smith, 2003; Kivisild, 2017).

Erinevatele geograafilistes piirkondades on omased teatud kindlad mtDNA ja Y kromosoomi variatsioonid (Torrioni et al., 2006; Wallace, 1995). Nende variantide klassifitseerimisel kasutatakse erinevaid mutatsioone millel on fülogeneesipuu harude moodustamisel erinev kaal (Karafet et al., 2008; van Oven ja Kayser, 2009). Monofüleetiline klaad ehk kõiki tütar tüüpe ja nende ühist eellast sisaldav grupp moodustab haplogrupi (Richards et al., 1998). Uurides mingi populatsiooni mtDNA ja Y kromosoomi erinevaid markereid ja määrates ära indiviidide haplogrupilise kuuluvuse ning haplogruppide esinemissageduse ühes populatsioonis, saab neid tulemusi võrrelda teiste populatsioonidega ja sellest lähtuvalt oletada, kuidas eri rahvad on aja jooksul liikunud (Jobling ja Tyler-Smith, 2003; Kivisild, 2017; Wallace et al., 1999).

### 1.1.2 Autosomaalne DNA

MtDNA ja Y kromosoom võimaldavad uurida fülogeneesi ainult ühe soo piires ning piiranguid seab ka nendes sisalduva info suhteliselt väike hulk kogu genoomi mitmekesisust arvestades. Selleks, et saada paremat ülevaadet kogu populatsioonist, uuritakse tänapäeval inimese demograafilise ajaloo rekonstrueerimiseks nende genoomi osade kõrval ka autosoome (Lazaridis et al., 2014; Mathieson et al., 2015). Tänu rekombinatsioonile kujutab autosomaalne DNA endast paljusid eri segmente, millel igal on oma ajalugu. Seega annab autosomaalse DNA uurimine informatsiooni kogu populatsiooni kohta (Colonna et al., 2011).

MtDNA, Y kromosoom ja autosomaalne DNA pakuvad demograafilise ajaloo uurimisel üksteist täiendavat informatsiooni. Kuna Y kromosoomi ja mtDNA pärandumise eripärad teevad nad nõ. vastuvõtlikuks juhuslikele sündmustele – geneetilise triivi tõttu võivad mõned liinid populatsioonis suhteliselt kiiresti kinnistuda või sealt kaduda – siis ainult tänapäevaste inimeste uurimisega ajalooliste liinide kohta kogu informatsiooni kätte ei saa. Siinkohal ongi abiks vana DNA uuringud, mis võimaldavad „üles leida“ ka tänaseks kadunud Y kromosoomi ja mtDNA liine (Colonna et al., 2011).

## 1.2 Eesti demograafiline ajalugu Euroopa kontekstis

### 1.2.1 Kütt-korilased

Esimesed anatoomiliselt moodsad inimesed (ingl. k. *anatomically modern humans* (AMH)) (*Homo sapiens*) jõudsid Euroopasse umbes 45 000 aastat tagasi ning asustasid üsna kiirelt kogu elamiskõlbliku ala Euroopas (Benazzi et al., 2011; Kriiska ja Tvauri, 2002). Vahepealsed külmaperioodid ja mandrijää laienemine sundis inimesed taanduma Lääne-Euroopasse ja Vene lavamaale. Pärast viimase jääaja lõppemist alustasid Euroopa taasasustamist Edela-Prantsusmaa Madeleine'i kultuuri kandjad, kes liikusid umbes 11 000 aastat tagasi Kesk-Euroopasse (Kriiska ja Tvauri, 2002).

Euroopas elanud kütt-korilasi on võimalik geneetiliselt jaotada kolme gruppi: ida (Haak et al., 2015; Mathieson et al., 2015; Mittnik et al., 2018), lääne (Lazaridis et al., 2014) ja Skandinaavia kütt-korilased (Skoglund et al., 2012). Euroopa kütt-korilaste seas oli valdavaks mtDNA haplogrupiks U (Fu et al., 2013; Krause et al., 2010; Posth et al., 2016). Tänapäeval võib seda kõige rohkem leida Põhja-Euroopas (Simoni et al., 2000). Y kromosoomide haplogruppide osas oli kütt-korilaste mitmekesisus suurem (Kivisild, 2017).

Läänemere idarannikule jõudsid esimesed inimeste grupid keskmisel kiviajal ehk mesoliitikumis umbes 9000 a eKr (Kriiska et al., 2017). Mesoliitikumiaegsed elupaigad asusid enamasti jõgede ja järvede kallastel ning peamiseks elatusviisideks olid küttimine, kalapüüdmine ja korilus. Umbes 2000 aastat hiljem, hilises mesoliitikumis alustati asulate rajamist ka rannikualadele ja hakkasid tekkima väikesed, kuni paarikümne inimesega kogukonnad (Kriiska, 2009; Kriiska et al., 2017).

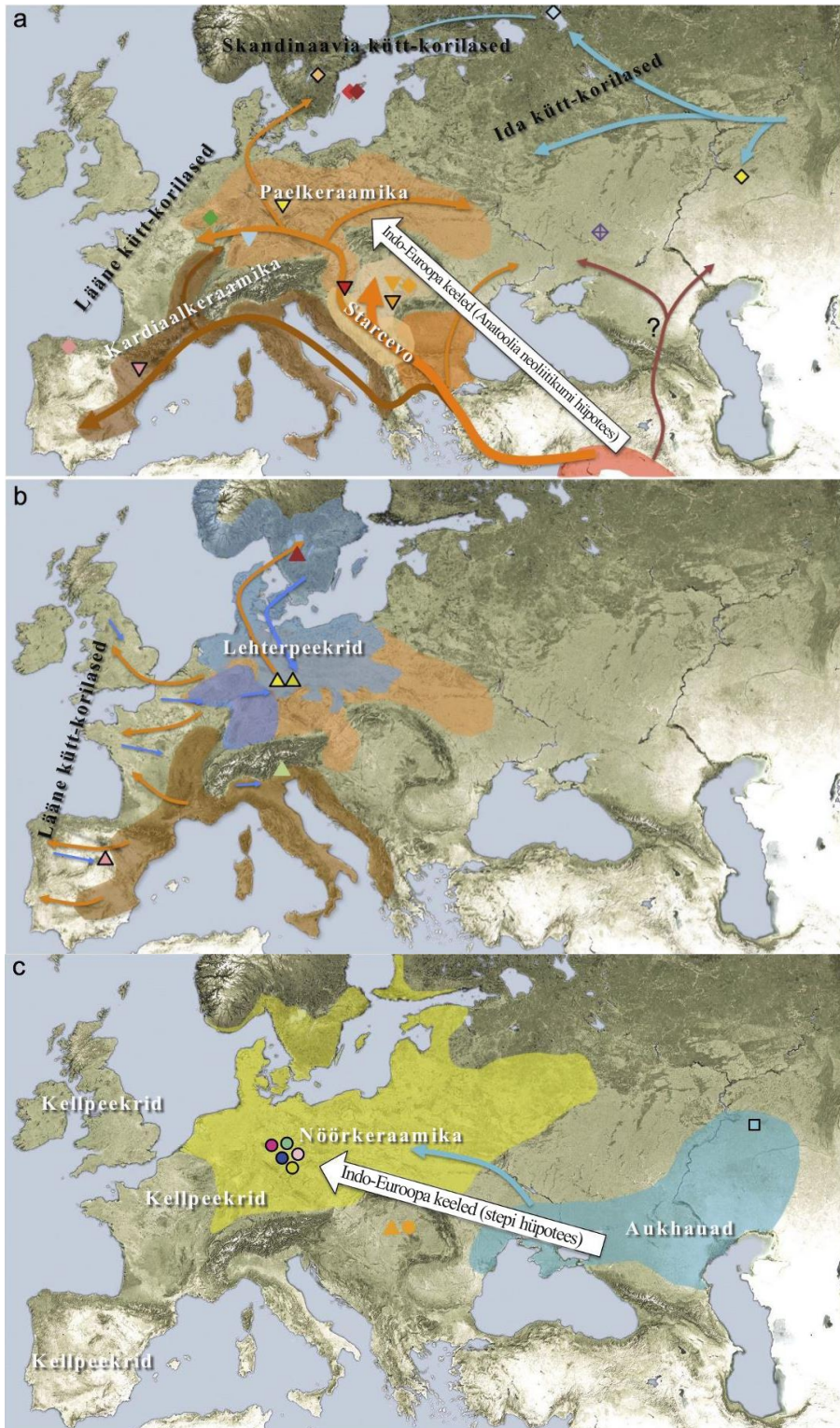
Umbes 4000 a eKr jõudis ida poolt Baltikumi aladele kammkeraamika kultuur, mis märkis ühtlasi mesoliitikumi lõppu ja noorema kiviaja ehk neoliitikumi algust (Kriiska, 2009; Kriiska et al., 2017). Uue kultuuri tulek idast on jättnud oma jälje ka siinsete alade geenipärandisse, kus on näha mõningast erinevust kahe ajaperioodi inimeste vahel. Kui mesoliitikumiaegsed Baltikumi kütt-korilased olid geneetiliselt sarnasemad Lääne-Euroopa kütt-korilastega, siis neoliitikumi kammkeraamika kultuuri indiviididel leiti rohkem sarnasusi Ida-Euroopa kütt-korilastega (Jones et al., 2017; Mittnik et al., 2018; Saag et al., 2017). Baltikumi mesoliitikumi ja neoliitikumi kütt-korilased kuulusid mtDNA haplogruppi U (Jones et al., 2017; Mathieson et al., 2018; Mittnik et al., 2018; Saag et al., 2017), mis on valdav emaliin kõigil Euroopa kütt-korilaste rühmadel (Posth et al., 2016). Y kromosoomi levinuimateks haplogruppideks Baltikumis olid R1 ja I2 (Jones et al., 2017; Mathieson et al., 2018; Mittnik et al., 2018; Saag et al., 2017).

### **1.2.2 Põlluharijad**

Põlluharimine ja karjakasvatus said alguse Lähis-Idas, Viljaka Poolkuu alal (Jobling et al., 2004) umbes 10 000 a eKr ning levisid sealt üle kogu Euroopa (Douglas Price, 2000; Fort, 2015; Pinhasi et al., 2005). Vaidlused selle üle, kas põlluharimise levikuga liikusid kaasa ka seda kultuuri kandvad inimesed, on olnud aktuaalne juba aastakümneid. Erinevad vana DNA uuringud on nüüdseks kinnitanud, et liikusid ka põlluharijad ise, mitte ainult kultuur (Lazaridis et al., 2016; Mathieson et al., 2018; Skoglund et al., 2012).

Umbes 7000 a eKr liikusid esimesed põlluharijad Balkani poolsaarele. Rändamine toimus nii mööda mandrit Kagu-Euroopasse, kui ka mööda Vahemere rannikut Lõuna-Euroopasse (joonis 2a), kust nad liikusid edasi põhja poole (joonis 2b) (Douglas Price, 2000). Ühe hüpoteesi kohaselt arvatakse, et koos esimeste põlluharijatega levisid Anatooliast Euroopasse ka indoeuroopa keeled (Anatoolia neoliitikumi hüpotees, joonis 2a) (Anthony, 2007; Renfrew, 1989). Teise hüpoteesi kohaselt seostatakse indoeuroopa keelte levikut ka hilisema stepimigratsioonide lainega (joonis 2c) (Anthony, 2007). Põhja-Euroopasse, Skandinaaviasse ja Briti saartele jõudis põlluharimine umbes 4000–3500 a eKr (Douglas Price, 2000; Fort, 2015). Põlluharimise levimist Euroopasse näitavad

ka taimede õietolmu põhjal tehtud uuringud, kus on muuhulgas näha ka metsade järkjärgulist vähenemist umbes 6000 aastat tagasi (Roberts et al., 2018).



**Joonis 2. Põlluharimiskultuuri levimine Euroopasse.** Varajaste põlluharijate levikuteed (a) Lähis-Idast Vahemere piirkondadesse (~9000–7000 aastat tagasi); (b) Põhja-Euroopasse (~7000–

5000 aastat tagasi) ja (c) stepi aladelt Euroopasse ~4500 aastat tagasi. Värviliste laikudena (valge kirjaga) on tähistatud erinevate kultuuride levilad. Valgete nooltega on tähistatud indoeuroopa keelte kaks võimaliku levikuteed Lähis-Idast Euroopasse: (a) Anatoolia neoliitikumi hüpotees ja (c) stepi hüpotees. Alajoonisel (c) tähistavad sinised nooled kütt-korilaste segunemist põlluharijatega. Kohandatud Haak et al., 2015 järgi.

Uurides Euroopa eri piirkondade neoliitikumiaegseid põlluharijaid, on näha suurt geneetilist sarnasust tänapäeva Vahemere piirkonna eurooplastega (eriti sardiinlastega) (Haak et al., 2010; Mathieson et al., 2018; Skoglund et al., 2012). Need inimesed erinesid geneetiliselt kohalikest kütt-korilastest (Skoglund et al., 2012), mis näitab, et põlluharimise kultuur ei ole tekkinud kohapeal, vaid on Euroopasse tulnud koos sissereändajatega. Rahvaste segunemisest ja liikumisest lõunast põhja poole annab märku ka neoliitiliste põlluharijate geneetilise pärandi vähenemine lõuna-põhja suunal (Skoglund et al., 2012). Enne põlluharijate tulekut oli Euroopas levinuimaks mtDNA haplogrupiks U (Fu et al., 2013; Krause et al., 2010; Posth et al., 2016). Põlluharijate tulekuga tuli juurde mitmeid erinevaid mtDNA haplogruppe: H, T, J (Bramanti et al., 2009; Haak et al., 2005, 2010; Skoglund et al., 2012, 2014).

Põlluharimise leviku teine laine jõudis Euroopasse umbes 2800 a eKr, neoliitikumi lõpus, kui stepialadelt saabusid aukhaudade (ingl. k. *Yamnaya*) kultuuri esindajad, kes panid aluse ka nõörkeraamika kultuurile (joonis 2c) (Allentoft et al., 2015; Haak et al., 2015). Taimestikumuutuste põhjal tehtud uuringud näitavad Põhja- ja Kesk-Euroopas laialehiste metsade vähenemist ja karjamaa/rohumaa alade suurenemist samal ajal, mil sinna jõudsid aukhaudade kultuuri esindajad. Samuti näitavad need uuringud, et Lõuna-Euroopasse jõudis aukhaudade kultuur hiljem kui Põhja- ja Kesk-Euroopasse (Racimo et al., 2020). Aukhaudade kultuuri seostamist mõnede indoeuroopa keelte leviku kandjana Euroopasse (stepi hüpotees, joonis 2c) (Anthony, 2007) toetavad ka erinevad geneetilised uuringud (Allentoft et al., 2015; Haak et al., 2015). Euroopas toimus segunemine kohalike varajaste põlluharijate ja lisandunud stepirahvaste vahel (Haak et al., 2015; Saag et al., 2017). Nn. *Yamnaya* geneetiline komponent ehk stepikomponent on tänapäeval leitav kõikjal Euroopas: lõunas vähem, põhjas rohkem (Allentoft et al., 2015; Haak et al., 2015). Ka eestlaste geenipärandis on seda täna umbes 60% (Tambets et al., 2018). Mitokondri haplogruppides suuri muutusi enam ei toimunud. Küll aga hakkas laialdaselt levima Y kromosoomi haplogrupp R1a (Allentoft et al., 2015; Haak et al., 2015; Mitnik et al., 2018), mis on sage ka tänapäeva Euroopas (Underhill et al., 2010).

Baltikumi jõudis nöörikeramika kultuur sarnaselt ülejäänud Euroopaga umbes 2800 a eKr, tuues kaasa olulise muutuse elustiilis ja pannes alguse edasisele intensiivsele põlluharimisele ja karjakasvatusele (Kriiska, 2000; Kriiska ja Tvauri, 2002). Baltikumi nöörikeramika indiviidide sarnasust stepialade populatsioonidega on näidatud mitmetes geneetilistes uurimistöodes (Jones et al., 2017; Mittnik et al., 2018; Saag et al., 2017). Soospetsiifilistes uuringutes on leitud, et Eesti aladele jõudis stepirahvaste geneetiline pärand peamiselt meesliini kaudu, samas kui eri indiviidides varieeruva proportsiooniga Euroopa varajaste põlluharijate komponenti esineb rohkem naisliinis (Saag et al., 2017). Nii nagu mujal Euroopas, oli ka kammkeramika kultuuri kütt-korilaste seas levinuimaks mtDNA haplogrupiks Baltikumis U. Nöörikeramika ajal leidis aga juba rohkem erinevaid haplogruppe: H5a, T2a ja J1c (Saag et al., 2017).

Nooremaks pronksiajaks (1200–400 a eKr) oli Baltikumis taaskord toimunud kütt-korilaste geneetilise komponendi sageduse tõus. Võrreldes omavahel nöörikeramika ja pronksiaja indiviidide pärilikkusainet, leiti, et viimastel on rohkem sarnasusi lääne kütt-korilastega kui nöörikeramika indiviididel (Saag et al., 2019). Samasugust lääne kütt-korilaste komponendi sisaldust on leitud ka hilisematel indiviididel ja see on ka tänapäeva eurooplaste seas kõrgeim just Baltikumis (Lazaridis et al., 2014; Mittnik et al., 2018; Saag et al., 2019). Nöörikeramika kultuuri ajal Euroopas laialdaselt levinud Y kromosoomi haplogrupp R1a (Allentoft et al., 2015; Haak et al., 2015) oli valdav ka pronksiaegses Baltikumis (Mittnik et al., 2018; Saag et al., 2019). Pronksiaja rauaajal Baltikumis esinenud mtDNA haplogrupid on leitavad ka tänapäeval siinsetel elanikel (Mittnik et al., 2018; Saag et al., 2019).

Eesti ja Ingerimaa rauaegsetel indiviididel (500 a eKr – 450 a pKr) on tuvastatud 2–4% ulatuses arvatavalt Siberi päritolu geneetilist komponenti, mille sagedus on täna kõrgeim Lääne-Siberis elavatel nganassaanidel ja mida Eesti pronksiaegsetel indiviididel ei leidu (Saag et al., 2019). „Siberi“ komponent on leitav ka tänapäeva eestlastel (5%) ja enamikul teistel uurali keeli kõnelevatel rahvastel (Tambets et al., 2018). Rauaajal võis siinsetel elanikel lisaks juba varasemalt esinenud Y kromosoomi R1a haplogrupile leida ka N3a haplogruppi (Saag et al., 2019), mis on iseloomulik just uurali keelte kõnelejatele ja Siberi populatsioonidele (Ilumäe et al., 2016; Tambets et al., 2018). Selle isaliini levimine toimus Kirde-Euroopasse viimase 5000 aasta jooksul idast (Ilumäe et al., 2016).

Keskaegse Eesti elanikkonna kohta on geneetilisi andmeid väga vähe ja seetõttu ei osata selle kohta täpseid järeldusi veel teha. Küll aga on uurimisega juba algust tehtud (Saag et al., 2019).

### **1.3 Eesti keskaeg**

Lõuna- ja Kesk-Euroopas loetakse keskajaks perioodi aastatel umbes 500–1500 pKr. Eestisse jõudis keskaeg palju hiljem, alles 13. sajandil, kui Eesti alad vallutati ristisõdade käigus, ja kestis kuni Liivi sõjani ehk 16. sajandi keskpaigani. Keskaja algul kuulus Lõuna-Eesti koos Põhja-Lätiga Liivimaa koosseisu ja Põhja-Eesti Taani kuningriigile, kuid 14. sajandil läks ka Põhja-Eesti Liivimaa koosseisu (Kriiska et al., 2017).

Läbi kogu keskaja elas Eesti aladel mitmeid eri rahvuste esindajaid: nii eestlasi kui ka sakslasi, rootslasi, soomlasi, taanlasi ja venelasi. Rootsused, soomlased ja taanlased elasid valdavalt Põhja- ja Lääne-Eestis. Suurema osa linnaelanike ülemkihist moodustasid sakslastest kaupmehed ja käsitöölised. Maapiirkondades elasid vaesemad talupidajatest eestlased. Üldiselt jaguneski keskaegne Eesti rahvas kaheks: nõ „sakslased“, kes olid staatusest kõrgemad, samas aga arvuliselt vähemuses, ning „mittesakslased“, kes olid põliskeelt kõnelevad kohalikud, enamjaolt talupojad. Tavaliselt määras inimese seisuse ära see, kuhu perre ta sündis. Mõningatel juhtudel oli võimalik ühest seisusest teise liikumine. Ristisõja järgselt kaotasid paljud kohalikud eestlastest ülikud oma kõrgema seisuse. Põhja- ja Lääne-Eestis siiski mingil määral nende seisused säilisid. Osad kõrgema seisuse eestlased hakkasid läbi käima ka saksa aadlikega (Kriiska et al., 2017; Selart et al., 2012).

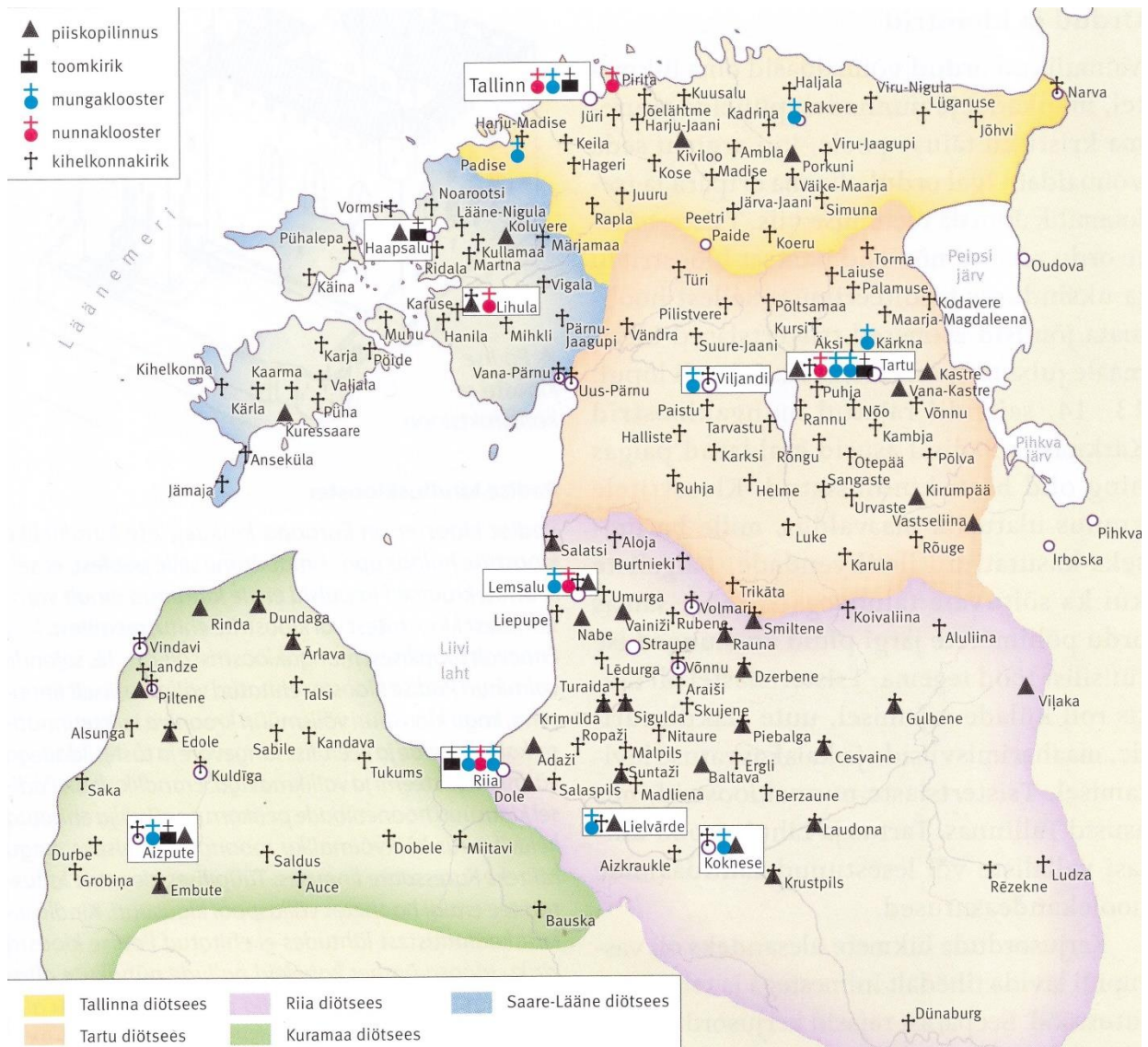
13. sajandi algul elas Eestis umbes 100 000–200 000 inimest ning 16. sajandi keskpaigaks oli see arv kasvanud ~250 000–300 000 inimeseni. Rahvaarv kasvas mingil määral ka sisserände tõttu, kuid suurem roll oli loomuliku iibe tõusul (Kriiska et al., 2017; Selart et al., 2012).

Keskaegne Eesti talupere koosnes tavaliselt talupidaja enda perest, sulastest ja vahel ka mõningatest kaugematest sugulastest. Keskmiselt elas taludes 6–8 täiskasvanut, vahel ka 10–12. Talumajapidamise juhiks oli taluperemees. Põhiliselt tegeleti karjakasvatuse ja põlluharimisega, vähesel määral ka küttimisega. Linnas käidi põhiliselt kauplemas (Kriiska et al., 2017).

#### **1.3.1 Eesti keskaegne kiriku- ja matmiskultuur**

Ristiusu tulekuga loodi Eesti aladele palju kirikuid ja kloostreid. Iga kihelkonna keskuses oli oma kirik (joonis 3), mis oli rahva mõtte- ja vaimumaaailma kujundajaks (Kriiska et al., 2017). Kirikute juurde rajati uued kristlikud kalmistud, kuid maapiirkondades jätkati matmist vanade kivikalmete alusel kujunenud külakalmistutele (Selart et al., 2012; Valk, 2003). Kirik oli osa inimese elust sünnist surmani, kuhu tavaliselt mindi kogu perega. Suuremate pühade ajal peeti kirikute juures ka laatu, kuhu tulid kohale nii talupojad kui ka linnarahvas (Kriiska et al., 2017).

