

TARTU ÜLIKOOL
LOODUS-JATEHNOLOOGIATEADUSKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
EVOLUTSIOONILISE BIOLOOGIA ÕPPETOOL

Anu Solnik

Wales´i populatsiooni geneetiline analüüs emaliinide ja autosoomsete markerite varieeruvuse alusel Lääne-Euraasia kontekstis

Magistritöö

Juhendajad
Ph.D. Ene Metspalu
Ph.D. Kristiina Tambets

Tartu 2013

SISUKORD

KASUTATUD LÜHENDID	3
MÕISTED.....	4
SISSEJUHATUS	5
1. KIRJANDUSE ÜLEVAADE	7
1.1 INIMESE MITOKONDER.....	7
1.1.1 Inimese mtDNA iseloomustus	7
1.1.2 Inimese mtDNA populatsioonigeneetilise markerina	8
1.1.3 Inimese mtDNA nomenklatuur	10
1.2 ÜLEGENOOMSED MARKERID POPULATSIOONIGENEETILISTES UURIMISTÖÖDES	12
1.3 LÄÄNE-EURAASIA ASUSTAMISEST KAASAEGSE INIMESE POOLT.....	14
1.3.1 Lääne-Euraasia asustamine arheoloogilisest vaatenurgast.....	14
1.3.2 Lääne-Euraasia asustamine geneetilisest vaatenurgast	17
1.3.2.1 Euroopa emaliinide geneetilise varieeruvuse kujunemine	19
1.4 BRITI SAARED JA WALES.....	25
1.4.1 Briti saarte ja Wales`i geograafiast, ajaloost ja arheoloogiast.....	25
1.4.2 Briti saarte ja Wales`i geneetilisest mitmekesisusest	27
2. EKSPERIMENTAALNE OSA	31
2.1 TÖÖ EESMÄRGID	31
2.2 MATERJAL JA METOODIKA	31
2.2.1 Katsematerjal.....	31
2.2.2 MtDNA analüüs	31
2.2.2.1 DNA amplifitseerimine	32
2.2.2.2 RFLP analüüs	34
2.2.2.3 Sekveneerimine	34
2.3 ANDMETÖÖTLUSMEETODID JA- PROGRAMMID.....	36
2.3.1 MtDNA andmete analüüs.....	36
2.3.2 Autosoomsete markerite analüüs.....	37
3. TULEMUSED JA ARUTELU	39
3.1 WALESI EMALIINIDE ANALÜÜS	39
3.1.1 MtDNA üldine varieeruvus Wales`i populatsioonis	39
3.1.2 Wales`i populatsiooni emaliinid Lääne-Euraasia kontekstis.....	47
3.1.3 MtDNA täisjärjestuste analüüs.....	48
3.2 AUTOSOOMSE DNA POLÜMORFISMIDE GENOTÜPISEERIMISANDMETE ANALÜÜS.....	52
3.3 PÕHITULEMUSED JA JÄRELDUSED	55
KOKKUVÕTE	56
SUMMARY	58
TÄNUAVALDUSED.....	60
KASUTATUD KIRJANDUS	61
LISAD.....	69
LIHTLITSENTS	82

KASUTATUD LÜHENDID

AMI	anatomiliselt moodne inimene
ap	aluspaar
at	aastat tagasi
dNTP	desoksüribonukleotiid
ddNTP	didesoksüribonukleotiid
EDTA-Na₂	naatriumetüleendiamiintetraatsetaat
e.m.a	enne meie aega
g	raskuskiirendus
hg	haplogrupp
HVS	hüpervarieeruv ala (<i>hypervariable segment</i>)
LD	aheldustasakaalutus (<i>linkage disequilibrium</i>)
LGM	viimase jääaja maksimum (<i>Last Glacial Maximum</i>)
m.a.j	meie aja järgi
mtDNA	mitokondriaalne DNA
np	nukleotiidipositsioon
NADH	redutseerunud nikotiinamiidadeniindinukleotiid
ND	NADH dehüdrogenaas
PCA	põhikomponentanalüüs (<i>Principal Component Analysis</i>)
rCRS	resekveneeritud Cambridge'i referentsjärjestus (<i>revised Cambridge Reference Sequence</i>)
RFLP	restriktsioonifragmentide pikkuse polümorfism (<i>Restriction Fragment Length Polymorphism</i>)
RSRS	inimese rekonstrueeritud referentsjärjestus (<i>Reconstructed Sapiens Reference Sequence</i>)
rpm	pööret minutis (<i>rotations per minute</i>)
SNP	ühenukleotiidne polümorfism (<i>single nucleotide polymorphism</i>)
U	ühik (<i>unit</i>)

MÕISTED

Asutajaefekt (e. rajajaefekt) - geenitriivi erijuht, kus ainult üksikud asutajapopulatsioonist eraldunud isendid panevad aluse uuele populatsioonile, mille tulemusel uue populatsiooni geneetiline struktuur erineb lähtepopulatsiooni omast. Selle mõju populatsioonile on samasugune, kui „pudelikaela“ efektil (vt. altpoolt).

Geneetiline triiv – alleelisageduste juhuslikud muutused populatsiooni järjestikustes põlvkondades.

Haplogrupp (e. monofüleetiline rühm haplotüüpe) – kõik ühise eellasvariandi järeltulijad, s.h ka eellasvariant ise.

Haplotüüp – DNA järjestusvariandid või alleelsed kombinatsioonid, mis enamasti grupeeruvad haplogruppideks. Käesolevas töös käsitletakse mõistet konkreetse mtDNA järjestusvariandi puhul.

Koalestsentsiaeg – aeg, mis kulub vaadeldava haplogrupi kõikide olemasolevate haplotüüpide keskmise mutatsioonilise kauguse saavutamiseks nende kõige viimasest ühisest eellasvariandist (*most recent common ancestor*).

Pudelikaelaefekt – geneetilise triivi erijuht, kus populatsiooni drastiline arvukuse vähenemine põhjustab alleelisageduste juhusliku ja järsu muutumise populatsioonis (vt. ka asutajaefekt).

Parafüleetiline rühm – rühm organisme (haplotüüpe), mille hulka kuulub nende kõigi viimane ühine eellane. Erinevalt monofüleetilisest rühmast ei kuulu siia kõik viimase ühise eellase järglased.

Parsimoonia- ehk säästuprintsiip – printsiip, mille puhul eelistatakse mitme võimaliku hüpoteesi hulgast lihtsaimat, vähima muutuste arvuga seletatavat hüpoteesi.

Sõsarrühmad – ühe evolutsioonipuu haru divergeerumisel tekkinud kaks uut monofüleetilist rühma.

Tähekujuline puu – fülogeneesipuu või selle haru, mille moodustavad haplotüübid, millest enamik, ideaaljuhul kõik, koalestseeruvad ühte ja samasse haplotüüpi.

SISSEJUHATUS

Molekulaarse evolutsioonibioloogia üheks suunaks on kaasaja inimkonna geneetilise varieeruvuse ja selle kujunemise uurimine. Aastakümneid on selleks kasutatud isaliinis päranduvat Y kromosoomi ja emaliinis päranduvat mitokondriaalset DNAd (mtDNA). Erinevalt rakutuuma DNAST, mis päritakse mõlemalt vanemalt ning mis rekombinatsiooni käigus igas põlvkonnas ümber korraldatakse, kanduvad mtDNA ja Y-kromosoom järglastele edasi uniparentaalselt ja ei rekombineeru – nad muutuvad ajas ja põlvkondade vahel ainult sedavõrd, kui palju neisse aja jooksul mutatsioone koguneb. Nende erinevate variantide sageduse ja levikumustrite uurimine populatsioonides annab geneetikutele - sarnaselt genealoogidega, kes koostavad perekonna ajaloo alusel sugupuid - võimaluse ema- ja isaliinide fülogeneetiliste puude ja võrgustike konstrueerimiseks. Nende alusel on võimalik kindlaks määrata erinevates populatsioonides leiduvate liinide asukoht globaalsel fülogeneetilisel puul, mõista selle hargnemiste suundi ja koalestsentsiteooriale toetudes hinnata ka hargnemiste ajalist sügavust, et heita selle kaudu valgust kaasaegse inimese tänasele geneetilisele mitmekesisusele viinud demograafilisele ajaloole.

Viimastel aastatel on eelpoolnimetatud markeritele lisandunud ka genoomide täisjärjestuste andmed. Kui täisgenoomide kasutamine on populatsioonigeneetilistes uuringutes veel algfaasis, siis üle genoomi jaotunud sadade tuhandete ühenukleotiidsete polümorfismide (SNP) genotüüpiseerimine on tänaseks saanud juba rutiiniks.

Lähiminevikus on jõudsalt arenenud ka vana DNA eraldamise ja analüüsi meetodid. Vana DNA analüüs annab meile võimaluse kontrollida kaasaegsete geeniproovide andmete põhjal püstitatud hüpoteese varieeruvuse tekkimise ja kujunemise kohta ning seeläbi paremini mõista inimeste demograafilist ajalugu. Terviklikuma pildi loomiseks tuleb lisaks geneetilistele andmetele kasutada ka arheoloogilisi, antropoloogilisi ning lingvistilisi andmeid, millele tuginedes on omakorda võimalik teha järeldusi erinevate maailma piirkondade asustamise kohta.

Antud töö kirjanduse ülevaade käsitleb teemasid nagu Lääne-Euraasia asustamine ning siinsete rahvaste kultuuriline ja geneetiline mitmekesisus. Kitsamalt keskendutakse Wales'i piirkondlikule ajaloolisele ja geneetilisele iseloomustamisele naaberpopulatsioonide kontekstis. Lisaks antakse ülevaade inimese mitokondri genoomist, selle ja üle genoomi jaotunud markerite varieeruvusest populatsioonigeneetilistes uuringutes.

Käesoleva töö eksperimentaalne osa annab lühikese ülevaate töös kasutatud meetoditest ja andmeanalüüsist: DNA resekveneerimisest, restriksioonanalüüsist, DNA järjestuste võrdlevast analüüsist, mediaanvõrgustike koostamisest, populatsioonide struktureerituse analüüsist jne.

Kirjanduses seni avaldatud Wales'i populatsiooni emaliinide andmed on mtDNA täisjärjestustest saadava maksimaalse fülogeneetilise resolutsiooniga võrreldes oluliselt madalama genotüüpiseerimise tasemega. Käesolev magistritöö aitab neid andmeid täiendada ja uurib Wales'i populatsiooni geneetilist mitmekesisust ja struktuuri, kasutades selle kirjeldamiseks mtDNA ja üle kogu genoomi jaotunud SNP-de genotüüpiseerimisandmestiku analüüsi. Saadud andmete paigutamine geograafiliselt lähedaste populatsioonide ning laiemalt Lääne-Euraasia konteksti annab võimaluse paremini mõista uuritava populatsiooni kujunemislugu ning kohta tänapäeva Lääne-Euroopa geneetilisel maastikul.

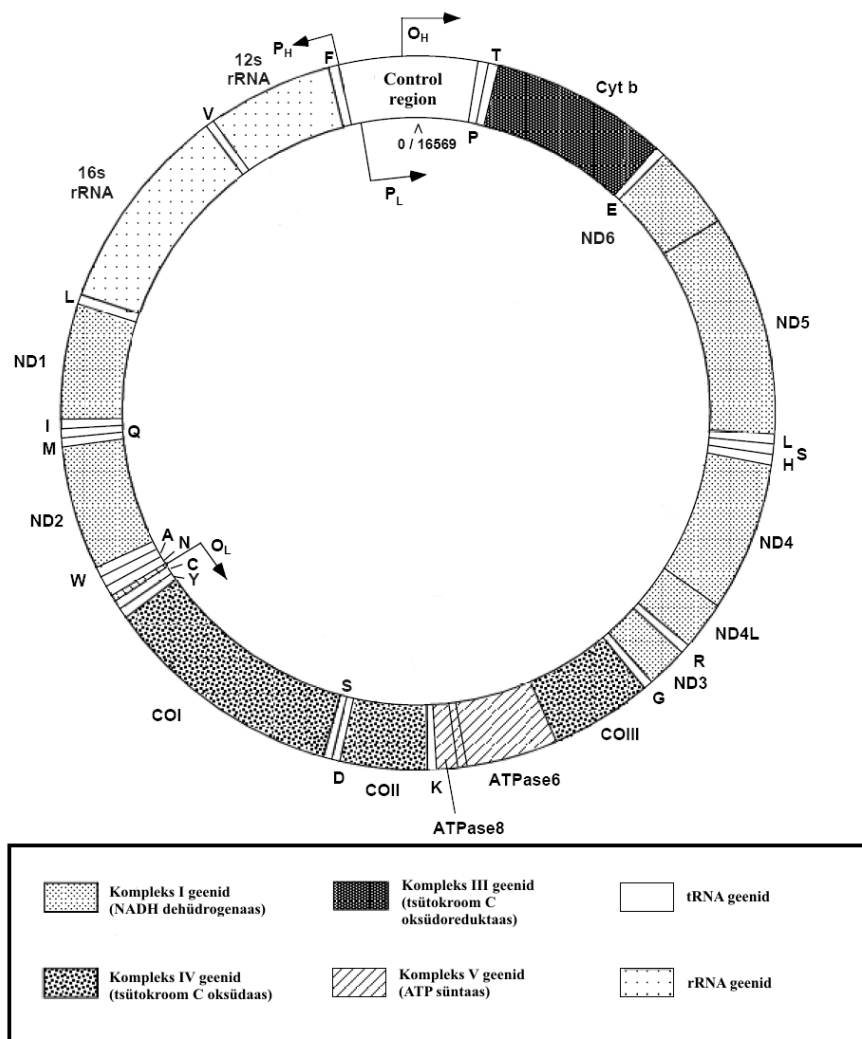
1. KIRJANDUSE ÜLEVAADE

1.1 Inimese mitokonder

1.1.1 Inimese mtDNA iseloomustus

Mitokondrid on kahekordse membraaniga ümbritsetud endosümbiootilise päritoluga eukarüootsete rakkude organellid, millede eelasteks peetakse α -proteobakterit (Margulis, 1975; Wallace et al., 1999). Nad on väga plastilised st. nende kuju võib kiiresti muutuda. Mitokondrid vastutavad rakule energia tootmise eest ning seetõttu viidatakse neile tihti kui raku “tööjaamadele”. Oksüdatiivse fosforüleerimise protsessi käigus toodavad nad adenosiin-trifosfaati (ATP), mis on peamine raku energia ülekandemolekul. Uued mitokondrid tekivad jagunemise teel, kusjuures nende jagunemine ja kasv ei ole seotud rakutsükliga.

Mitokondritel on oma genoom. Ühes mitokondris on leitud kuni 15 mtDNA koopiat, ühes rakus leidub neid ligikaudu 100 – 1000 (Cavelier et al., 2000; Ingman & Gyllensten, 2001) Evolutsiooni jooksul on enamus mitokondriaalseid geene liikunud tuuma (Wallace et al., 1999). Seetõttu on mitokondritele jäänud suhteliselt väike genoom, mille suurus liigiti varieerub. Imetajate mtDNA on tsirkulaarne kaheahelaline molekul, mille pikkus erinevatel liikidel on vahemikus 15 000 – 17 000 aluspaari (ap). Inimese mtDNA sekveneeriti esmakordselt täies mahus 1981. aastal ja see koosneb üldjuhul 16569 ap-st (Anderson et al., 1981; Andrews et al., 1999). Inimese mtDNA genoom on väga kompaktne. MtDNA on 93% ulatuses kodeeriv ning mitokondri geenid ei sisalda introneid (Anderson et al., 1981). Transkriptsioon toimub mõlemas suunas: raskelt ahelalt (H - *heavy*), mis sisaldab rohkem guaniini ning kergelt ahelalt (L - *light*), mis sisaldab rohkem tsütosiini. Enamus produktidest transkribeeritakse H-ahelalt (Wallace et al., 1999). Mitokondri genoom sisaldab 13 valku kodeerivat geeni, mis põhiliselt on seotud mitokondri sisemembraanil toimuva elektrontranspordi ja oksüdatiivse fosforüleerimise rajaga ehk siis rakusisese energia tootmise süsteemiga ja 22 tRNA-d, 2 rRNA-d kodeerivat geeni. Lisaks kodeerivale regioonile sisaldab mtDNA ka 1122 ap suurust kontrollregiooni (Anderson et al., 1981; Andrews et al., 1999), mis on mittekodeeriv ja moodustab terveist mitokondriaalsest genoomist 7% (Ingman & Gyllensten, 2001). Selles piirkonnas asuvad mtDNA raske ahela replikatsiooni alguspunkt ning promootorid mõlema ahela transkriptsiooniks (Wallace et al., 1999) (vt. joonis 1). Kontrollregioonis paiknevad mtDNA kolm kõrge mutatsioonisagedusega regiooni, mida nimetatakse hüpervarieeruvateks segmentideks (HVS-I, HVS-II ja HVS-III; vastavalt nps: 16024-16365, 73- 340 ja 438-574) (Greenberg et al., 1983; Vigilant et al., 1991; Lutz et al., 1998;)



Joonis 1. Inimese mtDNA (~16,6 ap) geenikaart, kus on ära märgitud oksüdatiivse fosforüleerimise kompleksid ja nendega seotud geenide asukohad: kompleks I (ND1, 2, 3, 4, 4L, 5, 6); kompleks III (Cytb); kompleks IV (COI, COII, COIII). Kaardil on näidatud 12S rRNA, 16S rRNA ja 22 tRNA geenide paiknemine ning kontrollregioon. MtDNA kergelt ja raskelt ahelalt transkribeeritavad geenid on joonisel märgitud vastavalt sisemisel või välimisel ringjoonel. Nooled märgivad transkriptsiooni promootoreid (P_H , P_L) ja replikatsiooni alguspunkte (O_H , O_L) asukohta. Allikas: <http://www.mitomap.org>

1.1.2 Inimese mtDNA populatsioonigeneetilise markerina

Fülogeneetiliste uuringute aspektist lähtudes on inimese mtDNA mitmetel põhjustel väga sobiv uurimismaterjal. Üheks neist põhjustest on mitokondri DNA uniparentaalne s. o. emaliinis päritavus (Giles et al., 1980). Kui imetajatel, kaasa arvatud inimestel, sisenevad viljastamisel munarakku koos spermiga ka selle viburis olevad umbes 100 mitokondrit, siis ootsüütides on mitokondreid kordades rohkem ja nende arv võib küündida koguni 100 000-ni. Lisaks sellele, et isapoolseid mitokondreid on ~1000 korda vähem kui emapoolseid, toimub rakujagunemise käigus ka isapoolsete mitokondrite hävitamine - need degradeeritakse

ubikvitinüleerimise protsessi abil proteosoomides ning lüsoosoomides (Sutovsky et al., 1999). MtDNA soospetsiifiline pärandumine annab meile võimaluse uurida inimpopulatsioonide emaliinide ajalugu.

Inimese mitokondri rõngasmolekul üldiselt ei rekombineeru (Merriwether et al., 1991; Olivo et al., 1983; Piganeau & Eyre-Walker, 2004), mis tähendab, et geneetiline materjal kandub edasi aheldunult ja järglasele pärandub üks kindel järjestus koos kindlate mutatsioonide kombinatsioonidega. Lisaks iseloomustab mtDNA-d suhteliselt kõrge mutatsioonide tekke kiirus, mis on tuumagenoomiga võrreldes keskmiselt viis kuni kümme korda suurem (Brown et al., 1979; Cann et al., 1987; Ingman & Gyllensten, 2001). See võimaldab uurida evolutsioonilises ajas suhteliselt hiljutisi sündmusi, näiteks anatoomiliselt moodsa inimese (AMI) kujunemist ja liikumisi. Suurema mutatsioonikiiruse põhjusteks võivad olla vabade radikaalide suurem sisaldus mitokondri maatriksis ja asjaolu, et mtDNA-l puuduvad reguleeriva ja kaitsva funktsiooniga histoonvalgud (Ingman & Gyllensten, 2001). Lisaks sellele on mtDNA reparatsioonisüsteem väiksema efektiivsusega ja täielikult töötavaid mehhanisme nukleotiidiasenduste ning valepaardumiste parandamiseks pole inimese mitokondris leitud (Bogenhagen, 1999).

Mutatsioonide tekkekiirus ja nende fikseerumine mitokondriaalses DNA-s sõltub nii piirkonnast kui ka mutatsiooni iseloomust. Põhilisteks mutatsioonideks mtDNAs on ühenukleotiidsed asendused, insertioonid ja deletsioonid. Transitsioonilised asendused (puriin asendub puriiniga, pürimidiin pürimidiiniga) on sagedasemad kui transversioonilised (puriin asendub pürimidiiniga ja vastupidi) (Aquadro & Greenberg, 1983). Mittekodeerivas kontrollregioonis on mutatsioonikiirus umbes 10 korda kõrgem kui kodeerivas alas (Howell et al., 2007; Pakendorf & Stoneking, 2005), kuid ka kodeerivas regioonis on kõrge mutatsioonitasemega nukleotiidipositsioone (np) (Soares et al., 2009).

Molekulaarse kella hüpoteesi kohaselt akumulereeruvad mutatsioonid kindlas DNA järjestuses suhteliselt konstantse kiirusega (Zuckerlandl & Pauling, 1962). Kalibreeritud molekulaarne kell võimaldab mutatsioonilise aja teisendada ajaks kalendriaastates. Kontrollregioonis HVS-I piirkonnas (np-de vahemikus 16090–16365) on põhiliselt olnud kasutusel mutatsioonikiirus $1,80 \cdot 10^{-7}$ transitsiooni nukleotiidi kohta aastas ehk üks transitsioon 20 180 aasta kohta (Forster et al., 1996; Saillard et al., 2000). Kodeeriva regiooni põhjal arvatud kella järgi toimub üks nukleotiidne asendus 5139 aasta kohta (Mishmar et al., 2003). Erinevad autorid on testinud mtDNA evolutsiooni neutraalsust mittesünonüümsete ja sünonüümsete asenduste fikseerumise erinevusi võrreldes ning leidnud, et fülogeneetiliste puude “nooremates” harudes on mittesünonüümseid asendusi rohkem kui

“vanades” harudes (Elson et al., 2004; Kivisild et al., 2006). Viimast seletab peaaegu neutraalse evolutsiooni teooria – nooremates harudes eksisteerib rohkem kergelt kahjulikke mutatsioone, sest puhastav valik pole jõudnud neid veel elimineerida (Elson et al., 2004; Ingman et al., 2000; Mishmar et al., 2003). Viimast arvestades on välja pakutud molekulaarse kella kalibreerimismeetod, mis kasutab vaid kodeeriva regiooni sünonüümsetest mutatsioonidest saadavat informatsiooni (Kivisild et al. 2006). Loogväli *et al.*, rek calibreeris mtDNA molekulaarset kella sünonüümseid mutatsioone kasutades ning sai tulemuseks, et keskmiselt tekib üks nukleotiidne asendus 7990 aasta kohta (Loogväli et al., 2009). Täna on enim kasutatav mtDNA mutatsioonikiirus $1,665 \cdot 10^{-8}$ asendust nukleotiidi kohta aastas ehk et keskmiselt tekib üks nukleotiidne asendus 3624 aasta jooksul, mis arvutati nii kodeeriva kui ka mittekodeeriva regiooni andmeid ja puhastava valiku mõju arvestades (Soares et al., 2009).

1.1.3 Inimese mtDNA nomenklatuur

Võttes arvesse eelpool toodud mtDNA omadused, on inimeste mtDNA haplotüüpide varieeruvus esitatav fülogeneesipuuna. Selle puu mistahes monofüleetilisi osi ehk klaade nimetatakse haplogruppideks (hg). Haplogruppe tähistatakse suurte tähtedega ning nende jaotamisel alamhaplogruppideks lisatakse tähele ka number (nt. J1). Järgmise taseme alamklaade eristatakse väikeste, alfabeetilises järjekorras olevate tähtedega (nt. J1a, J1b) ning edasi lisatakse taas number, täht jne. (Richards et al., 1998). Aja jooksul on kasutatud erinevaid haplogruppide defineerimise süsteeme s.h. (Torrioni et al., 1996; Richards et al., 1998; Macaulay et al., 1999). Tänapäevaks on haplogruppide nomenklatuur tänu üha põhjalikumatele fülogeneetilistele analüüsidele ühtseks muudetud (van Oven & Kayser, 2009). Internetilehele *Phylotree.org* on koondatud enamik inimese mitokondri genoomide järjestuste andmestikust tulenevast infost, mille alusel on koostatud nii kodeeriva kui mittekodeeriva regiooni mutatsioonidel põhinev inimese mtDNA ülemaailmne fülogeneesipuu (www.phylotree.org).

Haplogruppe defineeritakse mutatsioonide kaudu, võrreldes uuritavaid haplotüüpe kindla mtDNA järjestusega – referentsjärjestusega. Aastakümneid kasutati selleks esimesena sekveneeritud mitokondri täisjärjestust – nn. parandatud Cambridge'i referentsjärjestust (rCRS, ingl. k. *revised Cambridge Reference Sequence*) (Anderson et al., 1981; Andrews et al., 1999). Kuna rCRS pärineb Euroopa päritolu indiviidilt, ei ole see fülogeneetilise puu juurmine osa, vaid paigutub ülemaailmsel mtDNA fülogeneesipuul haplogruppi H2a2a1, mis on fülogeneesipuu üks tipmistest harudest (van Oven & Kayser, 2009). Hiljuti avaldatud neandertaallase (*Homo Neanderthalensis*) mtDNA genoomide järjestuste olemasolu andis

võimaluse inimese mtDNA fülogeneesipuud korrektselt juurida ning 2012 aastal pakuti välja uus mtDNA inimese rekonstrueeritud referentsjärjestus - RSRS (ingl k. *Reconstructed Sapiens Reference Sequence*) (Behar et al., 2012). Koos *PhyloTree* versiooni nr. 14 uuendamisega toimus ka üleminek rCRS referentsjärjestuselt RSRS-referentsjärjestusele (Behar et al., 2012).

Kuna rCRS on olnud kasutusel pika aja jooksul, siis võib eeldada, et üleminek uuele RSRS-referentsjärjestusele võib olla aeganõudev protsess. rCRS-lt RSRS-le põhinevale nomenklatuurile ülemineku hõlbustamiseks on loodud avalik internetiportaal *MtDNACommunity* (<http://www.mtDNAcommunity.org/>), mis pakub muudatuste tegemiseks vajalikku tarkvara (Behar et al., 2012) - seda kasutati ka antud töös.