Muutusid ka matmiskombed. Rauaajal oli põhiliseks matmisviisiks põletusmatus – surnute põletamine toimus väljaspool kalmet ning hiljem viidi põletatud luud ja kaasapandud esemed kalmu. Rauaaja lõpul hakkasid levima laibamatused, mis kujunesid keskajal valdavaks matmisviisiks (Kriiska ja Tvauri, 2002; Valk, 2003). Lisaks hakati surnuid matma peaga lääne või edela suunas, nagu see oli kombeks kristlikus Euroopas (Selart et al., 2012; Valk, 1995).



**Joonis 3. Piiskopkonnad, kloostrid ja kirikud 16. sajandi Liivimaal (Kriiska et al., 2017).**

### 1.3.1.1 Matmiskultuur küldes

Keskaegses Euroopas oli kirikaeda või kirikusse matmine ainuke õige viis (Sala, 2002), Liivimaal jätkati matmist ka külakalmistutele, kuid ka need kalmistud lasti kiriklikult pühitseda (Selart et al., 2012; Valk, 2003). Külakalmistud kujutasid endast tavaliselt lagendikel asuvaid väikseid puudega

kaetud alasid (Selart et al., 2012). Ristiusu tulekuga vähenes Saaremaal ning Lääne- ja Põhja-Eestis külakalmistute kasutamine, Lõuna-Eestis olid need aga endiselt levinud (Selart et al., 2012; Valk, 2003).

Lõuna-Eestis jätkati vana traditsiooniga, kus surnutele pandi kaasa erinevaid meeneid, peamiselt raha ja tarbeesemeid (Valk, 1994). Kristlikus Euroopas oli selline komme vähe levinud (Selart et al., 2012; Valk, 2003). Lisaks võis leida ka teisi erinevusi, näiteks võis Lõuna-Eestis leida matuseid, kus inimene ei olnud maetud peaga lääne või edela suunas, nagu oli kombeks kristlike matuste puhul (Selart et al., 2012; Valk, 1995).

Kui üldiselt tähendas kabel seost ristiusuga ning lahkunute andmist ristiusuga seotud pühakute kaitse alla, siis Tartumaal ja ka mujal Lõuna-Eestis võis kabel, lisaks kristlikule pühakojale, tähendada ka muid matusepaiku, kihelkonnasurnuaedadest loomakorjaste matmispaikadeni (Selart et al., 2012). Mõned sellised kabelid olid rajatud nõ. omavoliliselt ja nad ei kuulunud ametlikku kiriklikku ringkonda (Valk, 1994).

### **1.3.1.2 Matmiskultuur linnades**

Eesti aladel oli kolm piiskopkonda: Tartu, Tallinna ja Saare-Lääne piiskopkond (joonis 3). Igal piiskopkonnal oli oma toomkirik. Tartus ja Tallinnas võis leida ka 2 kihelkonnakirikut, väiksemates linnades oli neid tavaliselt üks (Kriiska et al., 2017).

Kirikute vahetus ümbruses paiknesid kihelkonnakalmistud. Tähtsamad ja jõukamad linnainimesed maeti kirikusse, ülejäänud aga kirikaedadesse (Selart et al., 2012). Nagu eelmises peatükis öeldud, jätkati mõningatel juhtudel vana traditsiooniga, kus surnutele pandi kaasa erinevaid meened, seda eelkõige Lõuna-Eestis (Valk, 1994). Rohkem tehti seda külates, kuid on leide ka linnadest. Meenete kaasapanek surnutele oli levinud nii kõrgemast kui alamast seisusest indiviidide puhul, olles taaskord erinev muust Euroopast, kus tavaliselt maeti inimesed vaid surilinnasse mähituna – erandiks võisid olla mõned ülikud ja piiskopid, kes maeti riietes (Selart et al., 2012; Valk, 2003). Sellised külakalmistutele omased ehete leiud kirikaedades viitavad tavaliselt eestlaste matustele (Selart et al., 2012; Valk, 2003) ja nende puudumine sakslastele (Valk, 2004). 16. sajandi alguses hakkas laialdasemalt levima komme panna surnutele kaasa mõni münt. Rohkem oli see levinud Lõuna-Eesti linnakalmistutel, nt Tartu Jaani kiriku kalmistul (Selart et al., 2012).

### 1.3.1.2.1 Tartu kirikud ja matused

Keskaegses Tartus oli kaks kihelkonnakirikut: Jaani kirik ja Maarja kirik (Altoa, 2009). Lisaks asus Tartus ka Tartu piiskopkonna toomkirik (Kriiska et al., 2017; Selart et al., 2012).

Tartu toomkiriku kalmistu oli matmispaigaks ülikutele ning teistele kõrgemast seisusest inimestele, kelleks võisid olla näiteks sõjamehed, vasallid või mõni aadelkonna liige (Malve, 2010). Piiskopkonna peakirikusse hauaplatsi ostmine oli kallid, seega said seda endale lubada ainult rikkamad (Malve ja Valk, 2009; Valk, 1995). Kuna enamik ülemklassi inimesi olid sakslased (Selart et al., 2012), siis võib oletada, et põhiliselt maeti sinna saksa rahvusest inimesi.

Püha Maarja kiriku kalmistu oli kogudusekalmistu, kuhu maeti igas vanuses mehi ja naisi. Uuritud keskaegsetest haudadest on leitud palju erinevaid hauapanuseid. Leidude seas oli erinevaid ehteid, sõlgesid, münte jm (Malve et al., 2011). Selliseid hauapanuseid on palju leitud erinevatelt Lõuna-Eesti külakalmistutelt (Valk, 1994, 1995), viidates seega eestlaste matustele. Eestlastele viitavad ka haudadest leitud sõled, mis olid nende seas levinud 13.–17. sajandil (joonis 4) (Malve et al., 2011; Valk, 1995), samas kui Saksamaal ja Skandinaavias oli nende kandmine 15. sajandiks juba lõppenud (Valk, 1995).



**Joonis 4. Maarja kiriku kalmistult leitud keskaegsed sõled.** 1) Hansasõlg (13.–14. saj.), 2) rööngassõlg (15.–17. saj.). Kohandatud Malve et al., 2011.

Jaani kiriku kalmistu esimeste arheoloogiliste ja demograafiliste uuringute käigus leiti erinevates vanustes nii meeste kui naiste matuseid ja ka laste matuseid (Kalling, 1995). Kuna ajaloolistes kirjalikes allikates ei leitud konkreetselt informatsiooni maetute päritolu või rahvuse kohta, siis

oletatavasti oli tegemist tavalise koguduse kalmistuga (Kalling, 1995; Malve, 2010). Sealt on leitud nii sakslaste (Kalling, 1995) kui eestlaste matuseid (Valk, 2004). Lähtudes kaasa pandud ornamentidest ja riietest on arvatud, et kirikusse olid maetud sakslased ning kirikuaeda eestlased (Valk, 2004). Meeste koljumõõtmete põhjal tehtud võrdlusuuringud näitavad, et Jaani kirikusse oli erinevatel ajaperioodidel maetud eri rahvusest inimesi. 13.–14. sajandi ning varauusaegsed (17.–18. sajandi) matused kuulusid põhiliselt läänest tulnud rahvastele, kes arvatavasti kuulusid ka kõrgemasse seisusse. 14. sajandi keskpaigast keskaja lõpuni võis aga rohkem leida eestlaste matuseid (Kalling, 1995).

## **2. Eksperimentaalosa**

### **2.1 Töö eesmärgid**

Antud töö eesmärgiks oli uurida ja iseloomustada Eesti keskaegse elanikkonna geneetilist struktuuri. Selleks tehti järgnevat:

1. Eraldati ja sekveneeriti DNA Eesti alal elanud keskaegsete indiviidide hammastest
2. Võrreldi saadud andmeid teiste vana DNA ja tänapäevaste populatsioonide andmetega, et
  - a. selgitada välja keskaegse Eesti elanikkonna geneetiline struktuur;
  - b. võrrelda erineva sotsiaalse taustaga linnaelanikke;
  - c. võrrelda keskaegse Eesti indiviidide sarnasust tänapäevaste rahvastega;
  - d. analüüsida hiljemalt varajasel rauaajal Eestisse jõudnud uurali keeli kõnelevatele rahvastele iseloomuliku „Siberi“ komponendi osakaalu ja jaotust keskaegses Eestis.
3. Uuriti keskaegsete indiviidide sugulussidemeid

### **2.2 Materjal ja metoodika**

#### **2.2.1 Bioarheoloogiline materjal vana DNA analüüsiks**

Uurimistöös kasutati Tartu Ülikooli ajaloo ja arheoloogia instituudi kogude arheoloogilist luumaterjali, mis pärineb erinevatelt arheoloogilistelt väljakaevamistelt üle Eesti (joonis 5, tabel 1). Meie koostööpartnerid selle projekti raames olid osteoloog Martin Malve ja arheoloog Dr. Heiki Valk Tartu Ülikooli ajaloo ja arheoloogia instituudist.



**Joonis 5. Keskaegsete indiviidide (n=30) leiukohtade kaart.** Kaardi tegemiseks kasutati programmi QGIS 3.10 (*QGIS Development Team (2020). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>*), aluskaart on pärit Maa-ameti geoportaali leheküljelt (Haldus- ja asustusjaotus: Maa-amet 01.03.2020).

DNA eraldati 30 hambast, mis olid pärit Eesti alal elanud keskaegsetelt indiviididelt ning üks ka varauusajal elanud indiviidilt (edaspidi viidatakse proovidele ühiselt kui keskaegsetele) (tabel 1).

**Tabel 1. Vana DNA uuringus kasutatud indiviidid, nende leiukohad (maakond, kalmistu) ja tüpokronoloogilised dateeringud.**

#	Indiviid	Leiukoht		Dateering
		Maakond	Kalmistu	
1	KAU003	Viljandimaa	Kaubi külakalmistu	16. sajandi IV veerand
2	KAU004	Viljandimaa	Kaubi külakalmistu	15. sajand
3	KAU005	Viljandimaa	Kaubi külakalmistu	16. sajand
4	KOD001	Tartumaa	Kodavere külakalmistu	13.-16. sajand
5	KOD002	Tartumaa	Kodavere külakalmistu	13. saj. - 15. saj. keskpaik
6	KOD003	Tartumaa	Kodavere külakalmistu	13. saj keskpaik - 15. saj keskpaik
7	KOD004	Tartumaa	Kodavere külakalmistu	13.-16. sajand
8	MAL003	Tartumaa	Mäletjärve külakalmistu	14. sajandi IV veerand
9	MAL004	Tartumaa	Mäletjärve külakalmistu	14. sajandi IV veerand

10	MAL005	Tartumaa	Mäletjärve külakalmistu	16. sajand
11	MAL006	Tartumaa	Mäletjärve külakalmistu	16. sajand
12	MAL007	Tartumaa	Mäletjärve külakalmistu	16. sajand
13	TAJ001	Tartumaa	Tartu Jaani kirik	13. sajand - 16. sajand
14	TAJ002	Tartumaa	Tartu Jaani kirik	13. sajand - 16. sajand
15	TAJ003	Tartumaa	Tartu Jaani kirik	13. sajand
16	TAJ004	Tartumaa	Tartu Jaani kirik	13. sajand
17	TAJ005	Tartumaa	Tartu Jaani kirik	13. sajand
18	TAJ006	Tartumaa	Tartu Jaani kirik	13. sajand
19	TAT006	Tartumaa	Tartu toomkirik	17. saj. lõpp - 18. saj algus
20	TUD001	Ida-Virumaa	Tudulinna külakalmistu	15.-16. sajand
21	TUD002	Ida-Virumaa	Tudulinna külakalmistu	15.-16. sajand
22	VAL001	Saaremaa	Valjala kirikaed	13. sajand
23	VAL002	Saaremaa	Valjala kirikaed	13. sajand
24	VAS001	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool
25	VAS002	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool
26	VAS003	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool
27	VAS004	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool
28	VNK003	Tartumaa	Vana-Kuuste külakalmistu	16. sajandi II pool
29	VNK004	Tartumaa	Vana-Kuuste külakalmistu	14. sajand
30	VNK005	Tartumaa	Vana-Kuuste külakalmistu	16. sajand

### 2.2.2 Materjali kataloogimine

Enne uuritavate hammastega tööle asumist, pildistati üles kõik vaadeldavad hambad, nii pealt kui külgvaatest (joonis 6). Saadud pildid salvestati Tartu Ülikooli genoomika instituudi (TÜ GI) Eesti Biokeskuse vana DNA labori andmebaasi, kust neid on võimalik kätte saada edasisteks uuringuteks.

### 2.2.3 DNA eraldamine

DNA eraldamine hammastest toimus 2019. aastal TÜ GI Eesti Biokeskuse vana DNA laboris TÜ GI vana DNA teaduri Lehti Saagi juhendamisel.

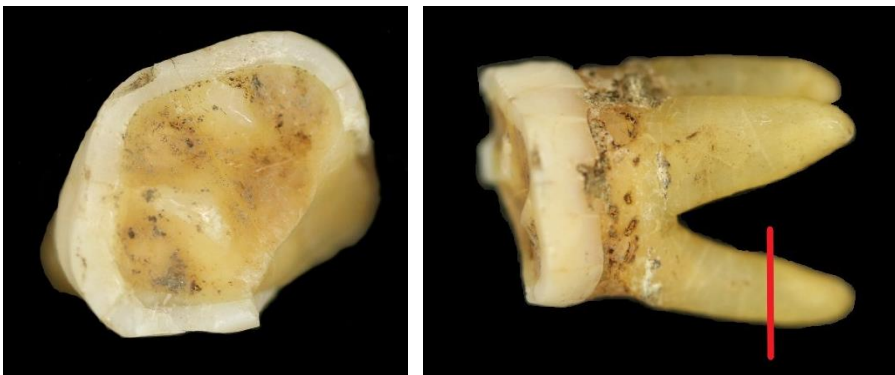
#### 2.2.3.1 Saastuse vältimine

Vana DNA laboris on saastuse vältimine väga oluline. Labori eesruumi sisenedes vahetati ära kõik riided ja jalanõud, et ei kannaks laborisse mustust väliskeskkonnast ega tänapäevase DNA labori produkte. Minnes edasi järgmisse vaheruumi, pandi selga spetsiaalne kaitseriietus, mis koosnes kaitseülkonnast, näomaskist, juuksevõrgust, kummikutest, kaitseprillidest ja kahest paarist kummikinnastest.

Enne tööle asumist puhastati kõik tööpinnad ja kasutatavad vahendid DNA-ExitusPlus lahusega. DNA eraldamiseks kasutatavad hambad tõsteti vaheruumis ümber puhastesse kottidesse, et vähendada väljast tuleva saastuse jõudmist laborisse ja proovidesse.

### 2.2.3.2 Proovi võtmine DNA analüüsiks

Hambaproov vana DNA analüüsiks võeti hambajuure tipust (joonis 6). Võimalusel kasutati selleks ees- või tagapurihammast. Hammaste puurimine toimus vana DNA laboris eraldi ruumis laminaar-tõmbekapi all. Töötasapinnad ning -vahendid puhastati DNA-ExitusPlusiga. Puuri lõikeketaste, pintsettide ja hambasondi puhastamiseks leotati neid 6% naatrium hüpokloriti (NaOCl) lahuses umbes 5 minutit, seejärel loputati neid Milli-Q (Merck Millipore) vee ja 70% etanooliga ning kuivatati puhta pabersalvrätikuga. Neid protseduure korrati pärast iga hamba puurimist. Hambatangide puhastamiseks kasutati DNA-ExitusPlusi. Töö hõlbustamiseks kasutati kaht komplekti puuri lõikekettaid ja pintsette.



**Joonis 6. Indiviidi TAJ002 purihamba pildid pealt- ja külgvaates. Punase joonega on tähistatud umbkaudne juure lõikamise koht.**

Tõmbekapi tasapind kaeti ühtlaselt fooliumiga. Töötasapinna keskele moodustati fooliumist L-kujuline sein, et vähendada puurimisel tekkiva tolmu lendumist tõmbekapi seintele. Fooliumseina sisse laotati veel üks kiht fooliumit. Puhastatud puuri lõikekettad, pintsetid, hambatangid ja kaalumisaluse asetati fooliumseinaga piiratud tööalale ning teisele poole seina pandi valmis puuritav hammas ja kotike proovi jaoks. Hambatangide otsad mähiti fooliumisse. Enne puurima asumist tuli kätte panna lisakäised ja puhtad kindad.

Puurimine teostati võimalikult sügaval tõmbekapi sees ja kaalumisaluse kohal, et vältida puurimisel lenduvate osakeste sattumist tõmbekapist väljapoole ja vähendamaks saastunud ala suurust. Hoides hammast kindlalt hambatangide vahel, lõigati hamba küljest ära tükk hambajuurt,

mis seejärel pandi eelnevalt proovi andmetega märgistatud puhtasse kotti. Kahelt hambalt, MAL007 ja VAS001, eemaldati hambasondi kasutades ka hambakivi, mis säilitati hilisemateks uuringuteks. Järelejäänud hammas pandi tagasi vanasse kotti. Kasutatud puuri lõikekettad ja pintsetid asetati 6% NaOCl lahusesse. Fooliumseinte vahel olev pealne fooliumkiht visati koos kaalumisaluse ja hambatangide otstes olnud fooliummähistega ära. Sama tehti ka lisakäiste ja pealmise kindakihiga, mis asendati uutega. Iga järgmise hamba käsitlemisel tarvikud kas asendati uutega või kasutati NaOCl lahuses puhastatud vahendeid.

### **2.2.3.3 Pinnasaastuse eemaldamine ja DNA eraldamine hambajuurtest**

Hambajuure pesemiseks ja saastava DNA eemaldamiseks võeti tuli iga proovi jaoks kaks 50 ml tuubi ning neisse valati vastavalt 5 ml 6% NaOCl lahust ja 5 ml 70% etanooli. Juuretükide kuivatamiseks valmistati ette pabersalvrätikutega kaetud alus, mis asetati tõmbekappi fooliumile UV valguse alla. Edasi asetati hambajuurte tükid, eelnevalt proovi numbriga tähistatud kaalumisalustele. Seejärel paigutati juuretükid aluselt 6% NaOCl lahusega tuubidesse 5 minutiks seisma, loksutades aeg-ajalt tuube mustuseosakeste eemaldamiseks juure pinnalt. Järgmisena pandi kraanist vesi jooksuma ja valati NaOCl ettevaatlikult kraanikaussi. Proove loputati 3 korda MilliQ (Merck Millipore) veega ja valati seejärel etanooliga tuubi 2 minutiks seisma. Pärast etanooli ära valamist asetati juuretükid eelnevalt tõmbekappi valmis pandud aluse peale ~3 tunniks UV alla kuivama.

DNA eraldamiseks hambajuurest kaaluti kõigepealt juuretükid, et välja arvutada vajalikud puhvrite kogused. Juuretükide kaalud jäid vahemikku 30–250 mg (lisa 1). Iga proovi jaoks võeti eraldi 5 ml tuub, kuhu pipeteeriti EDTA (0,5M, pH 8.0) puhvrit 2ml ja proteinaas K (18,182 mg/ml) lahust 50 µl 100mg juuretüki kohta [lõppkontsentratsioonid ( $C_{lõpp}$ ): EDTA 0,487 M, proteinaas K 0,443 mg/ml]. Seejärel lisati igasse tuubi vastav hambaproov ja suleti kaas tihedalt parafilmiga ning tuubid asetati aeglaselt roteeruvasse loksutisse (*Nutating Mixer, VWR International*) toatemperatuuril 72 tunniks.

### **2.2.3.4 DNA ekstrakti puhastamine**

Kõigepealt eemaldati tuubidelt parafilm ja tuubide välispinda puhastati DNA-ExitusPlusiga. Seejärel tsentrifugeeriti proove 5 minutit kiirusel 4000 pööret minutis (rpm) (*Centrifuge 5810 R, Eppendorf*). Samal ajal tõsteti 1,6 ml elueerimispuhvrit EB (*QIAquick PCR Purification Kit, QIAGEN*) pudelist ümber 5 ml tuubi. Tsentrifugeeritud proovide tuubidest eemaldati ettevaatlikult DNA-d sisaldav ekstrakt, samal ajal jälgides, et pipetiotsikusse ei tõmmataks sadet,

ja pandi see eelnevalt proovide nimedega (ID) märgistatud kontsentraatori tuubidesse (*Vivaspin Turbo 15 30000 MWCO PES*, Sartorius). Proove tsentrifugeeriti kiirusel 4000 rpm 20–40 minutit, kuni ekstrakt oli kontsentreerunud mahuni 250 µl. Niikaua kui proovid olid tsentrifugeeritud, pandi valmis suure mahutiga puhastuskolonnid (*High Pure Viral Nucleic Acid Large Volume Kit*, Roche) ja 1,5 ml tuubid, kirjutades neile peale proovide informatsiooni. Kolonnidesse lisati 2,5 ml sidumispuhvrit PB (*QIAquick PCR Purification Kit*, QIAGEN), mis aitab DNA maatriksile siduda. Seejärel lisati kontsentreeritud ekstrakt ja suspendeeriti pipetiga läbi. Proove tsentrifugeeriti 1 minut kiirusel 4000 rpm. Kolonnist läbi jooksnud vedelik valati ära ning maatriksile seondunud DNA puhastamiseks lisati proovidele 1 ml PE puhvrit (*QIAquick PCR Purification Kit*, QIAGEN) ja tsentrifugeeriti 1 minut kiirusel 4000 rpm. Kolonnid kanti ümber 1,5 ml kogumistuubidesse, suur puhvermahuti eemaldati ning vedelikuga kogumistuub visati ära. Tsentrifugeerimise rootor vahetati ümber 2 ml rootori vastu ja proove tsentrifugeeriti 1 minut kiirusel 13 000 rpm, et kolonn saaks kuivaks. Seejärel asetati kolonnid varasemalt valmis pandud puhastesse 1,5 ml tuubidesse ja lisati 100 µl EB elueerimispuhvrit (*MinElute PCR Purification Kit*, QIAGEN), mis vabastas DNA maatriksi küljest. Proovid asetati 10 minutiks termoblokki 37 °C juurde. Järgnevalt tsentrifugeeriti tuube 2 minutit kiirusel 13 000 rpm. Kolonnid visati ära, tuubide kaaned suleti parafilmiga ja proove säilitati sügavkülmas -20 °C juures kuni sekveneerimisraamatukogude valmistamiseni.

#### **2.2.4 DNA sekveneerimisraamatukogude koostamine**

DNA sekveneerimisraamatukogude koostamine ehk vajalike adapterite ja indeksite lisamine DNA-fragmentidele toimus TÜ GI Eesti Biokeskuse vana DNA laboris Lehti Saagi juhendamisel. Järgnevalt kirjeldatud etapid 1–3 viidi läbi vana DNA laboris. 4. etapi polümeraasi ahelreaktsiooni (PCR) segu valmistati vana DNA laboris, kuid DNA-fragmendid amplifitseeriti tänapäevase DNA laboris. Raamatukogude koostamiseks kasutati *NEBNext DNA Sample Prep Master Mix Set for 454* (E6070, *New England BioLabs*) komplekti. Antud komplekt sisaldas otste täitmiseks vajalikku puhvrit ja ensüümisegu, ligeerimispuhvrit ja ensüümi ning adapterite otste täitmise puhvrit ja ensüümi.

##### **2.2.4.1 DNA fragmentide üheaahelaliste otste täitmine ja reaktsioonisegust puhastamine**

DNA-fragmentide üheaahelaliste otste täitmiseks valmistati kõigepealt 1,5 ml tuubi reaktsioonisegu, mis koosnes ühe proovi kohta 12,5 µl steriilsest veest, 5 µl 10x otste täitmise puhvrist ja 2,5 µl otste täitmise ensüümisegust. Seejärel tõsteti reaktsioonisegu 20 µl kaupa 0,2 ml tuubidesse. Igasse tuubi lisati 30 µl DNA proovi, misjärel pandi tuubid PCR-i masinasse (*K640 Thermal Cycler*, *Heat Force*), kus neid inkubeeriti 20 °C juures 30 minutit.

Pärast seda puhastati proovid DNA üheaheelaliste otste täitmise reaktsioonisegust, kasutades *MinElute PCR Purification Kit* (QIAGEN) komplekti. Kolonnidesse pandi 550 µl PB puhvrit, mis aitab DNA-l matriksile seonduda, ja lisati 50 µl täidetud otstega DNA-d. Seejärel tsentrifugiti proove kiirusel 13 000 rpm 1 minut, läbi jooksnud vedelik valati ära ja tuubidesse lisati 650 µl PE puhvrit, mis puhastas DNA üheaheelaliste otste täitmise reaktsioonisegust. Proove tsentrifugiti uuesti 1 minut kiirusel 13 000 rpm, vedelik valati ära ja tuubid pandi veel minutiks tsentrifuugi, et kolonnid saaksid võimalikult kuivaks. Kolonnid tõsteti puhastesse 1,5 ml tuubidesse, lisati 30 µl EB puhvrit, et vabastada DNA matriksi küljest, ja tuubid asetati 10 minutiks termoblokki 37 °C juurde. Pärast seda tsentrifugiti proove 2 minutit kiirusel 13 000 rpm, kolonnid visati ära ning tuubide kaaned suleti.

#### **2.2.4.2 Adapterite lisamine ja reaktsioonisegust puhastamine**

Adapterite lisamiseks DNA-fragmentidele valmistati 1,5 ml tuubi reaktsioonisegu. Ühe proovi kohta sisaldas segu 10 µl 5x ligeerimispuhvrit, 5 µl 2,5 µM adapterite segu ( $C_{lõpp} = 0,25 \mu\text{M}$ ), 5 µl 5 U/µl T4 ligaasi ( $C_{lõpp} = 0,5 \text{ U}/\mu\text{l}$ ). Antud segu kanti 0,2 ml tuubidesse 20 µl kaupa. Seejärel lisati segusse 30 µl DNA proovi ja pandi tuubid 15 minutiks PCR-i masinasse 20 °C juurde inkubeerima.

DNA puhastamine adapterite liitmise reaktsioonisegust toimus samamoodi nagu punktis 2.2.4.1 kirjeldatud üheaheelaliste otste täitmise reaktsioonisegust puhastamine.

#### **2.2.4.3 Adapterite otste täitmine**

Adapterite otste täitmiseks tehti 1,5 ml tuubi valmis reaktsioonisegu, mis sisaldas ühe proovi kohta 5 µl 10x adapteri lünkade täitmise puhvrit, 2 µl Bst DNA polümeraasi, 13 µl steriilset vett. Igasse 0,2 µl tuubi lisati 20 µl reaktsioonisegu ja 30 µl proovi. Tuubid pandi PCR masinasse, kus neid inkubeeriti 30 minutit 37 °C juures ja siis 20 minutit 80 °C juures, et ensüüm inaktiveerida.

#### **2.2.4.4 Indeksite lisamine, DNA-fragmentide amplifitseerimine ja reaktsioonisegust puhastamine**

PCR-i segu valmistamine toimus vana DNA laboris. DNA fragmentide amplifitseerimine toimus aga TÜ GI tuumiklaboris, et vältida vana DNA labori saastamist amplifikatsiooniproduktidega.

Indeksite lisamiseks ja DNA-fragmentide amplifitseerimiseks valmistati PCR-i segu, mis ühe proovi kohta sisaldas 17 µl steriilset vett, 10 µl 10x PCR puhvrit (Eurogentec), 10 µl 25 mM MgCl<sub>2</sub> ( $C_{lõpp} = 2,5 \text{ mM}$ ), 5 µl 20 mg/ml BSA-d ( $C_{lõpp} = 1 \text{ mg}/\text{ml}$ ), 2 µl 10 µM universaalset praimerit (*NEBNext Multiplex Oligos for Illumina*, New England Biolabs) ( $C_{lõpp} = 0,2 \mu\text{M}$ ), 2 µl 4x10 mM

dNTP-sid (iga dNTP  $C_{l\ddot{o}pp} = 0,2 \text{ mM}$ ),  $2 \mu\text{l}$   $5 \text{ U}/\mu\text{l}$  *HGS Diamond Taq* polümeraasi (Eurogentec) ( $C_{l\ddot{o}pp} = 0,1 \text{ U}/\mu\text{l}$ ). Igasse  $0,2 \text{ ml}$  tuubi lisati  $48 \mu\text{l}$  reaktsioonisegu ja  $2 \mu\text{l}$   $10 \mu\text{M}$  indekseeritud praimerit ( $C_{l\ddot{o}pp} = 0,2 \mu\text{M}$ ) (iga proovi jaoks oli erinev indeks; lisa 1) (*NEBNext Multiplex Oligos for Illumina*, New England Biolabs). Seejärel lisati tuubidesse  $50 \mu\text{l}$  proovi ning viidi tänapäevase labori PCR-i masinasse (*Veriti 96 Well Thermal Cycler*, Applied Biosystems). PCR-i programm algas 5-minutilise eelkuumutamisega  $94 \text{ }^\circ\text{C}$  juures, sellele järgnes denaturatsioon  $94 \text{ }^\circ\text{C}$  juures 30 sekundit, 30 sekundit praimerite seondumist  $60 \text{ }^\circ\text{C}$  juures ja DNA süntees  $68 \text{ }^\circ\text{C}$  juures 30 sekundit. Denaturatsiooni, praimerite seondumise ja DNA sünteesi tsüklit korrati 18 korda. Programmi lõpus jäeti proovid veel 7 minutiks  $72 \text{ }^\circ\text{C}$  juurde, et DNA süntees saaks lõppeda.

Seejärel puhastati DNA proovid polümeraasi ahelreaktsiooni komponentidest samamoodi, nagu on kirjeldatud punktis 2.2.4.1.

Saadud sekveneerimisraamatukogude kontsentratsioonid määrati fluoromeetriliselt (Qubit, *Thermo Fisher Scientific*).

### **2.2.5 DNA sekveneerimine**

Enne sekveneerimist määras teise põlvkonna sekveneerimise spetsialist Anu Solnik sekveneerimisraamatukogude pikkused paralleel-kapillaarelektroforeesi meetodil (Fragment analyzer, *Agilent Technologies*) ja lahjendamise järel täpsed kontsentratsioonid kvantitatiivse PCR-iga (qPCR, *Applied Biosystems 7900HT Fast Real-Time PCR System*). DNA sekveneerimine toimus TÜ GI tuumiklaboris teise põlvkonna *single-end* meetodil Illumina NextSeq 500 platvormil.

### **2.2.6 Sekveneeritud järjestuse joondamine referentsjärjestusele**

Sekveneeritud järjestuste joondamine inimese referentsgenoomi järjestusele ning hilisem analüüsimine toimus Lehti Saagi juhendamisel TÜ Teadusarvutuste keskuse hallatavas TÜ GI Eesti Biokeskuse serveris, milles töötati programmi *Putty* (<https://www.putty.org/>) abil.

#### **2.2.6.1 Järjestuste ettevalmistamine joondamiseks**

Serverisse sisestatud sekveneerimise tulemusfailid *\*.fastq.gz* formaadis, kopeeriti serveris ümber vastavasse keskaja projekti kausta.

#### **2.2.6.2 Järjestuste referentsile joondamine**

Kõigepealt tuli vaadata, millised üksikmasinad klastris on kasutamiseks vabad ja milliseid ei saa kasutada. Selleks kasutati käsku *sinfo*. Mittetöötavad masinad märgiti üles töövoos käivitusskripti.

Järgmisena tuli valmis teha *\*.txt* fail, mis sisaldas sekveneerimisel saadud tulemusfailide asukohti ja proovide ID-sid väljendusfailide nimede jaoks. Järjestuste joondamiseks inimese referentsgenoomi järjestusele kasutati kolmest etapist koosnevat töövoogu. Iga etapi tulemus edastati järgmisele etapile lähteandmeteks.

Esmalt saadeti käivitusskripti abil kõigi kolme etapi tööd serveri järjekorda. Töövoos esimeses etapis toimus programmi *cutadapt* (Martin, 2011) abil adapterijärjestuste eemaldamine. Fragmentid, mis olid lühemad kui 28 ap, jättis programm edasisest analüüsist välja. Saadud fail *\*\_trimRI.fastq.gz* suunati edasi järjestuste joondamisele ja sorteerimisele. Järjestuste joondamisel kasutati referentsina GRCh37 (*Genome Reference Consortium Human Build 37*, hs37d5) järjestust. Joondamiseks ja sorteerimiseks kasutati programme *bwa mem* (H. Li ja Durbin, 2009) ja *samtools* (Li et al., 2009) ning lõpuks väljastati ainult inimese genoomile joondunud järjestusi sisaldav *\*.bam* fail. Illumina NextSeq 500 süsteemis toimub sekveneerimine paralleelselt 4 rajal, mistõttu väljastatakse iga indiviidi kohta 4 faili, mis läbivad joondamise töövoos esimese etapi igaüks eraldi.

Teises etapis pandi iga proovi neljal rajal sekveneeritud fragmentid kokku ühte faili ja toimus programmi *picard* (<http://broadinstitute.github.io/picard/>) abil PCR-i käigus tekkinud duplikaatide eemaldamine. Seejärel reastas programm *GATK* (McKenna et al., 2010) uuesti indelid. Kasutades *samtoolsi* (Li et al., 2009), filtreeriti välja need järjestused, millele *bwa mem* (H. Li ja Durbin, 2009) programmi poolt antud joondumiskvaliteedi hinnang jäi alla 10. Järgmisena produtseeriti *samtoolsi* (Li et al., 2009) ja TÜ GI vanemteadur Dr. Lauri Saagi loodud skripti abil statistikud, mis kolmandas etapis väljastati kõiki koos joondatud indiviide sisaldava tabelifailina. Väljastatud tabelis oli informatsiooni täisgenoomi ning eraldi X ja Y kromosoomi ja mitokondri genoomi kohta: sh keskmine genoomi katvus, fragmentide keskmine pikkus, fragmentide arv erinevates töövoos staadiumites jms.

### **2.2.7 Järjestuste *in silico* kvaliteedikontroll**

Alustuseks tuli üle vaadata klastris olevate üksikmasinate seisukord ja teha vajalikud muudatused nii nagu on kirjeldatud peatüki 2.2.6.2 alguses. Seejärel koostati *\*.txt* fail, kus olid kirjas *\*.bam* joondusfailide asukohad ja proovide ID-d väljendusfailide nimede jaoks.

MtDNA saastatuse uurimiseks kasutati meetodit *contamMix* (Fu et al., 2013). Esmalt eraldati *\*.bam* failist mitokondri informatsiooni sisaldavad järjestused. Järgnevalt liideti vastava mtDNA konsensusjärjestus võrdluspaneeliga, mis sisaldas 311 erinevat mtDNA järjestust, ning seejärel joondati järjestused omavahel. Saastatus määrati tulenevalt sellest, kui suur osa mtDNA

järjestustest sobitus paremini mõnele teisele võrdluspaneeli mtDNA genoomile kui uuritava proovi konsensusjärjestusele. Tulemusi loeti usaldusväärseks, kui mtDNA katvus oli  $>5x$ .

Kogu genoomi kahjustuste määramiseks kasutati programmi mapDamage (Jónsson et al., 2013), mis tuvastas, kui palju on fragmentide otstes toimunud nukleotiidide asendusi.

Soo määramiseks kasutati programmi sexing.py (Skoglund et al., 2013), mis tegi soomäärangud, tuginedes sellele, kui suur osa mõlemale sugukromosoomile kokku joondunud järjestustest on joondunud Y kromosoomile. Tulemusi loeti usaldusväärseks, kui genoomi keskmine katvus oli  $>0,005x$ .

Meessoost indiviidide X kromosoomi saastatuse uurimiseks kasutati Rasmussen et al. 2011 artiklis tutvustatud meetodeid, mis on implementeeritud programmis ANGSD (Korneliussen et al., 2014). Tulemusi loeti usaldusväärseks, kui X kromosoomi katvus oli  $>0,1x$ .

*Grep* käsuga tõmmati väljundfailist välja indiviidide soo, mitokondriaalse DNA ja meessoost indiviidide X kromosoomi saastatuse ning fragmentide 5' otste tsütosiini deaminatsiooni sageduste kohta infot andvad read ning need väljastati \*.txt failidena.

### 2.2.8 Genotüüpide määramine

Andmete edasiseks analüüsimiseks tuli määrata uuritavate proovide genotüübid. ANGSD programmi käsuga *doHaploCall* liikus programm sammhaaval mööda genoomi ning kirjutas *sites* käsuga ette antud nukleotiidsetes positsioonides üles juhuslikult valitud fragmendis olnud nukleotiidi. Need positsioonid pärinesid TÜ GI Eesti Biokeskuse ülemaailmsest genotüüpiseerimiskiibi andmestikust.

Seejärel konverteeriti failid edasistes analüüsidest kasutatavasse PLINK formaati, kasutades käsku *haploToPlink*. Kuna ANGSD (Korneliussen et al., 2014) programm kirjutab puuduvate alleelide kohale „N“, aga PLINK (Purcell et al., 2007) programmi formaadis on selle asemel „0“, siis selleks, et PLINK suudaks neid faile õigesti lugeda, tuli *sed* käsuga muuta andmestikes puuduvate alleelide tähised N tähest nulliks.

Järgmiseks kasutati programmi PLINK (Purcell et al., 2007), et muuta mittebinaarsed failid (\*.tfam, \*.tped) binaarseteks failideks (\*.bed, \*.bim, \*.fam), mis võtsid andmemahult vähem ruumi ja mida kasutati edasisteks analüüsideks.

Sama protseduur, kuid TÜ GI Eesti Geenivaramu täisgenoomiandmete informatiivsete positsioonidega, tehti läbi ka eraldi ainult Y kromosoomi jaoks, et saaks hiljem määrata Y kromosoomide haplogrupid (peatükk 2.2.9). Lisaks tehti sama protseduur läbi ka sugulusanalüüsi jaoks, millest lähemalt on juttu peatükis 2.2.11.2.

### **2.2.9 MtDNA ja Y kromosoomi haplogruppide määramine**

Mitokondri haplogruppide määramiseks kasutati esmalt programmide paketi *samtools* (Li et al., 2009) programmi *BCFtools 1.9*. Käsuga *mpileup* kirjutati välja iga indiviidi iga mitokondri positsiooni genotüüpide tõenäosused, käsuga *call* otsustati sellest lähtuvalt genotüübid ja positsioonid, kus esines varieeruvust, kirjutati üles \*.vcf faili. Haplogruppide teada saamiseks kasutati veebiprogrammi *HaploGrep 2.0* (van Oven ja Kayser, 2009; Weissensteiner et al., 2016), kuhu laeti üles saadud \*.vcf fail. Tulemusi loeti usaldusväärseks, kui mtDNA katvus oli >1x. Haplogrupid kontrollis käsitsi üle populatsioonigeneetika vanemteadur Dr. Ene Metspalu, toetudes *PhyloTreele* (van Oven ja Kayser, 2009).

Y kromosoomi haplogruppide uurimiseks tehti esmalt ANGSD-i väljundfailid PLINK failideks, nii nagu on kirjeldatud punktis 2.2.8. Käsuga *--recode vcf-iid* väljastatati \*.vcf tüüpi fail. Programmi *Bedtools 2.19* (Quinlan ja Hall, 2010) abil jäeti alles ainult need Y kromosoomide regioonid, mida on võimalik usaldusväärsetelt joondada. Tulemusi loeti usaldusväärseks, kui Y kromosoomi katvus oli >0,005x. Lõpliku haplogruppide määramise tegi populatsioonigeneetika vanemteadur Dr. Toomas Kivisild.

### **2.2.10 Võrdlusandmetega liitmine**

Edasiste analüüside tegemiseks tuli antud uurimistöös uuritavate indiviidide andmetele juurde liita varasemalt avaldatud andmed, millega uuritavaid proove saaks võrrelda. Tänapäevased võrdlusandmed pärinesid publikatsioonidest Behar et al., 2010, 2013; Fedorova et al., 2013; Kushniarevich et al., 2015; Li et al., 2008; Raghavan et al., 2014; Rasmussen et al., 2010; Tambets et al., 2018; Yunusbayev et al., 2012, 2015 ja vana DNA võrdlusandmed publikatsioonidest Lazaridis et al., 2016; Unterländer et al., 2017; Saag et al., 2017; Mitnik et al., 2018; Mathieson et al., 2018; Damgaard et al., 2018a; Damgaard et al., 2018b; Narasimhan et al., 2018; Saag et al., 2019.

Andmete liitmine toimus PLINK 1.9 (Purcell et al., 2007) programmi abil, mille *bmerge* käsuga liideti kokku uuritavad andmed ja võrdlusandmed. Kuna esines ka selliseid positsioone, kus alleelid andmestike vahel erinesid, siis selleks, et saaks neid liita, pöörati uuritavate indiviidide

andmestikus vastavates positsioonides alleelid omavahel ümber. Selleks kasutati käsku *flip*. Seejärel liideti saadud fail uuesti võrdlusandmestikuga. Positsioonid, kus oli jätkuvalt kolm või rohkem alleelivarianti, jäeti *exclude* käsuga andmestikust välja. Saadud fail liideti seekord õnnestunult võrdlusandmestikuga kokku ning *extract* käsuga jäeti alles ainult 524920 autosoomide ehk kromosoomide 1–22 positsiooni. Selle tulemusena tekkinud *\*.bed*, *\*.bim*, *\*.fam* failid olid aluseks edasistele analüüsidele.

Infoga katmata ühenukleotiidsete polümorfismide (SNP) arvu leidmiseks igas indiviidis kasutati käsku *missing*. Proovid, milles infoga kaetud SNP-de arv oli <10 000, jäeti analüüsist välja.

### **2.2.11 Populatsioonigeneetilised analüüsid**

Antud uurimistöös tehtavad analüüsid on praegusel hetkel vana DNA uuringutes kõige laialdasemalt kasutatavad. Peakomponentanalüüsiga vaadatakse populatsiooni struktuuri, f3 ja D statistikutega populatsioonide vahelisi kaugusi ning Admixture, qpAdm-i ja CP/NNLS-iga eellaskomponentide osakaalu. Sugulussidemeid uuriti READ meetodiga.

#### **2.2.11.1 Peakomponentanalüüs**

Peakomponentanalüüsi (ingl. k. *principal component analysis*, *PCA*) tegemiseks kasutati esmalt PLINK 1.9 (Purcell et al., 2007) programmi, kus *keep* käsuga jäeti eelnevalt punktis 2.2.10 valmis tehtud failidest alles ainult meid huvitavad indiviidid. Erineva resolutsiooni saavutamiseks tehti valmis kaks faili: üks, mis sisaldas informatsiooni tänapäevaste populatsioonide kohta Euroopas, Kaukaasias ja Lähis-Idas ning muinasaegsete paleoliitikumi kütt-korilaste, Kaukaasia ja Euroopa hilisemate kütt-korilaste, Anatoolia ja Euroopa varajaste põlluharijate ja stepipäritolu põlluharijate kohta, ning teine, mis sisaldas informatsiooni tänapäeva Põhja-, Kesk- ja Ida-Euroopa populatsioonide ning Eesti kivi-, pronksi-, raua- ja keskaegsete indiviidide kohta.

Seejärel konverteeriti saadud failid ümber PCA analüüsiks vajalikku *eigenstrat* formaati (*\*.geno*, *\*.snp*, *\*.ind*). Selleks kasutati paketi EIGENSOFT 7.2.0 (Patterson et al., 2006) programmi *convertf*. PCA viidi läbi sama paketi programmiga *smartpca*.

Graafikute tegemiseks kasutati programmi R 3.6.0 (R Core Team, 2019). Analüüsitulemuste (*eigenvectors*) failist maatriksina vektorisse lugemiseks kasutati käskude *as.matrix* ja *read.table*. Saadud vektori teisest ja kolmandast tulpast loodi omakorda esimese kahe peakomponendi väärtusi sisaldavad vektorid. Käsuga *scan* loeti vektoritesse indiviididele määratavate nimede ja värvide failid. Järgmisena jagati indiviidid gruppidesse, luues vektorid, mis sisaldasid numbreid, mis

põhinesid indiviidide järjekorral analüüsitulemuste faili esimeses tulbas. Erinevate graafikute tegemiseks kuvati ainult nende gruppide andmed, mida parasjagu vaadelda taheti. Antud uurimistöö jaoks tehti kolm erinevat graafikut. Nendes kasutatud tänapäevased- ja muinasajapopulatsioonid ning uuritavad proovid on toodud lisa 2, lisa 3).

Graafiku tegemiseks kasutati käsku *plot*. Seejärel kuvati *text* käsuga graafikule huvipakkuvad grupid. Valminud PCA graafikud väljastas programm *\*.pdf* formaadis.

Mõningaid muudatusi tuli teha sakslaste populatsioonis kasutatavate indiviidide osas. Kuna PCA tulemustest oli näha, et sakslaste populatsioon sisaldas ka neid indiviide, kes on segunenud muude rahvastega, siis täpsemate tulemuste saamiseks jäeti nad edasistest analüüsides välja ja kasutati ainult indiviide German1, German2, German6, German7, German9.

### **2.2.11.2 Sugulusanalüüs**

Sugulussidemete uurimiseks tehti meetodiga READ (Kuhn et al., 2018) sugulusanalüüs. Analüüsis kasutati ainult neid proove, mille keskmine katvus oli  $>0,03x$ . Antud uurimistöös uuritud 30 proovist kvalifitseerus sellesse analüüsi 15 proovi (lisa 3). Võrdlusandmeteks kasutati varasemalt vana DNA uurimisgrupi poolt uuritud 26 Eesti keskaegset proovi, millest analüüsi sobis 20 proovi (Lisa 3).

Sugulusanalüüsi tegemiseks kasutati eelnevalt punktis 2.2.8 ANGSD (Korneliussen et al., 2014) ja PLINK (Purcell et al., 2007) programmidega valmis tehtud mittebinaarseid *\*.tped* ja *\*.tfam* faile. Edasise analüüsi alamskriptide käivitamiseks kasutati programme Python 2.7.11 (Van Rossum ja Drake Jr, 1995) ja R 3.1.0 (R Core Team, 2019). READ meetodiga võrreldi, kui sarnased kõik indiviidid omavahel keskmiselt olid ja seejärel võrreldi iga indiviidide paari selle keskmisega. Saadud tulemustest väljastati graafik (lisa 5.1).

Lisaks üldisele, kõiki indiviide hõlmavale analüüsile tehti kaks analüüsi erineva geneetilise kuuluvusega indiviidide gruppidele. Lähtudes varasemalt tehtud PCA tulemustest, jagati Tartu indiviidid kahte gruppi: tänapäeva eestlastega sarnased (TPM001, TPM002, TPM003, TAJ003, TAJ004, TAJ005) ja tänapäeva keskeurooplastega (sakslastega) sarnased (TAT001, TAT002, TAT003, TAT005, TAJ002, TAJ006) (lisa 3). Kuna indiviidid TAT006 ja TAJ001 paiknesid PCA graafikul kahe suure grupi vahel, siis jäeti nad antud analüüsides kõrvale. Nende kahe analüüsi tegemiseks tehti uued lähtefailid: *\*\_Est.txt* fail sisaldas külasurnuaedade ja Tartu surnuaedade nn. eestlastest indiviide ning *\*\_Ger.txt* sisaldas ainult nn. sakslastest indiviide.

### 2.2.11.3 f3 ja D statistikud

F3 ja D statistikute arvutamiseks kasutati serveris programmi Python 2.7.6 (Van Rossum ja Drake Jr, 1995) ja EIGENSOFT 7.2.0 (Patterson et al., 2006). Kõigepealt pandi \*.txt failidesse kirja uuritavate populatsioonide ja võrdluspopulatsioonide nimed. Python programmiga tehti valmis failid kalkulatsioonides kasutatavate kombinatsioonidega, kus välisgrupina kasutati Aafrika populatsiooni joruubasid ning teise kahe populatsioonina kõiki võimalikke paare uuritavatest ja võrdluspopulatsioonidest.

Paketi EIGENSOFT programmiga *convertf* muudeti saadud failid *eigenstrat* formaati. Seejärel asendati *awk* käsuga \*.ind faili nn fenotüübi tulp populatsiooninimedega. Kasutades programmi ADMIXTOOLS 1.1 (Patterson et al., 2012), arvutati käskudega qp3Pop ja qpDstat f3 ja D statistikud. Tulemuste paremaks visuaalseks väljendamiseks tehti Microsoft Office Excelis (Microsoft Corporation, 2019) punktdiagrammid: x- ja y-teljel erinevate uuritavate populatsioonide ja kõigi võrdluspopulatsioonidega arvatud statistikute väärtused. Mida suurema statistiku väärtuse andis võrdluspopulatsioon, seda sarnasem oli see vastavale uuritavale populatsioonile. Kui mingi võrdluspopulatsiooni punkt jäi punktide pilve iseloomustavast sirgest ühe telje poole, oli see võrdluspopulatsioon sarnasem selle telje uuritava populatsiooniga, kui teise telje omaga.

Tartu indiviide jagati gruppidesse kahel erineval moel – esmalt matmispaikade kaupa ja hiljem ka lähtuval PCA tulemustest n-ö eestlaste ja sakslaste sarnasteks (lisa 3).

### 2.2.11.4 Admixture

Analüüsi tegemiseks kasutati programmi Admixture 1.30 (Alexander et al., 2009). PLINK (Purcell et al., 2007) programmi *keep* käsuga jäeti punktis 2.2.9 loodud failist alles tänapäevased eestlased ja muinasaegsed paleoliitikumi kütt-korilased, Kaukaasia ja Euroopa hilisemad kütt-korilased, Anatoolia ja Euroopa varajased põlluharijad ja stepi päritolu põlluharijad ning uuritavad ja võrdlemiseks kasutatavad Eesti keskaegsed indiviidid. Käsuga *extract* jäeti alles need SNPd, mida kasutas Lehti Saag varasemalt tänapäevaste indiviidide Admixture analüüsi tegemiseks. Tänapäevaste indiviidide analüüsi käigus programmi väljastatud statistiliste parameetrite alusel usaldusväärseimaks osutunud 9 komponendiga algse analüüsi alleelisageduste failid kopeeriti töökausta. Seejärel käivitati Admixture 1.30 (Alexander et al., 2009) programm, mis alustas vana DNA andmete tänapäevaste indiviidide põhjal arvatud struktuuri projitseerimist. Saadud tulemused kopeeriti Excelisse ja koostati andmete põhjal graafikud.

Graafikud tehti eraldi muinasaja ja keskaja proovide populatsioonide keskmistest väärtustest ja muinasaja ja keskaja iga indiviidi ning tänapäevaste populatsioonide keskmistest väärtustest.

#### **2.2.11.5 qpAdm ja CP/NNLS**

qpAdm analüüsi tegemiseks tuli välja jätta proovid, mille SNPde arv jäi alla 100 000 ja ChromoPainter/non-negative least squares (CP/NNLS) analüüsi jaoks proovid, millel SNPde hulk jäi alla 20 000.

qpAdm analüüsi tegemiseks kasutati AdmixTools/4.1 programmi käsku qpAdm (Patterson et al., 2012), mis proovib uuritavaid populatsioone kokku panna ette antud teistest populatsioonidest (*left* populatsioonid), milleks antud uurimistöös olid Lääne-Euroopa kütt-korilased (WHG), Kesk-Euroopa varajased põlluharijad (Central\_MN), stepi põlluharijad (Yamnaya\_Kalmykia) ja tänapäeva siberlased (Nganasan). Nendele omakorda olid välisgrupiks populatsioonid (*right* populatsioonid), mis on neist kauged ja/või eristavad neid üksteisest – joruubad (vähesegunenud aafriklased), Ust-Ishim (vanim euraaslane, 45 000 a), Mal'ta (ammune läänesiberlane, 24 000 a), Kostenki (vanim eurooplane, 36 000 a), Skandinaavia kütt-korilased (SHG), tänapäeva hiinlased (Han).

CP/NNLS analüüsi tegemiseks tuli kõigepealt uuendada kasutuses olevate serverite nimekirja, nagu on kirjeldatud punktis 2.2.6.2. Seejärel käivitati programm ChromoPainter (Lawson et al., 2012), mis jagas iga indiviidi genoomi komponentideks. Seejärel käivitati programmiga R 3.3.0 (R Core Team, 2019) NNLSALLcounts.R skript (Saag et al., 2019), millega leitakse, kui suur osa indiviidi genoomist pärineb teisest etteantud genoomist (antud analüüsis olid nendeks eelnevalt mainitud *left* populatsioonid). Andmed viidi Excelisse ja iga grupi kohta arvutati keskmised.