1.2 Ülegenoomsed markerid populatsioonigeneetilistes uurimistöodes

Kuigi mtDNA ja Y-kromosoom on inimese geneetilise mineviku rekonstrueerimisel väga väärtuslikud uurimisobjektid, pakuvad nad siiski vaid soospetsiifilist lähenemist. Viimastel aastatel on lisandunud ka genoomide täisjärjestuste uuringud, millest saadavad teadmised populatsioonide geneetilise struktuuri kohta on olulised mitte ainult AMI demograafilistes uuringutes, vaid samamoodi ka ülegenoomsetes assotsiatsiooniuringutes, mis on suunatud eeskätt jahtima „haiguste gene“ (Price et al., 2006).

Rekombinatsioonide tõttu on iga inimese genoom kui erineva minevikuga DNA segmentide andmekogum. Kromosoomi segmentides olevat spetsiifilist alleelide kombinatsiooni nimetatakse haplotüübiks ning selle mittejuhuslikku aheldatud pärandumist nimetatakse aheldustasakaalutuseks - LD (ingl k. *linkage disequilibrium*) (Ardlie et al., 2002). Tänu LD-le moodustuvad inimgenoomis piirkonnad, mis päranduvad alati koos. Selliseid piirkondi nimetatakse haplotüübi plokkideks ehk haploplokkideks. Viimased kanduvad edasi põlvest põlve, muutudes aeglaselt ainult läbi juhuslike rekombinatsioonide ja mutatsioonide ja olles tugevalt aheldatud ning seega heaks vahendiks populatsioonide päritolu uurimisel (Jobling et al., 2004; International & Consortium, 2005).

Populatsiooni tasemel vaadatuna on haplotüüpide mitmekesisust mõjutanud erinevad evolutsioonilised mehhanismid, sealhulgas looduslik valik ja geenitriiv. Üheks oluliseks evolutsiooniliseks protsessiks on nn. küütvalik (*genetic hitchhiking*), mis seisneb loodusliku valiku mõju all oleva SNP ja temaga aheldunud mutatsioonide sageduse kujundamises. Näiteks kui SNP fikseerub läbi positiivse valiku, on mõjutatud ka kõik teised SNP-d samas haplotüübis ning kogu haplotüüp fikseerub populatsioonis. Seega, üldine haplotüüpide varieeruvus populatsioonis väheneb. Seda nähtust nimetatakse valiku luuatõmbeks (*selective sweep*) (Jobling et al., 2004; Smith & Haigh, 2007). Lisaks kujundavad haplotüübisagedusi sotsiaalsed ja demograafilised protsessid nagu pudelikaela efekt, asutajaefekt, populatsiooni ekspansioon(id), isoleeritus, paarilise valik ja populatsioonide segunemised (Gabriel et al., 2002; Tishkoff & Verrelli, 2003; Wang et al., 2002; Jobling et al., 2004). Näiteks populatsiooni “pudelikael” võib vähendada haplotüüpide mitmekesisust, samas kui populatsioonide segunemine võib seda tõsta. Lisaks mõjutab haplotüüpide mitmekesisust ja pikkust aeg. Mida lühem on haploplokk, seda rohkem on toimunud rekombinatsioone ning seda kauem on see haploplokk eksisteerinud (Gabriel et al., 2002).

Inimese genoomi varieeruvust iseloomustav informatsioon saadakse kas genotüpiseerimise või kogu genoomi täieliku resekveneerimise teel. Üle genoomi sadade tuhandete SNP genotüpiseerimine on tänaseks laialdaselt kasutatav meetod, kuid tehnoloogia

kiire arenguga (Metzker, 2010) muutuvad ka kogu genoomi resekveneerimise projektid järjest taskukohasemaks ja laialdaselt kättesaadavamaks (Colonna et al., 2011).

Oktoobris 2002 kutsuti ellu rahvusvaheline organisatsioon – HapMap`i projekt (*HapMap Project*), mille põhieesmärk on kaardistada SNP-d, mis kanduvad edasi plokkidena, kirjeldada nende sagedusi inimpopulatsioonides ning koostada inimgenoomi haplotüüpide kaart. HapMap`i projekti viimane versioon (nr. 28) sisaldab teavet nelja algselt uuritud populatsiooni [Euroopa päritoluga Utahi ameeriklased (CEU), hanid Hiinast (CHB), jaapanlased (JPT) ja joruubalased Nigeeriast (YRI)] ~ 1,5 miljoni ja seitsme lisapopulatsiooni ~ 4 miljoni SNP kohta inimgenoomis (Altshuler et al., 2010) (<http://hapmap.ncbi.nlm.nih.gov/>).

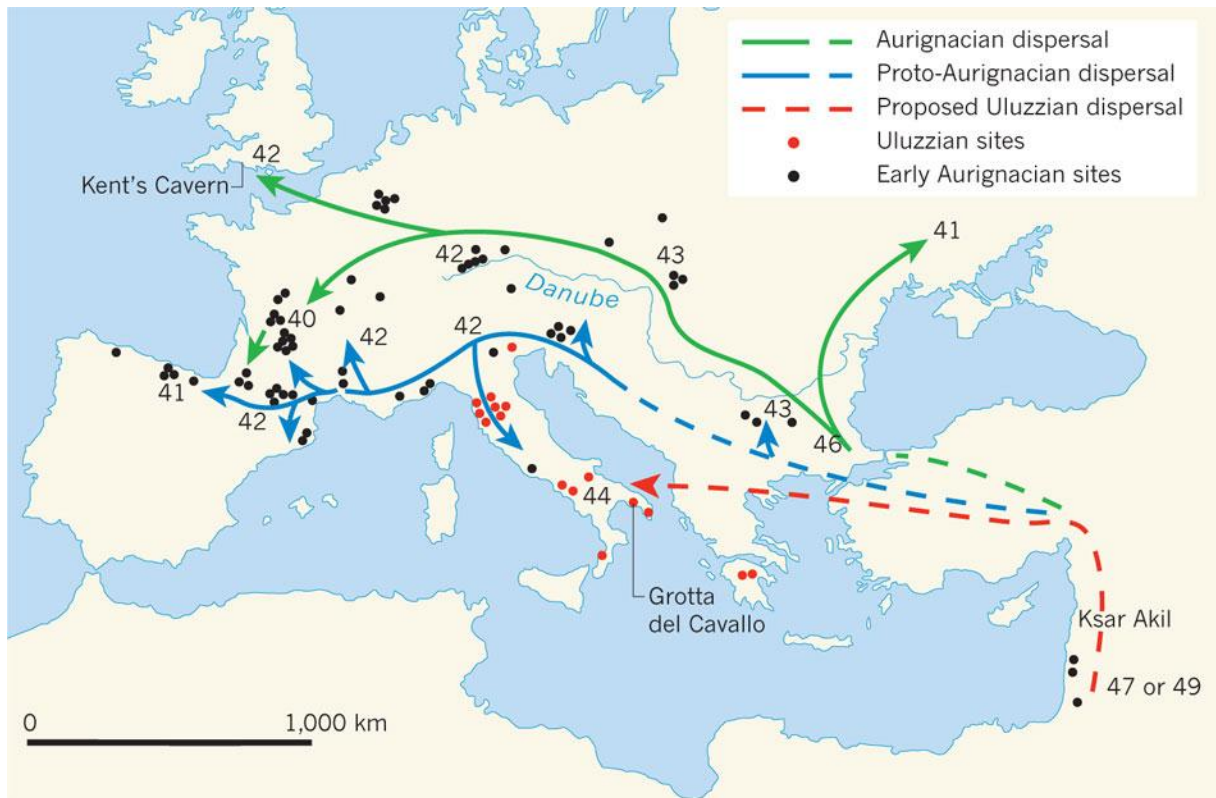
2008 aastal käivitatud 1000 Genoomi Projekti (*The 1000 Genomes Project*) eesmärgiks on kirjeldada populatsioonide geneetilist varieeruvust järjestusvariantide alusel, kaardistades üle 95% genoomist. Tänapäevaks on läbitud projekti pilootuuring, mis sisaldab andmeid 2-6-kordse katvusega sekveneeritud 170 Lääne-Aafrika, Euroopa, Hiina ja Jaapani indiviidi genoomi, keskmiselt 42-kordse katvusega kahte perekonda kuuluva kuue indiviidi genoomi (YRI ja CEU) ning seitsme populatsiooni 697 indiviidi eksoomi kohta (Abecasis et al., 2010) ning projekti esimene faas, mis sisaldab infot 14 populatsiooni 1092 indiviidi ülegenoomse 38 miljoni SNP kohta (Abecasis et al., 2012) (<http://www.1000genomes.org/>).

1.3 Lääne-Euraasia asustamisest kaasaegse inimese poolt

1.3.1 Lääne-Euraasia asustamine arheoloogilisest vaatenurgast

Palju kauem, kui geneetikal põhinevate meetoditega, on meie minevikku uuritud arheoloogilisi vahendeid kasutades. Üheks neist on radiosüsiniku dateerimistehnika, mille täiustumine (Hughen et al., 2004; Mellars, 2006; Ramsey et al., 2006), on tõstnud arheoloogiliste ajahinnangute täpsust ja usaldusväärsust. Anatoomiliselt kaasaegne inimene (AMI) on määratletud kolju morfoloogia põhjal ning selle vanim fossiilne jäänuk - kaks osalist pealuu, käe-, jala-, labajala ja vaagnaluu leidu Omo Kibish`ist, Etioopiast - on hinnatud 196 000 vanuseks (McDougall et al., 2005). Teine varane AMI leid - Herto fossiilid - pärineb samuti Etioopiast, nende hinnanguline vanus on 154 000 - 160 000 aastat (Stringer, 1992; Clark et al., 2003).

Euroopasse jõudis varaseim AMI kesk- ja ülempaleoliitikumi üleminekuajajärgul ja nende levimist üle kogu Euroopa on tihti seostatud sealsete neandertaallaste järk-järgulise kadumisega (Mellars, 2004). AMI ilmumine tõi endaga kaasa kultuurilisi muutusi hiliste neandertaallastega seostatud Moustier ja/või Aurignaci kultuurile eelnenud üleminekukultuuris (Bailey et al., 2009) (vt. tabel 1). AMI suhteliselt kiiret levimist Euroopas toetavad ka paar hiljutist tööd, mis kooskõlas kalibreeritud radiosüsiniku andmete ja arheoloogiliste tõenditega näitavad, et see võis toimuda veelgi kiiremini ja laiemalt, kui seda arvati varem (Mellars, 2004, 2006), ulatudes Lõuna-Inglismaale 44 200 – 41 500 aastat tagasi (at) (Higham et al., 2011) ja Lõuna-Itaaliasse 45 000 - 43 000 at (Benazzi et al., 2011). Aurignaci kultuuris viljeletud tehnoloogiad peegeldavad näiliselt järsku üleminekut selgelt eristatavale "kaasaegsele" kultuurilisele käitumisele, mille kohta annavad tunnistust keerukamaks muutunud luust ja sarvest tehtud tööriistad, eemaldatavate peadega harpuunid, mitmed isikupärased ehted, merekarpidest tehtud helmed, hammastest ripatsid jne. Keerukamaks olid muutunud ka abstraktne ja kujutatav kunst, mida tõestab värvide laialdasem kasutamine leitud esemetel ja koopamaalingud, millest tõenäoliselt suurejoonelisemad ja ühed vanimad leiud asuvad Lõuna Prantsusmaa Chauvet koobastes (Mellars, 1992, 2004; Conard & Bolus, 2003). See, et eelnevail Moustier kultuuri viljelevatel neandertaallastel puudusid sellised käitumuslikud tunnused (D'Errico 2003; Mellars, 2004), on viinud teadlased üsna üksmeelsele järeldusele, et just Aurignaci kultuuri saab seostada AMI jõudmisega tänapäeva Euroopasse (Mellars, 2005; Trinkaus, 2005; Bailey et al., 2009).



Joonis 2. Varajase AMI põhilised levikuteed Euroopa asustamisel (Mellars, 2011).

Aurignaci kultuuri arheoloogilistele tõenditele toetudes kasutas AMI Euroopasse jõudmiseks kahte põhilist teed. „Klassikaline“ tee läbi Kesk-Euroopa mööda Doonau kallast ja varajase Aurignaci (*Proto Aurignacian*) kultuuri kandjate levikutee piki Vahemere kallast (Mellars, 2004, 2005, 2011). Kolmas levikutee (Mellars, 2011) on välja pakutud seoses Uluzzia kultuuri ajajärku jäävate uute leidude avastamisega (vt. joonis 2). Uluzzia kultuurile iseloomulikke jooni peeti veel hiljuti kesk- ja ülempaleoliitikumi ajajärgu üleminekukultuuriks ja seostati Apenniini poolsaarel paiknenud neandertaallaste populatsiooniga, kes olid uudsemaid tehnoloogiaid üle võtnud sisserännanud ja algset Aurignaci kultuuri viljelenud AMI-lt (d’Errico et al., 1998; D’Errico, 2003; Mellars, 2005). Hiljutised uuringud Uluzzia tehnoloogilise kompleksiga seostatud leidudega näitavad, et see on siiski olnud pigem AMI kultuur (Benazzi et al., 2011) (vt. tabel 1).

Viimase jääaja maksimumi (LGM) eel levis Euroopas Gravette´i kultuur, mis on tuntud koopamaalingute, luudest ehete ning luust ja savist veenusekujukeste (väikesed naisekujukesed) poolest (Soares et al., 2010). LGM-i perioodil taandusid inimpopulatsioonid soojema kliimaga piirkondadesse - pelgupaikadesse ehk refuugiumitesse. Lõunapoolsematele aladele liikumine ei leidnud aset ilmselt üht kindlat teed pidi ja on oletatud, et paljudel juhtudel võis ümberasumise olulisemaks põhjuseks olla pigem toidunappus kui ebasobiv temperatuur. Teada on mitmeid arvatavaid refuugiume Edela-Euroopas, Vahemere rannikul,

Balkani poolsaarel ja Levantis, samuti Ida-Euroopa tasandikul ja ilmselt ka Musta mere idarannikul (Tambets, 2004 ja viited selles; Gamble et al., 2005;)

Tabel 1. Ülevaade peamistest *Homo* (H.) perekonna arheoloogilistest kultuuridest Euroopas. Kohandatud Soares et al., (2010) alusel.

Geoloogiline ajastik	Arheoloogiline ajastu	Kultuur	Periood*	Levikuala	Liigid
Pleistotseen	Kesk-paleoliitikum:	Moustier	300-30 at	Lääne-Euraasia	H.neanderthalensis, H.sapiens
		Uluzzia	45-43 at	Lõuna-Euroopa	H.sapiens
	Ülem-paleoliitikum (ÜP):	Aurignaci	45-30 at	Lääne-Euraasia	H.sapiens
		Châtelperroni	41-35 at	Lääne-Euroopa; Prantsusmaa	H.neanderthalensis, H.sapiens
		Gravette	32-23 at	Euroopa	H.sapiens
		Solutre	24-18 at	Lääne-Euroopa	H.sapiens
		Epigravette	20-10 at	Euroopa	H.sapiens
		Madeleine	17-12 at	Lääne- ja Kesk-Euroopa	H.sapiens
Pleistotseen/ Holotseen	Mesoliitikum:	Natufi	12,5 - 9,5 at	Lähis-Ida	H.sapiens
Holotseen	Ülem-paleoliitikum (ÜP):	Swidry	~9,5 at	Ida-euroopa	H.sapiens
	Mesoliitikum:	Kunda	8-5 at	Kirde-Euroopa	H. sapiens
	Neoliitikum	Paelkeraamika	7,5-6,9 at	Euroopa	H. sapiens

*tuhat aastat tagasi (at)

Inimese asualade killustumine Euroopa tasandikul tõi kaasa Gravette'i kultuuri regionaal-spetsiifiliste kultuuride tekke. Näiteks Solutre kultuur Lääne-Euroopas (24 000 – 18 000 at), mida tuntakse ennekõike arenenud tulekivitötluse poolest ja Epigravette'i Itaalias ning ilmselt ka Kesk- ja Ida-Euroopas (20 000 – 10 000 at). Hinnanguliselt 17 000 - 12 000 at eksisteerinud hilise jääaja Madeleine'i kultuuri on seostatud LGM-i järgsete populatsioonide ekspansiooniga ja see hõlmas põhiliselt Kesk- ja Lääne-Euroopa piirkondi (Soares et al., 2010) (vt. tabel 1). Kunstis esinevad koopareljeefid ja –joonised, millest tuntuimad asuvad Altamira koobastes Hispaanias ning Lascaux's Prantsusmaal. Soojad ajajärgud 15 000 – 13 000 at põhjustasid jää taganemise ja põhjapoolsetele alade taasasustamise (Forster, 2004).

Ida-Euroopas levinud Swidry kultuur märgib üleminekut hilispaleoliitikumist mesoliitikumi ajajärku (Dolukhanov, 1997) ja oli ilmselt eellaseks mesoliitikumi kultuuridele nagu näiteks Kunda kultuur, mida kutsutakse Swidry-järgseteks (*post-swiderian*) kultuurideks (Dolukhanov, 1997; Velichko et al., 2009). Mesoliitikum lõppes geograafiliselt erinevates piirkondades erinevatel aegadel seoses üleminekul neoliitikumi ehk põllumajanduse tekkimisega. Põlluharimiskultuur sai alguse Lähis-Idast umbes 12 000 - 11 000 at (Brown et al., 2009), kui uueks elatusviisiks said maaviljelus ja karjakasvatus. On väidetud, et selline põhjalik muutus võis olla tingitud lühikesest külmaperioodist *hilises driüüases* 12 800 - 11 500 at, mis tõi kaasa inimeste paikseksjäämise (Bar-Yosef, 1998). Seda arengut on nimetatud ka

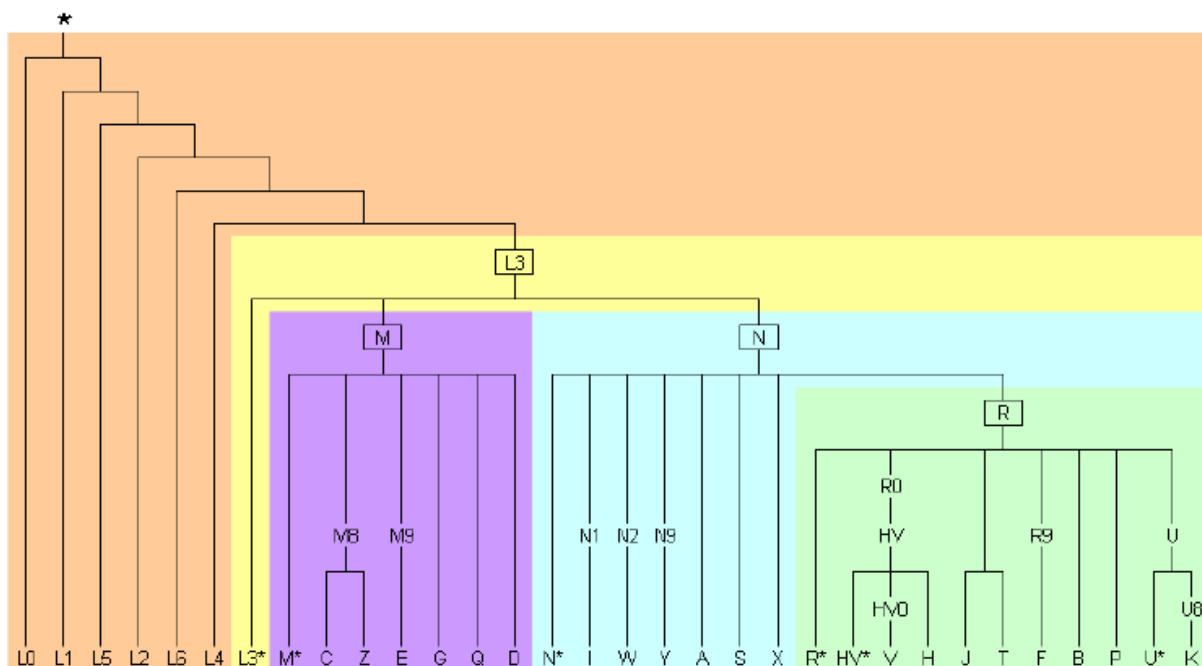
neoliitiliseks revolutsiooniks. Huvitaval kombel näib, et neoliitikumi ülemineku ajal ~10 000 - 6 000 aastat tagasi kasutasid ka esimesed põlluharimisega tegelevad kogukonnad Euroopasse jõudmiseks kahte umbes samasugust levikuteed nagu Aurignac'i kultuuride kandjad - põhjapoolset teed mööda Doonau orgu ja lõunapoolset teed mööda Vahemere rannikut (Mellars, 2004, 2006). Kesk-Euroopasse jõudis põllumajandus ~7 500 at. Samasse ajajärku jääb ka nn. paelkeraamika (*Linearbandkeramik*) kultuuri levik, mis ulatus Lääne-Ukrainast Prantsusmaani. Vaidlused selle üle, kas põllumajanduse levik tõi endaga kaasa peamiselt ülemineku uutele tehnoloogiatele (kultuuriline difusioon) või hõlmas see suuri migratsioone ja rahvastiku ulatuslikku vahetumist (deemiline difusioon), kestavad siiani (Cavalli-Sforza et al., 1994; Haak et al., 2005; M. Richards et al., 1996; Soares et al., 2010; Pinhasi et al., 2012).

Neoliitikumile järgnes eneoliitikum ehk vaseaeg, mida iseloomustab vasetöötlemise tehnoloogiate kasutuselevõtt, kusjuures peetakse tõenäolisemaks selle tehnoloogia sõltumatut arengut mitmetes erinevates piirkondades (Radivojević et al., 2010). Vasemaagi kaevandamise ja selle sulatamise tehnoloogiad levisid kiiresti ja muutusid eri piirkondades eri aegadel järk-järgult pronksiaja kultuurideks. Vaseaeg tõhustas ka kaubandusvõrgustike teket ja kasutamist ja ulatuslikke migratsioone (Roberts et al., 2009). Laiaulatuslikud liikumised Euroopa sees ja kaugemalegi jätkusid ka pronksi- ja rauaajal, mille näiteks võib tuua mitmeid sissetunge Briti saartele jne (Haywood, 1997).

1.3.2 Lääne-Euraasia asustamine geneetilisest vaatenurgast

Aurignaci kultuuri arheoloogilise leidude levikumustril põhinevat (Mellars, 2004, 2011) AMI levimist Aafrikast läbi Levanti koridori Euroopasse (Soares et al., 2010) toetavad ka kõik praegused geneetilised tõendid. Ekspansiooniprotsesside käigus tekkinud gradient, kus Aafrikast kaugenedes geneetiline mitmekesisus üha väheneb, on tingitud mitmest järjestikusest asutajaefekti sündmusest või/ja populatsiooni „pudelikaelast“ (Li et al., 2008; Amos & Hoffman, 2010). MtDNA täisjärjestuste põhjal konstrueeritud fülogeneesipuu alusel on näidatud, et Aafrikast väljunud liinid moodustavad kogu mtDNA haplogruppide mitmekesisusest vaid väikese osa ning puu sügavamad harud asuvad Aafrikas (Ingman et al., 2000; Torroni et al., 2006) (vt. joonis 3). Enamusele mtDNA haplogruppidele on omane spetsiifiline geograafiline levikuala – võib öelda, et suuremad haplogrupid on kontinenti-spetsiifilised (Ingman et al., 2000; Richards et al., 2000; Herrnstadt et al., 2002) ning erinevaid populatsioone iseloomustavad erinevad haplogruppide sagedusmustrid. Kogu emaliinide varieeruvuse väljaspool Aafrikat hõlmavad Aafrika-spetsiifilise haplogrupi L3

kaks haru - makrohaplogrupid M ja N (Kivisild et al., 2002; Metspalu et al., 2004). Erinevalt makrohaplogrupist M, mis on levinud Ida-Euraasias (Kivisild et al., 2002) ning sama hästi kui puudub Euroopas, on teine suur haru N selline fülogeneetiline makrohaplogrupp, mis koondab enda alla enamuse Lääne-Euraasia haplogruppidest nagu R (HV, H, V T, J, U koos alamharuga K), X, N1 koos alamharuga I ning N2 koos alamharuga W (vt. joonis 3 ja 4) (www.phylotree.org).



Joonis 3. Lihtsustatud mtDNA fülogeneesi puu. Puu juur, mis esindab kõikide inimeste viimast ühist emaliini esivanemat, on tähistatud tärniga. Superhaplogrupi L3 (kollane) harud M (lilla), N (sinine) ja R (roheline) hõlmavad varieeruvuse väljaspool Aafrikat. Tärniga märgistatud haplogrupi sümbolid esindavad kõiki teisi ülejäänud sellesse klaadi kuuluvaid liine (nt. hg HV* sisaldab endas alamharusid HV1, HV2, HV4 - HV17). Globaalne inimese mtDNA varieeruvus on kajastatud veebisaidil <http://www.phylotree.org> (joonis on pärit van Oven M & Kayser M., 2009).

Kui omavahel kõrvutada Euroopas ja Lähis-Idas levinud mtDNA haplogruppe, siis on näha, et need suures osas kattuvad. Diversiteediarvutused Euroopa ja Lähis-Ida haplogruppidele on näidanud, et mõlemas piirkonnas esindatud haplogruppide sisene mitmekesisus oli üldjuhul suurem Lähis-Idas, välja arvatud Euroopa-sisese tekkega haplogrupid (Richards et al., 2000). Lähis-Ida on seega osutunud teatud mõttes sõlmpunktiks, kust toimusid mitmed ränded Euroopa aladele (M. Richards et al., 2000; Tambets et al., 2003; Pala et al., 2012).