## 2.3 Tulemused ja arutelu

### 2.3.1 Uuritavate proovide kvaliteet ja sekveneerimise tulemused

Vana DNA ekstrakt ei sisalda ainult uuritava liigi, vaid ka teiste organismide DNA-d. Uuritava liigi DNA-d sellistes proovides nimetatakse endogeenseks DNA-ks (Poinar et al., 2006). Endogeense DNA sisaldust inimese proovis võib kokkuleppeliselt lugeda heaks, kui seda on vähemalt 10%. Antud uurimistöo proovide endogeense DNA sisaldused jäid vahemikku 0,06–84,85% (mediaanväärtus 15,26%) ning neist 18 proovi sisaldasid >10% endogeenset DNA-d (lisa 1). Lisaks määrati sekveneerimisraamatukogu efektiivsus, mis näitab lõplike analüüsis kasutatavate järjestuste osakaalu algselt sekvenaatorist väljuvatest andmetest pärast liiga lühikeste, inimese genomile mitte joondunud, duplikaatsete ja madala kvaliteediga fragmentide välja filtreerimist. Selle parameetri väärtused on tavaliselt umbes pool endogeense DNA osakaalust. Seega proov on hea, kui raamatukogu efektiivsus on >5%. Antud töö proovide efektiivsuse väärtused jäid vahemikku 0,02–54,36% (mediaanväärtus 7,40%) (lisa 1).

Üks võimalus, kuidas teha kindlaks vana DNA autentsus, on uurida fragmentide 5' otstes surmajärgselt toimunud tsütosiini deamineerimist (Briggs et al., 2007; Hofreiter, 2001). Need asendused viitavad, et tegemist on vana DNA-ga (lisa 1).

Vana DNA uuringute puhul on oluline ka saastatuse vältimine. Selleks, et kindlaks teha, kas proovid on saastunud, uuriti mtDNA ja meeste X kromosoomi võimalikku saastust proovis. Saastuse osakaalu loetakse proovi edasiseks kasutamiseks piisavalt madalaks, kui see on <3%. Vaadeldavas töös õnnestus mtDNA saastatuse tase määrata 15 proovil ehk 50% uuritavatest proovidest, kõikide nende proovide puhul jäi tase alla 3%. Meessoost indiviidi X kromosoomi saastatus õnnestus määrata ühel proovil ja see jäi samuti alla 3% (lisa 1).

Ühel NextSeq 500 75 ap kiibil sekveneeriti koos 15 proovi ja saadi seeläbi keskmiselt 25 miljonit järjestust proovi kohta (lisa 1). Raamatukogude efektiivsusest ja sekveneerimisel saadud järjestuste hulgast tulenevalt jäid uuritavate proovide ülegenoomsed katvused vahemikku 0,0001–0,2450 (mediaanväärtus 0,03) (lisa 1). Kogu genoomi keskmise katvuse lävend usaldusväärsete tulemuste jaoks oli 0,02 – seda ületavaid proove sai edasi kasutada autosomaalsete analüüside tegemiseks. Selliseid proove oli antud töös kokku 24 (lisa 1), seega 80% proovidest olid piisavalt head edasiste analüüside tegemiseks.

### 2.3.1.1 Soo määramine ning mtDNA ja Y kromosoomi haplogrupid

Uuritavast 30 indiviidist õnnestus sugu kindlaks määrata 23 indiviidil – neist 14 olid naised ja 9 mehed (tabel 2). Eelnevalt oli indiviidide sugu morfoloogia alusel määratud ka arheoloogide poolt. Morfoloogiliselt oli naise 17 ja mehi 9, lisaks üks meesindiviid, kelle soos ei olnud päris kindel, ja 3 määratlemata indiviidi. Nii morfoloogiliselt kui geneetiliselt õnnestus sugu määrata kokku 19 indiviidil ning nendest 17 indiviidil (12 naist, 5 meest) ehk umbes 89% mõlema kasutatud meetodi tulemused soo määrangul kattusid. Indiviidid, kes jäid arheoloogidel sooliselt määratlemata, said geneetiliselt määratletud (2 naist, 2 meest). Nende seas leidis kinnitust ka arheoloogide poolt soo osas küsimärgi alla seatud meesindiviid. Arheoloogilised ja geneetilised tulemused lahkesid kahe indiviidi puhul – KOD001 ja VNK004 – kes arheoloogide poolt määratleti naisteks, aga kes geneetiliselt olid mehed. 7 indiviidil, kelle sugu oli morfoloogiliselt määratletud, jäi see geneetiliselt kinnitamata.

**Tabel 2. Uuritavate indiviidide morfoloogiline sugu (morf. sugu), geneetiline sugu (gen. sugu), mtDNA haplogrupp (MT hg), Y kromosoomi haplogrupp (Y hg). NA – pole saadaval.**

#	Indiviid	Morf. sugu	Gen. sugu	MT hg	Y hg
1	KAU003	Naine	NA	NA	-
2	KAU004	Naine	NA	NA	-
3	KAU005	Naine	XX	U5a2a1	-
4	KOD001	Naine	XY	H7a1a/F1a1c1	I1-M253
5	KOD002	Naine	NA	U5a1h/U5a1a1	-
6	KOD003	Mees	NA	NA	NA
7	KOD004	Mees	NA	NA	NA
8	MAL003	Naine	XX	H1b2	-
9	MAL004	?	XX	H2a1	-
10	MAL005	Naine	XX	U5a2a1	-
11	MAL006	?	XY	H14a	NA
12	MAL007	Naine	XX	HV0g	-
13	TAJ001	Mees?	XY	HV0a	R1a-M417
14	TAJ002	Mees	XY	K1a3a1	R1b
15	TAJ003	Naine	XX	H1cj	-
16	TAJ004	Naine	XX	T1a1b	-
17	TAJ005	Mees	XY	U5a1b1d1	R1a1c-M558
18	TAJ006	Naine	XX	J1c3a1	-
19	TAT006	?	XX	U4d1	-
20	TUD001	Naine	XX	U5a2e	-
21	TUD002	Mees	NA	NA	NA

22	VAL001	Mees	XY	H1	N3
23	VAL002	Naine	XX	H1c1	-
24	VAS001	Mees	NA	U2e1f1	NA
25	VAS002	Naine	XX	U4d1b	-
26	VAS003	Naine	XX	J1c5d	-
27	VAS004	Mees	XY	NA	NA
28	VNK003	Mees	XY	T2 + 16189	N3a-L392
29	VNK004	Naine	XY	T2c1c2	R1a-Z645
30	VNK005	Naine	XX	H5a1a	-

MtDNA haplogrupid õnnestus kindlaks määrata 24 indiviidil 30-st (tabel 2). Neist 8 kuulusid haplogruppi U. See haplogrupp oli Euroopas valdav mesoliitikumi aegsetel Euroopa kütt-korilastel ning oli levinud ka hiljem (vt. pt. 1.2). Analüüsides on näha, et ka keskaegsetel Eesti indiviididel oli see üsna levinud – 33% 24 indiviidist kuulusid haplogruppi U. Lisaks mtDNA haplogrupp U-le määrati 10 indiviidil H, 3 indiviidil T, 2 indiviidil J ja ühel K haplogrupid. Need haplogrupid jõudsid Euroopasse arvatavasti koos varajaste põlluharijatega (vt. pt 1.2) ning on leitavad ka tänapäeva eurooplastel (Simoni et al., 2000). Uuritavate keskaegsete indiviidide seas kõige levinum haplogrupp H (41%), on valdav mtDNA haplogrupp ka tänapäeva Euroopas (Simoni et al., 2000).

Y kromosoomi haplogrupp õnnestus määrata 7 mehel üheksast (tabel 2). Kolm indiviidi kuulusid haplogruppi R1a. See haplogrupp hakkas laiemalt levima neoliitikumi ajal, kui stepialadelt tulid Euroopasse aukhaudade kultuuri esindajad. Samuti oli see valdav Y kromosoomi haplogrupp ka pronksiaegses Baltikumis (vt. pt. 1.2.2). Indiviidid VAL001 ja VNK003 kuulusid haplogruppi N3, mida leidis Eesti aladel ka juba varajase rauaaja indiviididel tarandkalmete kontekstis ning see isaliin on laialdaselt levinud ka tänapäeval Ida- ja Kirde-Euroopas, eelkõige uurali keeli rääkivate rahvaste seas, aga ka Baltikumis (vt. pt. 1.2.2). Indiviid KOD001 kuulus haplogruppi I. Need kolm haplogruppi on väga levinud ka tänapäeva Eestis – peaaegu 85% eesti meeste seas levinud isaliinidest kuulub ühte nimetatud haplogruppi (Tambets et al., 2018).

### 2.3.2 Populatsioonigeneetiline analüüs

Igas analüüsis kasutatud populatsioonid ning keskaegsed indiviidid on toodud lisades (lisa 2, lisa 3).

#### 2.3.2.1 Peakomponentanalüüs

Kokku tehti 3 erineva lahutusvõimega analüüsi. Esimeses PCA analüüsis vaadati uuritavate indiviidide sarnasust laiemas kontekstis tänapäevaste indiviididega Euroopast, Kaukaasiast ja



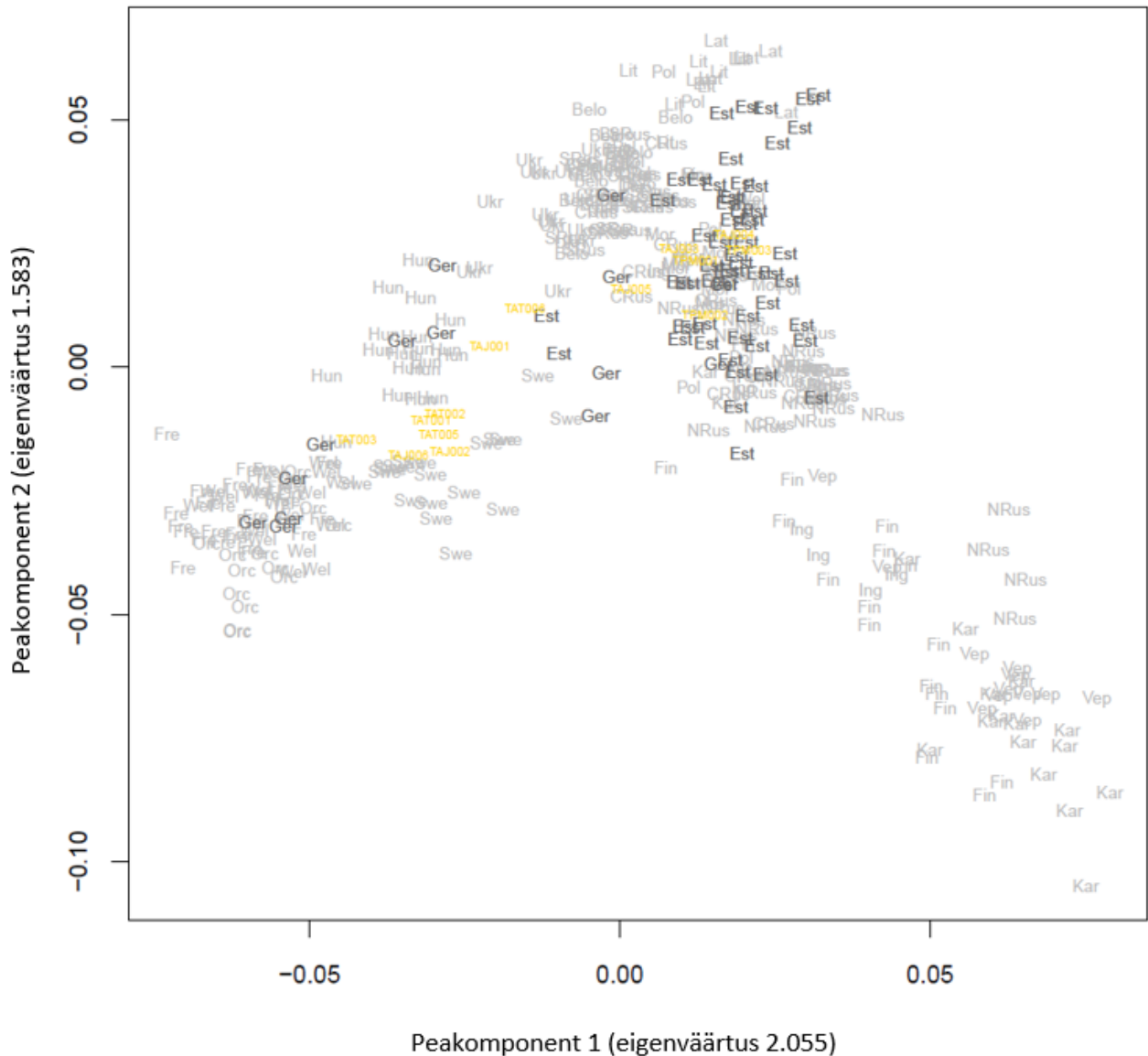
Keskaegsete indiviidide paiknemine pronksi- ja rauaaja indiviidide (joonisel helepruun ja oranž) läheduses on oodatav, kuna ka ajaliselt on keskaeg neile lähim. Keskaegsed indiviidid (joonisel kollane) klasterduvad uuritud tänapäevaste populatsioonide taustal (joonisel hall) nii, et kõik keskaegsed indiviidid on valdavalt koos tänapäeva eestlaste või sakslastega (joonisel tumehall). Nii nagu eelnevalt kirjeldati (vt. pt. 1.3), jagunesid Eesti alal keskajal elanud inimesed nõ eestlasteks ja sakslasteks. Sellist jagunemist on näha ka PCA-s.

Selleks, et saada paremat ülevaadet uuritavate indiviidide paiknemisest Euroopa geneetilisel kaardil, tehti eraldi PCA, kus olid välja toodud tänapäeva Ida-, Põhja- ja Kesk-Euroopa populatsioonid koos Eesti muinas- ja keskaja proovidega (lisa 4). Enamus uuritavaid keskaja indiviide paigutuvad graafikul tänapäevaste eestlaste lähedusse, erandiks on mõned Tartu proovid, mis on sarnasemad tänapäeva sakslastega ja üldisemalt ka teiste germaani rahvastega.

Kolmandas peakomponentanalüüsis uuriti lähemalt Tartu linna kirikumatuste indiviide (joonis 8). Antud analüüs aitab heita valgust kirikusse maetud inimeste võimalikule päritolule. Kõik Tartu Püha Maarja kiriku indiviidid (TPM001, TPM002, TPM003) paigutusid PCA joonisel samasse gruppi tänapäevaste eestlastega. Enamik Tartu toomkiriku indiviide (TAT001, TAT002, TAT003, TAT005) klasterdusid kokku tänapäevaste sakslastega. Nagu öeldud (vt. pt. 1.3.1.2.1), maeti toomkirikusse peamiselt ülemklassi inimesi, kes olid põhiliselt sakslased. Antud analüüs näitab samuti, et toomkirikusse maetud olid peamiselt nõ. sakslased. Seega on see vastavuses kirjanduses öelduga. Püha Maarja kirikust leitud indiviidid, kes geneetiliselt sarnanesid tänapäeva eestlastega, on samuti kooskõlas kirjandusest loetuga – Maarja kirik oli kihelkonnakirik ning sealt eestlaste matuste leidmine on ootuspärane (vt. pt. 1.3.1.2.1).

Jaani kiriku indiviidid jagunesid kahte gruppi: indiviidid TAJ003, TAJ004, TAJ005 klasterdusid tänapäevaste eestlaste populatsiooniga ning indiviidid TAJ002 ja TAJ006 tänapäevaste sakslaste populatsiooniga. Jaani kirik oli kihelkonnakirik ning kirjalikest allikatest on teada, et eri aegadel on kirikusse maetud nii ülem- kui alamklassi inimesi, seega nii eestlasi kui sakslasi (vt. pt. 1.3.1.2.1). Seda toetab ka antud analüüs, kus umbes pooled uuritavatest Jaani kiriku indiviididest grupeerusid eestlaste ja pooled sakslastega.

Kaks indiviidi, TAJ001 ja TAT006, paiknesid aga tänapäevaste eestlaste ja sakslaste populatsioonide vahepeal. Võib oletada, et tegemist oli kahe nõ rahvuse segunemisega.



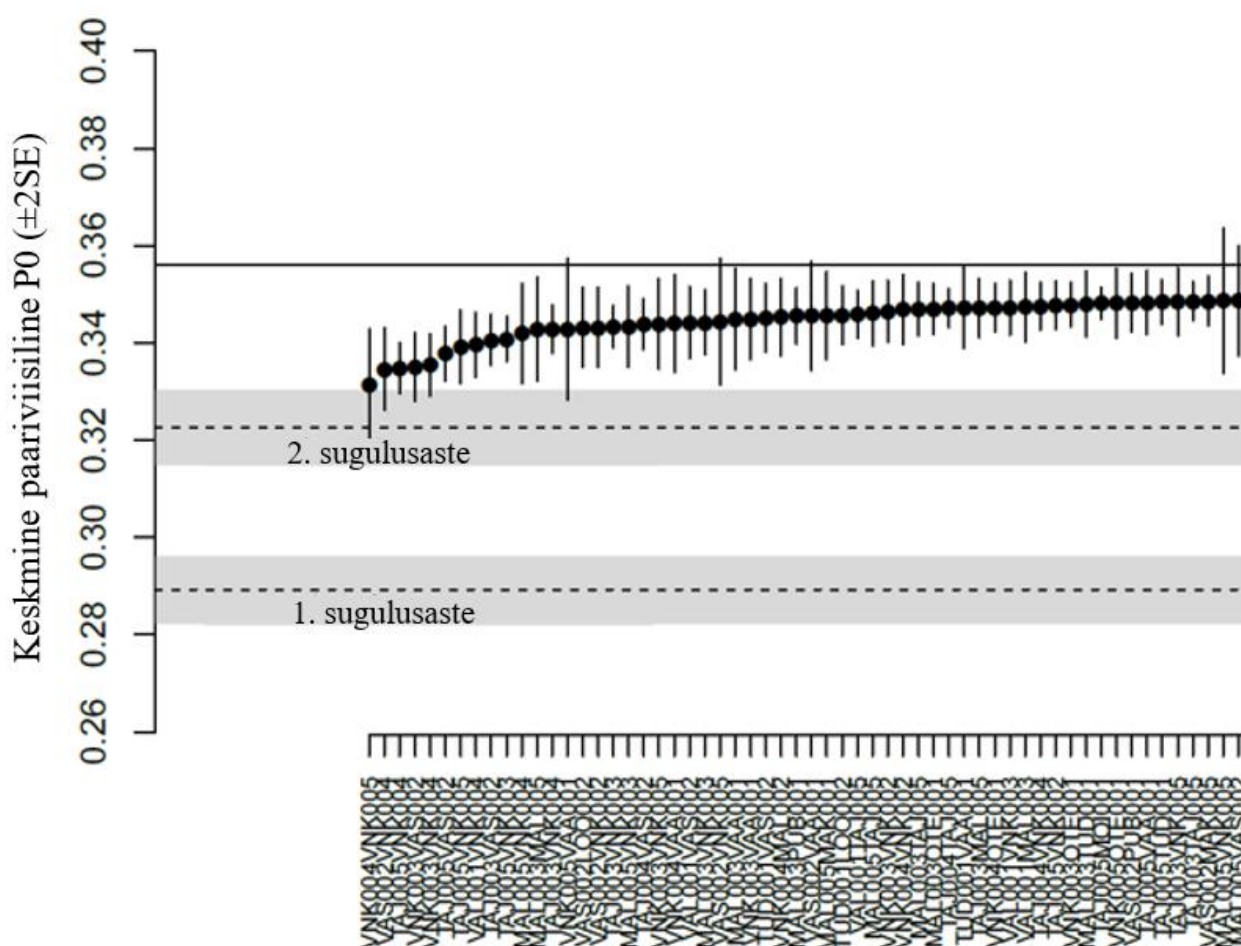
**Joonis 8. Peakomponentanalüüs Euroopa tänapäevaste indiviidide ja keskaegse Tartu kirikute indiviidide autosomaalse varieeruvuse põhjal.** Halliga on tähistatud tänapäeva Ida-, Põhja- ja Kesk-Euroopa indiviidid. Tumehalliga on tähistatud tänapäevaste eestlaste ja sakslaste populatsioonid. Kollasega on tähistatud Tartu kolme kiriku indiviidid: Tartu toomkirik (TAT), Tartu Püha Maarja kirik (TPM), Tartu Jaanikirik (TAJ).

### 2.3.2.2 Sugulusanalüüs

Sugulusanalüüsis vaadati, kas uuritavate indiviidide seas leidub esimese ja/või teise astme sugulasi. Kõigepealt võrreldi omavahel kõiki uuritavaid keskaegseid indiviide (lisa 5.1). Analüüs näitas, et sugulasi nende seas ei olnud.

Eelnevalt tehtud PCA analüüsides lähtudes (lisa 4, joonis 8), on meil alust arvata, et uuritavad individid ei olnud kõik ühest geneetilisest populatsioonist, vaid jagunesid kaheks grupiks, olles sarnasemad siis kas tänapäeva eestlaste või sakslaste populatsioonidega. Sugulussidemete täpsemaks uurimiseks tehti eraldi vastavate rühmade sisesed sugulusanalüüsid.

Ainult eestlastest koosnev uuring kindlaid sugulussidemeid ei näidanud (lisa 5.2). Tulemuste lähemalt vaatlemisel võivad individid VNK004 ja VNK005 olla 2. astme sugulased (joonis 9), aga kuna tulemus on väga piiripealne, siis ei saa seda kindlalt väita. Tegemist oli mehe ja naisega, kes ei jaganud mtDNA haplogruppi (tabel 2, lisa 1), mis tähendab, et need individid ei olnud naisliini kaudu suguluses.



**Joonis 9. Väljalõige keskaegsete nn. eestlaste sugulusanalüüsist READ (lisa 5.2). P0 – iga individide paari kohta arvatud mittekattuvate alleelide osakaal igas 1 megaaluspaarises genoomses aknas. Pidev joon tähistab uuritavate populatsioonide mediaanväärtust, katkendjooned tähistavad 1. ja 2. sugulusastme lävendeid ning neid ümbritsev hall ala tähistab 95% usaldusvahemikku, SE – standardviga ( $\pm 2$ ).**

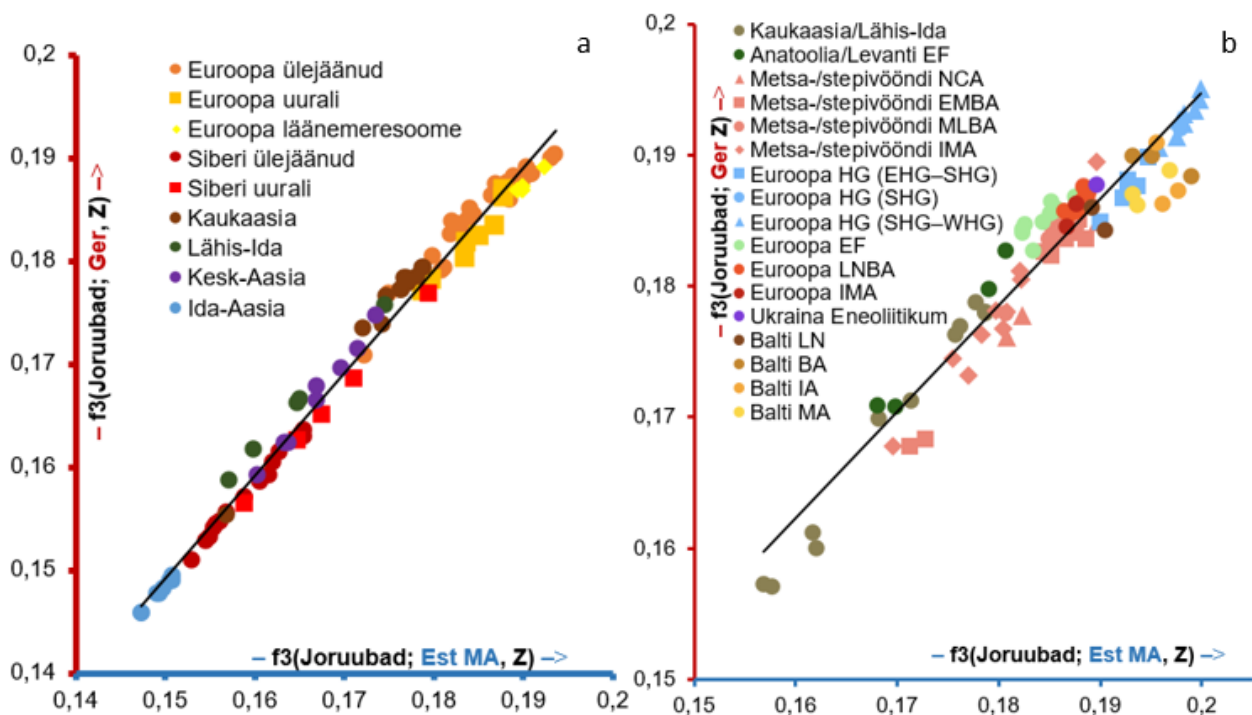
Eraldi uuring tehti ka nn. sakslaste kohta (lisa 5.3), kuhu kuulusid Tartu toomkiriku ja osad Tartu Jaani kiriku invidiidid. Ka selles analüüsis ei tuvastatud sugulasi.

Antud uurimistöös kasutatud indiviidide seast sugulaste leidmise tõenäosus oli väike, kuna proovid olid võetud Eesti eri paikadest. Suuremalt jaolt tehti see analüüs just Tartu indiviide silmas pidades. Tihtipeale maetakse ühe pere liikmed samasse matusepaika, seega tõenäosus ühe kalmistu piires sugulasi leida on suurem. Antud töös uuritud indiviidide vahel lähedasi sugulussidemeid ei tuvastatud, samas tuleb meeles pidada, et need invidiidid on ainult väike osa kalmistusse maetud inimestest. Selleks, et saada parem ülevaade kalmistutesse maetud võimalike sugulaste kohta, on vaja analüüsida rohkem proove ühelt või mitmelt üksteisele lähedal asuvalt kalmistult. Selle uurimisega on nüüdseks juba alustatud.

### **2.3.2.3 f3 ja D statistikud**

Populatsioonidevaheliste geneetiliste kauguste uurimiseks arutati f3 ja D statistikud, mis näitavad alleelisageduste korrelatsiooni populatsioonide vahel. Kuna f3 ja D analüüside tulemused on oma olemuselt võrdväärased, siis on edaspidi täpsemalt juttu f3 statistikute tulemustest. Vastavad D statistikute graafikud on toodud lisa 7).

Võrreldes Eesti keskaegseid indiviide (Est MA) või tänapäeva sakslaste populatsiooni (Ger) tänapäeva Euraasia populatsioonidega (Z), kasutades välisgrupina aafriklasti (joruuba), ning vastandades neid võrdlusi omavahel (joonis 10a), on näha keskaegsete Eesti ala indiviidide veidi suuremat sarnasust tänapäeva läänemeresoome (joonisel kollane) ja teisi uurali keeli (joonisel heleoranž) kõnelevate rahvastega ja sakslaste suuremat sarnasust Lähis-Ida (roheline) ja Kaukaasia (joonisel pruun) populatsioonidega.



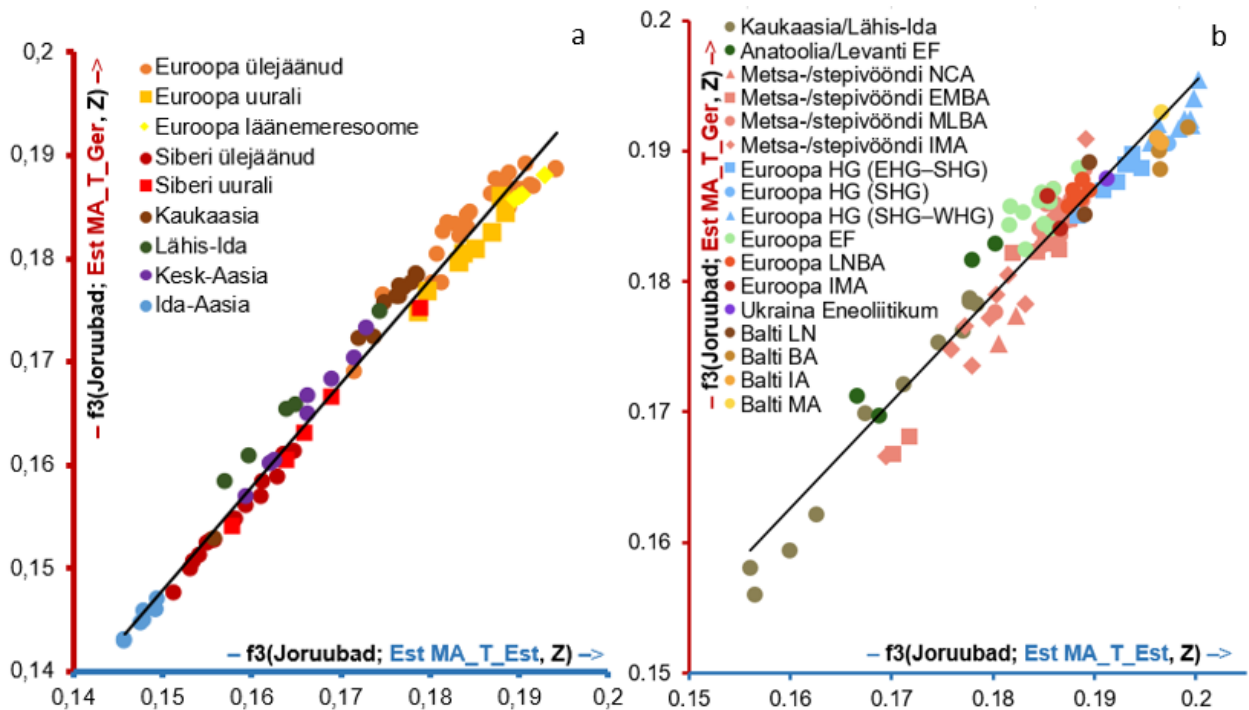
**Joonis 10. Keskaegsete eestlaste ja tänapäeva sakslaste geneetiline sarnasus  $f_3$ -statistiku alusel a) tänapäeva Euraasia populatsioonidega ja b) muinasaja populatsioonidega.**

Kummalgi paneelil on vastandatud kahe testi tulemused, kus nii Eesti keskaegseid indiviide (Est MA) kui tänapäeva sakslaste populatsiooni (Ger) on esmalt võrreldud a) tänapäeva Euraasia populatsioonidega, ja b) muinasaja populatsioonidega ning seejärel teise kihina võrreldud saadud väärtusi kahe uuritava populatsiooni (Est MA ja Ger) vahel. Lühendite tähendused: EF – varajased põlluharijad; NCA – neoliitikum/vase-kiviaeg; EMBA – vanem/keskmise pronksiaeg; IMA – raua-/keskaeg; HG – kütt-korilased (E – ida, W – lääne, S – Skandinaavia), LNBA – hilisneoliitikum/pronksiaeg, LN – hilisneoliitikum, BA – pronksiaeg, IA – rauaaeg, MA – keskaeg. Mida suuremad on punkti koordinaatide väärtused, seda sarnasem on vastav populatsioon uuritavatele populatsioonidele.

Võrreldes tänapäeva sakslaste ja keskaja eestlaste sarnasust ammuste populatsioonidega (joonis 10b) on näha, et tänapäeva sakslaste populatsioon on sarnasem Euroopa ja Anatoolia varajaste põlluharijatega (joonisel rohelised). See on kooskõlas meie varasemate teadmistega põlluharimise leviku kohta (vt. pt. 1.2.2), mille kohaselt esimesed põlluharijad liikusid Lähis-Idast Kesk-Euroopasse ja on seega jätnud märgatavama jälje sakslaste geenipärandisse. Keskaegsed eestlased on aga sarnasemad Baltikumi raua- ja pronksiaegsete indiviididega (joonisel oranž ja helepruun). See on samuti oodatav, kuna need ajastud paiknevad üksteisega lähestikku.

Tartu kirikute indiviidide uurimiseks jaotati nad lähtuvalt PCA-st nn. eestlasteks ja sakslasteks (joonis 8). Indiviidid TAJ001 ja TAT006 jäeti sellest analüüsist välja, kuna nad paiknesid PCA joonisel (joonis 8) tänapäevaste eestlaste ja sakslaste populatsioonide vahepeal ja seega ei saanud neid kumbagi uuritavasse gruppi otseselt rühmitada.

Võrreldes Tartu kirikute indiviidide sarnasust tänapäevaste populatsioonidega (joonis 11a), siis sakslastest keskaegsed indiviidid (Est MA\_T\_Ger) on veidi sarnasemad Lähis-Ida (joonisel roheline) ja Kaukaasia (joonisel pruun) populatsioonidega ning ülejäänud Euroopa populatsioonidega (joonisel tumeoranž) ja eestlased (Est MA\_T\_Est) läänemeresoome (joonisel kollane) ja teiste uurali keeli (joonisel heleoranž) kõnelevate rahvastega. Need tulemused kattuvad PCA tulemustega (joonis 8).



**Joonis 11. Tartu kirikute eestlastest ja sakslastest indiviidide geneetiline sarnasus a) tänapäeva Euraasia populatsioonidega ja b) muinasaja populatsioonidega f3 statistiku alusel.**

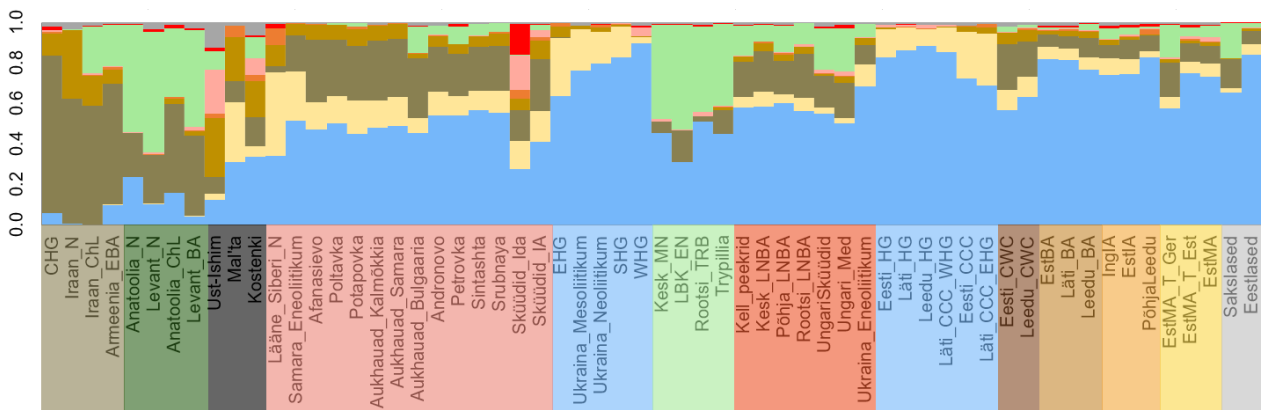
Kummalgi joonisel on vastandatud kahe testi tulemused, kus Tartu kirikute nõ. eestlastest (Est MA\_T\_Est) ja sakslastest (Est MA\_T\_Ger) indiviide on esmalt võrreldud a) tänapäeva Euraasia populatsioonidega, ja b) muinasaja populatsioonidega ning seejärel teise kihina võrreldud saadud väärtusi kahe uuritava populatsiooni (Est MA\_T\_Ger ja Est MA\_T\_Est) vahel. Lühendite tähendusi vt. joonise 10 allkirjast. Mida suuremad on punkti koordinaatide väärtused, seda sarnasem on vastav populatsioon uuritavatele populatsioonidele.

Võrreldes keskaja Tartu kirikute indiviide muinasaja eri perioodide indiviididega (joonis 11b), on jällegi näha mõningast erinevust: sakslastest indiviidid on sarnasemad Euroopa varajaste põlluharijatega (joonisel heleroheline) ja eestlastest indiviidid Euroopa kütt-korilastega (joonisel sinine).

Lisaks tehti analüüsid, kus võrreldi Tartu Maarja kiriku (Est MA\_TM), toomkiriku (Est MA\_TT) ja Jaani kiriku (Est MA\_TJ) indiviide (lisa 6) tänapäeva Euraasia populatsioonidega ning seejärel omavahel. Maarja kiriku ja toomkiriku indiviidide võrdlemisel on näha, et toomkiriku indiviidid sarnanesid rohkem Lähis-Ida (joonisel roheline) ja Kaukaasia (joonisel pruun) populatsioonidele ning ülejäänud Euroopa populatsioonidele (joonisel tumeoranž), sarnaselt seega nõ. sakslastele, nagu oli näha joonisel 11a. Maarja kiriku indiviidid sarnanesid rohkem läänemeresoome (joonisel kollane) ja teiste uurali keeli (joonisel heleoranž) kõnelevate rahvastele, sarnaselt nõ eestlastega (joonis 11a). Kui aga võrrelda Jaani kiriku indiviide nende kummagi kiriku indiviididega, siis paiknesid võrreldavad Euraasia populatsioonid enam vähem kõik ühel sirgel, viidates sellele, et Jaani kiriku indiviidid sarnanevad võrdselt nii toomkiriku kui Maarja kiriku indiviididega. See tulemus toetab kirjanduses öeldut, et Jaani kirikusse maeti erineva etnilise taustaga rahvast (vt. pt. 1.3.1.2.1). Samuti kinnitab see PCA analüüsi (joonis 8), kus Jaani kiriku indiviidide seas oli nii tänapäeva eestlastele kui sakslastele sarnaseid indiviide.

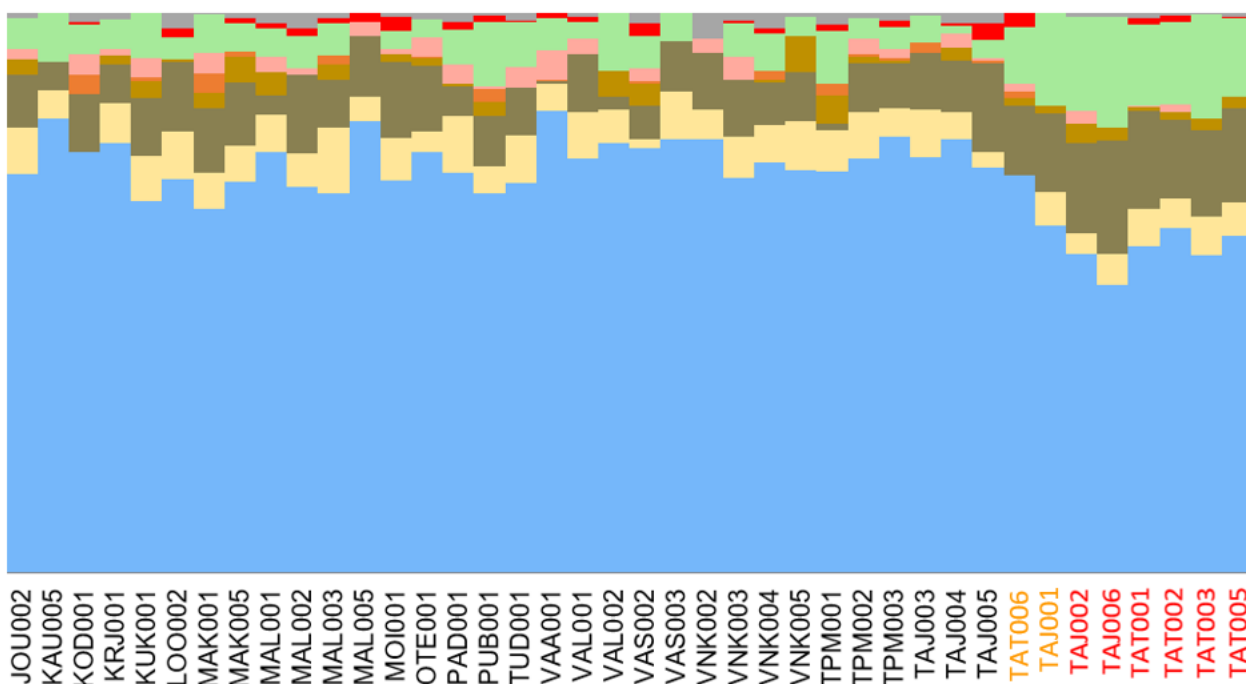
#### **2.3.2.4 Admixture analüüs**

Admixture analüüsiga uuriti geneetiliste eellaskomponentide osakaalu uuritavates ja võrdlusena kasutatud populatsioonides ning indiviidides. Lähtuvalt PCA tulemustest jaotati keskaegsed Eesti ala indiviidid nõ eestlasteks ja sakslasteks (lisa 4). Admixture analüüsist on näha, et Tartu linna keskaegsete sakslastest indiviidide populatsioon erineb selgelt sama ajajärgu eestlaste populatsioonist (joonis 12). Keskaegsed sakslastest indiviidid sarnanevad rohkem tänapäeva sakslaste populatsiooniga kui keskaegsete või tänapäevaste eestlastega.



**Joonis 12. Admixture analüüsi väljastatud geneetiliste komponentide osakaalud populatsioonide keskmistena.** X-telje värvide tähistused vasakult: Kaukaasia ja Iraani populatsioonid; Anatoolia ja Levanti esimesed põlluharijad; paleoliitilised kütt-korilased; stepi populatsioonid; Euroopa kütt-korilased, varajased põlluharijad, stepi päritolu põlluharijad; Läänemere idaranniku kütt-korilased, nöörikeramikud, pronksiaja populatsioonid, rauaaja populatsioonid, keskaja populatsioonid; tänapäeva sakslaste ja eestlaste populatsioonid. Lühendite tähendused: CHG – Kaukaasia kütt-korilased, N – neoliitikum, ChL – eneoliitikum, EBA – vanem pronksiaeg, BA – pronksiaeg, IA – rauaeg, EHG – ida kütt-korilased, SHG – Skandinaavia kütt-korilased, WHG – lääne kütt-korilased, MN – keskneoliitikum, LBK – paelkeramika kultuur, EN – varaneoliitikum, TRB – lehterpeekrite kultuur, LNBA – hilisneoliitikum/pronksiaeg, Med – keskaeg, HG – kütt-korilased, CCC – kammkeramika kultuur, CWC – nöörikeramika kultuur, MA – keskaeg/varauusaeg, T – Tartu, Ger – sakslased, Est – eestlased.

Vaadates lähemalt siinse uurimistöö indiviidide geneetilisi profile (joonis 13), on näha, et nõ eestlaste ja sakslaste erinevus geneetiliste eellaskomponentide osakaalus ei tulene ainult mõnest indiviidist, vaid on üldine. Sakslaste valim on küll väike, aga kõigil neist on näha rohkem varajaste põlluharijate komponenti (roheline) kui eestlastel.

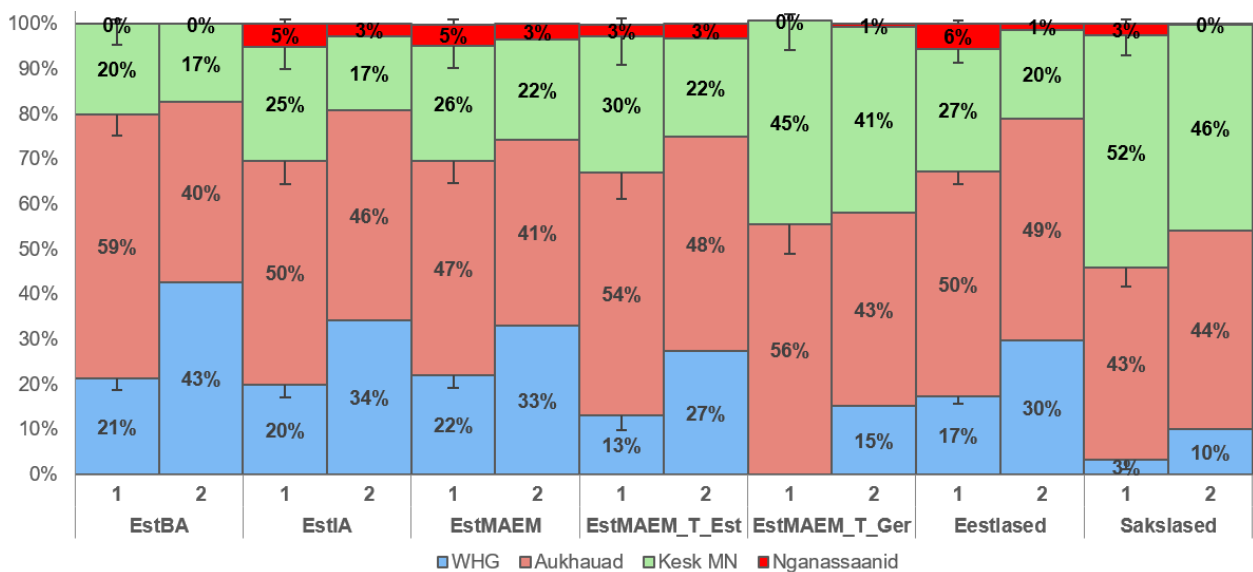


**Joonis 13. Antud uurimistöö keskaegsete Eesti ala indiviidide Admixture analüüsi tulemused.**

Punase kirjaga on märgitud geneetiliselt tänapäeva sakslastega sarnased individid ning oranžiga individid, kes jäid PCA joonisel 8 tänapäevaste sakslaste ja eestlaste populatsioonide vahepeale. Väljalõige jooniselt lisa 8.1.

### 2.3.2.5 qpADM ja ChromoPainter/NNLS

Eellaskomponentide osakaalu uurimiseks tehti lisaks veel qpADM ja ChromoPainter/NNLS analüüsid, kus on välja toodud lääne kütt-korilaste (joonisel sinine), stepi (roosa), varajaste põlluharijate (roheline) ja tänapäevase „Siberi“ komponendi (punane) osakaalud erinevates populatsioonides (joonis 14). Jooniselt on näha, et uuritud pronksiaja eestlastel „Siberi“ komponent puudus ning esimest korda võis seda näha rauaaegsetel indiviididel (vt. pt. 1.2.2). Uuritavatel keskaja indiviididel on see samuti säilinud. Küll aga on näha väikest erinevust Tartu keskaegsete nõ eestlaste ja sakslaste vahel – sakslastel (EstMAEM\_T\_Ger) „Siberi“ komponent praktiliselt puudub (0-1%).



**Joonis 14. Erinevate Eesti populatsioonide ja tänapäeva sakslaste geneetiliste eellaskomponentide osakaalud populatsioonipõhisest qpAdm analüüsi (1) ja indiviidipõhisest CP/NNLS analüüsi (2) tulemustest.** Lääne kütt-korilaste (WHG), stepist tulijate (Aukhauad), varajaste põlluharijate (Kesk MN) ja tänapäeva Siberi (Nganassaanid) komponentide osakaal pronksiaja eestlaste (EstBA), rauaaja eestlaste (EstIA), keskaja külasurnuaedade eestlaste (EstMA), keskaegsete Tartu kirikumatuste eestlaste (EstMA\_T\_Est) ning sakslaste (EstMA\_T\_Ger) ning tänapäeva eestlaste ja sakslaste populatsioonide seas.

Vaadates „Siberi“ komponendi jagunemist indiviidide seas lähemalt (lisa 8), on samuti näha, et võrreldes keskaegsete nõ. sakslastega, on eestlaste seas „Siberi“ komponent suhteliselt laialdaselt levinud. „Siberi“ komponendi selge olemasolu muinasaegsetel ja keskaegsetel eestlastel ning selle vähene esinemissagedus keskaegsetel nõ. sakslastel, kuid suurem esinemissagedus tänapäeva sakslastel, võib viidata aja jooksul toimunud liikumistele ja segunemistele, mille käigus on idast tulnud „Siberi“ komponent ka kaasaegsete sakslaste populatsiooni edasi kandunud.

## Kokkuvõte

Käesoleva töö eesmärgiks oli heita valgust keskaegse Eesti elanikkonna geneetilisele struktuurile. Selleks eraldati DNA 30 Eesti keskaegse indiviidi hammastest, mis olid saadud Tartu Ülikooli ajaloo ja arheoloogia instituudi kogudest.

Uuritavast 30 proovist 24 olid piisava genoomse katvusega ( $>0,02x$ ) edasiste populatsioonigeneetiliste analüüside tegemiseks. Sugu õnnestus kindlaks määrata 23 indiviidil, kellest 14 olid naised ning 9 mehed. Keskaegsete indiviidide seas olid emaliinis levinud mtDNA haplogrupid H, U, T, J ja K ning Y kromosoomi haplogruppidest olid esindatud R1a, N3 ja I. Samad haplogrupid on Eestis levinud ka täna.

Populatsioonigeneetiliste analüüsidega uuriti lähemalt Eesti keskaegse populatsiooni struktuuri. Analüüsid näitasid keskaegsete eestlaste suurimat sarnasust tänapäeva uurali keeli kõnelevate rahvaste ning Eesti pronksi- ja rauaaja populatsioonidega, mis oli ootuspärane. Lisaks tuvastati, et Tartu toomkirikusse ja osad Tartu Jaani kirikusse maetud indiviidid olid kõige sarnasemad hoopis tänapäeva sakslastega. Need tulemused toetasid varasemaid uurimistöid, mis tuginesid arheoloogia- ja ajalooallikatele ja kus leiti, et toomkiriku kalmistule on valdavalt maetud sakslaste, Maarja kirikusse eestlaste kogukonna liikmeid ning Jaani kiriku kalmistule mõlemaid.

Lisaks uuriti lähemalt ka uurali keeli rääkivate rahvaste seas esineva „Siberi“ komponendi olemasolu Eesti muinas- ja keskaja ning tänapäevaste eestlaste ja sakslaste populatsioonis. Keskaegsetel Eesti ala sakslastel „Siberi“ komponent peaaegu puudus, samas kui Eesti aladel oli see leitav juba rauaajal ning on olemas ka tänapäeva eestlastel. Samuti on antud komponent olemas ka tänapäeva sakslastel, kuid vähemal määral kui eestlastel.

Käesoleva magistr töö tulemused annavad esmavaate keskaegse Eesti elanikkonna geneetilisele struktuurile ning on kooskõlas ajaloolaste ja arheoloogide uuringutega Eesti keskaja rahvastiku kohta.

Magistr töö teostati TÜ genoomika instituudi projekti „Keskaegse populatsiooni struktuur, tervis ja haigused Eesti territooriumil vana DNA vaatenurgast“ (PP1GI19936) raames.

# **Population structure of Medieval Estonia through aDNA perspective**

Kadri Irdt

## **Summary**

Throughout time, people have always moved around and explored the world. This has left a mark in their culture and genes. Different genetic components have been passed down through many generations and most of these can be found in modern day populations. The components that have not reached the present can be tracked down using ancient DNA.

Medieval Estonia has been studied by archaeologist and historians amongst others but has not yet been studied through genetics. Throughout time, many people with different genetical backgrounds have lived in the area of Estonia. The most well-known are Germans who came here during the Crusades on 13<sup>th</sup> century AD and became part of the Estonian population.

The aim of this thesis was to understand the genetic structure of the people who lived in Estonia during the Middle Ages. For that, DNA was extracted from the teeth of 30 Medieval Estonian individuals. Out of these 30 samples, 24 had sufficient genomic coverage ( $>0,02x$ ) to be used in population genetic analyses. Genetically, 14 females and 9 males could be identified. The mitochondrial DNA (mtDNA) and Y chromosome (chrY) haplogroups found in medieval Estonia (mtDNA H, U, T, J and K; chrY R1a, N3 and I) also represent the present-day Estonian population well.

The genetic structure of the medieval Estonian population was studied further with population genetic analyses, using the autosomal aDNA data. These analyses revealed that medieval Estonians were most similar to modern Uralic-speaking populations and to Estonian Bronze and Iron Age populations, which was to be expected. However, it was seen that in Tartu, the people buried in the Cathedral and some buried in St. John's church were most similar to modern Germans. These results provide support for archaeological and historical studies which showed that mostly Germans were buried in the Cathedral, mostly Estonians in St. Mary's church and both in St. John's church.

One particular genetic component, the "Siberian" component present in Uralic speakers, was studied more closely in this thesis. This component was first found in Estonia during the Iron Age and can also be found in modern-day Estonians. However, in medieval Germans who lived in

Estonia, this component was almost non-existent. In modern Germans, the "Siberian" component can be found but in lower frequency than in modern Estonians.

The results of this thesis give a first look of the genetic structure of medieval Estonia and are in accordance with the results of archeological and historical studies.

This thesis is part of the UT Institute of Genomics project called "Population structure, health and disease in Medieval Estonia through aDNA perspective" (PP1GI19936).