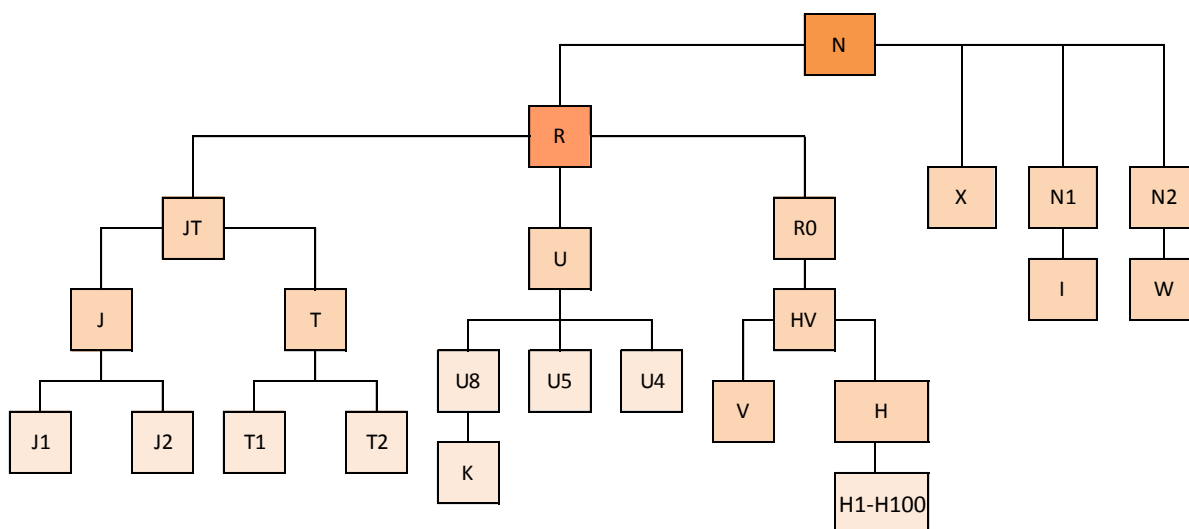
Euroopa ja Aasia populatsioonide ülegenoomsete andmete võrdlusel on tuvastatud, et kõikide teiste populatsioonide haplotüüpid olid 90% ulatuses Aafrika mandril esindatud

(Gabriel et al., 2002; Wang et al., 2002; Tishkoff & Verrelli, 2003; Jobling et al., 2004). Aafrika suurim haplotüübiline mitmekesisus ning lineaarselt vähenev haplotüübi heterosügootsus sahara-tagusest Aafrikast läbi Lähis-Ida Euroopani toetab mitme järjestikuse asutajaefektiga mudelit ning inimeste saabumist Euroopasse läbi Lähis/Kesk-Ida (Colonna et al., 2011; Li et al., 2008). Aafriklaste haploplokid (kuni 96 tuhat ap) on võrreldes Euroopa ja Aasia populatsioonidega (kuni 137 tuhat ap) üldiselt lühemad - rõhutades sellega kaasaegse inimese populatsioonide „Aafrika juuri“ (Gabriel et al., 2002).

Ülegenoomsete genotüüpide andmetel põhinevad põhikomponentanalüüsi kui ka populatsiooni struktuuri analüüsi andmed kajastavad samuti üldiselt sujuvat üleminekut Lähis/Kesk-Idast Euroopasse (Li et al., 2008; Behar et al., 2010;). Eurooplastel uuritud kogu genoomi genotüübi andmed näitavad peamiselt tihedat seost geograafilise ja geneetilise kuuluvuse vahel (Novembre et al., 2008; Nelis et al., 2009) ning struktuuranalüüsis on näha selget kogu Euroopale ühist eellaskomponenti, mis on olemas ka Lähis-Ida populatsioonides (Li et al., 2008; Abecasis et al., 2010; Behar et al., 2010; Rasmussen et al., 2010).

1.3.2.1 Euroopa emaliinide geneetilise varieeruvuse kujunemine

Euroopa on geneetiliselt üsna ühetaoline – geneetilised distantsid erinevate populatsioonide vahel on suhteliselt lühikesed, kuigi on ka mõned erandid nagu näiteks saamid, sardiinlased ja baskid (Cavalli-Sforza et al., 1994), mis kajastub ka mtDNA haplogruppide leviku- ja sagedusmuustrites. Sügavam fülogeneetiline analüüs on geneetiliselt ühetaolises Euroopas siiski välja toonud haplogruppide piirkondlikke levikumustrite erinevusi (Richards et al., 2000, 2003).



Joonis 4. Lihtsustatud mtDNA fülogeneesi puu, kus on välja toodud makrohaplogrupist N ja selle alamhargnevad põhilised Euroopas levinud haplogrupid (joonis on tehtud Torroni et al., 2006 alusel).

Haplogrupist R hargnevaist Euroopale iseloomulikest mtDNA hg-st (vt. lk 19) kirjeldatakse selles peatükis hg H, V, T, J, U ja nende põhilisi alamharusid, mis hõlmavad kokku ligikaudu 80 % kogu Euroopa emaliinide varieeruvusest (LISA 5).

Euroopa vanimad emaliinid kuuluvad haplogruppi U – need on hg-d U5 ja U8, mille vanuseks on arvatud vastavalt ~ 37 000 ja 50 000 aastat (Soares et al., 2010). Hg U8b alamharu hg K on üks esimesi Euroopa populatsioonis kirjeldatud haplogruppe, mille kuuluvust hg-sse U näitas Hofmann et al., (1997). Hg K arvatakse pärinevat Lähis-Idast umbes 30 000 aastat tagasi (Soares et al., 2010) ning see on U8 alamharudest kõige laialdasema levikuga ning kõrgema sagedusega haplogrupp. Euroopas moodustab hg K emaliinide mitmekesisusest enamasti vähem kui 10%, olles sagedasem Lääne- ja haruldasem Ida-Euroopas (Richards et al., 2000; Herrnstadt et al., 2002). Kõrgeim hg-i K sagedus (30%) on leitud Ashkenazi juutidel (Behar et al., 2004). Hg U5 sagedus jääb kaasaegses Euroopas keskmiselt 5 - 7% piiresse (Malyarchuk et al., 2010). Hg U5 alamharude U5a ja U5b vanuseks on arvatud vastavalt ~16 000 - 20 000 ja ~20 000 - 24 000 aastat, mis jäävad LGM ajajärku. Mõlemad hg U5 harud on esindatud peaaegu kõikides Euroopa populatsioonides (Tambets et al., 2004; Achilli et al., 2005). Kui U5a on levinud rohkem Euroopa lõuna- ja vähemal määral loodeosas (Richards et al., 1998), siis U5b ja tema alamharud on levinud madalal sagedusel üle kogu Euroopa - Ibeeriast kuni Uurali mägedeni (Richards et al., 1998; Sajantila et al., 1995; Tambets et al., 2004). Erandina tuleb välja tuua ühe U5b alamhaplogrupi U5b1b haru – U5b1b1, mis moodustab suure osa Põhja-Skandinaavia saamide emaliinidest (26% Rootsi, 41% Soome ja 57% Norra saamid). Hg U5b1b suurem diversiteet Lääne-Euroopas toetab teooriat, mille kohaselt algas U5b1b ekspansioon Euroopa edelaosast, Franco-Cantabria refuugiumist, kuid jõudis Skandinaavia põhjaossa läbi Ida-Euroopa, kus hg U5b1b1 on ka levinud, ehkki suhteliselt madala sagedusega (Tambets et al., 2004). Hg U5b alamharu U5b3 arvatakse aga pärinevat Lõuna-Euroopa regioonist, sealsest võimalikust Apenniini poolsaare refuugiumist (Pala et al., 2009). Tõendeid selle kohta, et hg U5 oli LGM järgsel perioodil Kesk – ja Põhja-Euroopas suhteliselt sage, on saadud ka vana DNA uuringutest, kus võrreldi omavahel nende Euroopa piirkondade hilisest ülempaleoliitikumist ja mesoliitikumist pärinevate küttide ja korilaste (22 mtDNA proovi) ning samast piirkonnast pärit tänapäevaste (484 mtDNA proovi) ja varasest neoliitikumist pärit säilmetest eraldatud mtDNA andmeid [25 proovi (Haak et al., 2005)] (Bramanti et al., 2009). Kõik kolm uuritud gruppi olid hg-de esinemismustri poolest erinevad. Näidati, et koguni 82% küttide ja korilaste mtDNA-dest kuulusid hg-sse U, enamus selle alamhg-sse U5, samas kui kaasaegses Euroopa valimis moodustab hg U harva enam kui veerandi emaliinidest.

Hg U puudus varasemalt uuritud varajaste põlluharijate valimis aga täielikult, ja vastupidi – neoliitikumi valimis sage hg N1a puudus mesoliitikumi ja ülempaleoliitikumi valimist. Töö autorid järeldasid, et tõendeid selle kohta, et esineks otsene geneetiline järjepidevus küttide ja korilaste ja Kesk-Euroopa varajaste põlluharijate vahel, ei leitud. Rõhutati täiendavate uuringute olulisust ja seda, et tõenäoliselt mõjutas põlluharimise levik ja sellega kaasnenud rahvastiku liikumine Lähis-Idast Euroopasse erinevaid Euroopa piirkondi erinevalt (Bramanti et al., 2009).

Levinud on seisukoht, et hg H laienes Euroopasse Lähi-Idast enne LGM (~22 000 at). Samas on oluline märkida, et enamuse alamhaplogruppe hakkasid levima seoses jääaja lõpu ning jää all olnud alade taasasustamisega refuugiumidest lahkuvate küttide-korilaste poolt (Pereira et al. 2005; Soares et al. 2010). Hg H ja tema rohked haplotüübid on tänaseks jaotunud ligikaudu 100 erineva haru vahel (http://www.phylotree.org/tree/subtree_R0.htm), olles Lääne-Euraasias kõige sagedasemad mtDNA tüübid ja moodustades keskmiselt 45% kaasaegsete Euroopa populatsioonide emaliinidest (Richards et al., 2000; Achilli et al., 2004; Loogväli et al., 2004; Pereira et al., 2005; Roostalu et al., 2007). Hg H suurim tütarharu H1 hõlmab ligi 30% hg H-st ja 13% kogu Euroopa mtDNA geenitiigist (Loogväli et al., 2004). Korduvalt on näidatud, et Euroopa-sisese päritoluga hg-de H1 ja H3 sagedus Euroopas langeb edelast ida suunas (Torroni et al., 1998, 2001; Achilli et al., 2004; Richards, 2003). Need kaks hg-d moodustavad ~65% Pürenee poolsaare, ~46% Loode-Euroopa ja ~ 27% Kesk- ja Ida-Euroopa hg H emaliinidest (Pereira et al., 2005). Erinevad uuringud on andnud alust arvata, et mtDNA hg-d H1 ja H3 pärinevad tõenäoliselt Pürenee poolsaarel asunud refuugiumist ja levisid Euroopa sisemaale pärast viimase jääaja lõppu (Achilli et al., 2004, Loogväli et al., 2004; Pereira et al., 2005; Torroni et al., 2006). Ülejäänud H alamklaadid on väiksema sagedusega, mis teeb ka nende geograafiliste levikumustrite leidmise raskemaks. Kui hg H5 on madala sagedusega levinud üle Lääne-Euraasia (Loogväli et al., 2004), siis tema alamharu H5a on enamjaolt levinud just Euroopas. Hg H5a väga madalat esinemissagedust Lähis-Ida ja Kaukasuse populatsioonides on seostatud tema võimaliku Euroopa-sisese tekkega (Pereira et al., 2005). Hg-d H6 ja H2 on sagedasemad idapoolses Euroopas ja Kaukasuses, kuid kumbagi pole leitud Lähis-Idast. Hg H4, H7 ja H13 leidub nii Euroopas kui ka Lähis-Idas, neist viimane on esindatud ka Kaukasuses (Pereira, et al., 2005; Roostalu et al., 2007).

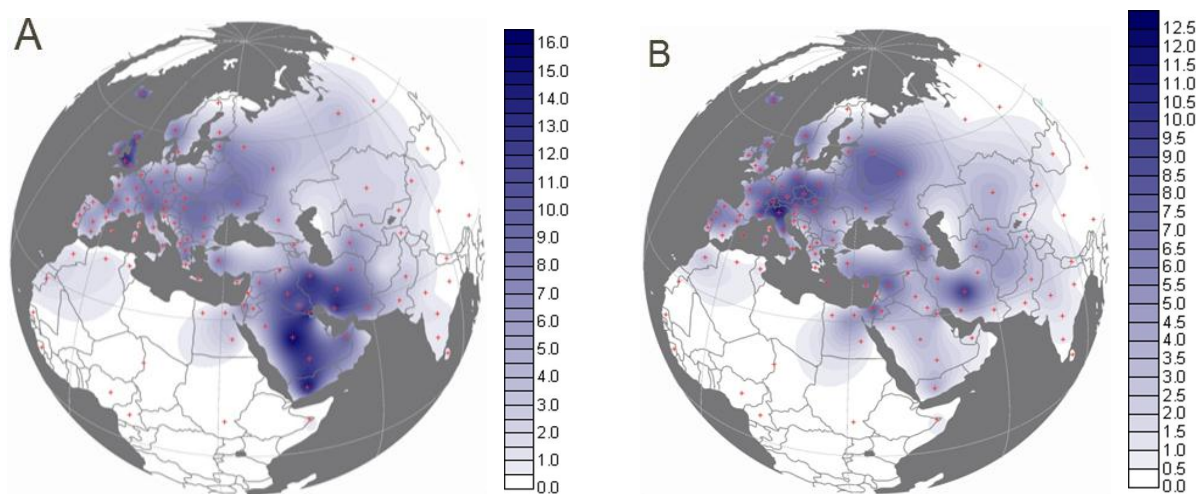
Hiljuti avaldati uurimustöö hg H kunagiste levikumustrite kohta (Brotherton et al. 2013). Töö põhines Kesk-Euroopast Saksamaalt neoliitikumi erinevatest perioodidest pärineva luumaterjali DNA analüüsil. Hg-sse H kuuluvate emaliinide täisgenoomide (39 proovi) andmete analüüs näitas, et uuritud arheoloogilise materjali kollektsiooni lõikes on

sellele haplogrupile, nii nagu tänapäevalgi, iseloomulik suur mitmekesisus ja vanuseline kihistumine. Leiti, et vanemaid neoliitilisi hg H emaliinivariante leidub tänapäeva Euroopa populatsioonides vähe (H16, H23, H26), need on kas kadunud või pole valimitesse sattunud (H46b, H88, H89). Samas, nooremaid variante (näiteks H3, H4, H6, H7, H11 ja H13) leiab tänapäeva Euroopast piisavalt. Lisaks tuvastati, et vanimad uuritud neoliitikumi proovid kuuluvad haplogrupi H sellistesse alamharudesse, mis liigituvad kaasaegsete proovide taustal kokku Lähis-Ida ja Anatoolia populatsioonidega - piirkonnaga, kus põlluharimiskultuur tekkis ja kust see Euroopa poole levima hakkas. Seega Euroopas aset leidnud neolitiseerumise protsess tõi ilmselt kaasa lisaks kultuurilisele üleminekule ka rahvastiku sisserände Lähis-Idast ja kõige esimesed neoliitilised sisserändajad ei ole tänaste eurooplastega tõenäoliselt seotud otsese järjepidevuse alusel, vaid et vahepeal, umbes 6000 aasta eest, on toimunud seoses rahvastiku kiire kasvuga ühtede liinide asendumine teistega, millel omakorda on järjepidevus kaasaegsete eurooplastega. Keskmises neoliitikumis asetleidnud võtmetähtsusega demograafilised muutused olid ka ilmselt aluseks sellele, et haplogrupp H tänapäeval Euroopas nii sage on (Brotherton et al., 2013).

H haplogrupile fülogeneetiliselt lähedase haplogrupi V esinemissagedus enamikes Euroopa populatsioonides jääb 1 - 6 % piiridesse (Torroni et al., 1998, 2001), madalam on see Euroopa idapoolsetes populatsioonides ning kõrgem esinemissagedus on tuvastatud Pürenee poolsaarel, eriti sealsetel baskidel ja kataloonlasel (vastavalt 12,4% ja 24%). Erandiks idapoolsete populatsioonide hulgas on saamid, kuna just neil on leitud kõige suurem hg V sagedus [Rootsi (68%) Soome (37%) ja Norra saamid (33%)] (Tambets et al., 2004). Samas on leitud, et saamid on hg V varieeruvus minimaalne, millest võib järeldada, et hg V kõrge sagedus on seotud geenitriiviga (Torroni et al., 1998). Enamlevinud teooria kohaselt pärineb ka mtDNA hg V (Torroni et al., 1998, 2001) tõenäoliselt Pürenee poolsaarelt. Hg-de H1, H3 ja V ekspansioon toimus ~11 000 – 11 500 at, mis langeb kokku jääaja järgse hilise drüüase külmaperioodi lõpuga (Soares et al., 2010), kui Franco-Cantabria refuugiumis külmaperioodi eest varjus olnud inimesed hakkasid taasasustama Euroopa sisemaad (Torroni et al., 1998, Torroni 2001, 2006; Achilli et al., 2004, Loogväli et al., 2004; Pereira et al., 2005).

Neoliitiliste populatsioonide ekspandeerumise kandidaatmarkeriteks on varasemate uuringute põhjal pakutud mtDNA haplogruppe T1 ning J (Richards et al., 1996, 2000; Soares et al., 2010). Uued mtDNA täisjärjestustel põhinevad andmed näitavad aga mtDNA hg J ja T (ja nende põhiliste alamklaadide J1, J2, T1 ja T2) võimalikku levikut Lähis-Ida refuugiumist Euroopasse juba LGM lõpus või koheselt pärast seda, 19 000 - 12 000 at (Pala et al., 2012).

Hg JT vanus - ~58 000 aastat - on võrreldav hg U vanusega [(~55 800) (Soares et al., 2010)], mistõttu võiks olla usutav teooria, et need haplogrupid pärinevad ühest ja samast varajaste väljarändajate grupist, kes on eellasteks nii Lähis-Ida kui ka Euroopa elanikele. Hg-d J ja T on täisjärjestuste põhjal tehtud arvutuste kohaselt tekkinud vastavalt ~40 000 ja ~30 000 aastat tagasi ning nende levimine Euroopasse algas pärast LGM-i. Hg J moodustab ~9% Euroopa ja ~13% Lähis-Ida emaliinide valimist ning selle sagedasem alamharu J1 omakorda ~80% kogu hg J-st. Hg T sagedus eurooplastel jääb ~10% piiresse ja on ~8% Lähis-Idas, selle alamklaad T2 on rohkem levinud Lääne-Euroopas kattes sealsest emaliinide varieeruvusest ~8% ja ~80% kogu hg T-st (Pala et al., 2012) (joonis 5).



Joonis 5. Haplogruppide J1 (A) and T2 (B) sageduse jaotuse kaardid (Pala et al., 2012). Värviskaala näitab sageduste väärtusi (%) ja punased punktikesed uuritud populatsioonide asukohta kaardil.

Viimase jääaja hilisel perioodil (16 000 - 12 000 at) arvatakse olevat Euroopasse levinud alamharud J1c, J1b1a, J2a1, T1a1, T2a1b ja T2f1, vahetult jääaja järgsel perioodil aga hg-d T2b ja T2e. Hg J1 alamklaad J1c (peaaegu 80% hg J1-st) on põhiliselt esindatud Kesk-Euroopas, eriti selle keskosas, ning ka Ukrainas ja Balkanil. Hg J2a1 ja J1b1a päritolukohaks on pakutud Euroopat, kuid nende levik üle Kesk-Euroopa ning Vahemere piirkonnas toimus erinevatel ajaperioodidel: vastavalt ~11 000 at (alamklaadina J2a1a) ning ~8000 at. Hg J2b1 on rohkem leitud Vahemere,- Kesk- ja Lääne-Euroopa (Atlandi-Euroopa) piirkonnast, kusjuures tema alamklaad J2b1a on esindatud peaaegu ainult Euroopas (Pala et al., 2012).

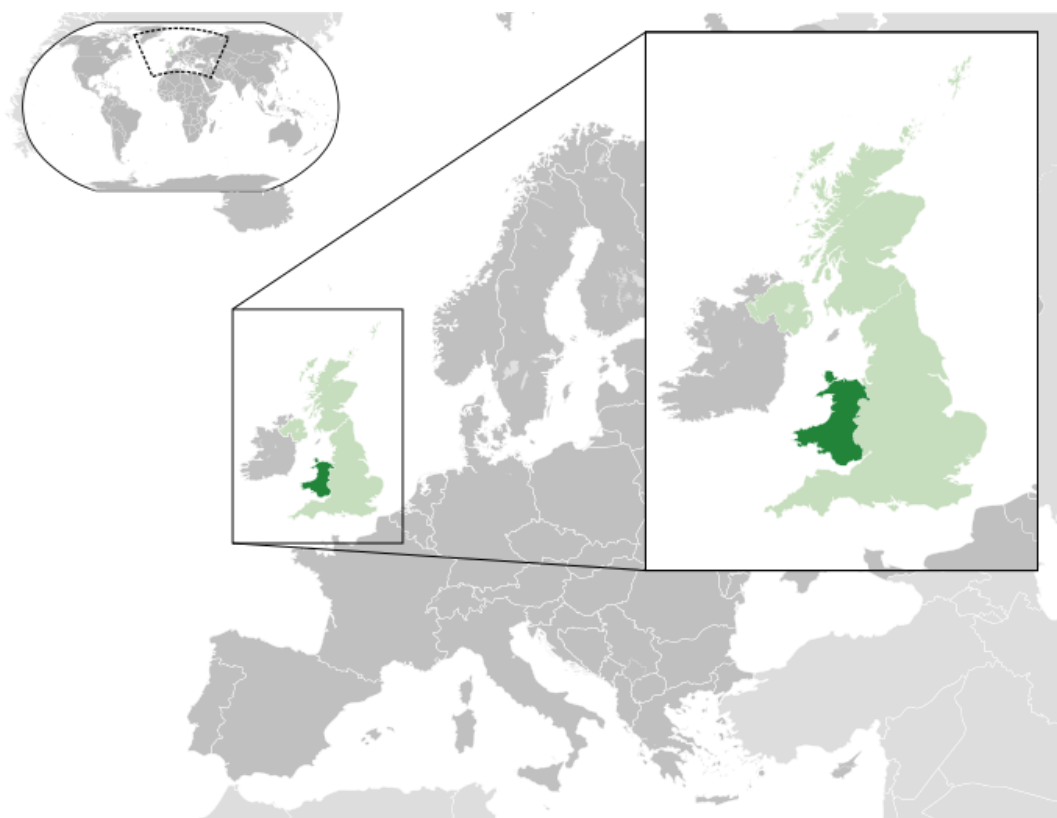
Haplogrupp T1 varieeruvusest katab umbes 90% tema alamklaad T1a. Kõige selgemini vastab neoliitikumis toimunud ekspansioonile ilmselt Euroopas tekkinud alamharu T1a1a1 (~7000 aastat) levikumuster. Enamus hg T2 alamharude fülogeneesipuid, sh T2b, on täheksjulise struktuuriga ning dateeritud ~10 000 aasta vanuseks. Viimase võimalikku levikut varajasel neoliitikumi perioodil üle Euroopa ning oletatavat Euroopa päritolu toetavad ka

mitmed vana DNA uuringute põhjal neoliitikumi perioodi dateeritud T2b liinid (Haak et al., 2005, 2010; Lacan et al., 2011a;b). Kui alamhg T2f (vanusega ~17 000 at) on ilmselt pärit Euroopast ja põhiliselt seal ka levinud, siis alamklaad T2a tekkis ilmselt samal ajal Lähis-Idas ja selle ekspandeerumine Euroopasse on toimunud mitme erineva lainena, näiteks T2a1a ~6 000 at neoliitikumis ning T2a1b ~13 500 at hilisel jääaja perioodil (Pala et al., 2012).

1.4 Briti saared ja Wales

1.4.1 Briti saarte ja Wales`i geograafiast, ajaloost ja arheoloogiast

Wales on Suurbritannia ajaloolis-geograafiline piirkond, mis asub Suurbritannia saare edelaosas (joonis 6). Lõunast põhja hõlmab Wales ligikaudu 274 km ja idast läände 97 km, Wales`i pindala on umbes 20,779 km². Wales piirneb idas Inglismaaga, teistes suundades merega: põhjas Iiri merega, läänes Saint George`i väinaga, lõunas Bristol'i lahega ning edelas Keldi merega. Wales`i koosseisu kuulub üle 50-ne saare, suurim neist on loodes asuv Anglesey (*The Times*, 2004). Waleslased ehk kõmrid erinevad Suurbritannia põhielanikkonnast oma keele- ja kultuuriloolise tausta poolest. Kõmri (Cymru) keel on Indo-Euroopa keldi keelte hulka kuuluv keel (<http://www.ethnologue.com>), mis on lähedasem sugulane Mandri-Euroopa saksa, prantsuse ja hispaania keelega kui inglise keelega. Elanikke on Wales`is ligikaudu 3 miljonit ning kõmri keele kasutajaid (oskajaid) sellest ligikaudu 22%.



Joonis 6. Wales`i geograafiline paiknemine Euroopa kaardil.

Allikas: http://en.wikipedia.org/wiki/File:Wales_in_the_UK_and_Europe.svg

1927. aastal Suurbritannia lõunaosast Kenti koopast avastatud AMI arheoloogilised leiud on uuemate andmete põhjal dateeritud 44 200 - 41 500 aastasteks, mis on ühtlasi üheks vanimaks arheoloogiliseks leiuks Loode-Euroopast (Higham et al., 2011). Esimene teadaolev

AMI fossiilne leid Wales'ist pärineb aastast 1823, kui Lõuna-Wales'i Goweri poolsaare Pavilandi lubjakivi koobastest leiti nn. "Red Lady" luustik. Nimetus tuleneb sellest, et skelett oli üleni ookripunast värvi (Jacobi & Higham, 2008). Hiljem sai selgeks, et esmased tähelepanekud luustiku kohta olid olnud osaliselt eksitavad ja et "Red Lady" on hoopis noor mees (Trinkaus & Holliday, 2000). Leitud luude vanuseks hinnati ~34 000 aastat, mis on ühtlasi vanim arheoloogiline tõend AMI kohta Wales'is. "Red Lady't" on seostatud ka vanima Lääne-Euroopas teadaoleva tseremoonilise matmistavaga, mida iseloomustavad koopast avastatud rohked ehteleid (elevandiluust helmed ja merekarbid) ja mis sobituvad nii iseloomulike tunnuste kui ka vanuse poolest varajase Gravette'i kultuuri perioodi. Samadest koobastest pärinevad ka ühed Suurbritannia rikkalikumad leiud (~4 000 kivitööriista, hambad, luud, helmed) ülempaleoliitikumi Aurignaci kultuurist (Jacobi & Higham, 2008).

Wales'i asustamine on toimunud ilmselt periooditi, see on sõltunud kliima külmenemise ja soojenemise ajajärkudest, jääkatte tekkimisest ning taganemisest. LGM-i ajal ja vahetult pärast seda ~21 000 kuni 13 000 aastat tagasi oli Wales ilmselt hüljatud ala. Viimase jääaja lõpuperioodist on leitud arheoloogilisi tõendeid Wales'i põhjaosast. Llandudnost, Great Orme kaljult, Kendrick'u koopast leiti haud inimeste ja loomade jäänustega, tulekivist tööriistu, looma hammastest tehtud mulgustatud helmeid jms. (Lynch et al., 2000). Pidevalt on see ala asustatud olnud alates viimase jääaja lõpust (12 000 - 10 000 aastat tagasi), mil Suurbritanniasse hakkas Loode- ja Kesk-Euroopast rändama mesoliitikumi kütte-korilasi. Kuni 7 000 aastani enne meie aega (e.m.a) oli inimestel võimalik Mandri-Euroopast Suurbritanniasse kõndida jalgsi. Liustike sulades tõusis aga merevee tase ja Suurbritanniast sai saar. Tekkisid La Manche'i väin ning Wales'i ja Iirimaa eraldav Iiri meri. Arvatakse, et asutajate teine suurim laine tuli Prantsusmaalt Atlandi rannikut mööda ja ilmselt veeteed pidi Inglismaa lääneossa, Walesi ning Iirimaa (Keeling, 2013). Wales'ist on leitud mitmeid mesoliitikumi leide, varaseima dateeringuga mesoliitiline leiukoht asub Wales'i edelaosas, Pembrokeshire's, Nab Head'is, mis on dateeritud ~9 200 aasta vanuseks. Mõningad varajased kunstileiud kaunistatud kivide näol on leitud Põhja-Wales'ist Rhuddlan'ist (Lynch et al., 2000).