## **Tänuõnad**

Kõigepealt tahan tänada oma juhendajaid Lehti Saagi ja Kristiina Tambetsit. Tänu Kristiinale saan olla osa TÜ genoomika instituudi projektist „Keskaegse populatsiooni struktuur, tervis ja haigused Eesti territooriumil vana DNA vaatenurgast“ (PP1GI19936) ning tema toetus kogu selles ettevõtmises on olnud suur. Lisaks tahan tänada Martin Malvet ja Heiki Valku, ilma kellela ei oleks mul olnud isegi proove, mida analüüsida. Suurima tänu tahan öelda aga Lehtile, kellel jätkus kannatust mind sellel kriisiajal interneti teel serveri kasutamises juhendada ning kes alati oli valmis kõigile minu küsimustele vastama. Ilma Lehtita ei oleks ma praegu siin, kus ma olen: tema oli esimene, kes mulle seda teadusala tutvustas ning suunas mind sellele põnevale teele juba bakalaureuseõppes, olles ka seal mulle toredaks juhendajaks. Lisaks tahan tänada kõiki evolutsioonibioloogia osakonna ja eriti vana DNA töögrupi inimesi, kes on alati olnud väga toredad ja toetavad.

## **Kasutatud kirjanduse loetelu**

- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Allentoft, M. E., Sikora, M., Sjögren, K. G., ... Willerslev, E. (2015). Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), 167–172. <https://doi.org/10.1038/nature14507>
- Alttoa, K. (2009). Die Tartuer Marienkirche und die Frage der Bauschule in Tartu im 14. Jahrhundert. *Baltic Journal of Art History*, 7–30.
- Anderson, S., Bankier, A. T., Barrell, B. G., ... Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*. <https://doi.org/10.1038/290457a0>
- Anthony, D. W. (2007). *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppe shaped the Modern World*. Princeton and Oxford: Princeton University Press.
- Behar, D. M., Metspalu, M., Baran, Y., ... Rosenberg, N. A. (2013). No Evidence from Genome-Wide Data of a Khazar Origin for the Ashkenazi Jews. *Hum Biol.* 85(6), 859–900. <https://doi.org/10.3378/027.085.0604>
- Behar, D. M., Yunusbayev, B., Metspalu, M., ... Villems, R. (2010). The genome-wide structure of the Jewish people. *Nature*, 466(7303), 238–242. <https://doi.org/10.1038/nature09103>
- Benazzi, S., Douka, K., Fornai, C., ... Weber, G. W. (2011). Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature*, 479(7374), 525–528. <https://doi.org/10.1038/nature10617>
- Bramanti, B., Thomas, M. G., Haak, W., ... Burger, J. (2009). Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, 326(5949), 137–140. <https://doi.org/10.1126/science.1176869>
- Briggs, A. W., Stenzel, U., Johnson, P. L. F., ... Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *PNAS*, 104(37), 14616–14621. <https://doi.org/10.1073/pnas.0704665104>
- Colonna, V., Pagani, L., Xue, Y., and Tyler-Smith, C. (2011). A world in a grain of sand: Human history from genetic data. *Genome Biol.* 12:234, <https://doi.org/10.1186/gb-2011-12-11-234>

- Damgaard, P. B., Marchi, N., Rasmussen, S., ... Willerslev, E. (2018). 137 ancient human genomes from across the Eurasian steppes. *Nature*, 557(7705), 369–374. <https://doi.org/10.1038/s41586-018-0094-2>
- Damgaard, P. B., Martiniano, R., Kamm, J., ... Willerslev, E. (2018). The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science*, 360(6396). <https://doi.org/10.1126/science.aar7711>
- Douglas Price, T. (2000). *Europe's First Farmers*. Cambridge university Press. <https://doi.org/10.1017/cbo9780511607851>
- Fedorova, S. A., Reidla, M., Metspalu, E., ... Villems, R. (2013). Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): Implications for the peopling of Northeast Eurasia. *BMC Evol Biol.* 13(1), 127. <https://doi.org/10.1186/1471-2148-13-127>
- Fort, J. (2015). Demic and cultural diffusion propagated the Neolithic transition across different regions of Europe. *J R Soc Interface*, 12(106), 20150166. <https://doi.org/10.1098/rsif.2015.0166>
- Fu, Q., Mittnik, A., Johnson, P. L. F., ... Krause, J. (2013). A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol.* 23(7), 553–559. <https://doi.org/10.1016/j.cub.2013.02.044>
- Gilbert, M. T. P., Kivisild, T., Grønnow, B., ... Willerslev, E. (2008). Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. *Science*, 320(5884), 1787–1789. <https://doi.org/10.1126/science.1159750>
- Haak, W., Balanovsky, O., Sanchez, J. J., ... Cooper, A. (2010). Ancient DNA from European early Neolithic farmers reveals their near eastern affinities. *PLoS Biol.* 8(11), 1000536. <https://doi.org/10.1371/journal.pbio.1000536>
- Haak, W., Forster, P., Bramanti, B., ... Burger, J. (2005). Ancient DNA from the first European farmers in 7500-year-old neolithic sites. *Science*, 310(5750), 1016–1018. <https://doi.org/10.1126/science.1118725>
- Haak, W., Lazaridis, I., Patterson, N., ... Reich, D. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), 207–211. <https://doi.org/10.1038/nature14317>

- Handt, O., Richards, M., Trommsdorff, M., ... Pääbo, S. (1994). Molecular genetic analyses of the Tyrolean ice man. *Science*, 264(5166), 1775–1778. <https://doi.org/10.1126/science.8209259>
- Herrmann, B., and Hummel, S. (1994). *Ancient DNA. Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical, and Forensic Specimens*. New York: Springer-Verlag. <https://doi.org/10.1007/978-1-6412-4318-2>
- Hofreiter, M. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res.* 29(23), 4793–4799. <https://doi.org/10.1093/nar/29.23.4793>
- Ilumäe, A. M., Reidla, M., Chukhryaeva, M., ... Rootsi, S. (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. <https://doi.org/10.1016/j.ajhg.2016.05.025>
- Jobling, M. A., and Tyler-Smith, C. (2003). The human Y chromosome: An evolutionary marker comes of age. *Nat Rev Genet.* 4(8), 598-612. <https://doi.org/10.1038/nrg1124>
- Jobling, M., Hollox, E., Hurles, M., Kivisild, T., and Tyler-Smith, C. (2014). *Human Evolutionary Genetics* (2nd ed.). New York and London: Garland Science, Taylor & Francis Group, LLC.
- Jobling, M., Hurles, M., and Tyler-Smith, C. (2004). *Human Evolutionary Genetics: Origins, Peoples & Disease*. Garland Science, Taylor & Francis Group, LLC.
- Jones, E. R., Zarina, G., Moiseyev, V., Lightfoot, E., Nigst, P. R., Manica, A., Pinhasi, R., and Bradley, D. G. (2017). The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Curr Biol.* 27(4), 576–582. <https://doi.org/10.1016/j.cub.2016.12.060>
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., and Orlando, L. (2013). MapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13), 1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>
- Kalling, K. (1995). Paleoantropoloogilisi andmeid Tartu Jaani kiriku kalmistu 13.-14. sajandi matuste kohta. *Tartu Arheoloogias Ja Vanemast Ehitusloost*, Tartu Ülikooli Arheoloogia Kabineti Toimetised, 8, 47–58.

- Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., and Hammer, M. F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res*, 18(5), 830–838. <https://doi.org/10.1101/gr.7172008>
- Karmin, M., Saag, L., Vicente, M., ... Kivisild, T. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res*, 25(4), 459–466. <https://doi.org/10.1101/gr.186684.114>
- Kivisild, T. (2017). The study of human Y chromosome variation through ancient DNA. *Hum Genet*. 136(5), 529-546. <https://doi.org/10.1007/s00439-017-1773-z>
- Korneliussen, T. S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, 15(1), 1–13. <https://doi.org/10.1186/s12859-014-0356-4>
- Krause, J., Briggs, A. W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., and Pääbo, S. (2010). A Complete mtDNA Genome of an Early Modern Human from Kostenki, Russia. *Curr Biol*. 20(3), 231–236. <https://doi.org/10.1016/j.cub.2009.11.068>
- Kriiska, A. (2000). Settlements of coastal Estonia and maritime hunter-gatherer economy. *Lietuvos Archeologija*, 19, 153–166.
- Kriiska, A. (2009). The beginning of farming in the Eastern Baltic area. The East European Plain on the Eve of Agriculture. *Brit Archaeol Rep In*, 159–179.
- Kriiska, A., Mäesalu, A., Selart, A., Põltsam-Jürjo, I., Piirimäe, P., Seppel, M., Andresen, A., Pajur, A., ja Tannberg, T. (2017). *Eesti Ajalugu*. Tallinn: Avita.
- Kriiska, A., ja Tvauri, A. (2002). *Eesti muinasaeg*. Avita. lk 54-84.
- Kuhn, J. M. M., Jakobsson, M., and Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. *PLOS ONE*, 13(4), e0195491. <https://doi.org/10.1371/journal.pone.0195491>
- Kushniarevich, A., Utevska, O., Chuhryaeva, M., ... Balanovsky, O. (2015). Genetic Heritage of the Balto-Slavic Speaking Populations: A Synthesis of Autosomal, Mitochondrial and Y-Chromosomal Data. *PLOS ONE*, 10(9), e0135820. <https://doi.org/10.1371/journal.pone.0135820>

- Lazaridis, I., Nadel, D., Rollefson, G., ... Reich, D. (2016). Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), 419–424. <https://doi.org/10.1038/nature19310>
- Lazaridis, I., Patterson, N., Mittnik, A., ... Krause, J. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), 409–413. <https://doi.org/10.1038/nature13673>
- Lawson, D. J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), e1002453. <https://doi.org/10.1371/journal.pgen.1002453>
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, J. Z., Absher, D. M., Tang, H., ... Myers, R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319(5866), 1100–1104. <https://doi.org/10.1126/science.1153717>
- Malve, M. (2010). Kaks luustikku Tartu toomkirikust: arheoloogia ja osteoloogia andmed. *Õpetatud Eesti Seltsi Aastaraamat*, 65–84.
- Malve, M., Roog, R., and Tvauri, A. (2011). Preliminary Results of the Rescue Excavation in St Mary's Churchyard and its surroundings in Tartu 2010-2011. *Arheoloogilised Välitööd Eestis = Archeological Fieldwork in Estonia*, 137–150.
- Malve, M., and Valk, H. (2009). Archaeological Excavations in the Counties of Harjumaa. *Arheoloogilised Välitööd Eestis = Archeological Fieldwork in Estonia*, 2008, 141–150.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17, 10–12. <https://doi.org/https://doi.org/10.14806/ej.17.1.200>
- Mathieson, I., Alpaslan-Roodenberg, S., Posth, C., ... Reich, D. (2018). The genomic history of southeastern Europe. *Nature*, 555(7695), 197–203. <https://doi.org/10.1038/nature25778>

- Mathieson, I., Lazaridis, I., Rohland, N., ... Reich, D. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), 499–503. <https://doi.org/10.1038/nature16152>
- McKenna, A., Hanna, M., Banks, E., ... DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Microsoft Corporation. (2019). Microsoft Excel. <https://office.microsoft.com/excel>
- Mittnik, A., Wang, C.-C., Pfrengle, S., ... Krause, J. (2018). The genetic prehistory of the Baltic Sea region. *Nat Commun*, 9:442. <https://doi.org/10.1038/s41467-018-02825-9>
- Narasimhan, V., Patterson, N., Moorjani, P., ... Reich, D. (2018). The Genomic Formation of South and Central Asia. *BioRxiv*, 292581. <https://doi.org/10.1101/292581>
- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. *Genetics*, 192(3), 1065–1093. <https://doi.org/10.1534/genetics.112.145037>
- Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet*, 2(12), 2074–2093. <https://doi.org/10.1371/journal.pgen.0020190>
- Pinhasi, R., Fort, J., and Ammerman, A. J. (2005). Tracing the origin and spread of agriculture in Europe. *PLoS Biol*, 3(12), 1–9. <https://doi.org/10.1371/journal.pbio.0030410>
- Poinar, H. N., Schwarz, C., Qi, J., ... Schuster, S. C. (2006). Metagenomics to paleogenomics: Large-scale sequencing of mammoth DNA. *Science*, 311(5759), 392–394. <https://doi.org/10.1126/science.1123360>
- Posth, C., Renaud, G., Mittnik, A., ... Krause, J. (2016). Pleistocene mitochondrial genomes suggest a single major dispersal of non-africans and a late glacial population turnover in Europe. *Curr Biol*. 26(6), 827–833. <https://doi.org/10.1016/j.cub.2016.01.037>
- Purcell, S., Neale, B., Todd-Brown, K., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, 81(3), 559–575. <https://doi.org/10.1086/519795>
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>

- R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.r-project.org>
- Racimo, F., Woodbridge, J., Fyfe, R. M., Sikora, M., Sjögren, K. G., Kristiansen, K., and Linden, M. Vander. (2020). The spatiotemporal spread of human migrations during the European Holocene. *P Natl Acad Sci USA*, 117(16), 8989–9000. <https://doi.org/10.1073/pnas.1920051117>
- Raghavan, M., Skoglund, P., Graf, K. E., ... Willerslev, E. (2014). Upper palaeolithic Siberian genome reveals dual ancestry of native Americans. *Nature*, 505(7481), 87–91. <https://doi.org/10.1038/nature12736>
- Rasmussen, M., Guo, X., Wang, Y., ... Willerslev, E. (2011). An aboriginal Australian genome reveals separate human dispersals into Asia. *Science*, 334(6052), 94–98. <https://doi.org/10.1126/science.1211177>
- Rasmussen, M., Li, Y., Lindgreen, S., ... Willerslev, E. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*, 463(7282), 757–762. <https://doi.org/10.1038/nature08835>
- Renfrew, C. (1989). The Origins of Indo-European Languages. *Sci Am*. 261(4), 106–115. <https://doi.org/10.2307/24987446>
- Richards, M. B., Macaulay, V., Bandelt, H. J., and Skyes, B. C. (1998). Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet*, 62(3), 241–260. <https://doi.org/10.1046/j.1469-1809.1998.6230241.x>
- Roberts, N., Fyfe, R. M., Woodbridge, J., ... Leydet, M. (2018). Europe’s lost forests: A pollen-based synthesis for the last 11,000 years. *Sci Rep*, 8(1), 1–8. <https://doi.org/10.1038/s41598-017-18646-7>
- Saag, L., Laneman, M., Varul, L., ... Tambets, K. (2019). The Arrival of Siberian Ancestry Connecting the Eastern Baltic to Uralic Speakers further East. *Curr Biol*. 29(10), 1701-1711.e16. <https://doi.org/10.1016/j.cub.2019.04.026>
- Saag, L., Varul, L., Scheib, C. L., ... Metspalu, M. (2017). Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Curr Biol*. 27(14), 2185-2193.e6. <https://doi.org/10.1016/j.cub.2017.06.022>

- Sala, M. (2002). Igaviku tee. Eesti matusetraditsioonid ja tänapäev ning hauaplatside kujundamine. (R. Tammsaar,Ed.). Tallinn: Maalehe Raamat. lk 57
- Selart, A., Kala, T., Kaljundi, L., ... Valk, H. (2012). Eesti ajalugu II, Eesti keskaeg. Tartu: Tartu Ülikooli ajaloo ja arheoloogia instituut.
- Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J., and Barbujani, G. (2000). Geographic patterns of mtDNA diversity in Europe. *Am J Hum Genet*, 66(1), 262–278. <https://doi.org/10.1086/302706>
- Skoglund, P., Malmström, H., Omrak, A., ... Jakobsson, M. (2014). Genomic diversity and admixture differs for stone-age Scandinavian foragers and farmers. *Science*, 344(6185), 747–750. <https://doi.org/10.1126/science.1253448>
- Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M. T. P., Götherström, A., and Jakobsson, M. (2012). Origins and genetic legacy of neolithic farmers and hunter-gatherers in Europe. *Science*, 336(6080), 466–469. <https://doi.org/10.1126/science.1216304>
- Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci*, 40(12), 4477–4482. <https://doi.org/10.1016/j.jas.2013.07.004>
- Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H. J., Torroni, A., and Richards, M. B. (2010). The Archaeogenetics of Europe. *Curr Biol*. <https://doi.org/10.1016/j.cub.2009.11.054>
- Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A., Salas, A., Oppenheimer, S., Macaulay, V., and Richards, M. B. (2009). Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *Am J Hum Genet*, 84(6), 740–759. <https://doi.org/10.1016/j.ajhg.2009.05.001>
- Tambets, K., Yunusbayev, B., Hudjashov, G., ... Metspalu, M. (2018). Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biol*. 19(1), 139. <https://doi.org/10.1186/s13059-018-1522-1>
- Torroni, A., Achilli, A., Macaulay, V., Richards, M., and Bandelt, H. J. (2006, June 1). Harvesting the fruit of the human mtDNA tree. *Trends Genet*. <https://doi.org/10.1016/j.tig.2006.04.001>

- Underhill, P. A., Myres, N. M., Rootsi, S., ... Kivisild, T. (2010). Separating the post-Glacial coancestry of European and Asian y chromosomes within haplogroup R1a. *Eur J Hum Genet*, 18(4), 479–484. <https://doi.org/10.1038/ejhg.2009.194>
- Unterländer, M., Palstra, F., Lazaridis, I., ... Burger, J. (2017). Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nat Commun*, 8(1), 1–10. <https://doi.org/10.1038/ncomms14615>
- Valk, H. (1994). Neighbouring but distant: Rural burial traditions of Estonia and Finland during the Christian Period. *Fennoscandia Archaeologica*, 11, 61–76.
- Valk, H. (1995). Tartu toomkiriku kalmistust ja ümbruskonna varasemast asustusest. *Tartu Arheoloogiast Ja Vanemast Ehitusloost*, Tartu Ülikooli Arheoloogia Kabineti Toimetised, 8, 59–80.
- Valk, H. (2003). From Paganism to Christianity The Christianisation of Estonia and Changes in Burial Customs. *Wachholz Verlag Neumünster*, 215–222.
- Valk, H. (2004). Grave Goods in Estonian Urban Churchyards: a Reflection of Different Traditions and Ideologies. *Culture Clash or Compromise. CCC Papers*, 8, 103–112.
- van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*, 30(2), E386–E394. <https://doi.org/10.1002/humu.20921>
- Van Rossum, G., and Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.
- Wallace, D. C. (1995). Mitochondrial DNA Variation in Human Evolution, Degenerative Disease, and Aging. *Am J Hum Genet*, 57(2), 201–223.
- Wallace, D. C., Brown, M. D., and Lott, M. T. (1999). Mitochondrial DNA variation in human evolution and disease. *Gene*, 238(1), 211–230. [https://doi.org/10.1016/S0378-1119\(99\)00295-4](https://doi.org/10.1016/S0378-1119(99)00295-4)
- Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H. J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.*, 44(W1), W58–W63. <https://doi.org/10.1093/nar/gkw233>

Yunusbayev, B., Metspalu, M., Järve, ... Villems, R. (2012). The caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol*, 29(1), 359–365. <https://doi.org/10.1093/molbev/msr221>

Yunusbayev, B., Metspalu, M., Metspalu, E., ... Villems, R. (2015). The Genetic Legacy of the Expansion of Turkic-Speaking Nomads across Eurasia. *PLOS Genet.*, 11(4), e1005068. <https://doi.org/10.1371/journal.pgen.1005068>

### **Kasutatud veebiaadressid**

<https://qgis.org/en/site/> (17.03.2020)

[https://geoportaal.maaamet.ee/index.php?lang\\_id=1&page\\_id=119](https://geoportaal.maaamet.ee/index.php?lang_id=1&page_id=119) (17.03.2020)

<https://haplogrep.i-med.ac.at/app/index.html> (25.03.2020)

<https://cutadapt.readthedocs.io/en/stable/> (03.2020)

<https://broadinstitute.github.io/picard/> (03.2020)

<https://www.putty.org/> (03.2020)

# Lisad

## Lisa 1. Koondinfo töös kasutatud keskaegsete indiviidide (n=30) kohta

Lisa 1.1. Uuritavate keskaegsete indiviidide arheoloogiline taustainfo ja genoomsete raamatukogude valmistamise määrdikud.  
Konts. – genoomse raamatukogu kontsentratsioon.

#	Indiviid	Leiukoht		Dateering	Morfoloogiline sugu	Proovi kaal (mg)	Raamatukogu indeks	Sekveneritud järjestusi	Konts. ng/ul
		Maacond	Kalmistu						
1	KAU003	Viljandimaa	Kaubi küla kalmistu	16. sajandi IV veerand	Naine	37	35	25537342	18.5
2	KAU004	Viljandimaa	Kaubi küla kalmistu	15. sajand	Naine	114	36	24839672	13.9
3	KAU005	Viljandimaa	Kaubi küla kalmistu	16. sajand	Naine	181	37	28082967	11.3
4	KOD001	Tartumaa	Kodavere küla kalmistu	13.-16. sajand	Naine	88	41	27508693	19.8
5	KOD002	Tartumaa	Kodavere küla kalmistu	13. saj. - 15. saj. keskpaik	Naine	43	38	25310087	20
6	KOD003	Tartumaa	Kodavere küla kalmistu	13. saj keskpaik - 15. saj keskpaik	Mees	90	39	26951584	15.9
7	KOD004	Tartumaa	Kodavere küla kalmistu	13.-16. sajand	Mees	30	40	24254297	18
8	MAL003	Tartumaa	Mäletjärve küla kalmistu	14. sajandi IV veerand	Naine	192	25	26215065	22.2
9	MAL004	Tartumaa	Mäletjärve küla kalmistu	14. sajandi IV veerand	?	72	26	23806086	18.5
10	MAL005	Tartumaa	Mäletjärve küla kalmistu	16. sajand	Naine	96	27	25545814	20.2
11	MAL006	Tartumaa	Mäletjärve küla kalmistu	16. sajand	?	115	28	26588738	12.9
12	MAL007	Tartumaa	Mäletjärve küla kalmistu	16. sajand	Naine	115	29	26549566	17.9
13	TAJ001	Tartumaa	Tartu Jaani kirik	13. saj - 16. sajand	Mees?	78	19	25159885	17.6
14	TAJ002	Tartumaa	Tartu Jaani kirik	13. saj - 16. sajand	Mees	95	20	26343632	10
15	TAJ003	Tartumaa	Tartu Jaani kirik	13. sajand	Naine	104	15	27045579	16.4
16	TAJ004	Tartumaa	Tartu Jaani kirik	13. sajand	Naine	127	16	26033838	7.02
17	TAJ005	Tartumaa	Tartu Jaani kirik	13. sajand	Mees	88	17	25453472	16.2
18	TAJ006	Tartumaa	Tartu Jaani kirik	13. sajand	Naine	81	18	25429547	19.1
19	TAT006	Tartumaa	Tartu Toomkirik	17. saj. lõpp - 18. saj. algus	?	80	46	28433309	4.5
20	TUD001	Ida-Vinumaa	Tudulinna küla kalmistu	15.-16. sajand	Naine	153	21	25076793	16.3
21	TUD002	Ida-Vinumaa	Tudulinna küla kalmistu	15.-16. sajand	Mees	115	22	23187482	14.6
22	VAL001	Saaremaa	Väljala kirikaed	13. sajand	Mees	49	24	25422755	2.9
23	VAL002	Saaremaa	Väljala kirikaed	13. sajand	Naine	74	23	21391461	25.2
24	VAS001	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool	Mees	84	42	23018571	15.6
25	VAS002	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool	Naine	90	43	25871027	13.9
26	VAS003	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool	Naine	167	44	27881101	16.9
27	VAS004	Võrumaa	Vastseliina alevikalmistu	16. sajandi II pool	Mees	202	45	23003247	14.4
28	VNK003	Tartumaa	Vana-Kuuste küla kalmistu	16. sajandi II pool	Mees	73	32	25799646	14.1
29	VNK004	Tartumaa	Vana-Kuuste küla kalmistu	14. sajand	Naine	250	33	24281039	13.2
30	VNK005	Tartumaa	Vana-Kuuste küla kalmistu	16. sajand	Naine	118	34	23117252	15

**Lisa 1.2. Uuritavate keskaegsete indiviidide genoomsete raamatukogude sekveerimistulemuste statistikud.** Endog. DNA – endogeense DNA sisaldus; sugu – genoomne sugu; C>T – 5' otstes esinevate C→T asenduste hulk; MT – mitokondriaalne DNA; krom. – kromosoom; hg – haplogrupp; SNP – ühe nukleotiidi polümorfism; NA – pole saadaval.