Neoliitikumist pärinevad paljud hauakambrid, millest silmapaistvaimad asuvad Loode-Wales'is Anglesey saarel. Megaliitseid kalmeid on Wales'ist leitud mitut erinevat tüüpi, neil on sarnaseid jooni Iirimaa ja Euroopa Atlandi piirkonnast leitud (Lynch et al., 2000). Wales'ist on avastatud ka mõningad neoliitikumi perioodi kuulunud elamute jäänukeid. Kõige silmapaistvam neist on Clegyr Boia Pembrokeshire's St. Davidi lähedal. Esemelistest

leidudest domineerivad poleeritud kivist kirvepead ning keraamikal on leitud sarnasusi Iirimaa leidudega (Lynch et al., 2000).

43. aastal m.a.j tungisid roomlased Suurbritanniasse ja varaseim ajalooline dokument Wales'i kohta, mis kirjeldab roomlaste sissetungi Wales'i, pärineb 48 aastast. Roomlaste impeerium Briti saartel püsis üle 300 aasta (~410 aastani) ning sel ajal oli Suurbritannia asustatud erinevate hõimude poolt. Kuigi roomlased valitsesid suurema osa Britannia üle, polnud nad kunagi võimelised alistama saare põhja- ja läänepoolseid tippe, Cornwalli, Wales'i (Woolf, 2010). ~450 aastal m.a.j hakkasid Britanniat idast lääne suunas asustama germaani hõimud - anglid ja saksid (anglosaksid). 500. aastal oli tänapäeva Wales jaotatud mitmeteks kuningriikideks. Kogu anglosakside territoorium aga kandis nime *Angle-land*, mida aja jooksul hakati tarvitama kogu maa kohta - *England* (King, 2008). Lisaks poliitilisele, keelelisele ja kultuurilisele eraldatusele, eraldas Anglo-Saksi ja Wales'i kuningriike ka ~240 km pikkuse kraaviga tähistatud piir (*Offa's Dyke*). Perioodi Euroopa ajaloos, mis kestis hilisest 8.saj-st m.a.j. kuni varajase 11.saj-ni m.a.j. on üldistavalt iseloomustatud kui viikingite aega. Viikingid ründasid Skandinaavia ida-, lõuna- ja läänealadelt kogu Mandri-Euroopa läänerrannikut. Umbes aastast 800 said viikingite rüüsteretked Lääne-Euroopas tavaliseks nähtuseks. Kõige esimesena (780 m.a.j.) koloniseerisid nad Põhja-Atlandi saared - Orkney, Shetlandi, Šotimaa Läänesaared ning Fääri saared - ning 793. aastal ründasid viikingid Wales'i Anglesey saart. Viimasena vallutati Islandi saar. Erinevalt Britanniast ja Iirimaa rannikupaikadest (nagu nt. Dublin), pole Wales'ist leitud jälgi (võimubaasilaadsed kohad ja rikkalike panustega hauad) viikingite pikemaajalisest asustusest. Üksikud viikingitega seostatud esemelised leiud ja hauad on leitud Põhja-Walesist ja Anglesey saarelt. Wales'i asukoht, sealsete brittide tugev kogukond ning lisaks sellele ka hiljem sõlmitud vaherahu Taani ja Norra viikingitega aitas vastu seista jätkuvatele anglosakside rünnakutele, mistõttu püsis Wales'i riik iseseisvana veel keskajalgi (Roesdahl, 2007; Hall, 2008; Stevenson, 2003), kuni 1536. aastani, mil Wales täielikult Inglismaa võimule allutati (Teepel, 2004). Arvatakse, et seoses südi vastupanuga, mida waleslased erinevatel ajalooperioodidel vallutajate vastu osutasid, on nende puhul tegemist ühtede „vanimate“ Briti saarte elanikega, kes kannavad endas saarestiku algupärast geneetilist pärandit (Keeling, 2013).

1.4.2 Briti saarte ja Wales'i geneetilisest mitmekesisusest

Suurbritannia ja Põhja-Atlandi saared on ühed Euroopa neist viimastest piirkondadest, kuhu kaasaegne inimene pärast jääaega jõudis. Suurbritannia ja Põhja-Atlandi saarte populatsioone ja nende ema- ja isaliine on käsitletud põhiliselt sealse piirkonna vallutamise

kontekstis. Lähtepopulatsioonide kandidaatiteks on väljapakutud mitmeid erinevaid: anglosaksi eelsed britid ehk rooma-britid, anglosaksid, skandinaavlased (Wilson *et al.*, 2001). Ajaloolased on arvamusel, et rohked ekspansioonid ja vallutused, mis ajalooliselt Briti saartele osaks said (nt. skandinaavlaste vallutused), on enamjaolt olnud meessugu hõlmavad (Helgason *et al.*, 2001 viited seal sees; Lacan *et al.*, 2011b). Seetõttu on ka selle piirkonna populatsioonigeneetiliste tööde tähelepanu olnud keskendunud valdavalt just meesliinide (Y kromosoomi) uurimisele (Wilson *et al.*, 2001; Capelli *et al.*, 2003; Helgason *et al.*, 2000; Weale *et al.*, 2002;). Wales'i emaliine on seni põgusalt käsitletud üle-euroopalises kontekstis (Richards *et al.*, 1996). Hiljuti on varasematele soo-spetsiifilistele uuringutele lisandunud ka mõned autosoomsete andmete analüüsil põhinevad tööd, mis kirjeldavad Suurbritannia, Iirimaa ja sealsete väikesaarte populatsioonide struktuuri (Winney *et al.*, 2012; O'Dushlaine *et al.*, 2010).

Tööde tulemused, kus uuriti Skandinaavia populatsioonide ja keltide (Šotimaa sisemaa ja Iirimaa sega- populatsioon) emaliinide proportsioone tänapäeva Põhja-Atlandi piirkonna (Orkney, Island, Skye saarte, Šotimaa Läänesaarte ja looderanniku) populatsioonide geenitiigis, näitasid üldiselt head korrelatsiooni geograafiaga. Leiti, et Skandinaaviast kaugenedes suureneb Põhja-Atlandi populatsioonides briti ja väheneb skandinaavia päritoluga naisliinide osakaal (Helgason *et al.*, 2001). Kõik Põhja-Atlandi saarte populatsioonid sarnanevad oma emaliinide poolest rohkem keltidega ja vähem skandinaavlastega. See sarnasus oli pisut madalam islandlastel ning Orkney saarte elanikel, kuid siiski sarnanevad ka nende emaliinid oluliselt rohkem keltide kui skandinaavlaste omadele (Helgason *et al.*, 2001).

Skandinaavia päritolu emaliinide ja isaliinide proportsioonid Shetlandi (vastavalt 44,5% ja 43%), Orkney (vastavalt 30,5% ja 31%) ja Šotimaa looderanniku (vastavalt 14,5% ja 15%) populatsioonide geenitiigis on peaaegu identsed. Selline sümmeetria viitab sellele, et Shetlandi, Orkney ja Šotimaa looderanniku asutajapopulatsioonides esines umbes võrdne arv skandinaavia mehi ja naisi ehk siis teisisõnu toimus perepõhine asustatamine (Goodacre *et al.* 2005). Seevastu skandinaavia emaliinide hinnanguline panus (11%) Šotimaa Läänesaarte ja Skye geenitiiki moodustab ligikaudu poole sealsete Skandinaavia meesliinide panusest (22,5%). Selline lahknevus on veelgi tugevamalt esindatud islandlastel (vastavalt 34% ja 75%), mida on seletatud teooriaga, et Islandit asustanud efektiivne populatsioon koosnes ilmselt suuresti Briti/Iiri sünnipäraga naistest ja Skandinaavia päritolu meestest (Helgason *et al.* 2000, 2001).

Wilson *et al.*, (2001) näitasid mitmete Lääne-Euroopa populatsioonide (waleslased, baskid, Orkney saare elanikud, iirlased, friislased, norralased) mtDNA analüüside alusel, et

kõikidel uuritud populatsioonidel on sarnane mtDNA haplogrupiline jaotus. Põhikomponentanalüüsis eristusid teistest kõige selgemini vaid baskid. Autorid järeldasid, et kõikide teiste populatsioonide vahel on emaliinide geenivool olnud intensiivne ja see on viinud varieeruvuse homogeniseerumisele. Isaliinide andmed näitasid, et waleslased, iirlased ja baskid on isaliinide haplotüübilise mitmekesisuse ja sageduse poolest väga sarnased ning pärinevad arvatavalt ühise päritoluga isaliinide eellaspopulatsioonist ning et skandinaavlased on jätnud märkimisväärse geneetilise jälje just Orkney saarte populatsiooni Y-kromosoomi varieeruvusse (Wilson et al., 2001). Sellisele tulemusele jõudsid oma töös ka Capelli et al., (2003), kes leidsid sama ka Shetlandi isaliinide kohta. Kesk-Inglismaa ja Põhja-Wales'i populatsioonide Y-kromosoomide analüüs ajaloost teadaolevate Norra viikingite ja Friisimaa anglosakside vallutuste kontekstis näitas, et anglosakside isaliinide märkimisväärne panus kajastub küll uuritud Kesk-Inglismaa piirkonna isaliinide varieeruvuses, kuid mitte Põhja-Wales'i populatsiooni omas (Weale et al., 2002). See on kooskõlas ajaloolisest andmestikust teadaoleva pikaajalise konfliktiperioodiga Anglo-Saksi ja Wales'i kuningriikide vahel. Geneetiliste uuringute analüüside tulemused annavad alust arvata, et geenivool nende kahe regiooni vahel on olnud pika aja [umbes 1500 (vahemikul 500-2000) aasta)] jooksul piiratud. Kui ajalooliste ja arheoloogiliste andmete alusel on leidnud poolehoidu teooria mõnede Suurbritannia osade võimaliku asustamise kohta viikingite poolt, siis geneetilised andmed kohalike populatsioonide märkimisväärset väljavahetamist ei toeta (Richards et al., 2000; Weale et al., 2002).

Vana DNA uurimustöös (Töpf et al., 2007), kus uuriti ajalooliste (vanusega ligikaudu 300 -1000 m.a.j, roomlaste ja anglosakside vallutuste periood) brittide emaliine (48 mtDNA-d) ja võrreldi saadud tulemusi nii kaasaegse Inglismaa kui ka Euroopa populatsioonide emaliinidega, tuvastati vana DNA proovidel kaasaegsele Euroopa geenitiigile iseloomulik haplotüübilise mitmekesisuse muster. Esindatud olid kõik kaasaegses Euroopas leiduvate emaliinide haplogrupid (U5a, U5b, J, T1, T2, V, K, W, I). Autorid näitasid, et ajalooliste brittide DNA proovide emaliinide geneetiline mitmekesisus on isegi suurem kui kaasaegsetel, mis viitab kas valiku ja/või siis geenitriivi mõjul toimunud diversiteedi vähenemisele viimase aastatuhandete vältel. Oma osa selles võis olla ka katkuepidemiatega 14. ja 17. sajandil m.a.j. (Töpf et al., 2007).

Ülegenoomse varieeruvuse uuringud, mis põhinesid ~350 000 SNP genotüüpiseerimisandmestikul ja suurel valimil (3654 indiviidi, Iirimaa, Šotimaa, Inglismaa, Bulgaaria, Portugali, Rootsi ja Euroopa päritoluga Utah ameeriklased), näitasid, et Suurbritannia ja Iirimaa haplotüübiline mitmekesisus on Mandri-Euroopa populatsioonidega

võrreldes madalam, seda iseäranis Šotimaal ja Irimaal, vähem Lõuna-Inglismaal - üle Euroopa kagust loodesse ulatuv haplotüübilise mitmekesisuse kahanev kliin laieneb ka Briti saartele (O'Dushlaine et al., 2010). Autorid leidsid, et viimased tulemused on kooskõlas varasema kogu genoomi haplotüübilise mitmekesisuse uuringuga (Auton et al., 2009) ning Y-kromosoomi (Semino et al., 2000) kui ka mtDNA (Richards et al., 2002) üldiste levikumustritega ning toetavad teadaolevat, summaarsetel alleelisagedustel põhinevat kliini, mida populatsioonigeneetilistes töödes on varasemalt seostatud Lähis-Idast lähtuva neoliitiliste põlluharijate levikuga üle Euroopa (Ammerman & Cavalli-Sforza, 1984; Chikhi et al., 2002).

2004 aastal alguse saanud projekti raames *"People of the British Isles"* (PoBI), mis uurib Briti saare populatsioonide geneetilist struktureeritust, oli 2011 aastaks ~4 000 indiviidi DNA proovi genotüpiseerimisandmestiku alusel loodud mahukas andmebaas (<http://www.peopleofthebritishisles.org>). Projekti eesmärgiks on kasutada neid andmeid nii brittide päritolu paremaks mõistmiseks, kuid ka haiguste assotsiatsiooniuringutes kontrollpopulatsiooni andmestikuna. Pilootturing, mis viidi läbi põhiliselt geograafilise jaotuse ja ajalooliste info baasil välja valitud 1057 DNA proovi ja 9 geneetilist markerit kasutades, ei tuvastanud Suurbritannia erinevate regioonide vahel olulisi erinevusi. Põhjusena toodi välja uuringus kasutatud proovide ja markerite vähesus (Winney et al., 2012). Sama projekti raames on autorid lubanud peagi avaldada autosomaalsete markerite analüüsi alusel detailse kaardi Suurbritannia eri osade varieeruvuse kohta (<http://www.peopleofthebritishisles.org/nl5.pdf>).

2. EKSPERIMENTAALNE OSA

2.1 Töö eesmärgid

Käesoleva magistritöö laiemaks eesmärgiks oli heita valgust waleslaste populatsiooni kujunemisloole ning kohale tänase Lääne-Euroopa geneetilisel maastikul. Töö kitsamateks eesmärkideks olid:

1. Uurida waleslaste emaliinide mitmekesisust mtDNA varieeruvuse alusel;
2. Analüüsida populatsiooni geneetilist struktuuri autosomaalse varieeruvuse alusel;
3. Analüüsida saadud tulemusi Lääne-Euraasia populatsioonide geneetilise varieeruvuse taustal.

2.2 Materjal ja meetodika

Wales'i populatsiooni geneetilise mitmekesisuse kirjeldamiseks kasutati mtDNA ja üle kogu genoomi jaotunud 660 000 SNP genotüpiseerimisandmestiku analüüsi.

2.2.1 Katsematerjal

Käesolevas töös kasutatud 176 waleslase DNA proovid on kogutud Suurbritannia Loughborough Ülikoolis töötava Dr. Sarabit S. Mastana ja tema üliõpilaste poolt 2003 aastal ja olnud alates sellest ajast Eesti Biokeskuse (EBK) biopangas. Kõikidel Wales'i proovidel oli eelnevalt EBK teaduri Erwan Pennaruni poolt sekveneeritud mtDNA HVS-I piirkond, tema poolt oli antud ka esialgne hinnang proovide haplogrupilise kuuluvuse kohta. Autosomaalsete markerite analüüsiks kasutatud 24 populatsiooni 418 (sh. 20 Wales'i) indiviidi olid eelnevalt EBK tuumiklaboris Illumina 610K/650K/660K SNP kiibiga (*Human 660W-Quad v1.0 DNA Analysis BeadChip Kit, Illumina, Inc.*) genotüpiseeritud. Võrdlusena kasutatud populatsioonide valik teostati avaldatud autosomaalsete andmete seast alusel, et see kataks võimalikult hästi Lääne-Euroopale iseloomuliku geneetilise varieeruvuse mustri.

2.2.2 MtDNA analüüs

Kasutades informatsiooni mtDNA kodeeriva regiooni defineerivate np-de kohta (www.phylotree.org) analüüsiti töös kõiki mtDNA-sid kindlate haplogruppide ja nende alamharude markerite suhtes. Kasutati nn. hierarhilist lähenemist, mille korral analüüsitakse mitte kõiki teadaolevaid markereid, vaid lähtuvalt HVS-I järjestusest ja haplogruppide Euroopa-sisese sagedusjaotusest alustatakse kõige tõenäolisemalt haplogrupilist kuuluvust näitavate markerite genotüpiseerimisega. Kuna hg H on üks sagedasemaid haplogruppe

Euroopas (keskmiselt 45%), siis alustati just selle haplogrupi tüpiseerimisega. Esmalt kasutati selleks restriksioonanalüüsi (RFLP) (vt. tabel 3), mille puhul mutatsioone detekteeritakse teatud restriiktaasi äratundmissaidi tekke või kao kaudu. RFLP meetodiga kontrolliti 102 DNA proovi kuuluvust hg H. Analüüsi tulemused on ära toodud lisa 2. Kogu edasine mtDNA haplotüüpide klassifitseerimine toimus sekveneerimise teel, kus mtDNA nukleotiidses järjestuses varieeruvuse kindlakstegemiseks kasutati mtDNA osalist või täielikku resekveneerimist. Viimast valikut otsustati kasutada haplogrupi H sisestruktuuri täpsustamisel sõelale jäänud 9 mtDNA puhul (tabel 2), mille täpse asukoha leidmiseks fülogeneetilisel puul ei piisanud HVS-I järjestuse ja hg H defineerivate põhiharude määramisest. Lisaks neile, esines antud valimis kahel proovil (L76, BK142; LISA 4d ja e) ühesugne HVS-I motiiv, mis HVS-I järjestuste andmebaasis seni puudus ning mille alusel polnud võimalik määrata nende asukohta globaalsel fülogeneesipuul, mistõttu otsustati ka ühel neist mtDNA täielikult sekveneerida (tabel 2).

Tabel 2. Töös kasutatud waleslaste DNA proovid, mille mtDNA sekveneeriti täielikult.

Proovi ID	Hg	HVS-I*
BK46	H	16256 16325
BK49	H	16172
BK88	H	16129 16316
BK102	H	16192
BK134	H	CRS
CHF13	H	CRS
CHF121	H	16093
DR161	H	16263
L44	H	16291
L76	LMN?	16039 16188 16189 16223 16290 16319 16356 16362

* HVS-I mutatsioonid on antud rCRS suhtes

MtDNA täisjärjestused sekveneeriti osade kaupa, 800 - 1 000 bp pikkuste fragmentidena. Kõigepealt toimus fragmentide amplifitseerimine polümeraasi ahelreaktsiooniga (PCR), seejärel nende resekveneerimine ja andmeanalüüs.

2.2.2.1 DNA amplifitseerimine

MtDNA amplifitseerimiseks PCR-i abil kasutati kirjanduses avaldatud primereid (Rieder et al., 1998) (vt. LISA 1). DNA amplifitseerimine toimus kahe alljärgnava (A ja B) protokollil alusel, olenevalt amplifitseeritava DNA produkti suurusel. Et PCR reaktsiooni produkti jätkuks kogu piirkonna (ja/või kordus-) sekveneerimiseks, valiti reaktsioonimahuks vastavalt 15 µl ja 25 µl. Protokollis on komponentide juures olevates sulgudes välja toodud tüvilahuse alg- ja lõppkontsentratsioon.

A) Kuni 1500 ap amplifitseeritava piirkonna puhul olid PCR-i reaktsiooni komponendid ja amplifikatsiooni tingimused järg-mised:

5x FIREPol® Master Mix (*Solis BioDyne*) [FIREPol® DNA polümeraas, 5x Reaktsiooni Puhver B (0,4 M Tris-HCl, 0,1 M (NH₄)₂SO₄, 0,1% w/v Tween-20); 12,5 mM MgCl₂ (1x PCR lahus – 2,5 mM MgCl₂), 2 mM dNTP-d (1x PCR lahus – 200 μM dATP, 200 μM dCTP, 200 μM dGTP ja 200 μM dTTP), sinine ja kollane indikaatorvärv]

3 μl (5X) (1X)

- *Forward* praimer
0,3 μl (10 pmol/μl) (0,2 μM)
- *Reverse* praimer
0,3 μl (10 pmol/μl) (0,2 μM)
- DNA
1 μl
- MilliQ H₂O
10,4 μl

Amplifitseerimisprogramm:

1. Esmane denaturatsioon 94°C, 1 min
30sek
2. Denaturatsioon 94°C, 15 sek
3. Praimerite seondumine 56°C, 15 sek
4. Süntees 72°C, 20 sek
5. Lõplik süntees 72°C, 1 min

30- 33 tsüklit (vahemikus 2. – 4.)

B) Alates 1500 ap amplifitseeritava piirkonna puhul olid PCR-i reaktsiooni komponendid ja amplifikatsiooni tingimused järg-mised:

- *10x Long PCR Buffer* (*Thermo Scientific*) 2,5 μl (10X) (1X)
- MgCl₂
2,5 μl (25mM) (2,5mM)
- dNTP Mix
0,25 μl (10mM) (0,1mM)
- *Forward* praimer
0,5 μl (10 pmol/ μl) (0,2 μM)
- *Reverse* praimer
0,5 μl (10 pmol/ μl) (0,2 μM)
- *Long PCR Enzyme Mix* (*Thermo Scientific*)
0,25 μl (5 U/ μl) (0,05 U/ μl)
- DNA
2 μl
- MilliQ H₂O
16,5 μl

Amplifitseerimisprogramm:

1. Esmane denaturatsioon 94°C, 1min
30sek
2. Denaturatsioon 94°C, 15 sek
3. Praimerite seondumine 56°C, 30 sek
4. Süntees 68°C, 1000 ap
= 1min (vastavalt amplifitseeritava produkti suurusele)
5. Lõplik süntees 68°C, 1000 ap
= 1min + 1min (vastavalt amplifitseeritava produkti suurusele)

36-38 tsüklit (vahemikus 2. – 4.)

Kõik PCR-i reaktsioonid viidi läbi kaanesoojendusega termoblokis “Biometra UNO II”. Igas reaktsiooniseerias oli alati negatiivne kontroll - reaktsioonisegu ilma DNA-ta, et kontrollida kasutatud reagentide puhtust. Saadud PCR-i produktide olemasolu kontrolliti agroosgeelelektrofooresil 0,5 kordses TBE puhvril (45mM Tris-boraat ja 0,05M EDTA-Na₂). Olenevalt amplifitseeritud DNA pikkusest kasutati erineva kontsentratsiooniga agarosgeeli (1,2 - 2%). DNA visualiseerimiseks ultraviolet (UV)-valguses lisati geelile etiidumbromiidi lõppkontsentratsiooniga 0,5 µg/ml. PCR-i produkti kanti geelile 4 µl. Kasutatud pinge oli 180V, elektroodide distants 32 cm. Tulemused jäädvustati UV-valguses 260 nm juures *UVIPro Gold* masinaga *UVI Pro v12.5* tarkvara abil.

2.2.2.2 RFLP analüüs

Restriktaasid on endonukleaase aktiivsusega ensüümid, mis lõikavad kaheaheelist DNA-d teatud kindla järjestusmotiivi esinedes. Mutatsioonid selles järjestuses võivad põhjustada restriktioonisaidi kadumise või tekke.

Tabel 3. MtDNA haplogrupi H määramine restriktioonianalüüsil. Tabelis olevad tähistused F ja R tähendavalt vastavalt kerge ja raske ahela primereid; +/- tähistavad vastavalt restriktioonisaidi teket või kadumist. Positsioon 2078 on tähistatud punasega.

Haplo-grupp	Positsioon	Definieeriv mutatsioon*	Praimerid**		Restriktiooni ensüümi äratundmis-motiiv ja lõikekoht ()	PCR produkti pikkus, ap	RFLP tulemus		Reaktsiooni-puhver***
H	7028	C→T	5' AAGCAATATGAAATGATCTG 3'	F	5'...A G C T...3'	242	+	137+30+75	Tango
			5' CGTAGGTTTGGTCTAGG 3'	R			-	167+75	

* Referetsjärjestuseks on rCRS

** Torroni et al., 1996

*** Nii ensüüm (*AluI*) kui ka Tango puhver on toodetud Fermentas (Thermo Scientific) poolt.

Restriktioonireaktsiooniks lisati 25 µl-le PCR-i produktile vastavat restriktiooniensüümi 0,05 µl (0,5 U) (*Fermentas*) ja 2,5 µl sellele sobivat 10x reaktsioonipuhvrit (tabel 3). Restriktioonireaktsioon toimus 37°C juures minimaalselt kaks tundi kuni üleöö. Restriktioonifragmendid eraldati 2% agroosgeel-elektrofooresil. Tulemused jäädvustati UV-valguses 260 nm juures *UVIPro Gold* masinaga *UVI Pro v12.5* tarkvara abil.

2.2.2.3 Sekveneerimine

MtDNA järjestuse määramiseks kasutati Sangeri didesoksüterminatsiooni meetodit, mille puhul DNA sünteesitakse *in vitro* DNA polümeraasi poolt, mis kasutab matriitsina üheaheelist DNA-d (Sanger et al., 1977). Reaktsioonisegu sisaldab kindlas kontsentratsioonis

didesoksünukleotiidtrifosfaate (ddNTP-e), mille kaasamisel sünteesireaktsiooni lõppeb konkreetse DNA ahela süntees. Eri pikkusega DNA fragmentide detekteerimiseks on ddNTP-d seotud nelja erineva fluorestseeruva värviga. Sekveneerimine viidi läbi *Applied Biosystems BigDye® Terminator v3.1 Cycle Sequencing Kit*'i *ABI PRISM® 3730xl DNA 20* analüsaatoril (*Applied Biosystems*).

Sekveneerimisreaktsiooni komponendid (10 µl) :

- | | |
|---|---------|
| • MilliQ H ₂ O | 4,65 µl |
| • Big Dye® Terminator v3.1 reagentide eelsegu (<i>Applied Biosystems</i>) | 0,75 µl |
| • Big Dye® Terminator v3.1 5x sekveneerimispuhver (<i>Applied Biosystems</i>) | 2 µl |
| • Sekveneerimispraimer (1 pmol/ µl) | 1,6 µl |
| • PCR-i produkt | 1 µl |

Sekveneerimistingimused:

- | | | |
|-------------------------|--------------|----------|
| • Esmane denaturatsioon | 94°C, 10 sek | } tsükli |
| • Denaturatsioon | 94°C, 20 sek | |
| • Praimerite seondumine | 52°C, 15 sek | |
| • Süntees | 60°C, 1 min | |

DNA sadestamiseks lisati 10 µl sekveneerimisreaktsiooni produktile:

- 2 µl NH₄Ac/dekstraan segu (700kDa punane dekstraan 10mg/ml; 7,5 M ammooniumatsetaat) ning 30 µl 96% (-20°C) etanooli
- hoiti -20°C juures umbes 10 min
- tsentrifuugiti 15 min 13000 rpm ($g=18\ 890$, *Hettich Zentrifugen MIKRO 22*)
- eemaldati supernatant
- pesti 200 µl külma 70% (-20°C) etanooliga
- tsentrifuugiti 10 min 13000 rpm ($g=18\ 890$, *Hettich Zentrifugen MIKRO 22*)
- eemaldati supernatant
- pesti uuesti 200 µl külma 70% (-20°C) etanooliga
- tsentrifuugiti 5 min 13000 rpm ($g=18\ 890$, *Hettich Zentrifugen MIKRO 22*)
- eemaldati supernatant
- lasti sademel kuivada +37°C ca 10 min
- sademele lisati 10 µl MegaBACE geelile kandmise lahust (70% formamiid, 1 mmol/l EDTA) (*Amersham Pharmacia Biotech*), lasti toatemperatuuril ca 10 min lahustuda ja tõsteti sekveneerimisplaadile.

2.3 Andmetöötlusmeetodid ja- programmid

2.3.1 MtDNA andmete analüüs

Sekveneeritud mtDNA järjestused joondati referentsjärjestuse rCRS-i suhtes (Andrews et al., 1999) programmi *ChromasPro* (1.7)-ga ning lisaks rCRS-ile ka inimese rekonstrueeritud referentsjärjestuse (RSRS) suhtes (Behar et al., 2012). Selleks kasutati internetiportaali *mtDNACommunity* rakendust FASTmtDNA (<http://www.mtcommunity.org/>).

MtDNA puhul kasutatakse kõige säästlikumate fülogeneesipuude konstrueerimisel enamasti mediaanvõrgustiku meetodit (Bandelt et al., 1999). Olukorras, kus säästlikke võimalusi on enam kui üks, kujutatakse alternatiivseid võimalusi retikulatsioonidena. Nii jääb puu topoloogia kujutamisel alles kogu informatsioon. Paljud retikulatsioonid lahenevad täiendavate uuringute käigus - uute informatiivsete positsioonide kindlakstegemisel. Antud töö mediaanvõrgustikud konstrueeriti programmi NETWORK 4.611 ja selle lisaprogrammi NETWORK PUBLISHER abil (A. Röhl; www.fluxus-engineering.com).

Hg H HVS-I ja kodeeriva regiooni järjestuste alusel koostatud fülogeneetilise puu konstrueerimisel kasutati mediaanvektorite ühendamise (*median joining*) algoritmi (Bandelt et al., 1995; Bandelt et al., 1999). Mediaanvõrgustiku arvutamisel anti np-dele mutatsioonikiiruse järgi kaalud (Soares et al., 2009) ja jagati need nelja klassi: kiiresti muteeruvatele np-dele (>30 esinemist) anti kaaluks üks, keskmise mutatsioonikiirusega np-dele (10-30 esinemist) kaks ning aeglaselt muteeruvatele ja individuaalsetele np-dele (0-10 esinemist) kolm. Transversioonidele ja teadaolevatele hg H alamhaplogruppe defineerivatele mutatsioonidele määrati kaaluks kümme (vt. LISA 3a).

Hg-de T ja J HVS-I ja kodeeriva regiooni järjestuste alusel koostatud mediaanvõrgustike konstrueerimisel kasutati taandatud mediaanvektorite (*reduced median*) algoritmi (Bandelt et al., 1995; Bandelt et al., 1999). Nukleotiidipositsioonid jagati mutatsioonikiiruse alusel kolme klassi: kiiresti muteeruvad np-d (>30 esinemist) said kaaluks ühe, keskmise või aeglase mutatsioonikiirusega ja individuaalsed np-d (0-30 esinemist) kolme ning teadaolevad hg J ja T defineerivad mutatsioonide said kaaluks kümme (Soares et al., 2009) (vt. LISA 3b ja 3c).

Hg H täisjärjestuste fülogeneesipuu (joonis 15) koostati käsitsi, järgides suurima säästlikkuse printsiipi ja kasutades *mtDNACommunity* ülemaailmse mtDNA fülogeneesipuu täisjärjestuste andmeid (<http://www.mtcommunity.org>) kooskõlas käibeloleva nomenklatuuriga (www.phylotree.org).

MtDNA haplogruppide sagedustel põhinev põhikomponentanalüüs (PCA) teostati programmiga POPSTR (<http://harpending.humanevo.utah.edu/popstr/>). MtDNA haplogruppide sageduste 95%-lised usalduspiirid arvutati dr. Vincent Macaulay poolt koostatud programmi SAMPLING abil.

2.2.2 Autosoomsete markerite analüüs

Töö teine osa hõlmas Wales'i populatsiooni üle kogu genoomi jaotunud 660 000 SNP genotüpiseerimisandmestiku analüüsi Lääne-Euraasia kontekstis. Üle kogu genoomi jaotunud autosoomsete SNP-de genotüübiandmete analüüsiks väljavalitud 418 indiviidist (mis kuulusid 24 populatsiooni) (vt. tabel 4) koosnevat valimit kontrolliti esmalt võimalike duplikaatide ja suguluses olevate indiviidide välistamiseks tarkvaraga KING (version 1.4) (Manichaikul et al., 2010). Pärast kahe esimese ja teise astme suguluses olevate indiviidide eemaldamist jäi edasisse analüüsi 416 proovi. Tarkvara PLINK 1.05 (Purcell et al., 2007) abil filtreeriti valimi andmekogumit nii, et see sisaldaks genotüübiinfot 22 autosomaalse kromosoomi kohta ning võtaks arvesse info ainult nende SNP-de kohta, mille minoorse alleeli sagedus on $> 1\%$ ja nende indiviidide kohta, kelle genotüpiseerimise edukus SNP-de osas oli $> 97\%$. Pärast kvaliteedikontrolli jäi kasutusse 507 734 SNP-d. LD mõju arvestamiseks kärbiti andmekogumit aheldunud SNP-de osas, eemaldades ühe SNP tugevas LD-s (korrelatsioon $r^2 > 0,4$) olevast SNP-de paarist. Selleks vaadeldi SNP-sid 200 SNP pikkuses lõigus 25 SNP kaupa liikuvast genoomses "aknas". Lõplik valim sisaldas 199 782 SNP-d.

Populatsiooni struktureerituse määramiseks kasutati iga uuritava indiviidi geneetilises profiilis eellaskomponentide osakaalu kindlaksmääramisel põhinevat ja suurima tõepära meetodit kasutavat analüüsi ADMIXTURE, (Alexander et al., 2009). Põhikomponentanalüüs viidi läbi tarkvaraprogrammiga EIGENSOFT, (Patterson et al., 2006). Arvutused tehti Tartu Ülikooli teadusarvutuste keskuse arvutifarmis. Programmide käivitamiseks ja tulemuste analüüsiks vajalikud skriptid on koostatud EBK-s dr. Mait Metspalu poolt. Tarkvaraprogramm ADMIXTURE võimaldab eellaskomponentide osakaalu järgi uuritud indiviide (ja populatsioone) omavahel võrrelda. Potentsiaalsete eellaskomponentide (K) arv sisestatakse algoritmi käivitamisel. Siin kasutati selleks eellaskomponentide arvu vahemikus $K=2$ kuni $K=15$. Tulemuste statistiliseks hindamiseks korrati analüüsi iga K puhul 100 korda. Madala K väärtuse juures saavutasid kõik sada arvutust väga sarnase *LogLikelihood* skoori (LLs), kõrgemate K väärtuste juures olid LL-id varieeruvad (LISA 6). Kõige parema ennustusvõimega loetakse olevat selline K mudel, mille LLs on võimalikult kõrge ja mille 100 erineva analüüsi hajuvus on võimalikult väike.

Tabel 4. Autosoomsete andmete analüüs (ADMIXTURE, PCA) kasutatud valimi iseloomustus.

Populatsioon	Indiviidide arv	Riik	Kirjanduse allikas	Genotüüpiseeritud SNP-de arv
Orkney elanikud	15	Suurbritannia	Li et al., 2008	650 000
waleslased	20	Suurbritannia	Avaldamata andmed	660 000
hispaanlased	12	Hispaania	Behar et al., 2010	610 000
prantslased	28	Prantsusmaa	Li et al., 2008	650 000
Pranstuse baskid	24	Prantsusmaa	Li et al., 2008	650 000
sardid	28	Itaalia	Li et al., 2008	650 000
põhjaitaallased	13	Itaalia	Li et al., 2008	650 000
toskaanlased	8	Itaalia	Li et al., 2008	650 000
küproslased	12	Küpros	Behar et al., 2010	660 000
bulgaarlased	13	Bulgaaria	Yunusbayev et al., 2012	610 000
ungarlased	19	Ungari	Behar et al., 2010	610 000
rumeenlased	16	Rumeenia	Behar et al., 2010	660 000
ukrainlased	20	Ukraina	Yunusbayev et al., 2012	660 000
valgevenelased	9	Valgevene	Behar et al., 2010	610 000
venelased	2	Eesti	Behar et al., 2010	650 000
	25	Venemaa	Li et al., 2008	660 000
leedulased	10	Leedu	Behar et al., 2010	610 000
mordvalased	15	Venemaa	Yunusbayev et al., 2012	610 000
tšuvašid	17	Venemaa	Behar et al., 2010	610 000
marokolased	9	Maroko	Behar et al., 2010	660 000
armeenlased	35	Armeenia	Behar et al., 2010	610 000
lesgiinid	18	Venemaa	Behar et al., 2010	610 000
jordaanlased	20	Jordania	Behar et al., 2010	660 000
türklased	19	Türgi	Behar et al., 2010	610 000
mongolid	9	Mongoolia	Rasmussen et al., 2010	660 000

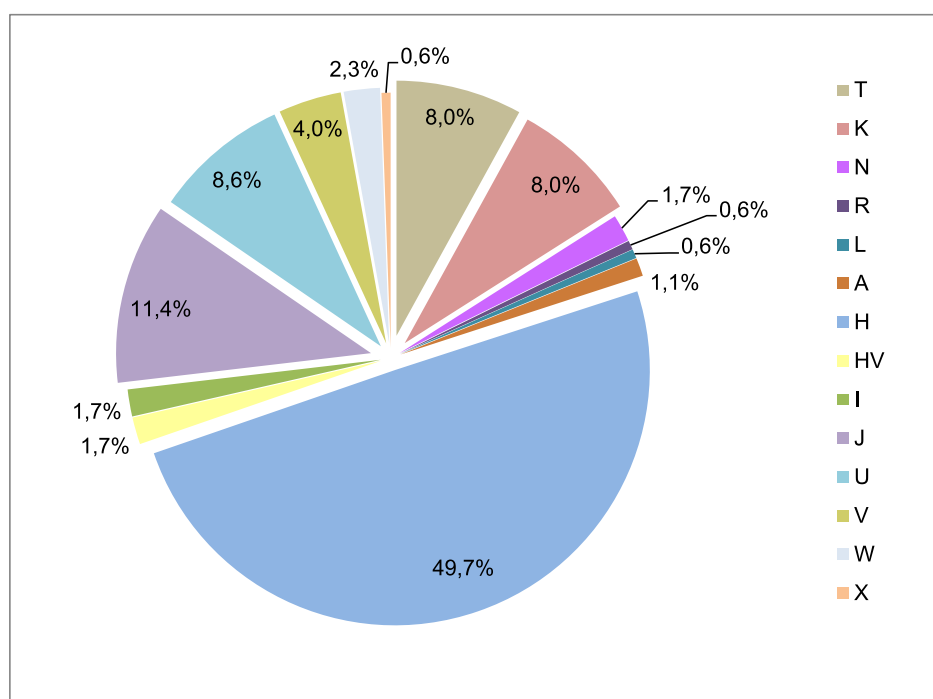
3. TULEMUSED JA ARUTELU

3.1 Walesi emaliinide analüüs

3.1.1 MtDNA üldine varieeruvus Wales`i populatsioonis

MtDNA haplogrupeerimine põhines pikka aega peamiselt mtDNA mittekodeeriva piirkonna polümorfsusel ja üksikutel kodeeriva regiooni defineerivate np-de seisundi määramisel. Siinsesse valimisse kuulunud 176 waleslase mtDNA-dele oli eelnevalt HVS-I piirkonna alusel antud esialgne hinnang nende haplogrupilise kuuluvuse kohta, kuid kodeeriva regioonist saadava info puudumise tõttu ei paigutunud paljud mtDNA haplotüübid ühtegi kindlasse haplogruppi ega alamhaplogruppi.

MtDNA HVS-I piirkonna ning kodeeriva regiooni defineerivate np-de sekveneerimistulemuste analüüsi põhjal määrati valimi kõigi mtDNA-de haplogrupiline jaotuvus (joonis 9). Kõik genotüpiseerimise andmed on antud rCRS-i (Andrews et al., 1999) ja RSRS (Behar et al., 2012) suhtes LISA 4-s.

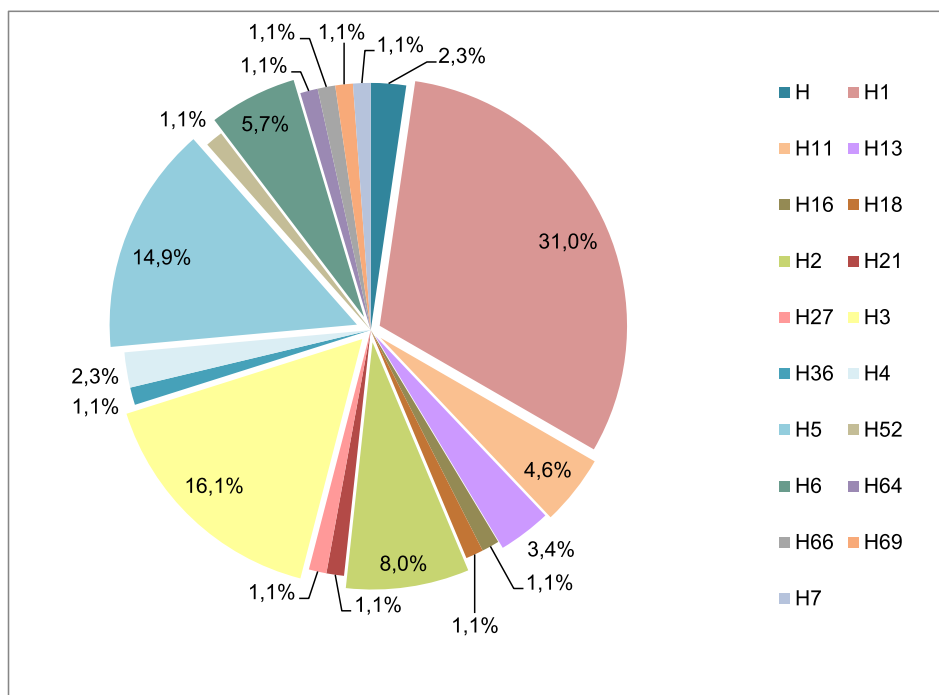


Joonis 9. Wales`i populatsiooni mtDNA (n=176) haplogrupiline jaotuvus.

Antud töös uuritud waleslaste valimi emaliine iseloomustab Euroopa populatsioonidele tüüpiline haplogrupiline jaotuvus, milles on esindatud valdav enamus Lääne-Euraasiale iseloomulikke mtDNA haplogruppe. 176 waleslaste emaliinid jaotusid 14 erineva mtDNA haplogrupi vahel (joonis 9). Kõige sagedasem oli hg H, kuhu kuulusid

ligikaudu pooled waleslaste emaliinidest (49,7%). Suhteliselt sagedased olid ka hg U (16,6%) ja selle üks alamharu K (8% - koguvallimist), lisaks hg J (11,4%) ning hg T (8%). Hg V hõlmas 4% ja hg W 2,3% ning kõik ülejäänud hg-d (N, X, II, HV, R1a, L3e, A) moodustasid vallimist < 2%.

Haplogrupp H rohketest alamharudest on waleslaste seas suurim ja mitmekesisem tütarharu H1, mis hõlmab 31% hg H-st ja 15% kogu vallimi emaliinidest (joonis 10). Samasugust jaotuvust on näidatud ka varem Euroopa andmestikku hõlmavates töodes, kus leiti, et H1 moodustab ~13% kogu Euroopa mtDNA geenitiigist ja ~30% hg H sisest varieeruvusest (Loogväli et al., 2004), olles kõrgeim Pürenee poolsaarel (~46%) (Pereira et al., 2004; Quintans et al., 2004). Hg H1 kuulunud 27 haplotüübist kaks kuulusid H1 juurhaplotüüpi (H1ⁿ) (n -, „nodal“), kaks jäid hg H1* tasemele ning ülejäänud jagunesid erinevate hg H alamharude vahel (vt joonis 11).



Joonis 10. Walesi populatsiooni mtDNA hg H (n=87) alamhaplogruppide sagedused.

Hg H2 kõik analüüsitud haplotüübid peale ühe (hg H2a1) kuulusid alamharusse H2a2. Varasemates töodes on näidatud, et H2a (nagu ka H1b) esineb sagedamini Euroopa ida- kui läänepiirkonnas (Loogväli et al., 2004), kuid nn. „St. Luke motiiviga“ 16235-16293 haplotüübid (Vernesi et al. 2001), mis kuuluvad hg H2 on eriti sagedased just Šotimaal ja Saksamaal (Helgason et al., 2001). Selles töös uuritud Walesi hg H2 haplotüüpidest (n=7) esines selline HVS-I motiiv kahel proovil, mille lisaks tuvastatud kodeeriva piirkonna np-de

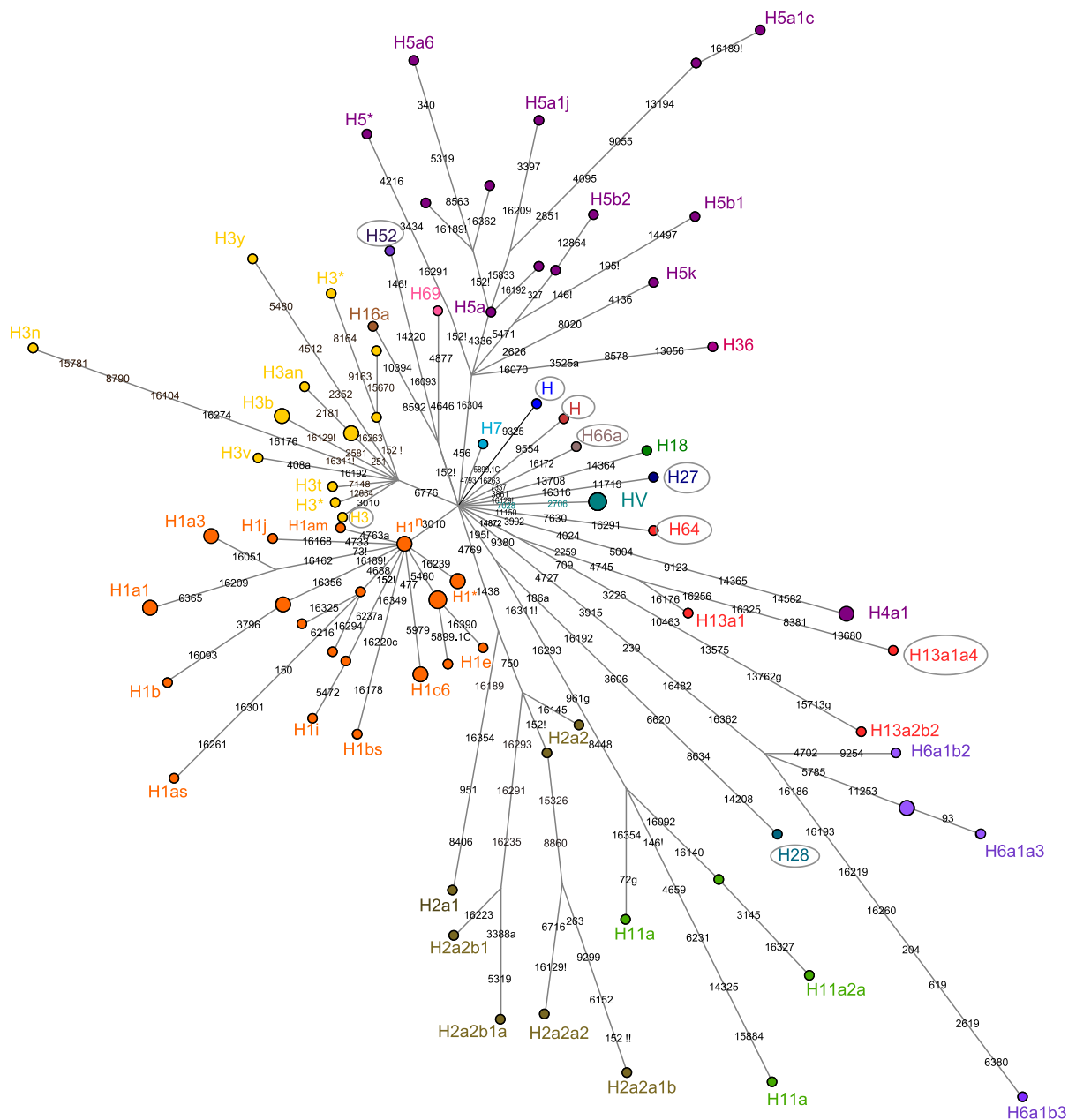
asendused paigutasid need haplotüübid mtDNA täisjärjestuste fülogeneesipuul alamhg-sse H2a2b1 ja H2a2b1a (joonis 11; LISA 4a).

Alamhg-d H3 ja H5 on samuti sagedased hg H harud, hõlmates kokku sama palju uuritud populatsiooni hg H emaliinidest kui hg H1 (vastavalt 16% ja 15%). Antud tulemused on kooskõlas varem avaldatud andmetega, kus on näidatud, et hg H3 on kõige sagedasem Lääne-Euroopas, eriti just Pürenee poolsaarel, kus see hg hõlmab hg H-st ~17% (Pereira et al., 2004; Quintans et al., 2004). Pooled hg H3 haplotüübid paigutusid olemasolevatesse alamhaplogruppidesse (H3an, H3b, H3n, H3t, H3v, H3y) ja pooled hg-i H3* ehk siis antud resolutsiooni juures ei leidnud kohta üheski kindlas hg H3 alamhaplogrupis. Hg H5 13 haplotüübist enamus (8) kuulusid alamhg-i H5a, ning ülejäänud alamhg-sse H5b, H5k ja H5* (joonis 11). Samaselt hg H3 on hg H5 seostatud jääaja järgse ekspansiooniga Edela-Euroopa piirkonnast (Soares et al., 2009) ning selle alamhg-i H5a kõrgemat esinemissagedust on näidatud Euroopa keskosas (Pereira et al., 2005).

Hg H6 ja hg H13 peetakse ühtedeks vanemateks hg H alamhaplogruppideks Kaukaasias ja Lähis-Idas (Roostalu et al., 2007). Waleslaste hg H6 haplotüübid kuulusid kõik H6a1 erinevatesse alamharudesse ning hg H13 emaliinid selle haplogrupi suurimasse alamhg-i H13a (joonis 11). Varasemalt on näidatud, et kõik Euroopas seni uuritud H13 emaliinid (Herrnstadt et al., 2002; Brandstätter et al., 2006) kuuluvad samuti alamharusse H13a (Roostalu et al., 2007).

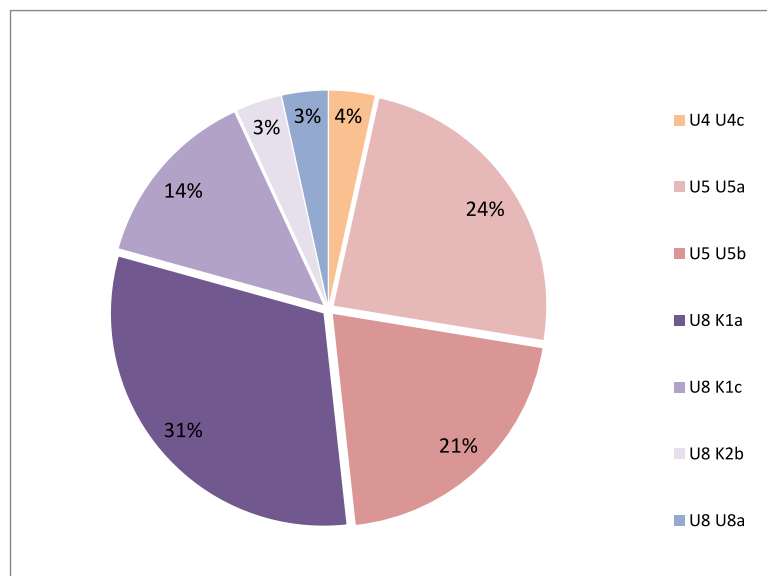
Kesk-Euroopast pärineva luumaterjali DNA andmete põhjal kirjeldatud kunagistest neoliitikumi varase perioodi hg H emaliinivariantidest (Brotherton et al., 2013) esinesid waleslastel tänapäeva Euroopa populatsioonide seas harva esinev alamharu H16. Teistest hg H alamharudest oli esindatud minoorne haru H5b ning sagedased hg H1 alamharud H1j ja H1e, millest viimane oli oma erinevate alamharudega esindatud ka neoliitikumi keskmise ja hilise perioodi mtDNA andmetes. Varasele neoliitikumile järgneva perioodi emaliinidest esinesid veel waleslaste valimis ka H3b, H4a1, H5a, H6a1a, H7, H11a, H13a1a.

Walesi valimis tuvastatud ülejäänud hg H minoorsete alamharude (nt. H18, H28, H64 jne) esinemise kohta Euroopas on oluliselt vähem infot, mis teeb ka nende fülogeograafilise leviku kirjeldamise keeruliseks. Lisaks - kahe (BK134, DR161) (2,3%) hg H kuuluva täielikult sekveneeritud haplotüübi puhul on tegu uute hg H alamharudega (joonis 11; LISA 4e).