#	Indiviid	Endog. DNA	Raamatukogu efektiivsus	Keskmine genoomne katvus	Sugu	C>T	MT katvus	MT	Keskmine X krom. katvus	Meesoost indiviidi X krom. saastatus	Mees	Keskmine Y krom. katvus	Y hg	Võrdlusandmestiku SNPdest kaetud
1	KAU003	0.47%	0.23%	0.000862	NA	13%	0.518	NA	0.0008	NA	NA	0.00002	-	531
2	KAU004	1.54%	0.75%	0.00274	NA	9%	0.573	NA	0.0027	NA	NA	0.00007	-	1562
3	KAU005	15.85%	6.96%	0.028	XX	6%	2.400	NA	0.0275	-	U5a2a1	0.00036	-	15792
4	KOD001	11.80%	6.92%	0.0299	XY	5%	3.530	NA	0.0153	NA	H7a1a/F1a1c1	0.00882	I1-M253	18065
5	KOD002	1.00%	0.57%	0.00227	NA	9%	1.440	NA	0.0022	NA	U5a1h/U5a1a1	0.00003	-	1319
6	KOD003	0.06%	0.02%	0.0001	NA	11%	0.194	NA	0.0001	NA	NA	0.00049	NA	55
7	KOD004	0.89%	0.44%	0.00163	NA	11%	1.080	NA	0.0008	NA	NA	0.00074	-	966
8	MAL003	27.10%	15.38%	0.0662	XX	6%	11.800	0.30%	0.0644	-	H1b2	0.00074	-	37776
9	MAL004	3.49%	1.94%	0.00675	XX	8%	1.250	NA	0.0064	-	H2a1	0.00010	-	3826
10	MAL005	16.92%	9.29%	0.0409	XX	9%	9.390	0.35%	0.0391	-	U5a2a1	0.00045	-	23412
11	MAL006	2.30%	1.32%	0.0063	XY	8%	4.140	NA	0.0032	NA	H14a	0.00186	NA	3732
12	MAL007	4.76%	2.76%	0.0118	XX	7%	2.500	NA	0.0115	-	HV0g	0.00013	-	6827
13	TAI001	68.41%	42.36%	0.176	XY	5%	19.900	0.10%	0.0887	NA	HV0a	0.05000	R1a-M417	95223
14	TAI002	78.71%	49.30%	0.219	XY	8%	28.400	0.35%	0.1110	0.31%	K1a3a1	0.06320	R1b	116180
15	TAI003	81.06%	51.23%	0.235	XX	4%	20.500	2.52%	0.2290	-	H1c1	0.00224	-	120740
16	TAI004	84.85%	54.36%	0.245	XX	4%	25.700	0.10%	0.2400	-	T1a1b	0.00233	-	125430
17	TAI005	76.75%	44.84%	0.186	XY	8%	34.900	0.53%	0.0865	NA	U5a1b1d1	0.04830	R1a1c-M558	96659
18	TAI006	76.18%	46.39%	0.205	XX	5%	33.300	0.07%	0.1970	-	J1c3a1	0.00200	-	105926
19	TAT006	62.68%	36.34%	0.191	XX	3%	30.500	0.08%	0.1820	-	U4d1	0.00180	-	96936
20	TUD001	41.98%	25.53%	0.111	XX	10%	15.400	0.13%	0.1090	-	U5a2e	0.00117	-	60872
21	TUD002	6.74%	3.60%	0.0128	NA	7%	1.220	NA	0.0066	NA	NA	0.00354	NA	7714
22	VAL001	41.91%	24.02%	0.101	XY	5%	20.300	0.12%	0.0516	NA	H1	0.02930	N3	57249
23	VAL002	11.82%	6.52%	0.0218	XX	5%	4.540	NA	0.0191	-	H1c1	0.00024	-	12712
24	VAS001	0.82%	0.48%	0.00205	NA	12%	2.600	NA	0.0010	NA	U2e1f1	0.00060	NA	1300
25	VAS002	26.81%	15.04%	0.0595	XX	8%	8.490	2.99%	0.0569	-	U4d1b	0.00067	-	33884
26	VAS003	9.22%	5.07%	0.0249	XX	6%	5.640	0.37%	0.0241	-	J1c5d	0.00024	-	14932
27	VAS004	3.82%	1.76%	0.00633	XY	11%	2.050	NA	0.0032	NA	NA	0.00186	NA	3727
28	VNK003	50.80%	30.01%	0.116	XY	11%	6.800	0.29%	0.0590	NA	T2 + 16189	0.08430	N3a-L392	65055
29	VNK004	39.68%	20.02%	0.0742	XY	12%	7.860	1.55%	0.0378	NA	T2c1c2	0.02160	R1a-Z645	42406
30	VNK005	14.68%	7.84%	0.0305	XX	7%	3.750	NA	0.0296	-	H5a1a	0.00036	-	17780

## Lisa 2. Populatsioonigeneetilistes analüüsid kasutatud populatsioonid.

**Lisa 2.1. Analüüsid kasutatud tänapäevased ja muinasaja populatsioonid.** PCA – peakomponentanalüüs; W – Lääne-Euraasia; Eur – Euroopa; f3, D – f3 ja D statistikud, ADM – Admixture analüüs.

Tänapäevased populatsioonid	Viited	PCA			f3, D	ADM
		1-W	2-Eur	3-Tartu		
Abkhassians	(Yunusbayev et al., 2012)	+			+	+
Adygei	(Li et al., 2008)	+			+	+
Altaians	(Raghavan et al., 2014; Rasmussen et al., 2010)				+	+
Armenians	(Yunusbayev et al., 2012)	+			+	+
Azeris	(Yunusbayev et al., 2015)	+			+	+
Balkars	(Yunusbayev et al., 2012)	+			+	+
Balochi	(Li et al., 2008)					+
Bantus	(Li et al., 2008)					+
Bashkirs	(Yunusbayev et al., 2015)				+	+
Basques	(Li et al., 2008)	+			+	+
Belorussians	(Behar et al., 2010; Kushniarevich et al., 2015)	+	+	+	+	+
Brahui	(Li et al., 2008)					+
Bulgarians	(Yunusbayev et al., 2012)	+			+	+
Burusho	(Li et al., 2008)					+
Buryats	(Rasmussen et al., 2010)				+	+
Chechens	(Yunusbayev et al., 2012)	+			+	+
Chukchis	(Fedorova et al., 2013; Rasmussen et al., 2010)				+	+
Chuvash	(Behar et al., 2010)				+	+
Cypriots	(Behar et al., 2010)	+			+	+
Daur	(Li et al., 2008)				+	+

Dolgans	(Fedorova et al., 2013; Rasmussen et al., 2010)				+	+
Erza	(Yunusbayev et al., 2012)					+
Estonians	(Kushniarevich et al., 2015; Raghavan et al., 2014; Tambets et al., 2018)	+	+	+	+	+
Ethiopians	(Behar et al., 2010)					+
Evenkis	(Rasmussen et al., 2010)				+	+
Evens	(Fedorova et al., 2013; Raghavan et al., 2014)				+	+
Finns	(Tambets et al., 2018)	+	+	+	+	+
French	(Li et al., 2008)	+	+	+	+	+
Gagauzes	(Tambets et al., 2018; Yunusbayev et al., 2012)	+			+	+
Germans	(Tambets et al., 2018; Yunusbayev et al., 2012)	+	+	+	+	+
Han	(Li et al., 2008)				+	+
Hazara	(Li et al., 2008)					+
Hezhen/Nanaj	(Li et al., 2008)				+	+
Hungarians	(Behar et al., 2010)	+	+	+	+	+
Ingrian_Finns	(Tambets et al., 2018)	+	+	+	+	+
Iranians	(Behar et al., 2010)				+	+
Japanese	(Li et al., 2008)				+	+
Jordanians	(Behar et al., 2010)	+			+	+
Kalmyks	(Yunusbayev et al., 2015)				+	+
Karelians	(Tambets et al., 2018;	+	+	+	+	+

	Yunusbayev et al., 2012)					
Kazakhs	(Raghavan et al., 2014)				+	+
Kets	(Fedorova et al., 2013; Rasmussen et al., 2010)				+	+
Khakases	(Fedorova et al., 2013)				+	+
Khanty	(Tambets et al., 2018)				+	+
Komis	(Yunusbayev et al., 2015)				+	+
Koryaks	(Fedorova et al., 2013; Rasmussen et al., 2010)				+	+
Kumyks	(Yunusbayev et al., 2012)	+			+	+
Kurds	(Yunusbayev et al., 2012)				+	+
Kyrgyzians	(Raghavan et al., 2014)				+	+
Latvians	(Kushniarevich et al., 2015; Tambets et al., 2018)	+	+	+	+	+
Lebanese	(Behar et al., 2010)	+			+	+
Lezgins	(Behar et al., 2010)	+			+	+
Lithuanians	(Behar et al., 2010)	+	+	+	+	+
Makrani	(Li et al., 2008)					+
Malayan	(Behar et al., 2010)					+
Mandenkas	(Li et al., 2008)					+
Mansis	(Tambets et al., 2018)				+	+
Maris	(Raghavan et al., 2014; Tambets et al., 2018)				+	+
Moksha	(Yunusbayev et al., 2012)					+
Mongola	(Li et al., 2008)				+	+

Mongolians	(Rasmussen et al., 2010)				+	+
Mordovians	(Yunusbayev et al., 2012)	+	+	+	+	+
Moroccans	(Behar et al., 2010)					+
Mozabites	(Li et al., 2008)					+
Nenets	(Yunusbayev et al., 2015)				+	+
Nganasans	(Rasmussen et al., 2010)				+	+
Nivkhs	(Fedorova et al., 2013)				+	+
Nogais	(Yunusbayev et al., 2012)				+	+
North_Italians	(Li et al., 2008)	+			+	+
North_Kannadi	(Rasmussen et al., 2010)					+
North_Ossetians	(Yunusbayev et al., 2012)	+			+	+
northwest_Mingrelians	(Behar et al., 2010)	+			+	+
Orcadians	(Li et al., 2008)	+	+	+	+	+
Oroqens	(Li et al., 2008)				+	+
Palestinians	(Li et al., 2008)	+			+	+
Paniya	(Behar et al., 2010)					+
Pathan	(Li et al., 2008)					+
Poles	(Behar et al., 2013)	+	+	+	+	+
Portugese	(Tambets et al., 2018)	+			+	+
Romanians	(Behar et al., 2010)	+			+	+
Russians_Central	(Behar et al., 2010; Yunusbayev et al., 2015)	+	+	+	+	+
Russians_North	(Behar et al., 2010; Kushniarevich et al., 2015; Li et al., 2008; Yunusbayev et al., 2015)	+	+	+	+	+
Russians_South	(Yunusbayev et al., 2015)	+	+	+	+	+

Saami_Kola	(Tambets et al., 2018)				+	+
Saami_SWE	(Tambets et al., 2018)				+	+
Sakilli	(Behar et al., 2010)					+
Sardinians	(Li et al., 2008)	+			+	+
Saudis	(Behar et al., 2010)					+
Selkups	(Raghavan et al., 2014; Rasmussen et al., 2010)				+	+
Shors	(Fedorova et al., 2013)				+	+
Sindhi	(Li et al., 2008)					+
Slovaks	(Kushniarevich et al., 2015)	+			+	+
Spaniards	(Behar et al., 2010)	+			+	+
Swedes	(Behar et al., 2013; Tambets et al., 2018)	+	+	+	+	+
Syrians	(Behar et al., 2010)	+			+	+
Tajiks	(Yunusbayev et al., 2012)				+	+
Tatars	(Tambets et al., 2018; Yunusbayev et al., 2012)				+	+
Tu	(Li et al., 2008)				+	+
Turkmens	(Yunusbayev et al., 2012)				+	+
Turks	(Behar et al., 2010)	+			+	+
Tuscans	(Li et al., 2008)	+			+	+
Tuvinians	(Rasmussen et al., 2010; Yunusbayev et al., 2015)				+	+
Udmurts	(Yunusbayev et al., 2015)				+	+
Ukrainians	(Yunusbayev et al., 2012)	+	+	+	+	+
Uzbeks	(Behar et al., 2010;				+	+

	Raghavan et al., 2014)					
Uygurs	(Li et al., 2008)				+	+
Vepsas	(Tambets et al., 2018; Yunusbayev et al., 2012)	+	+	+	+	+
Welsh	(Tambets et al., 2018)	+	+	+	+	+
Xibo	(Li et al., 2008)				+	+
Yakuts	(Fedorova et al., 2013; Li et al., 2008)				+	+
Yemenese	(Behar et al., 2010)					+
Yorubas	(Li et al., 2008)				+	+
Yukaghirs	(Rasmussen et al., 2010)				+	+

Muinasaja populatsioonid	Viited	PCA			f3, D	ADM
		1	2 Eur	3 Tartu		
Afanasievo	(Lazaridis et al., 2016; Narasimhan et al., 2018)	+			+	+
Aktogai_MLBA	(Narasimhan et al., 2018)	+			+	+
Alan	(Damgaard et al., 2018)	+			+	+
Anatolia_ChL	(Lazaridis et al., 2016)	+			+	+
Anatolia_N	(Lazaridis et al., 2016)	+			+	+
Andronovo	(Damgaard et al., 2018; Lazaridis et al., 2016)	+			+	+
Armenia_ChL	(Lazaridis et al., 2016)	+			+	+
Armenia_EBA	(Lazaridis et al., 2016)	+			+	+
Armenia_MLBA	(Lazaridis et al., 2016)	+			+	+
Bell_Beaker	(Lazaridis et al., 2016)	+			+	+

Bichon	(Lazaridis et al., 2016)	+				+	+
Botai	(de Barros Damgaard et al., 2018)	+				+	+
Cardial_EN	(Lazaridis et al., 2016)	+				+	+
Central_LNBA	(Lazaridis et al., 2016)	+				+	+
Central_MN	(Lazaridis et al., 2016)	+				+	+
CentralSteppe_EMBA	(de Barros Damgaard et al., 2018)	+				+	+
CHG	(Lazaridis et al., 2016)	+				+	+
Dali_MLBA	(Narasimhan et al., 2018)	+				+	+
Dashti_Kozy_BA	(Narasimhan et al., 2018)	+				+	+
EHG	(Lazaridis et al., 2016; Mittnik et al., 2018)	+				+	+
EstBa	(Saag et al., 2019)	+	+			+	+
EstIA	(Saag et al., 2019)	+	+			+	+
EstMA	Avaldamata	+	+			+	+
EstMA_7	(Saag et al., 2019)	+	+			+	+
EstMA_T*	Avaldamata	+	+	+		+	+
Estonia_CCC	(Mittnik et al., 2018; Saag et al., 2017)	+	+			+	+
Estonia_CWC	(Mittnik et al., 2018; Saag et al., 2017)	+	+			+	+
Estonia_HG	(Mittnik et al., 2018; Saag et al., 2017)	+	+			+	+
Globular_Amphora	(Mathieson et al., 2018)	+				+	+
HallstattBylany	(Damgaard et al., 2018)	+				+	+
Hungarian_Med	(Damgaard et al., 2018)	+				+	+

HungarianScythian	(Damgaard et al., 2018)	+			+	+
Hungary_BA	(Lazaridis et al., 2016)	+			+	+
Hungary_EN	(Lazaridis et al., 2016)	+			+	+
Iberia_BA	(Lazaridis et al., 2016)	+			+	+
Iberia_ChI	(Lazaridis et al., 2016)	+			+	+
Iberia_EN	(Lazaridis et al., 2016)	+			+	+
Iberia_MN	(Lazaridis et al., 2016)	+			+	+
IngIA	(Saag et al., 2019)	+	+		+	+
Iran_ChL	(Lazaridis et al., 2016)	+			+	+
Iran_HotuIIIb	(Lazaridis et al., 2016)	+			+	+
Iran_LN	(Lazaridis et al., 2016)	+			+	+
Iran_N	(Lazaridis et al., 2016)	+			+	+
Iron_Gates_HG	(Mathieson et al., 2018)	+			+	+
Kairan_MLBA	(Narasimhan et al., 2018)	+			+	+
Kangju	(Damgaard et al., 2018)	+			+	+
Karagash_MLBA	(Narasimhan et al., 2018)	+			+	+
Kashkarchi_BA	(Narasimhan et al., 2018)	+			+	+
Kazakh_His	(Damgaard et al., 2018)	+			+	+
Kazakh_Mys_MLBA	(Narasimhan et al., 2018)	+			+	+
Kostenki	(Lazaridis et al., 2016)					+
Krasnoyarsk_MLBA	(Narasimhan et al., 2018)	+			+	+
Latvia_BA	(Mittnik et al., 2018)	+			+	+
Latvia_CCC_EHG	(Mathieson et al., 2018)	+			+	+

Latvia_CCC_WHG	(Mathieson et al., 2018)	+			+	+
Latvia_HG	(Mathieson et al., 2018)	+			+	+
LBK_EN	(Lazaridis et al., 2016)	+			+	+
LchashenMetsamor	(Damgaard et al., 2018)	+			+	+
Levant_BA	(Lazaridis et al., 2016)	+			+	+
Levant_N	(Lazaridis et al., 2016)	+			+	+
Lisakovskiy_MLBA_Alakul	(Narasimhan et al., 2018)	+			+	+
Lithuania_BA	(Mittnik et al., 2018)	+			+	+
Lithuania_CWC	(Mittnik et al., 2018)	+			+	+
Lithuania_HG	(Mittnik et al., 2018)	+			+	+
Maitan_MLBA_Alakul	(Narasimhan et al., 2018)	+			+	+
Mal'ta	(Lazaridis et al., 2016)					+
Natufian	(Lazaridis et al., 2016)	+			+	+
NLithuania	(Damgaard et al., 2018)	+			+	+
Northern_LNBA	(Lazaridis et al., 2016)	+			+	+
Okunevo_EMBA	(de Barros Damgaard et al., 2018)	+			+	+
Oy_Dzhaylau_MLBA	(Narasimhan et al., 2018)	+			+	+
Petrovka	(Narasimhan et al., 2018)	+			+	+
Poltavka	(Lazaridis et al., 2016; Narasimhan et al., 2018)	+			+	+
Poprad	(Damgaard et al., 2018)	+			+	+
Potapovka	(Lazaridis et al., 2016; Narasimhan et al., 2018)	+			+	+

Remedello	(Lazaridis et al., 2016)	+				+	+
Romania_HG	(Mathieson et al., 2018)	+				+	+
Russia_EBA	(Lazaridis et al., 2016)	+				+	+
SaltovoMayaki	(Damgaard et al., 2018)	+				+	+
Samara_Eneolithic	(Lazaridis et al., 2016)	+				+	+
Sarmatian	(Damgaard et al., 2018)	+				+	+
Sarmatian_SU	(Unterländer et al., 2017)	+				+	+
Scythian_East	(Unterländer et al., 2017)	+				+	+
Scythian_IA	(Lazaridis et al., 2016)	+				+	+
SHG	(Lazaridis et al., 2016; Mittnik et al., 2018)	+				+	+
SidelkinoEHG_ML	(de Barros Damgaard et al., 2018)	+				+	+
Sintashta	(Lazaridis et al., 2016; Narasimhan et al., 2018)	+				+	+
Srubnaya	(Lazaridis et al., 2016; Narasimhan et al., 2018)	+				+	+
Sweden_LNBA	(Mittnik et al., 2018)	+				+	+
Sweden_TRB	(Mittnik et al., 2018)	+				+	+
Zevakinskiy_LBA	(Narasimhan et al., 2018)	+				+	+
Tagar	(Damgaard et al., 2018)	+				+	+
Trypillia	(Mathieson et al., 2018)	+				+	+
Ukraine_Eneolithic	(Mathieson et al., 2018)	+				+	+
Ukraine_Mesolithic	(Mathieson et al., 2018)	+				+	+

Ukraine_Neolithic	(Mathieson et al., 2018)	+			+	+
Ust-Ishim	(Lazaridis et al., 2016)					+
West_Siberia_N	(Narasimhan et al., 2018)	+			+	+
WHG	(Lazaridis et al., 2016; Mathieson et al., 2018)	+			+	+
Wusun	(Damgaard et al., 2018)	+			+	+
Yamnaya	(de Barros Damgaard et al., 2018)	+			+	+
Yamnaya_Bulgaria	(Mathieson et al., 2018)	+			+	+
Yamnaya_Kalmykia	(Lazaridis et al., 2016)	+			+	+
Yamnaya_Samara	(Lazaridis et al., 2016)	+			+	+
Yamnaya_Ukraine	(Mathieson et al., 2018)	+			+	+

**Lisa 3 . Uuritavad keskaegsed proovid ja võrdlusproovid, mida erinevates analüüsides kasutati.**

**Lisa 3.1. Uurimistöös analüüsitud indiviidid ja erinevates populatsioonigeneetilistes analüüsides (vt. 2.2.11) kasutatud grupid.** Katvus – keskmine genoome katvus; SNP – kaetud võrdlusandmestiku ühenukleotiidsete polümorfismide arv; grupp 1 – uuritavate indiviidide jaotumine matmispaiga järgi – keskaegsed Eesti külakalmistute indiviidid (EstMAEM), Tartu Jaani kiriku (TJ) ja Tartu toomkiriku (TT) indiviidid; grupp 2 – Tartu indiviidide jaotumine lähtuvalt geneetilisest sarnasusest – geneetiliselt tänapäeva eestlastele (T\_Est) või saklastele (T\_Ger) sarnased Tartu indiviidid; PCA – peakomponentanalüüs, W – Lääne-Euraasia, Eur – Euroopa; ADM – Admixture analüüs. Iga analüüsi nime all on toodud proovide analüüsi kaasamise tingimused.

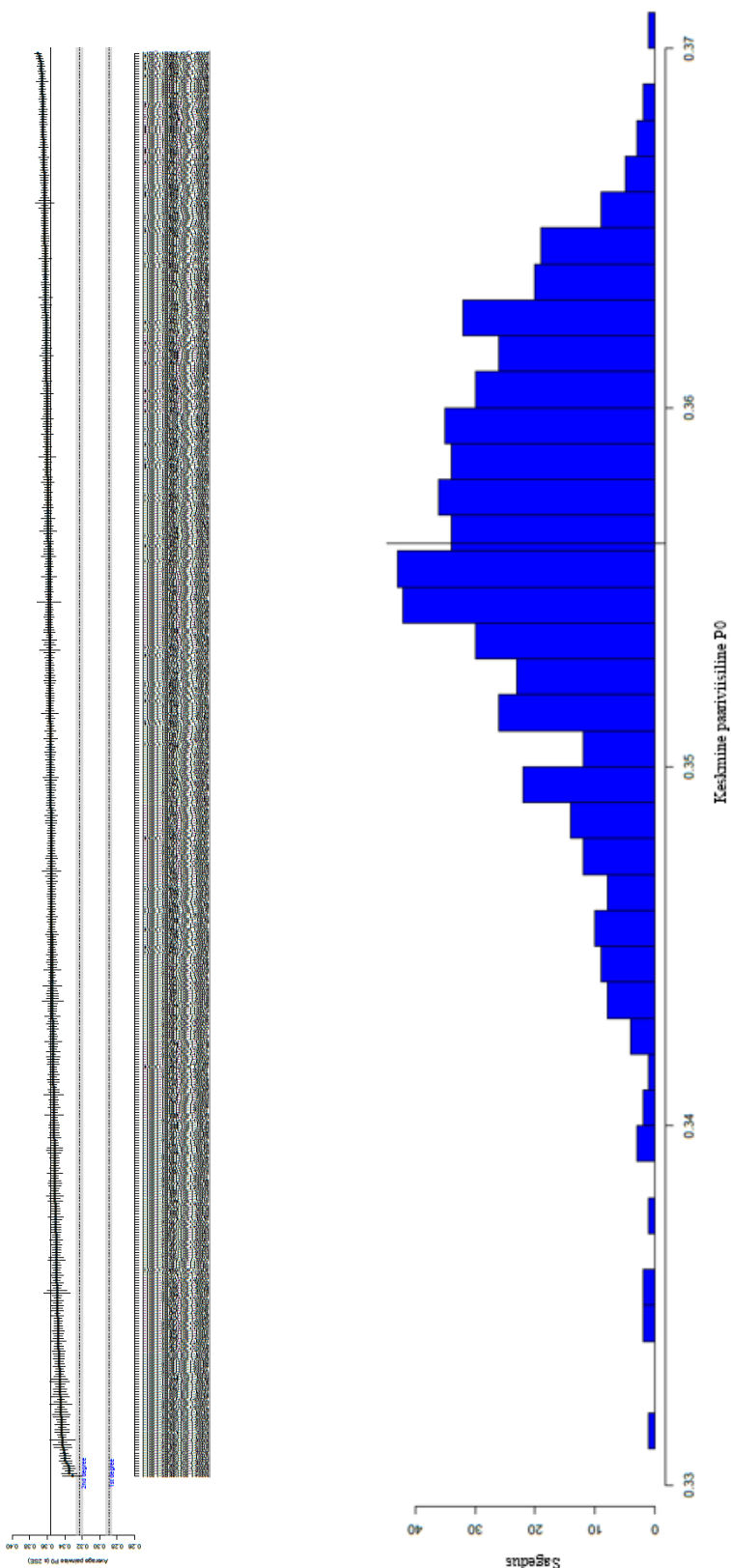
#	Indiviid	Katvus	SNP	Grupp 1	Grupp 2	PCA		READ Katvus > 0,03	f3, D		ADM SNP > 10 000	qpAdm SNP > 100 000	CP/NNLS SNP > 20 000
						SNP > 10 000 1-W/2-Eur/3-Tartu	SNP > 10 000 Grupp 1		SNP > 10 000 Grupp 2				
1	KAU003	0.0009	531										
2	KAU004	0.0027	1562										
3	KAU005	0.0280	15792	EstMAEM	EstMAEM	+	+		+	+			
4	KOD001	0.0299	18065	EstMAEM	EstMAEM	+	+		+	+			
5	KOD002	0.0023	1319										
6	KOD003	0.0001	55										
7	KOD004	0.0016	966										
8	MAL003	0.0662	37776	EstMAEM	EstMAEM	+	+	+	+	+			+
9	MAL004	0.0068	3826										
10	MAL005	0.0409	23412	EstMAEM	EstMAEM	+	+	+	+	+			+
11	MAL006	0.0063	3732										
12	MAL007	0.0118	6827										
13	TAU001	0.1760	95223	EstMAEM_TJ	EstMAEM_TJ	+	+	+	+	+			+
14	TAU002	0.2190	116180	EstMAEM_TJ	EstMAEM_T_Ger	+	+	+	+	+	+		+
15	TAU003	0.2350	120740	EstMAEM_TJ	EstMAEM_T_Est	+	+	+	+	+	+		+
16	TAU004	0.2450	125430	EstMAEM_TJ	EstMAEM_T_Est	+	+	+	+	+	+		+
17	TAU005	0.1860	96559	EstMAEM_TJ	EstMAEM_T_Est	+	+	+	+	+	+		+
18	TAU006	0.2050	105926	EstMAEM_TJ	EstMAEM_T_Ger	+	+	+	+	+	+		+
19	TAT006	0.1910	96936	EstMAEM_TT	EstMAEM_TT	+	+	+	+	+	+		+
20	TUD001	0.1110	60872	EstMAEM	EstMAEM	+	+	+	+	+	+		+
21	TUD002	0.0128	7714										
22	VAL001	0.1010	57249	EstMAEM	EstMAEM	+	+	+	+	+	+		+
23	VAL002	0.0218	12712	EstMAEM	EstMAEM	+	+	+	+	+	+		+
24	VAS001	0.0021	1300										
25	VAS002	0.0595	33884	EstMAEM	EstMAEM	+	+	+	+	+	+		+
26	VAS003	0.0249	14332	EstMAEM	EstMAEM	+	+	+	+	+	+		+
27	VAS004	0.0063	3727										
28	VNK003	0.1160	65055	EstMAEM	EstMAEM	+	+	+	+	+	+		+
29	VNK004	0.0742	42406	EstMAEM	EstMAEM	+	+	+	+	+	+		+
30	VNK005	0.0305	17780	EstMAEM	EstMAEM	+	+	+	+	+	+		+

**Lisa 3.2. Võrdlusena kasutatud keskaegsed indiviidid ja erinevates populatsioonigeneetilistes analüüsid (vt 2.2.11) kasutatud grupid.** Katvus – keskmine genoomne katvus; SNP – kaetud võrdlusandmestiku ühenukleotiidsete polümorfsimide arv; grupp 1 – uuritavate indiviidide jaotumine matrispaiga järgi – keskaegsed Eesti külakalmistute indiviidid (EstMAEM), Tartu toomkiriku (TT) ja Tartu Maarja kiriku (TM) indiviidid; grupp 2 – Tartu indiviidide jaotumine lähtuvalt geneetilise sarnasusest – geneetiliselt tänapäeva eestlastele (T\_Est) või sakslastele (T\_Ger) sarnased Tartu indiviidid; PCA – peakomponentanalüüs, W – Lääne-Euraasia, Eur – Euroopa; ADM – Admixture analüüs. Iga analüüsi nime all on toodud proovide analüüsi kaasamise tingimused.