Joonis 11. Haplogrupi H mediaanvõrgustik, mis põhineb käesoleva töö Wales'i populatsiooni mtDNA andmetel (n=87). Sõõrid tähistavad proove/indiviide ja nende suurus on kooskõlas haplotüüpi kuuluvate indiidide arvuga. Erinevate värvidega on tähistatud hg H alamharud kuhu antud proovid kuuluvad. 87-st haplotüübist 85 jaotuvad hg H alamhaplogruppide vahel ning ainult kaks järjestust ei kuulunud ühtegi seni defineeritud hg H klaadi (tähistatud kui "H"). H1ⁿ (n -, *nodal*) tähistab juurhaplotüüpi. Välisgrupiks on kolm haplogruppi HV kuulunud indiviidi, mida kasutati puu juurimiseks. Märkide tähendused: "!" – tagasipöördunud mutatsioon; väike täht nukleotiidse positsiooni taga – transversioon (nt. 186a); "d" – deletsioon; ".1C" – insertioon; täielikult sekveneeritud haplotüübid on tähistatud ringiga.

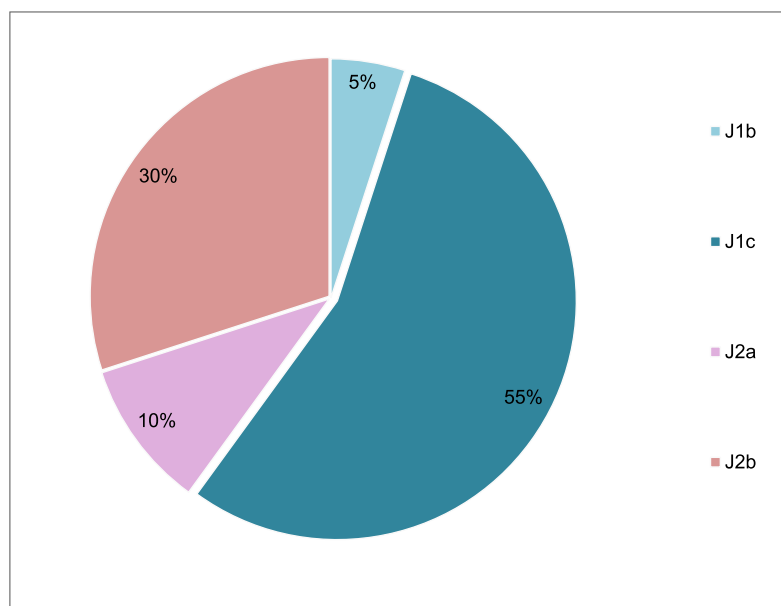
Hg U moodustas uuritud waleslaste emaliinide varieeruvusest 16,6 %, olles seega sageduselt selles valimis teine hg. Hg U sisene mitmekesisus jaotus kolme erineva alamhgi vahel: U8 (~ 52%), U5 (~ 45%) ja U4 (~ 3%).



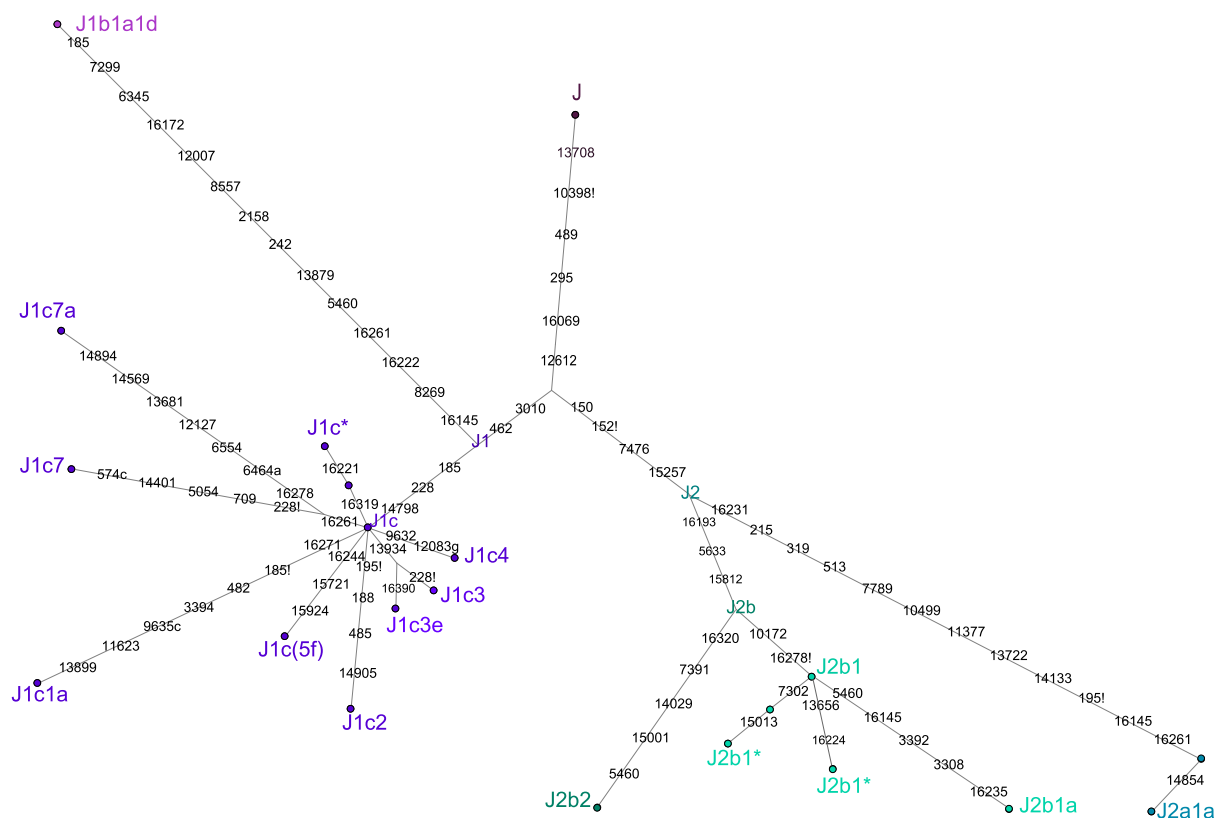
Joonis 12. Wales`i populatsiooni hg U (n=29) alamhgruppide sagedused.

Hg U8 alamharudest domineeris hg K, mis moodustas 8% kogu valimi emaliinidest ja ligi 50% hg U sisemisest varieeruvusest, olles kooskõlas varasemate andmetega, kus on näidatud, et Euroopas moodustab hg K emaliinide mitmekesisusest enamasti vähem kui 10% (Richards et al., 2000; Herrnstadt et al., 2002). Hg K sisemine varieeruvus jagunes põhiliselt selle suurima alamhgi K1 alamharude K1a ja K1c vahel, kusjuures neist esimene oli tunduvalt sagedasem kui teine (joonis 12). Esindatud oli ka hg K alamhgi K2 alamharu K2b näol. Hg U4 on üldiselt sagedasem Ida-Euroopas (Bermisheva et al. 2002). Lääne-Euroopas jääb selle hg sagedus paari % juurde (vt LISA 5). Euroopa üks vanimaid mtDNA hg U5 (Soares et al., 2010) jaguneb kaheks alamharuks - nooremaks U5a-ks ja vanemaks U5b-ks (Malyarchuk et al., 2010; Pereira et al., 2005; Richards et al., 2000), mis siinse Wales`i populatsiooni valimi hg U sisemisest varieeruvusest moodustasid peaaegu võrdse osa. Hg-de U5a ja U5b alamharudest olid antud valimis esindatud nii U5a1 ja U5a2 kui ka U5b1 ja U5b2 (LISA 4b), mida on on väga kõrge sagedusega (~80%) leitud Kesk-Euroopa mesoliitiliste DNA-de seast (Bramanti et al., 2009).

Sageduselt kolmas (11,4%) oli valimis hg J, mille sisemine varieeruvus jaguneb kahe selle alamharu J1 ja J2 vahel (joonis 13). Hg J1 moodustab kogu Wales`i populatsiooni emaliinidest 6,9% ja hg J2 4,5%.



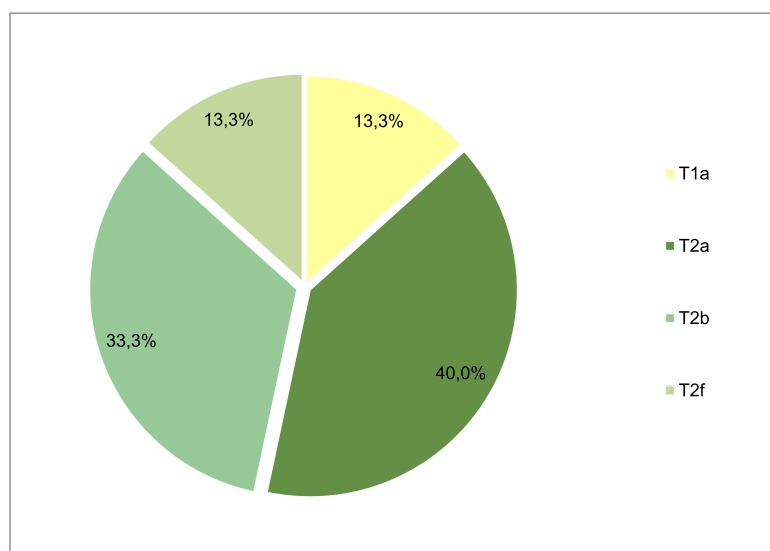
Joonis 13. Wales'i populatsiooni hg J (n= 20) alamhaplogruppide sagedused.



Joonis 14. Haplogrupi J taandatud mediaanvõrgustik, mis põhineb käesoleva töö Wales'i populatsiooni mtDNA andmetel (n=20). Sõõrid tähistavad proove/indiviide ja nende suurus antud võrgustikul ei varieeru st. iga haplotüüpi esindab üks indiviid. Erinevate värvidega on tähistatud hg J alamharud kuhu antud proovid kuuluvad. Märkide tähendused: “!” - tagasipöörduv mutatsioon; väike täht nukleotiidse positsiooni taga – transversioon (nt. 574a); “1C” – insertioon.

Alamhg J1c on haru, mis moodustab suurema osa (~80%) kõikidest J1 liinidest ning mis levis arvatavalt Euroopasse viimase jääaja hilisel perioodil (Pala et al., 2012). Ka Wales`i valimis tuvastatud J1 liinid paigutusid peaaegu kõik alamhg-i J1c, v.a üks haplotüüp, mis kuulus alamhg-i J1b1a. Tuvastatud hg J2 haplotüübid kuulusid alamhg-sse J2b ja J2a, täpsemalt alamharudesse J2b1a ja J2a1a (joonis 14). Alamhg J2b1a on peaaegu üksnes Euroopale omane mtDNA klaad ning on näidatud, et tema emaklaad (hg J2b1) on rohkem levinud Vahemere, Kesk- ja Atlandi-Euroopa piirkonnas. Hg J2a1a põhiline levimine Euroopa keskosas toimus tõenäoliselt jääaja lõpuperioodil (Pala et al., 2012). Seega - enamikku antud töös tuvastatud waleslaste hg J emaliine iseloomustavad Euroopale omased alamharud, mille põhiline ekspansioon Euroopas toimus viimase jääaja hilisel perioodil või vahetult pärast seda (LISA 4c).

Hg J sõsarklaad hg T moodustab Wales`i populatsiooni emaliinidest 8%. Hg T jaotub kaheks suuremaks alamharuks T1 ja T2, millest viimane on tunduvalt sagedasem kui esimene (joonis 15). See on kooskõlas kirjanduses avaldatud andmetega, kus on näidatud, et hg T2 on kordades sagedasem kui T1 ja seda just eriti Kesk- ja Lääne-Euroopa piirkonnas, hõlmates neist viimases isegi ligi 80% hg T sisest varieeruvusest (Pala et al., 2012).

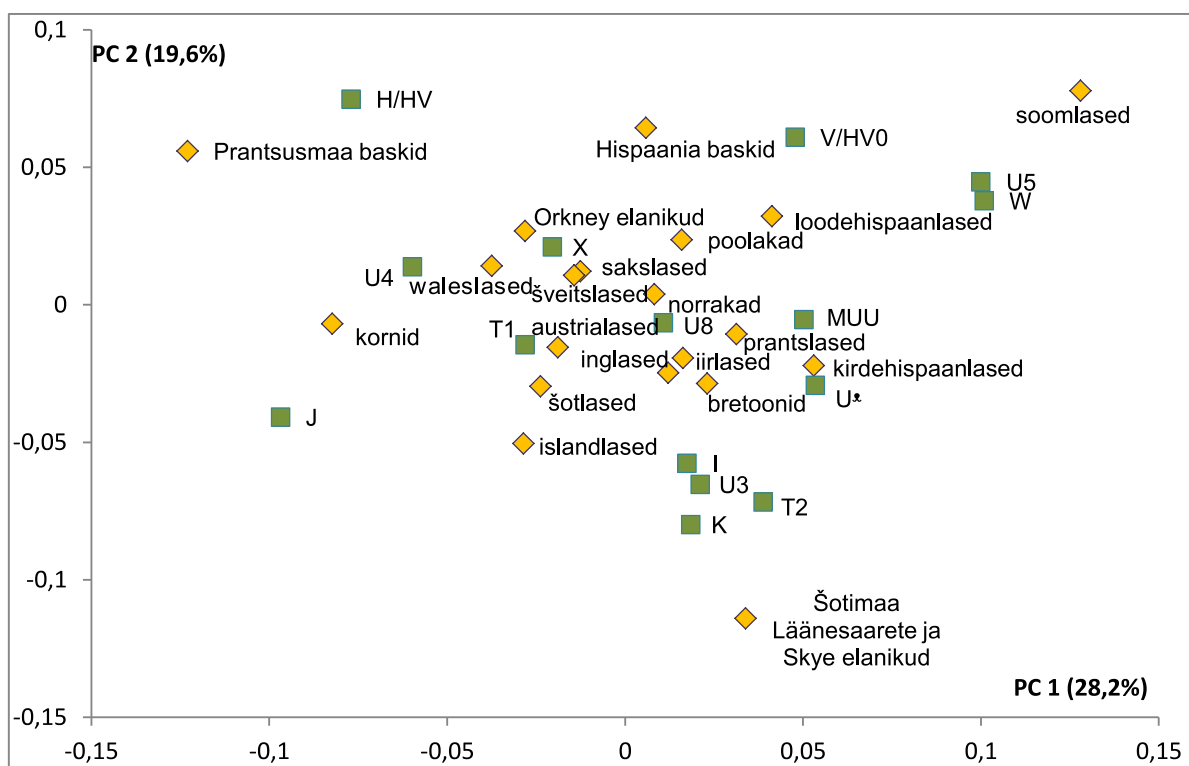


Joonis 15. Wales`i populatsiooni hg T (n=15) alamhaplogruppide sagedused.

Antud töö Wales`i populatsiooni valimi hg T sisene mitmekesisus jaotub selle alamhg-de T2a, T2b ja T2f vahel, mis moodustavad kokku ~87% hg T varieeruvusest (joonis 15). Kui T2b arvatakse olevat varajases neoliitikumi perioodis levinud Euroopasisese päritoluga alamhaplogrupp, siis tema sõsarharu T2a pärineb ilmselt Lähis-Idast ja on Euroopasse saanud mitme erineva lainena. Näiteks suurima alamhg-i T2a1 alamharu T2a1a näol neoliitikumi perioodil ning hilisel jääaja perioodil T2a1b (Pala et al., 2012). Mõlemad

3.1.2 Wales`i populatsiooni emaliinid Lääne-Euraasia kontekstis

Selleks, et vaadelda waleslaste emaliinide varieeruvust Lääne Euroopa teiste populatsioonide mtDNA mitmekesisuse taustal, tehti haplogruppide sagedusjaotuse põhjal uuritava piirkonna ja selle naaberalade andmestiku alusel põhikomponentanalüüs, kus kaks esimest komponenti kirjeldavad peaaegu 50% kogu varieeruvusest (joonis 17). Võrdlusmaterjalina kasutatud andmed teiste Euroopa populatsioonide mtDNA kohta olid kättesaadavad oluliselt madalama genotüpiseerimise resolutsiooniga, kui on seda käesolevas töös uuritud Wales`i populatsiooni andmed, mistõttu tuli ka antud töö proovide andmed viia PCA-ks samale tasemele. Et hinnata erinevate haplogruppide sageduste keskväärtusi, arvutati kõikide populatsioonide haplogruppide sagedustele ka 95%-lised usalduspiirid (LISA 5).



Joonis 17. Wales`i populatsioonis esinenud mtDNA varieeruvusel põhinev põhikomponentanalüüs Lääne- ja Kesk-Euroopa populatsioonide kontekstis. Hg HV on esitatud kombineeritult hg-ga H, hg HV0 kombineeritult hg-ga V, MUU hõlmab endas väga madala sagedusega (<1,6%) esinenud haplogruppe: L1, L2, L3, L3*, M1, M*, D/G, N1, N*, F, B, R0, C, A ja hg U*: U1, U2, U6, U7 (<1,2%).

PCA analüüsil kasutatud valimi emaliinide poolest sobitub Wales`i populatsioon üldiselt hästi teiste Euroopa populatsioonide hulka, sarnanedes kõige rohkem Orkney saarte populatsioonile ja Mandri-Euroopa populatsioonidest sakslastele ja šveitslastele (joonis 17). Kõikidest teistest kerkivad selgelt erinevatena esile Prantsusmaa baskid, soomlased ning Šotimaa Läänesaarete ja Skye elanikud. Haplogruppide V/HV0 kombineeritud kõrgem

esinemissagedus loodehispaanlastel ja Hispaania baskidel eristab neid Prantsusmaa baskidest. Viimaseid mõjutavad veel hg-de U5 ja U4 erinevad sagedused: U5 esineb rohkem Hispaania baskidel ja vähem Prantsusmaa baskidel ning hg U4 puhul vastupidi (Richard et al., 2007). Hg U4 on üldiselt sagedasem Ida-Euroopas (Bermisheva et al., 2002), siis Lääne-Euroopas jääb selle hg-i sagedus paari protsendi juurde (LISA 5). Siinses valimis esines võrreldes uuritavatega hg U4 kõrgem sagedus ($p < 0,05$) lisaks Prantsuse baskidele ka sakslastel ja poolakatel ning kornidel. Soomlasi nihutab kõikidest populatsioonidest eemale eelkõige kõrge hg U5 ja W sagedus ($p < 0,05$), kuid nad erinevad muudest siin analüüsitud populatsioonidest veel lisaks hg J ja H/HV madalama sageduse poolest (LISA 5). Kõrgeimat U5 esinemissagedust on Põhja-Euroopa rahvastel nagu eestlased, karjalased, soomlased ja saamid ka varasemalt korduvalt näidatud kirjanduses ilmunud andmetes (Richards et al., 1996; Sajantila et al., 1995; Tambets et al., 2003, 2004). Võrreldes waleslastega esineb oluliselt madalam hg J sagedus ($p < 0,05$) ka loodehispaanlastel, Hispaania baskidel, bretoonidel ja prantslastel. Hg J kõrgeim sagedus antud populatsioonide kontekstis esines kornidel. Šotimaa Läänesaarte ja Skye populatsioonid nihutab kõikidest populatsioonidest kaugemale eeskätt madalam hg H/HV sagedus ($p < 0,05$) ja waleslastest eelkõige kõrgem hg-de K ning I sagedus. Ka mõned waleslastele geograafiliselt lähedasemad populatsioonid nagu nt. šotlased, prantslased ja kirdehispaanlased on waleslastest eemaldunud ilmselt hg-de H/HV kombineeritud madalama sageduse poolest ($p < 0,05$), inglased ja iirlased lisaks kõrgema hg T2 ning viimased kindlasti ka kõrgema hg K sageduse poolest.

3.1.3 MtDNA täisjärjestuste analüüs

MtDNA täisjärjestuste alusel koostatud fülogeneesipuud annavad emaliinide fülogeneesist kõige parema ülevaate ja kõige täpsema topoloogiaga puu. Haplogrupi H sisestruktuuri täpsustamisel sõelale jäänud üheksa hg H täielikult sekveneeritud mtDNA proovid on esitatud tabelis 2. Saadud tulemuste ja *mtDNACommunity* andmebaasis (<http://www.mtDNAcommunity.org/>) avaldatud 117 mtDNA täisjärjestuse sünteesi alusel koostati hg H fülogeneesipuu (joonis 18). Antud töös täielikult sekveneeritud hg H mtDNA-d kuulusid alamhg-sse H3, H13a1a4, H27, H28, H52, H64, H66a ning kahe täielikult sekveneeritud mtDNA-d jäid seni defineerimata hg H alamharudesse (H) (vt. jooniseid 11 ja 18) (LISA 4b).

Lisaks antud töös täielikult sekveneeritud Walesi proovile (BK88), mis paigutus hg-i H27, on ka selle hg teistes alamharudes (H27a - H27d) veel mitmeid Atlandi-Euroopa piirkonnast pärit indiviidide proove, s.h üks Wales'i päritoluga. Antud hetkel on globaalsel

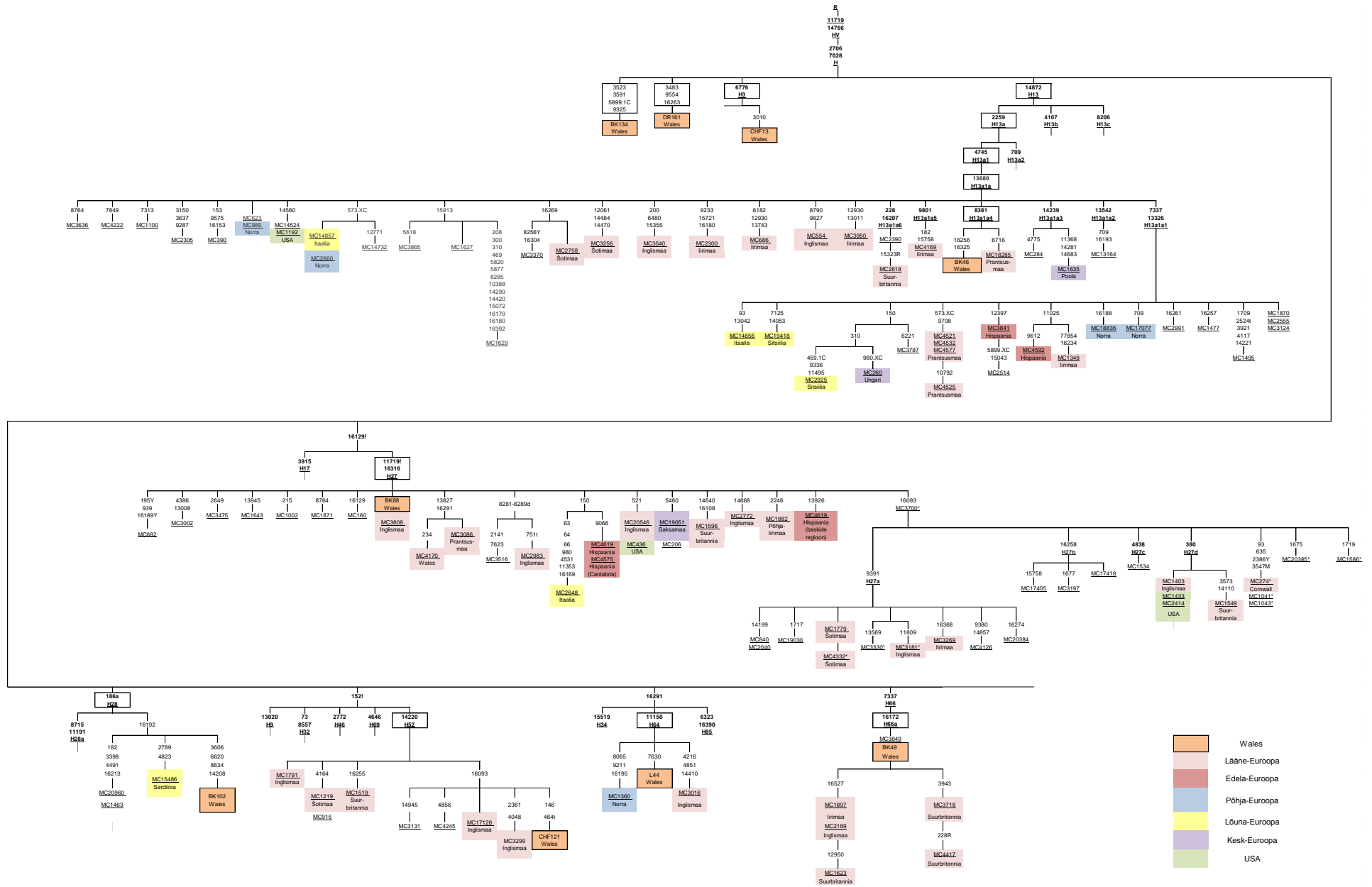
mtDNA fülogeneesipuul neis hg-des esindatud 47 täisjärjestust, millest 15 pärinevad Lääne- ja kolm Edela-Euroopa ning kaks Euroopa kesk- ja lõunaregiooni päritoluga indiviidilt. Kolme puhul on tegemist Euroopa päritolu ameeriklastega ja 24 täisjärjestuse päritolu kohta andmed puuduvad (vt joonis 18).

Samasse hg H13 alamharusse (H13a1a4), kuhu paigutus üks täielikult sekveneeritud Walesi mtDNA proov (BK46), kuulub lisaks ka üks Prantsusmaa päritoluga indiviidi haplotüüp. Hg-s H13a1a ja selle teistest alamharudest leiab veel mitmeid samast regioonist pärinevaid emaliine: 51-st täisjärjestusest kümme on Lääne-, seitse Edela-Euroopa ning neli Lõuna-Euroopa päritoluga. Kaks kuulub Kesk-, neli Põhja-Euroopa ja üks Euroopa päritoluga Ameerika indiviidile ning ülejäänud 23 täisjärjestuse päritolu kohta info puudub (vt joonis 18).

Hg H52 ja H66a, kuhu paigutusid kaks täielikult järjestatud waleslase mtDNA-d (CHF121 ja BK49), on samuti huvitavad, kuna antud mtDNA fülogeneesipuu harudesse kuulub lisaks veel vaid vastavalt kaheksa ja kuus mtDNA täisjärjestust, mis enamuse pärinevad Inglismaa, Suurbritannia, Iirimaa ja Šotimaa indiviididelt ja paari ülejäänud päritolu pole teada. Minoorsesse hg H64 harusse kuulus lisaks Walesi proovile (L44) veel üks Norra ja üks Inglismaa päritoluga indiviidi mtDNA täisjärjestus (joonis 18).

Kui arvestada vaadeldud hg-de H27 ja H13a1a vaid teadaoleva päritoluga emaliine, siis moodustavad Lääne-Euroopa päritoluga emaliinid antud hg-desse kuuluvatest täisjärjestusest küllalt suure osakaalu: vastavalt 65% ja 34,5% ning hg H52 ja H66a on sellisel juhul täielikult Lääne-Euroopa spetsiifilised hg H harud. Seega võib eelneva põhjal oletada, et need konkreetset hg H (alam)harud on rohkem levinud just Euroopa lääne- ja edelapiirkonnas.

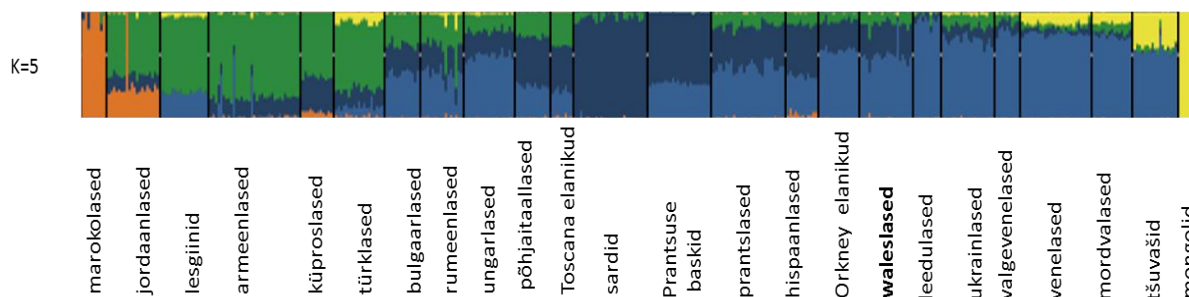
Täielikult sekveneeritud ebatüüpilise HVS-I motiiviga (tabel 2) mtDNA paigutus hg-sse A4b. Hg A on Ida-Euraasia spetsiifiline hg ja selle alamharu A4 on eelkõige üldlevinud populatsioonides Põhja-Siberist kuni Iraani ja Lõuna-Vietnamini. Waleslaste populatsioonis on see ilmselt üks neist Ida-Euraasiast pärit emaliinidest, mida Euroopa populatsioonides juhukülalistena leida võib ja mille osakaal populatsioonis ei ületa tavaliselt 1-5% (Richards et al., 2000).



Joonis 18. MtDNA hg H täisjärjestuste fülogeneesipuu, mis on koostatud 9 antud töös sekveneeritud hg H täisjärjestuse ja 117 mtDNA täisjärjestuste andmebaasis (*mtDNACommunity*) (*Phylotree*) avaldatud täisjärjestuse alusel. Fülogeneesipuu koostamisel kasutati vaid neid ülemaailmse mtDNA fülogeneesipuu harusid, kuhu antud töös täielikult järjestatud Wales`i proovid kuulusid. Avaldatud järjestused on puul tähistatud *MtDNACommunity* andmebaasi ID-ga. Neil mtDNA-del, mille kohta oli teada ka neid kandva indiviidi päritolu, on see ID-le lisatud. Märkide tähendused: “!” - tagasipöördunud mutatsioon; väike täht np-i taga – transversioon (nt. 464t); suur täht np-i taga (nt. 16189Y) – heteroplasmia; “d” – deletsioon; “.1C” - insertioon; “.XC” - polüC järjestus. Kui ID-le on lisatud tärn (*), viitab see neil haplotüüpidel esinevale heteroplasmiale np-s 16093. Parema resolutsiooniga täisjärjestuste fülogeneesipuu saab alla laadida siit: <http://www.evolutioon.ebc.ee/wales/>

3.2 Autosoomse DNA polümorfismide genotüpiseerimisandmete analüüs

Autosoomide varieeruvuse osas uuritud valim on toodud tabelis 4. ADMIXTURE programmiga läbiviidud populatsiooni struktureerituse analüüsi tulemused on toodud joonisel 19 ja lisa 6. Sobivaimateks mudeliteks loeti 5 - 7 eellaskomponendiga mudelid, eraldi on tulemustes välja toodud mudel K = 5 (joonis 19).



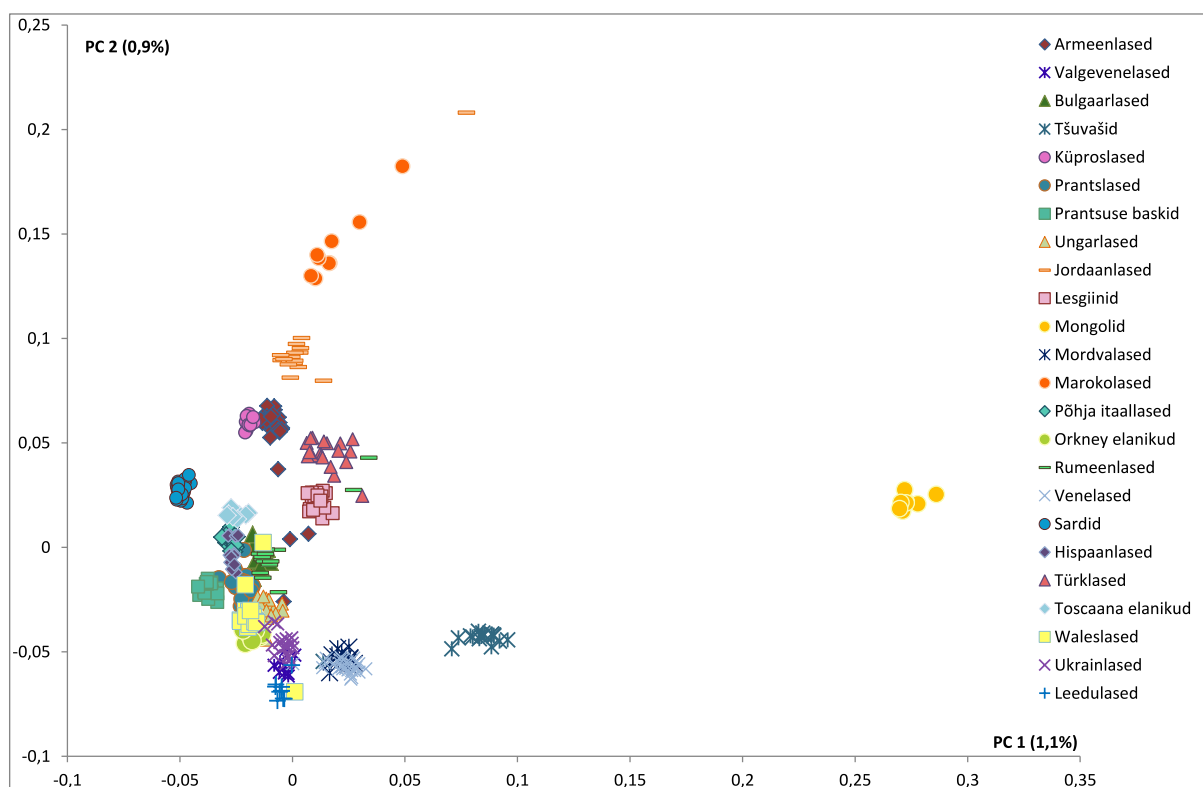
Joonis 19. ADMIXTURE analüüsi abil ja 660 000 autosoomsete SNP põhjal tuletatud Wales'i populatsiooni struktuur Lääne-Euraasia populatsioonide kontekstis viie eellaskomponendi puhul (K = 5). Hüpoteetilistest eellaspopulatsioonidest pärit komponendid on tähistatud eri värvidega ja esitatud vertikaalselt paigutatud tulpadena, millest igaüks vastab ühele kindlale uuritud indiviidile. Erinevatesse populatsioonidesse kuuluvad individid on eraldatud musta vertikaalse joonega.

K=5 juures iseloomustab uuritava Wales'i populatsiooni valimit 3 põhilist geneetilist eellaskomponenti (sinine, tumesinine ja roheline). Proportsiooniliselt on see kolmik esindatud rohkemal või vähemal määral enamikes uuritud Euroopa populatsioonides ja seda alates K=4-st kuni vaadeldud K=9-ni (vt. LISA 6a)

Lähis-Ida/Kaukaasia eellaskomponent (roheline) on enim esindatud just Kagu- ja Lõuna-Euroopa populatsioonides ning kõige vähem Ida-Euroopa ning Volga-Uurali piirkonna populatsioonides (venelased, mordvalased, tšuvašid). Tõenäolise Euroopa päritoluga eellaskomponentidest (sinine ja tumesinine) on viimane ülekaalus Lääne- ja Lõuna-Euroopas ning esimene Ida-Euroopas. Aafrika ja Ida-Aasia mõjutusi on Lääne-Euroopas üldiselt väga vähe või üldse mitte. Analüüs näitas, et erinevate eellaskomponentide proportsioonid sarnanevad geograafilistel lähinaabritel mõnevõrra rohkem (näiteks venelaste ja mordvalaste või põhja-itaallaste ja toskaanalaste profiilid) kui teineteisest kaugel asuvatel populatsioonidel. See kehtib ka antud töö uurimisobjektiks olnud waleslaste kohta, kes sarnanevad kõige rohkem Orkney saarte elanikega (joonis 19) ja mandriosa eurooplastest prantslastega. Geneetiline sarnasus prantslastega sobitub varajase Briti saarte koloniseerimismustriga viimase jääaja järgsel perioodil, kui esmaasustajad tulid Loode-Euroopast Suurbritanniasse üle maasilla, mis ulatus sel ajal Suurbritanniani ja asustajate teine suurem laine tuli Prantsusmaalt Atlandi rannikut mööda ja ilmselt veeteele pidi Inglismaa

lääneossa, Walesi ning Iirimaale (Keeling, 2013). 7-nda hüpoteetilise eellaspopulatsiooni juures lisandub Orkney populatsiooni spetsiifiliselt iseloomustav komponent (helesinine), mis rohkem kui pooltel sellest valimist muutub domineerivaks komponendiks (LISA 6a). Y-kromosoomi ja mtDNA andmete analüüsi põhjal tuvastatud mõnevõrra tugevam skandinaavia päritolu Orkney saarte elanikel võrreldes teiste Atlandi-Euroopa piirkonna populatsioonidega (Wilson et al., 2001; Helgason et al., 2001; Goodacre et al., 2005) võiks selgitada ka nende eristumist waleslastest. Kahjuks puudub selle kohta hetkel võrdlusmaterjal autosoomide osas.

Üheks erandiks on sardid (sardiinlased), kellel alates $K=3$ -st domineerib terves populatsioonis põhiliselt üks eellaskomponent (tumesinine) ja teiseks erandiks Prantsusmaa baskid, kellel alates 6-nda hüpoteetilise eellaspopulatsiooni lisamisest muutub domineerivaks uus, ainult nende populatsioonile omane komponent (kirkas sinine) (vt. LISA 6a). Mõlemad populatsioonid on teadaolevalt nn. isoleeritud populatsioonid (Cavalli-Sforza et al., 1994), mille geneetiline varieeruvus on mitmete demograafiliste (keelelisest ja geograafilisest eripärast tulenev isolatsioon) ja sellest tulenevate evolutsiooniliste sündmuste (geenitriiv) tagajärjel kahanenud ning mida on ka mitmete varasemate geneetiliste uuringute - nii näiteks mtDNA (Richard et al., 2007) kui ka ülegenoomsete markerite (Nelis et al., 2009) puhul varem näidatud.



Joonis 20. 660 000 ülegenoomse jaotusega SNP-del põhinev põhikomponentanalüüs. Esitatud on esimene (PC1) ja teine (PC2) põhikomponent koos nende poolt kaetava geneetilise varieeruvuse osakaaluga (%).

Autosoomsete andmete põhikomponendianalüüsil (joonis 20) eristusid antud valimis kõige selgemini Euroopa populatsioonidest geneetiliselt kõige kaugemad populatsioonid, nagu mongolid ja marokolased, kõige ühtlasem ja ühetaolisem on aga Euroopa. Üldiselt on populatsioonide geneetiline klasterdumine kooskõlas nende geograafilise jaotusega. Lähis-Ida ja Kaukaasia populatsioonid (jordanlased, armeenlased, küproslased, türklased, lesgiinid) moodustavad omaette asetseva grupi ning Euroopa idapoolsetest populatsioonidest on koondunud enam leedulased, valgevenelased, ukrainlased ning põhja-venelased ja mordvalased. Täiesti eraldiseisva klasteri moodustavad siinses valimis Volga-Uurali piirkonnas elavad ja turgi keeli rääkivad tšuvašid, kelle geenitiigis on juba märgatav osakaal Ida-Euraasiale iseloomulikke genotüüpe (vt. ka joonis 19).

Antud valimi Euroopa populatsioonid moodustavad suhteliselt ühtse grupi. Üldiselt koonduvad kõik ühe populatsiooni individid kokku ja on kõige sarnasemad just üksteisega, kuid on ka mõningad sellised populatsioonid, kus geneetiline varieeruvus "hajub" rohkem - näiteks rumeenlased, kellest osa on nihkunud türklaste poole, armeenlased ning jordanlased, kellest esimesed on ilmselt Kesk-Euroopa ja teised selgelt Aafrika mõjutustega. Waleslased paigutuvad Orkney elanike ja prantslaste vahele, mis on kooskõlas nende populatsioonide geograafilise asetusega Euroopa läänepoolses osas. Ka waleslaste hulgas on paar indiviidi, kelle isiku enesemääratlus ei sobitu sellele rahvusele geneetiliselt üldomase profiiliga. Juba ADMIXTURE analüüsis teistest erinevatena esile kerkinud sardid ja Prantsuse baskid paiknevad teistest eraldi ka PCA analüüsil. MtDNA analüüsis kasutatud sakslaste ja šveitslaste võrdlusandmed autosoomide osas puuduvad.

3.3 Põhitulemused ja järeldused

1. Uuritud waleslaste emaliine iseloomustab Lääne-Euroopa populatsioonidele iseloomulik mtDNA haplogrupiline jaotuvus: haplogrupid H, U, J, T ja V moodustavad ligikaudu 90% Walesi mtDNA varieeruvusest. Ida-Euraasia ja Aafrika päritoluga emaliinid moodustavad vaid tühise osa (< 2%);
2. Autosoomsete andmete analüüs näitas, et waleslaste populatsioonil on eurooplastele iseloomulik geneetiline struktuur, milles põhiosa on euroopapärastel komponentidel, vahesel määral esineb Lähis-Ida/Kaukaasia komponenti. Aafrika ja Ida-Aasia mõjutusi on kas väga vähe või üldse mitte. Uuritud waleslaste puhul ei ole kasutatud meetodite puhul täheldatav geenitriivi märgatavat mõju (nagu näiteks baskide või sardiinlaste puhul);
3. Waleslaste paigutumine geneetilisel kaardil on kooskõlas nende geograafilise paiknemisega, seda iseäranis autosoomsete andmete osas. Võrdlusmaterjalina kasutadaoleva emaliinide andmestiku madalam resolutsioonitase ei võimalda waleslaste mtDNA variantide puhul fülogeograafilisi detaile rõhutada, kuid saadud tulemused on üldiselt kooskõlas autosoomsete andmetega. Maksimaalse resolutsiooniga emaliinide sisestruktuuri uuringus tuvastatud harude hulgas oli mitmeid selliseid (nt. H27, H52, H66), mis tunduvad olevat sagedasemad just Euroopa lääne- ja edelapiirkonnas;
4. Nii autosoomide kui emaliinide varieeruvuse alusel on siinse valimi põhjal waleslastele geneetiliselt kõige lähedasemad Orkney elanikud, Mandri-Euroopa populatsioonidest aga prantslased (autosoomide analüüs) ning sakslased ja šveitslased (emaliinide analüüs).

KOKKUVÕTE

Suurbritannia ja Põhja-Atlandi saared on ühed Euroopa viimastest piirkondadest, kuhu kaasaegne inimene pärast jääaega jõudis. Suurbritannia edelaosas asuv ajaloolis-geograafiline piirkond Wales ja sealsed elanikud waleslased, kes on käesoleva magistratöö uurimisobjektiks, erinevad Suurbritannia põhielanikkonnast oma keele- (kõmri keel) ja kultuuriloolise tausta poolest ning on arvatud, et just waleslased võivad kanda Briti saarte algupärast geneetilist pärandit.

Käesolev magistratöö uurib Wales'i populatsiooni geneetilist mitmekesisust ja struktuuri, kasutades selle kirjeldamiseks mtDNA ja üle kogu genoomi jaotunud 650 000 SNP genotüpiseerimisandmestiku analüüsi.

Uurimustöö emaliinide analüüsi valimisse kuulunud kõik 176 waleslase mtDNA-d paigutusid globaalsel mtDNA fülogeneesipuul haplogruppidesse (hg) ja alamhaplogruppidesse viisil, mille põhjal võib järeldada, et waleslaste emaliine iseloomustab Euroopa populatsioonidele tüüpiline haplogrupiline jaotuvus. Esindatud on valdav enamus Lääne-Euraasiale iseloomulikke mtDNA haplogruppe (H, V, J, T, U jne.) ning Ida-Euraasia ja Aafrika päritoluga emaliinid moodustavad vaid tühise osa (<2%). Detailsemalt uuritud mtDNA haplogrupi H sisestruktuuri täpsustamisel sõelale jäänud 9 hg H mtDNA-d sekveneeriti täielikult. Neist seitse kuuluvad hg H alamhaplogruppidesse H3, H13a1a4, H27, H28, H52, H64, H66a ning kaks seni kirjeldamata hg H uutesse harudesse. Hg-d H27, H52 ja H66 on suhteliselt harva esinevad hg H harud, kuid globaalse fülogeneesipuu (*PhyloTree* ja *mtDNACommunity*) andmete põhjal kuulub neisse ja nende haplogruppide teistesse harudesse veel mitmeid Wales'i lähipiirkonnast pärit haplotüüpe ja võib oletada, et need haplogrupid on sagedasemad just Euroopa lääne- ja edelapiirkonnas.

MtDNA (PCA) ja autosomaalsete andmete analüüs (PCA, ADMIXTURE) näitas, et Wales'i populatsioon sobitub geneetilise varieeruvuse poolest üldiselt hästi teiste Euroopa populatsioonide hulka, sarnanedes mõnevõrra rohkem geograafiliste lähinaabritega – nagu Orkney elanikud. Euroopa mandriosast sarnanevad waleslased geneetiliselt eelkõige sakslaste ja šveitslastega (emaliinide puhul) ning prantslastega (autosoomide puhul, siinsesse analüüsi sakslasi ja šveitslasi kaasatud polnud). Laiemalt vaadeldes võib täheldada, et populatsioonide geneetiline klasterdumine on kooskõlas nende geograafilise asetusega. Autosoomsete andmete põhjal tehtud põhikomponentanalüüsil moodustavad uuritud Euroopa populatsioonid esimese ja teise põhikomponendi vastandamisel suhteliselt ühtse grupi. Kõige selgemini eristusid Euroopa populatsioonidest geneetiliselt need populatsioonid, mis siinses valimis asuvad neist geograafiliselt kõige kaugemal (nt. marokolased, jordanlased, mongolid). Uuritud waleslaste

populatsioon on autosoomide profiililt suhteliselt ühetaoline - kõik selle inividid koonduvad üldiselt kokku ja on kõige sarnasemad eelkõige üksteisega.

Genetic analyses of the population of Wales based on maternal lineages and autosomal markers in the context of West-Eurasia.

Anu Solnik

SUMMARY

Several major demographic processes have shaped the gene pool of the European populations – arrival of first settlers in the Upper Palaeolithic, re-colonization of the continent from several refugia after the Last Glacial Maximum (LGM), Neolithic migrations from the Near East and minor migrations on the continental trade routes. One of the proposed major events shaping the diversity of the current European maternal gene pool is the re-colonization of northern regions of the continent due to the warmer climate after LGM around 15 000 years ago. Recent studies of ancient DNA (aDNA) have also stressed the importance of Neolithic migrations in these formations (Brotherton et al., 2013; Soares et al., 2009; Bramanti et al., 2009; Richards et al., 2000).

Great Britain and the islands of the North Atlantic are among of the last regions in Europe where modern humans reached after the LGM. Wales is a country in the southwestern part of the island of Great Britain, bordered by England to its east and the Atlantic Ocean and Irish Sea to its west. The objects of the current Master thesis are the native people of Wales – the Welsh or Cymry who differ from the rest of the inhabitants of the Great Britain in their language – Cymry (Welsh) language and cultural heritage. It has been proposed the Welsh carry the original genetic legacy of the British Isles.

The maternal lineages of the Welsh population published so far are with much lower resolution known now in the light of mtDNA complete sequences. The purpose of the present Master thesis is to add more detailed data and to study the genetic diversity and structure of Welsh population by analysing the mtDNA and genotyping data of SNPs covering the whole the whole genome. Analysing the data from this study together with the neighbouring populations and in the wider context of West-Eurasian populations helps to understand the evolution of the Welsh population and its place on the current West-European genetic landscape.

The analysis of mtDNA show clearly that the knowledge of the nucleotide variation in the rapidly evolving HVS-I region alone is often not sufficient to clearly define the branches of the phylogenetic tree. So it can be presumed that genotyping the mutations in the coding region is crucial for the exact definition of maternal haplogroups.

All 176 Welsh mtDNAs in the current study fitted into haplogroups and subhaplogroups on the phylogenetic tree of global mtDNAs. The haplogroup distribution is similar to a typical European population. The prevalent majority of West-Eurasian maternal haplogroups (H, V, J, T, U etc.) are represented and haplogroups with East-Eurasian and African origin are present at minor levels (<2%). The Welsh carried also sub-branches of haplogroup H that are rare in the current European genetic pool, but based on aDNA studies characteristic to the early period of Central-European Neolithic. These are H16, minor branch H5b, rare sub-branches of the otherwise frequent H1 haplogroup – H1j and H1e and many other maternal lineages from the later period of Neolithic.

In the course of the detailed study of mtDNA haplogroup H internal structure 9 mtDNAs sequenced completely. Seven of them belong to H sub-haplogroups H3, H13a1a4, H27, H28, H52, H64, H66a and two the to the new H sub-groups not described so far. H27, H52 and H66a are relatively rare branches of haplogroup H, but on the global phylogenetic tree (*Phylotree* and *mtDNACommunity*) there are one haplotype from Wales and more from its neighbouring regions belonging to this haplogroups and its sub-branches. Thus, it can be speculated that these haplogroups are more frequent in Western and South-Western Europe.

The present analyses of mtDNA [principal component analyses (PCA)] and autosomal markers (PCA, ADMIXTURE) revealed that the genetic variation of the Welsh population fits well with other European populations, being more similar to the geographic neighbours, e.g. the inhabitants of Orkney Islands. In the European mainland the genetically most similar to the Welsh are the Germans and the Swiss (in the case of maternal lineages) and French (autosomal analyses which did not include the Swiss and the Germans). Thus, in the wider perspective, the populations cluster according to their geographic location. The studied European populations form a relatively uniform cluster on the PCA plot of the first and second component describing the genetic variation of the autosomal markers. Genetically, the most distinct of the studied European populations in the current sample were geographically the most far ones (Moroccans, Jordanians, Mongols). The profile of autosomal markers in the studied Welsh populations is relatively uniform – all individuals cluster together and are the most similar to each other.

TÄNUAVALDUSED

Soovin südamest tänada oma juhendajaid Ene Metspalu ja Kristiina Tambetsit põhjaliku juhendamise, kulutatud aja ning abivalmiduse eest. Suur aitäh Dr. Sarabit S. Mastanale, tänu kellele Wales'i DNA proovid evolutsioonilise bioloogia õppetooli jõudsid ja käesolev uurimustöö võimalikuks sai. Samuti soovin tänada Erwan Pennaruni, kes Wales'i populatsiooni mtDNA esialgseid tulemusi minuga jagas ning Monika Karminile tõlkeabi eest. Lisaks tänan eranditult kõiki meie õppetooli inimesi, kelle sõbralik ja rõõmsameelne hoiak on loonud meeldiva ja töötegemist soosiva õhkkonna.

Lõpetuseks tänan oma pere liikmeid, eeskätt oma tütart Nora Johannat, kes on olnud toetav ja ülimalt kannatlik kogu mu magistriõpingute perioodi jooksul.

Käesolev magistritöö on saanud tuge Eesti Teadusfondi uurimistoetustelt nr. 6040, 7858 ja 8973.

KASUTATUD KIRJANDUS

- Abecasis, G.R. et al., 2010. A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061–1073.
- Abecasis, G.R. et al., 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422), 56–65.
- Achilli, A. et al., 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet*, 75(5), 910–918.
- Achilli, A. et al., 2005. Saami and berbers--an unexpected mitochondrial DNA link. *Am J Hum Genet*, 76(5), 883–886.
- Alexander, D.H., Novembre, J. & Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*, 19(9), 1655–1664.
- Altshuler, D.M. et al., 2010. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311), 52–58.
- Amos, W. & Hoffman, J.I., 2010. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proceedings. Biological sciences / The Royal Society*, 277(1678), 131–137.
- Anderson, S. et al., 1981. Sequence and organization of the human mitochondrial genome. *Nature*, 290(5806), 457–465.
- Andrews, R.M. et al., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*, 23(2), 147.
- Aquadro, C.F. & Greenberg, B.D., 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics*, 103(2), 287–312.
- Ardlie, K.G., Kruglyak, L. & Seielstad, M., 2002. Patterns of linkage disequilibrium in the human genome. *Nature reviews. Genetics*, 3(4), 299–309.
- Auton, A. et al., 2009. Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Res*, 19(5), 795–803.
- Bailey, S.E., Weaver, T.D. & Hublin, J.-J., 2009. Who made the Aurignacian and other early Upper Paleolithic industries? *Journal of human evolution*, 57(1), 11–26.
- Bandelt, H.J. et al., 1995. Mitochondrial portraits of human populations using median networks. *Genetics*, 141(2), 743–753.
- Bandelt, H.J., Forster, P. & Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution*, 16(1), 37–48.
- Bar-Yosef, O., 1998. The Natufian culture in the Levant, threshold to the origins of agriculture. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), 159–177.
- Behar, D.M. et al., 2004. MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Eur J Hum Genet*, 12(5), 355–64.
- Behar, D.M. et al., 2010. The genome-wide structure of the Jewish people. *Nature*, 466(7303), 238–242.
- Behar, D.M. et al., 2012. A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet*, 90(4), 675–684.