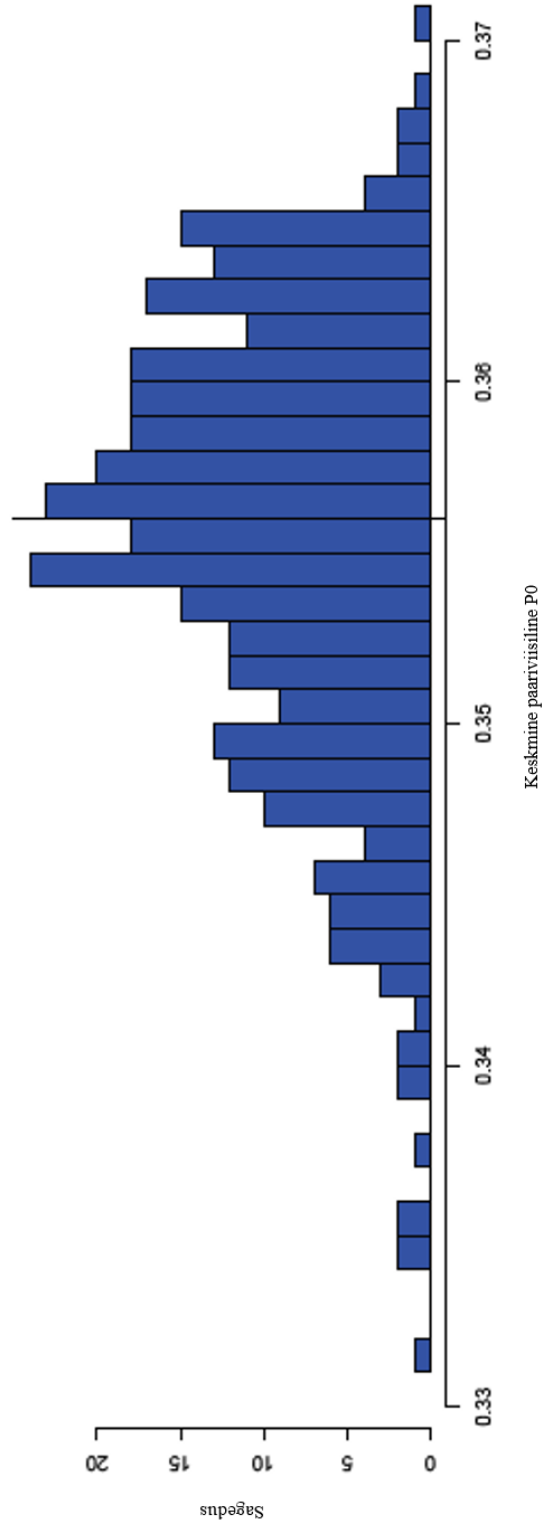
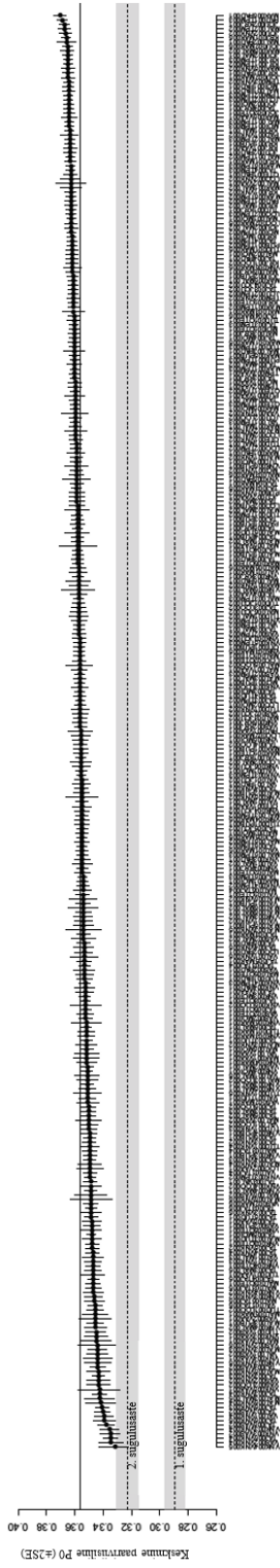
#	Indiviid	Katvus	SNP	Grupp 1	Grupp 2	PCA		READ	f3, D		ADM		qpAdm		CP/NNLS	
						SNP > 10 000	SNP > 10 000		Katvus	SNP > 10 000	SNP > 10 000	SNP > 10 000	SNP > 10 000	SNP > 10 000	SNP > 10 000	SNP > 20 000
						1-W	2-Eur	3-Tartu	Grupp 1	Grupp 2						
1	JOU002	0.316	150243	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
2	KRJ001	0.734	276849	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
3	KUK001	0.0696	39335	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
4	LIN001	0.0184	9937													
5	LOO002	0.0834	46940	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
6	MAK001	0.0884	49596	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
7	MAK005	0.283	152923	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
8	MAL001	0.332	162153	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
9	MAL002	0.0676	39114	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
10	MKS001	0.00027	144													
11	MOJ001	0.542	216859	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
12	OTE001	0.206	106825	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
13	PAD001	0.102	56566	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
14	PUB001	0.172	91993	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
15	TAB001	0.00389	2237													
16	TAT001	0.177	95797	EstMAEM_TT	EstMAEM_T_Ger	+	+	+	+	+	+	+	+	+	+	+
17	TAT002	0.413	187891	EstMAEM_TT	EstMAEM_T_Ger	+	+	+	+	+	+	+	+	+	+	+
18	TAT003	0.175	91862	EstMAEM_TT	EstMAEM_T_Ger	+	+	+	+	+	+	+	+	+	+	+
19	TAT005	0.222	362061	EstMAEM_TT	EstMAEM_T_Ger	+	+	+	+	+	+	+	+	+	+	+
20	TPM001	0.0219	30139	EstMAEM_TM	EstMAEM_T_Est	+	+	+	+	+	+	+	+	+	+	+
21	TPM002	0.522	224056	EstMAEM_TM	EstMAEM_T_Est	+	+	+	+	+	+	+	+	+	+	+
22	TPM003	11.4	517687	EstMAEM_TM	EstMAEM_T_Est	+	+	+	+	+	+	+	+	+	+	+
23	VAA001	0.0413	23780	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+
24	VIS001	0.00891	5260													
25	VLG001	0.0044	2643													
26	VNK002	0.0525	46165	EstMAEM	EstMAEM	+	+	+	+	+	+	+	+	+	+	+



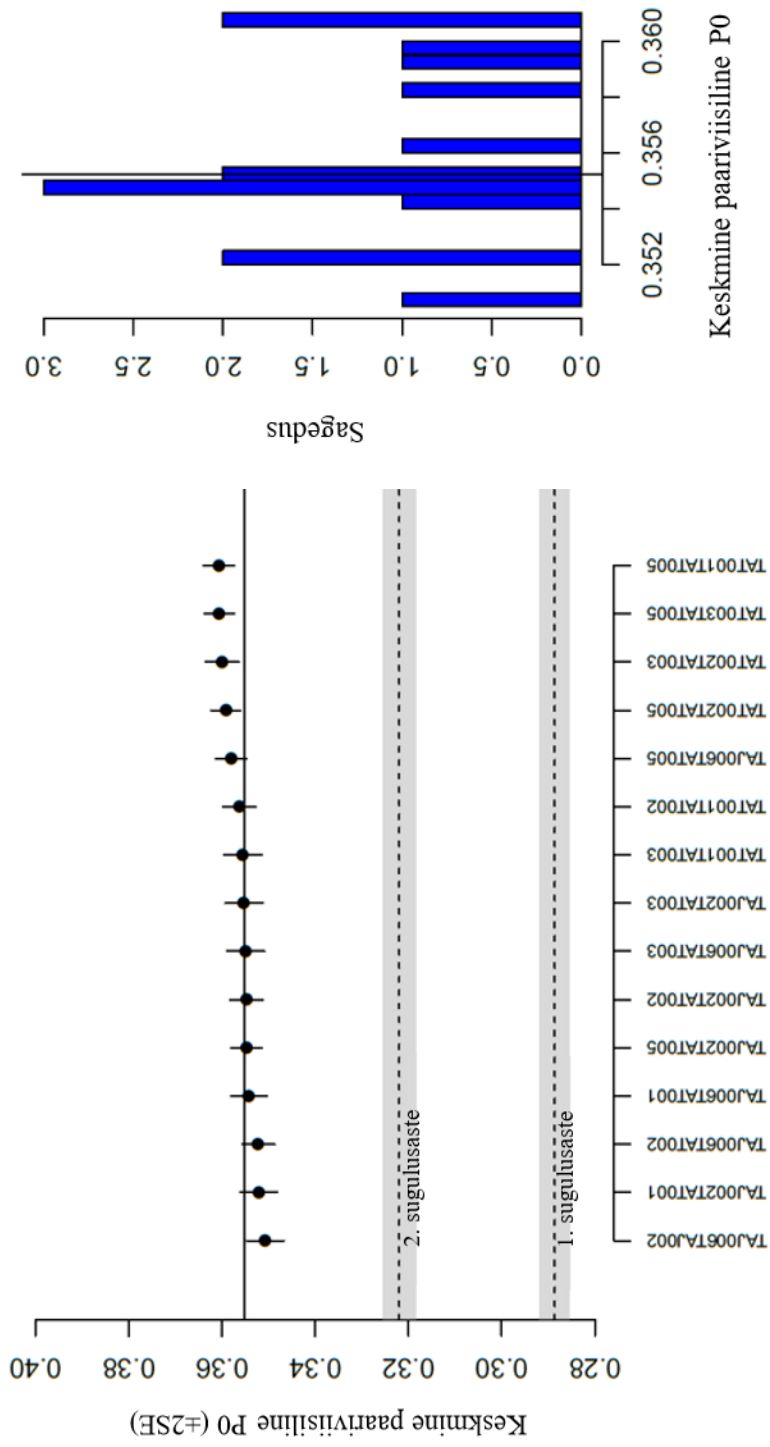
## Lisa 5. Sugulusanalüüsi READ tulemused.



**Lisa 5.1. Keskaegsete Eesti ala indiviidide sugulusanalüüs READ. P0** – iga indiviidide paari kohta arvutatud mittekattuvate alleelide osakaal igas 1 megaaluspaarises genoomeses aknas. Pidev joon tähistab uuritavate populatsioonide mediaanväärtust, katkendjooned tähistavad 1. ja 2. sugulusastme lävendeid ning neid ümbritsev hall ala tähistab 95% usaldusvahemikku, SE – standardviga ( $\pm 2$ ). Tulpdiagrammi tulbad tähistavad mittenormaliseeritud P0 keskmiste väärtuste sagedusi.

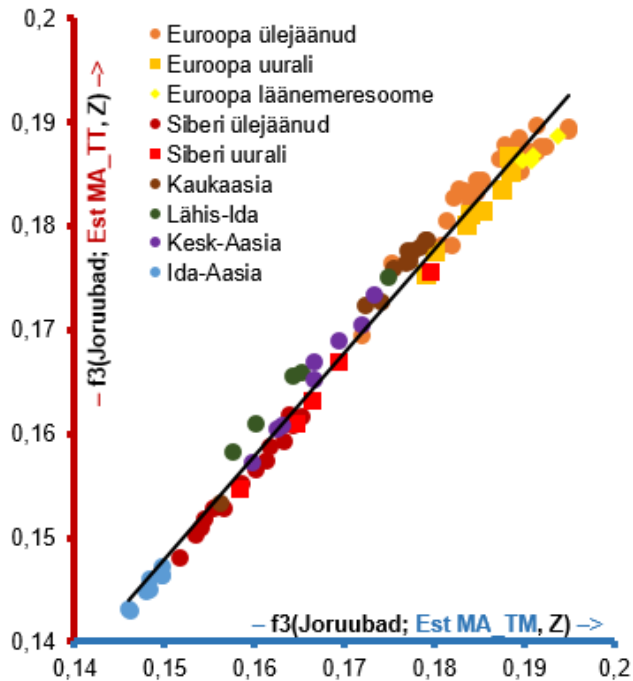


Lisa 5.2. Sugulusanalüüs READ uuritavate keskaegsete nõ. eestlastest indiviididega. Joonise selgitust vt lisa 5.1.

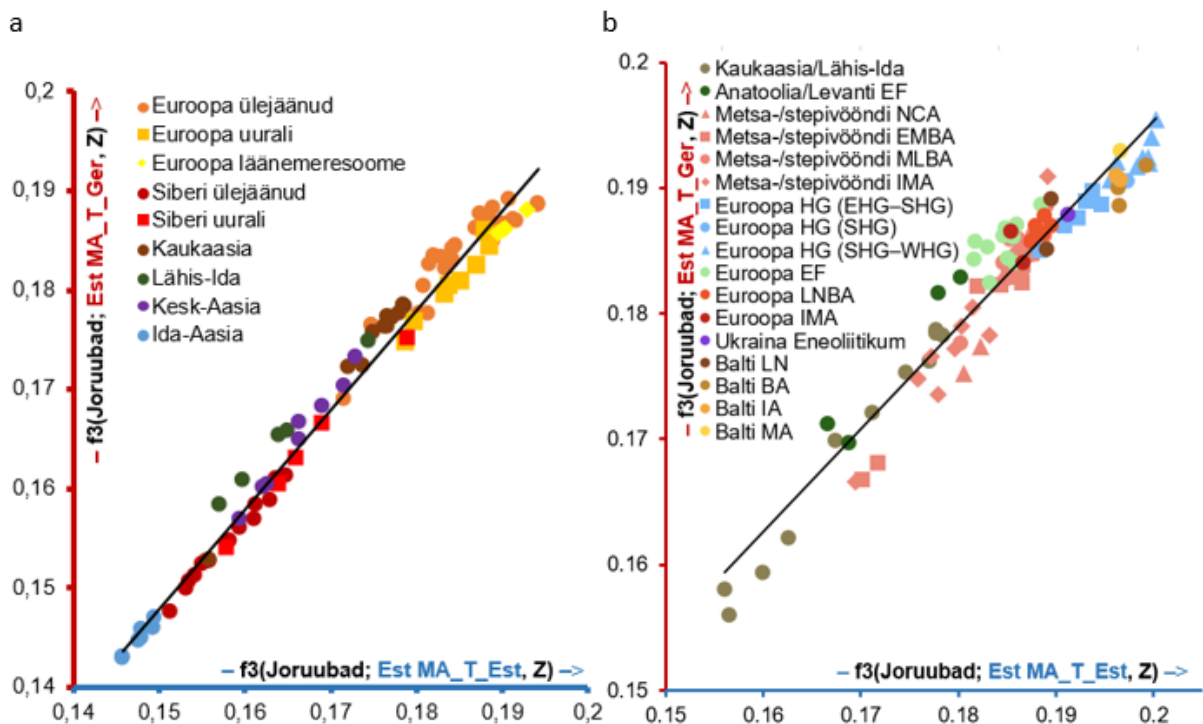


Lisa 5.3. Sugulusanalüüs uuritavate keskaegsete nõ. sakslastest indiviididega. Joonise selgitust vt lisa 5.1.

**Lisa 6 . f3 statistikute tulemused Tartu kirikute indiviidide kohta.**

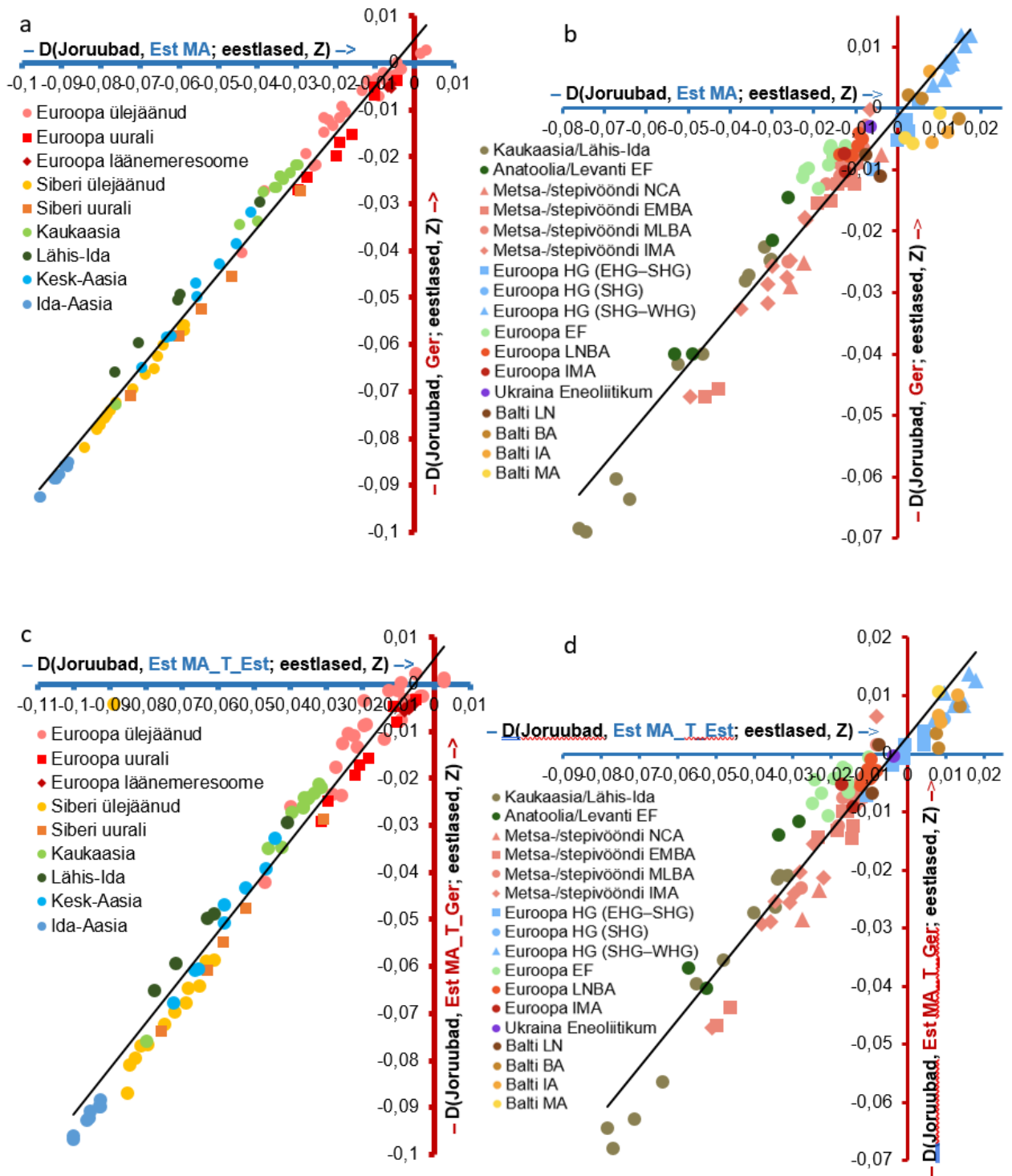


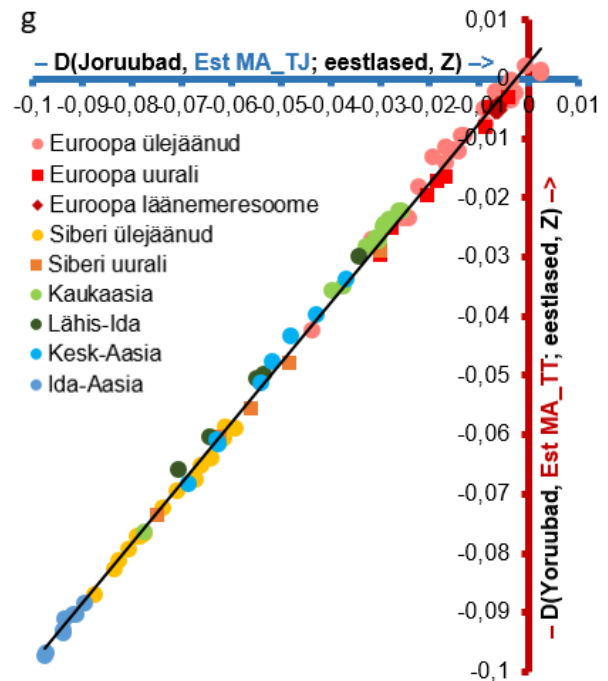
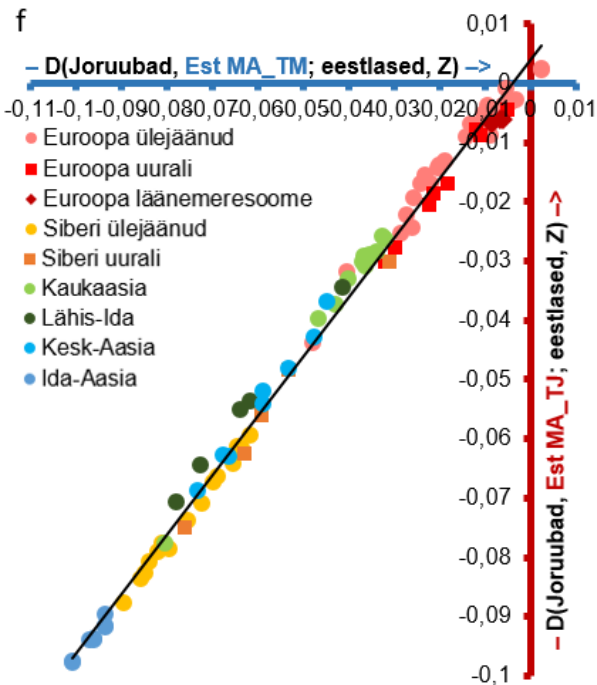
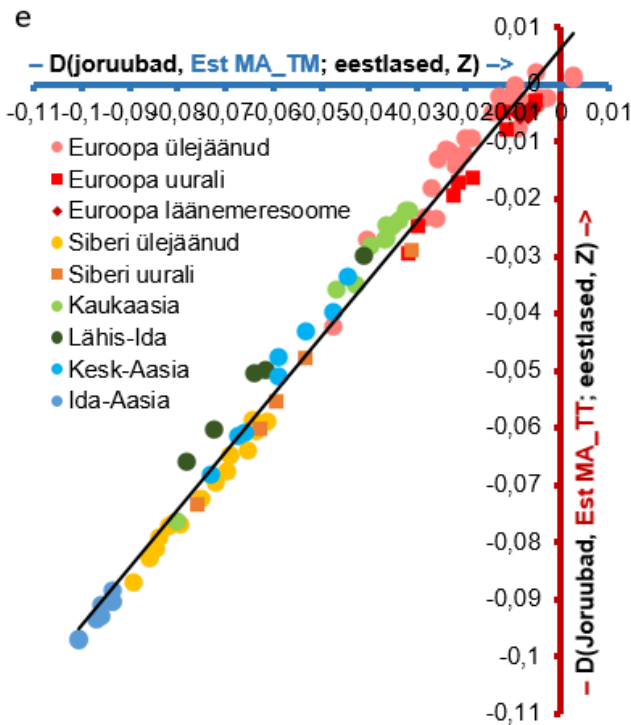
**Lisa 6.1. f3 statistikute tulemused. Tartu Püha Maarja kiriku (Est MA\_TM) ja Tartu toomkiriku (Est MA\_TT) indiviidide võrdlus tänapäeva Euraasia populatsioonidega.** Joonisel on vastandatud kahe testi tulemused, kus nii Tartu Püha Maarja kiriku indiviide (Est MA\_TM) kui toomkiriku indiviide (Est MA\_TT) on esmalt võrreldud tänapäeva Euraasia populatsioonidega ning seejärel teise kihina võrreldud saadud väärtusi kahe uuritava populatsiooni (Est MA\_TM ja Est MA\_TT) vahel. Mida suuremad on punkti koordinaatide väärtused, seda sarnasem on vastav populatsioon uuritavatele populatsioonidele.



**Lisa 6.2. f3 statistikute tulemused. Tartu Jaani kiriku (Est MA\_TJ) ja a) Tartu toomkiriku (Est MA\_TT) indiviidide ning b) Püha Maarja kiriku (Est MA\_TM) indiviidide võrdlus tänapäeva Euraasia populatsioonidega.**

## Lisa 7. D statistikute tulemused





Lisa 7.1. D statistikute tulemused. Keskaegsete eestlaste (EstMA) ja tänapäeva sakslaste (Ger) geneetiline sarnasus a) tänapäeva Euraasia populatsioonidega ja b) muinasaja populatsioonidega; Tartu kirikute eestlastest (EstMA\_T\_Est) ja sakslastest (EstMA\_T\_Ger) indiviidide geneetiline sarnasus c) tänapäeva Euraasia populatsioonidega ja d) muinasaja populatsioonidega; e) Tartu Püha Maarja kiriku (EstMA\_TM) ja Tartu toomkiriku (EstMA\_TT) indiviidide geneetiline sarnasus tänapäeva Euraasia populatsioonidega; Tartu

**Jaani kiriku (EstMA\_TJ) ja f) Tartu toomkiriku indiviidide ning g) Püha Maarja kiriku indiviidide geneetiline sarnasus tänapäeva Euraasia populatsioonidega.** Välisgrupiks on kasutatud joruubasid ning 0-punktiks on tänapäeva eestlased. D statistiku väärtus on negatiivne, kui omavahel on sarnasemad uuritav populatsioon ja tänapäeva eestlased, kui kumbi neist populatsiooniga Z. Väärtus on 0, kui sarnasemad on eestlased ja Z, kui kumbki neist uuritava populatsiooniga. Väärtus on positiivne, kui sarnasemad on uuritav populatsioon ja Z, kui kumbki neist eestlastega.





## **Lihtlitsents**

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Kadri Irdt,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Eesti keskaegse populatsiooni struktuur vana DNA vaatevinklist“, mille juhendajad on PhD Lehti saag ja PhD Kristiina Tambets, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, alates 11.06.2023 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.

Kadri Irdt

11.06.2020