- Benazzi, S. et al., 2011. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature*, 479(7374), 525–528.
- Bermisheva, M. et al., 2002. Diversity of mitochondrial DNA haplotypes in ethnic populations of the Volga-Ural region of Russia. *Mol Biol (Mosk)*, 36(6), 990–1001.
- Bogenhagen, D.F., 1999. Repair of mtDNA in vertebrates. *Am J Hum Genet*, 64(5), 1276–1281.
- Bramanti, B. et al., 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, 326(5949), 137–140.
- Brandstätter, A. et al., 2006. Dissection of mitochondrial superhaplogroup H using coding region SNPs. *Electrophoresis*, 27(13), 2541–2550.
- Brotherton, P. et al., 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nature communications*, 4, 1764.
- Brown, T. a et al., 2009. The complex origins of domesticated crops in the Fertile Crescent. *Trends in ecology & evolution*, 24(2), 103–109.
- Brown, W.M., George Jr., M. & Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*, 76(4), 1967–1971.
- Cann, R.L., Stoneking, M. & Wilson, A.C., 1987. Mitochondrial DNA and human evolution. *Nature*, 325(6099), 31–36.
- Capelli, C. et al., 2003. A y chromosome census of the british isles. *Curr Biol*, 13(11), 979–984.
- Cavelier, L. et al., 2000. MtDNA substitution rate and segregation of heteroplasmy in coding and noncoding regions. *Human genetics*, 107(1), 45–50.
- Chikhi, L. et al., 2002. Y genetic data support the Neolithic demic diffusion model. *Proc Natl Acad Sci U S A*, 99(17), 11008–11013.
- Clark, J.D. et al., 2003. Stratigraphic, chronological and behavioural contexts of Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature*, 423(6941), 747–752.
- Colonna, V. et al., 2011. A world in a grain of sand: human history from genetic data. *Genome biology*, 12(11), 234.
- Conard, N.J. & Bolus, M., 2003. Radiocarbon dating the appearance of modern humans and timing of cultural innovations in Europe: new results and new challenges. *Journal of human evolution*, 44(3), 331–371.
- d'Errico, F. et al., 1998. Neanderthal acculturation in western Europe? A critical review of the evidence and its interpretation. *Current Anthropology*, 39, S1–S44.
- D'Errico, F., 2003. The invisible frontier. A multiple species model for the origin of behavioral modernity. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(4), 188–202.
- Dolukhanov, P.M., 1997. The Pleistocene-Holocene transition in northern Eurasia: Environmental changes and human adaptations. *Quaternary International*, 181–191.
- Elson, J.L., Turnbull, D.M. & Howell, N., 2004. Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet*, 74(2), 229–238.
- Forster, P., 2004. Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. *Philos Trans R Soc Lond B Biol Sci.*, 359(1442), 255–264.

- Forster, P. et al., 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet*, 59(4), 935–945.
- Gabriel, S.B. et al., 2002. The structure of haplotype blocks in the human genome. *Science (New York, N.Y.)*, 296(5576), 2225–2229.
- Gamble, C. et al., 2005. The Archaeological and Genetic Foundations of the European Population during the Late Glacial: Implications for “Agricultural Thinking.” *Cambridge Archaeological Journal*, 15(2), 193–223.
- García, O. et al., 2011. Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity*, 106(1), 37–45.
- Giles, R.E. et al., 1980. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci U S A*, 77(11), 6715–6719.
- Goodacre, S. et al., 2005. Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity*, 95(2), 129–135.
- Greenberg, B.D., Newbold, J.E. & Sugino, A., 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene*, 21(1-2), 33–49.
- Haak, W. et al., 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*, 310(5750), 1016–1018.
- Haak, W. et al., 2010. Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biol*, 8(11), e1000536.
- Helgason, A. et al., 2001. mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *American journal of human genetics*, 68(3), 723–737.
- Helgason, A. et al., 2000. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet*, 67(3), 697–717.
- Helgason, A. et al., 2001. mtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am J Hum Genet*, 68(3), 723–737.
- Herrnstadt, C. et al., 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet*, 70(5), 1152–1171.
- Higham, T. et al., 2011. The earliest evidence for anatomically modern humans in northwestern Europe. *Nature*, 479(7374), 521–524.
- Hofmann, S. et al., 1997. Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. *Hum Mol Genet*, 6(11), 1835–1846.
- Howell, N. et al., 2007. Relative Rates of Evolution in the Coding and Control Regions of African mtDNAs. *Mol Biol Evol*, 24(10), 2213–2221.
- Hughen, K. et al., 2004. ¹⁴C activity and global carbon cycle changes over the past 50,000 years. *Science (New York, N.Y.)*, 303(5655), 202–207.
- Ingman, M. et al., 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*, 408, 708–713.
- Ingman, M. & Gyllensten, U., 2001. Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. *J Hered*, 92(6), 454–461.

- International, T. & Consortium, H., 2005. A haplotype map of the human genome. *Nature*, 437(7063), 1299–1320.
- Jacobi, R.M. & Higham, T.F.G., 2008. The “Red Lady” ages gracefully: new ultrafiltration AMS determinations from Paviland. *Journal of human evolution*, 55(5), 898–907.
- Keeling Judith, 2013. Oxford Today - Volume 25 No 2. *Oxford today*, 25(2), 27–28.
- Kivisild, T. et al., 2002. The emerging limbs and twigs of the East Asian mtDNA tree. *Mol Biol Evol*, 19(10), 1737–1751.
- Kivisild, T. et al., 2006. The role of selection in the evolution of human mitochondrial genomes. *Genetics*, 172(1), 373–387.
- Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Duranthon, F., et al., 2011a. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proceedings of the National Academy of Sciences of the United States of America*, 108(24), 9788–9791.
- Lacan, M., Keyser, C., Ricaut, F.-X., Brucato, N., Tarrús, J., et al., 2011b. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, 108(45), 18255–18259.
- Li, J.Z. et al., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319(5866), 1100–1104.
- Loogväli, E.-L. et al., 2004. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol Biol Evol*, 21(11), 2012–2021.
- Loogväli, E.-L. et al., 2009. Explaining the imperfection of the molecular clock of hominid mitochondria. *PloS one*, 4(12), e8260.
- Lutz, S. et al., 1998. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. *Int J Legal Med*, 111(2), 67–77.
- Macaulay, V.A. et al., 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet*, 64(1), 232–249.
- Malyarchuk, B. et al., 2010. The peopling of Europe from the mitochondrial haplogroup U5 perspective. *PloS one*, 5(4), e10285.
- Manichaikul, A. et al., 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics (Oxford, England)*, 26(22), 2867–2873.
- Margulis, L., 1975. Symbiotic theory of the origin of eukaryotic organelles; criteria for proof. *Symposia of the Society for Experimental Biology*, (29), 21–38.
- McDougall, I., Brown, F.H. & Fleagle, J.G., 2005. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature*, 433(7027), 733–736.
- Mellars, P.A., 1992. Archaeology and the population-dispersal hypothesis of modern human origins in Europe. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 337(1280), 225–234.
- Mellars, P., 2004. Neanderthals and the modern human colonization of Europe. *Nature*, 432(7016), 461–465.
- Mellars, P., 2005. The impossible coincidence. A single-species model for the origins of modern human behavior in Europe. *Evolutionary Anthropology: Issues, News, and Reviews*, 14(1), 12–27.
- Mellars, P., 2006. A new radiocarbon revolution and the dispersal of modern humans in Eurasia. *Nature*, 439(7079), 931–935.

- Mellars, P., 2011. Palaeoanthropology: the earliest modern humans in Europe. *Nature*, 479(7374), 483–485.
- Merriwether, D.A. et al., 1991. The structure of human mitochondrial DNA variation. *Journal of molecular evolution*, 33(6), 543–555.
- Metspalu, M. et al., 2004. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC genetics*, 5, 26.
- Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature reviews. Genetics*, 11(1), 31–46.
- Mishmar, D. et al., 2003. Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1), 171–176.
- Nelis, M. et al., 2009. Genetic structure of Europeans: a view from the North-East. *PloS one*, 4(5), e5472.
- Novembre, J. et al., 2008. Genes mirror geography within Europe. *Nature*, 456(7218), 98–101.
- O'Dushlaine, C.T. et al., 2010. Population structure and genome-wide patterns of variation in Ireland and Britain. *European journal of human genetics : EJHG*, 18(11), 1248–1254.
- Olivo, P.D. et al., 1983. Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. *Nature*, 306(5941), 400–402.
- Van Oven, M. & Kayser, M., 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*, 30(2), 386–394.
- Pakendorf, B. & Stoneking, M., 2005. Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet*, 6, 165–183.
- Pala, M. et al., 2009. Mitochondrial haplogroup U5b3: a distant echo of the epipaleolithic in Italy and the legacy of the early Sardinians. *American journal of human genetics*, 84(6), 814–821.
- Pala, M. et al., 2012. Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *American journal of human genetics*, 90(5), 915–924.
- Patterson, N., Price, A.L. & Reich, D., 2006. Population structure and eigenanalysis. *PLoS Genet*, 2(12), e190.
- Pereira, L. et al., 2004. Subdividing mtDNA haplogroup H based on coding-region polymorphisms—a study in Iberia. *Int Congr Ser*, 1261, 416–418.
- Pereira, L. et al., 2005. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome research*, 15(1), 19–24.
- Piganeau, G. & Eyre-Walker, A., 2004. A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. *Heredity*, 92(4), 282–288.
- Price, A.L. et al., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38(8), 904–909.
- Purcell, S. et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, 81(3), 559–575.
- Quintans, B. et al., 2004. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Sci Int*, 140(2-3), 251–257.

- Radivojević, M. et al., 2010. On the origins of extractive metallurgy: new evidence from Europe. *Journal of Archaeological Science*, 37(11), 2775–2787.
- Ramsey, C.B. et al., 2006. Developments in Radiocarbon Calibration for Archaeology. *Antiquity*, 80(310), 783.
- Rasmussen, M. et al., 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*, 463(7282), 757–672.
- Richard, C. et al., 2007. An mtDNA perspective of French genetic variation. *Ann Hum Biol*, 34(1), 68–79.
- Richards, M. et al., 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am J Hum Genet*, 59(1), 185–203.
- Richards, M.B. et al., 1998. Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet*, 62(3), 241–260.
- Richards, M. et al., 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet*, 67(5), 1251–1276.
- Richards, M. et al., 2002. In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet*, 71(5), 1168–1174.
- Richards, M., 2003. The Neolithic invasion of Europe. *ANNU REV ANTHROPOLOGY*, 32, 135–162.
- Rieder, M.J. et al., 1998. Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Res*, 26(4), 967–973.
- Roberts, B.W., Thornton, C.P. & Pigott, V.C., 2009. Development of metallurgy in Eurasia. *Antiquity*, 83(June), 1012–1022.
- Roostalu, U. et al., 2007. Origin and Expansion of Haplogroup H, the Dominant Human Mitochondrial DNA Lineage in West Eurasia: The Near Eastern and Caucasian Perspective. *Mol Biol Evol*, 24(2), 436–448.
- Saillard, J. et al., 2000. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *Am J Hum Genet*, 67(3), 718–726.
- Sajantila, A. et al., 1995. Genes and languages in Europe: an analysis of mitochondrial lineages. *Genome Res*, 5(1), 42–52.
- Semino, O. et al., 2000. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science*, 290(5494), 1155–1159.
- Smith, J.M. & Haigh, J., 2007. The hitch-hiking effect of a favourable gene. *Genetical research*, 89(5-6), 391–403.
- Soares, P. et al., 2009. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet*, 84(6), 740–759.
- Soares, P. et al., 2010. The archaeogenetics of Europe. *Curr Biol*, 20(4), 174–183.
- Stringer, C.B., 1992. Reconstructing recent human evolution. *Philos Trans R Soc Lond B Biol Sci*, 337(1280), 217–224.
- Sutovsky, P. et al., 1999. Ubiquitin tag for sperm mitochondria. *Nature*, 402(6760), 371–372.
- Zuckerandl, E. & Pauling, L., 1962. Molecular disease, evolution, and genetic heterogeneity. In M. Kasha & B. Pullman, eds. *Horizons in Biochemistry*. New York: Academic Press, 189–225.

- Tambets, K. et al., 2003. Complex signals for population expansions in Europe and beyond. In P. Bellwood & C. Renfrew, eds. *Examining the farming/language dispersal hypothesis*. Cambridge: Cambridge University Press, 449–458.
- Tambets, K. et al., 2004. The Western and Eastern Roots of the Saami--the Story of Genetic “Outliers” Told by Mitochondrial DNA and Y Chromosomes. *Am J Hum Genet*, 74(4), 661–682.
- Tambets, K., 2004. *Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogenetic approach*. Tartu: University of Tartu.
- Tishkoff, S.A. & Verrelli, B.C., 2003. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual review of genomics and human genetics*, 4, 293–340.
- Torrioni, A. et al., 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 144(4), 1835–1850.
- Torrioni, A. et al., 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet*, 62(5), 1137–1152.
- Torrioni, A. et al., 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet*, 69(4), 844–852.
- Torrioni, A. et al., 2006. Harvesting the fruit of the human mtDNA tree. *Trends Genet.*, 22(6), 339–345.
- Trinkaus, E., Holliday, T.W., et al., 2000. The human remains from Paviland Cave. In: Aldhouse-Green, S. (Ed.), *Paviland Cave and the “Red Lady”*: A Definitive Report. Western Academic & Specialist Press Limited, Bristol, 141–204.
- Trinkaus, E., 2005. EARLY MODERN HUMANS. *Annual Review of Anthropology*, 34(1), 207–230.
- Töpf, a L. et al., 2007. Ancient human mtDNA genotypes from England reveal lost variation over the last millennium. *Biology letters*, 3(5), 550–553.
- Wallace, D.C., Brown, M.D. & Lott, M.T., 1999. Mitochondrial DNA variation in human evolution and disease. *Gene*, 238(1), 211–230.
- Wang, N. et al., 2002. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *American journal of human genetics*, 71(5), 1227–1234.
- Weale, M.E. et al., 2002. Y chromosome evidence for Anglo-Saxon mass migration. *Mol Biol Evol*, 19(7), 1008–1021.
- Velichko, A.A., Kurenkova, E.I. & Dolukhanov, P.M., 2009. Human socio-economic adaptation to environment in Late Palaeolithic, Mesolithic and Neolithic Eastern Europe. *Quaternary International*, 1–9.
- Vernesi, C. et al., 2001. Genetic characterization of the body attributed to the evangelist Luke. *Proc Natl Acad Sci U S A*, 98(23), 13460–13463.
- Vigilant, L. et al., 1991. African populations and the evolution of human mitochondrial DNA. *Science*, 253(5027), 1503–1507.
- Wilson, J.F. et al., 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci U S A*, 98(9), 5078–83.

- Winney, B. et al., 2012. People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *European journal of human genetics : EJHG*, 20(2), 203–210.
- Yunusbayev, B., Metspalu M., Järve M. et al., 2012. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Mol Biol Evol* 29, 359–365.

Kasutatud raamatud

- Ammerman, A.J. & Cavalli-Sforza, L.L., 1984. *The Neolithic transition and the genetics of populations in Europe*, Princeton: Princeton University Press.
- Ammerman, A.J. & Cavalli-Sforza, L.L., 1984. *The Neolithic transition and the genetics of populations in Europe*, Princeton University Press, Princeton.
- Cavalli-Sforza LL, Menozzi P, Piazza A., 1994. “*The History and Geography of Human Genes*“. Princeton University Press, Princeton.
- Hall Richard, 2008. “Avastusretk Viikingite maailma“. Kirjastus KOOLIBRI.
- Haywood John, 1997. “*Atlas of World History*“. Barnes&Noble Books, New York.
- Jobling, M. A., M. Hurles, 2004. “*Human evolutionary genetics: origins, peoples & disease*“. New York, Garland Science.
- King Russell, 2007. “Inimkonna rändeatlas“. Kirjastus ASKirjastus Ilo.
- Lynch, F., Aldhouse-Green, S., Davies, J L., 2008. “*Prehistoric Wales*“. Sutton Pub Ltd.
- Roesdahl Else, 2007. “Viikingite maailm“. Kirjastus Tänapäev.
- Stevenson John, 2003. “Euroopa ajalugu“. Kirjastus Varrak.
- Teeple John B., 2004. “Maailma ajalugu“. Kirjastus Varrak.
- The Times, 2004. “Atlas of the World“. Times Books, London.
- Wolf Alex, 2010. “Lühike maailma ajalugu“. Kirjastus ERSEN.

Kasutatud veebiaadressid

- <http://www.mitomap.org>
- <http://phylotree.org/>
- <http://www.mtdnacommunity.org/>
- <http://www.1000genomes.org/>
- <http://hapmap.ncbi.nlm.nih.gov/>
- http://en.wikipedia.org/wiki/File:Wales_in_the_UK_and_Europe.svg
- <http://www.ethnologue.com>
- <http://www.peopleofthebritishisles.org>
- www.fluxus-engineering.com
- <http://harpending.humanevo.utah.edu/popstr/>

LISAD

LISA 1. MtDNA amplifitseerimiseks ja sekveneerimiseks kasutatud praimerite iseloomustus (Rieder et al., 1998).

Nr	Praimeri nimi	3' positsioon	Praimeri järjestus 5'-3' suunas	Praimeri pikkus	Praimeri nimi	3' positsioon	Praimeri järjestus 5'-3' suunas	Praimeri pikkus	Produkti pikkus, ap
1	1F	611	CTCCTCAAAGCAATACACTG	20	1R	1411	TGCTAAATCCACCTTCGACC	20	840
2	2F	1245	CGATCAACCTCACCACCTCT	20	2R	2007	TGGACAACCAGCTATCACCA	20	802
3	3F	1854	GGACTAACCCTATACCTTCTGC	23	3R	2669	GGCAGGTCAATTCCTGGT	20	858
4	4F	2499	AAATCTTACCCCGCTGTTT	20	4R	3346	AGGAATGCCATTGCGATTAG	20	887
5	5F	3169	TACTTCACAAAGCGCCTTCC	20	5R	3961	ATGAAGAATAGGGCGAAGGG	20	832
6	6F	3796	TGGCTCCTTTAACCTCTCCA	20	6R	4654	AAGGATTATGGATGCGGTTG	20	898
7	7F	4485	ACTAATTAATCCCCTGGCCC	20	7R	5420	CCTGGGGTGGGTTTTGTATG	20	975
8	8F	5255	CTAACCGGCTTTTTGCC	18	8R	6031	ACCTAGAAGGTTGCCTGGCT	20	814
9	9F	5855	GAGGCCTAACCCCTGTCTTT	20	9R	6642	ATTCCGAAGCCTGGTAGGAT	20	827
10	10F	6469	CTCTTCGTCTGATCCGTCCT	20	10R	7315	AGCGAAGGCTTCTCAAATCA	20	886
11	11F	7148	ACGCCAAAATCCATTTCACT	20	11R	8095	CGGGAATTGCATCTGTTTTT	20	987
12	12F	7937	ACGAGTACACCGACTACGGC	20	12R	8797	TGGGTGGTTGGTGAAATGA	20	900
13	13F	8621	TTTCCCCTCTATTGATCCC	20	13R	9397	GTGGCCTTGGTATGTGCTTT	20	816
14	14F	9230	CCCACCAATCACATGCCTAT	20	14R	10130	TGTAGCCGTTGAGTTGTGGT	20	940
15	15F	9989	TCTCCATCTATTGATGAGGGTCT	23	15R	10837	AATTAGGCTGTGGGTGGTTG	20	891
16	16F	10672	GCCATACTAGTCTTTGCCGC	20	16R	11472	TTGAGAATGAGTGTGAGGCG	20	840
17	17F	11314	TCACTCTCACTGCCCAAGAA	20	17R	12076	GGAGAATGGGGGATAGGTGT	20	802
18	18F	11948	TATCACTCTCCTACTTACAG	20	18R	12772	AGAAGGTTATAATTCCTACG	20	864
19	19F	12571	AAACAACCCAGCTCTCCCTAA	21	19R	13507	TCGATGATGTGGTCTTTGGA	20	977
20	20F	13338	ACATCTGTACCCACGCCTTC	20	20R	14268	AGAGGGGTCAGGGTTCATTC	20	970
21	21F	14000	GCATAATTAACCTTTACTTC	20	21R	14998	AGAATATTGAGGCGCCATTG	20	1038
22	22F	14856	TGAAACTTCGGCTCACTCCT	20	22R	15978	AGCTTTGGGTGCTAATGGTG	20	1162
23	23F	15811	TCATTGGACAAGTAGCATCC	20	23R	5	GAGTGGTTAATAGGGTGATAG	21	804
24	24F	16420	CACCATTCTCCGTGAAATCA	20	24R	775	AGGCTAAGCGTTTTGAGCTG	20	964

LISA 2. MtDNA RFLP analüüsi tulemused. Kui uuritavas mtDNA järjestuses on positsioonis 7028 nukleotiid C, siis restriктаas ei lõika (-) ja kui on T siis lõikab (+).

ID	HVS-I rCRS suhtes	HVS-I RRSRS suhtes	H AluI- 7028CT
BK43	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK46	C16256T T16325C	A16129G T16187C C16189T T16223C G16230A C16256T T16278C C16311T T16325C C16519T	-
BK47	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK49	T16172C	A16129G T16172C T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK58	A16240G T16298C	T16187C C16189T T16223C A16240G G16230A T16278C T16298C C16311T C16519T	+
BK66	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK68	C16104T C16176T G16274A T16311C	C16104T A16129G C16176T T16187C C16189T T16223C G16230A G16274A T16278C C16519T	-
BK81	C16111T T16311C	C16111T A16129G T16187C C16189T T16223C G16230A T16278C C16519T	+
BK88	G16129A A16316G	T16187C C16189T T16223C G16230A T16278C C16311T A16316G C16519T	-
BK89	C16179T T16356C	C16179T T16187C C16189T T16223C G16230A T16278C C16311T T16356C C16519T	+
BK096	C16223T A16235G C16291T A16293G	A16129G T16187C C16189T G16230A A16235G T16278C C16291T A16293G C16311T C16519T	-
BK098	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK102	C16192T	A16129G T16187C C16189T C16192T T16223C G16230A T16278C C16311T C16519T	-
BK104	G16129A	T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-

BK106	C16192T T16304C	A16129G T16187C C16189T C16192T T16223C G16230A T16278C T16304C C16311T C16519T	-
BK117	C16239T	A16129G T16187C C16189T T16223C G16230A C16239T T16278C C16311T C16519T	-
BK118	C16291T T16304C	A16129G T16187C C16189T T16223C G16230A T16278C C16291T T16304C C16311T C16519T	-
BK120	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
BK130	A16293G T16311C C16354T	A16129G T16187C C16189T T16223C G16230A T16278C A16293G C16354T C16519T	-
BK133	T16189C T16356C	A16129G T16187C T16223C G16230A T16278C C16311T T16356C C16519T	-
BK134	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK137	T16189C C16256T	A16129G T16187C T16223C G16230A C16256T T16278C C16311T C16519T	+
BK143	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
BK148	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
BK149	G16390A	A16129G T16187C C16189T T16223C G16230A T16278C C16311T G16390A C16519T	-
BK155	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
BK156	G16129A T16198C	T16187C C16189T T16198C T16223C G16230A T16278C C16311T C16519T	+
CHF12	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF13	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF15	C16069T T16126C	C16069T T16126C A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	+

CHF22	T16298C	A16129G T16187C C16189T T16223C G16230A T16278C T16298C C16311T C16519T	+
CHF35	T16362C A16482G	A16129G T16187C C16189T T16223C G16230A T16278C C16311T T16362C A16482G C16519T	-
CHF44	T16092C T16140C A16293G T16311C	T16092C A16129G T16140C T16187C C16189T T16223C G16230A T16278C A16293G C16519T	-
CHF56	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF63	T16298C	A16129G T16187C C16189T T16223C G16230A T16278C T16298C C16311T C16519T	+
CHF64	T16325C	A16129G T16187C C16189T T16223C G16230A T16278C C16311T T16325C C16519T	-
CHF65	C16168T	A16129G C16168T T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF76	T16092C T16140C A16293G T16311C C16327T	T16092C A16129G T16140C T16187C C16189T T16223C G16230A T16278C A16293G C16311T C16327T C16519T	-
CHF79	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF103	A16162G T16209C	A16129G A16162G T16187C C16189T T16209C T16223C G16230A T16278C C16311T C16519T	-
CHF110	T16311C	A16129G T16187C C16189T T16223C G16230A T16278C C16519T	+
CHF111	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF121	T16093C	T16093C A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF143	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	+
CHF144	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF152	T16304C T16362C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T T16362C C16519T	-

CHF156	T16362C A16482G	A16129G T16187C C16189T T16223C G16230A T16278C C16311T T16362C A16482G C16519T	-
CHF158	T16178C A16220c A16349G	A16129G T16178C T16187C C16189T A16220c T16223C G16230A T16278C C16311T A16349G C16519T	-
CHF161	C16176T	A16129G C16176T T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF170	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
CHF176	T16362C A16482G	A16129G T16187C C16189T T16223C G16230A T16278C C16311T T16362C A16482G C16519T	-
CHF181	C16259T C16261T T16298C 16343t	A16129G T16187C C16189T T16223C G16230A C16259T C16261T T16278C T16298C C16311T 16343t C16519T	+
CHF188	G16129A T16298C	T16187C C16189T T16223C G16230A T16278C T16298C C16311T C16519T	+
CHF202	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16394T C16519T	-
CHF219	C16239T	A16129G T16187C C16189T T16223C G16230A C16239T T16278C C16311T C16519T	-
CHF221	G16145A	A16129G G16145A T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR56	T16311C	A16129G T16187C C16189T T16223C G16230A T16278C C16519T	+
DR64	A16051G A16162G	A16051G A16129G A16162G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR68	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR77	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR105	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR108	T1620C9 T16304C	A16129G T16187C C16189T T16209C T16223C G16230A T16278C T16304C C16311T C16519T	-

DR110	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR115	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR116	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR119	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR120	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR121	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR122	T16189d C16354T	A16129G T16187C 16189dT16223C G16230A T16278C C16311T C16354T C16519T	-
DR134	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR135	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
DR137	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR147	G16129A	T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR148	T16298C G16390A	A16129G T16187C C16189T T16223C G16230A T16278C T16298C C16311T G16390A C16519T	+
DR151	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR152	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
DR161	T16263C	A16129G T16187C C16189T T16223C G16230A T16263C T16278C C16311T C16519T	-
DR172	T16311C	A16129G T16187C C16189T T16223C G16230A T16278C C16519T	+

DR175	A16070G T16304C	A16070G A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
DR180	T16189C T16356C	A16129G T16187C T16223C G16230A T16278C C16311T T16356C C16519T	-
DR181	T16093C T16189C	T16093C A16129G T16187C T16223C G16230A T16278C C16311T C16519T	-
DR185	C16192T	A16129G T16187C C16189T C16192T T16223C G16230A T16278C C16311T C16519T	-
DR187	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
L17	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L40	C16261T C16301T	A16129G T16187C C16189T T16223C G16230A C16261T T16278C C16301T C16311T C16519T	-
L44	C16291T	A16129G T16187C C16189T T16223C G16230A T16278C C16291T C16311T C16519T	-
L45	A16162G T16209C	A16129G A16162G T16187C C16189T T16209C T16223C G16230A T16278C C16311T C16519T	-
L49	A16051G A16162G	A16051G A16129G A16162G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L53	C16186T C16193T A16219G C16260T T16362C A16482G	A16129G C16186T T16187C C16189T C16193T A16219G T16223C G16230A C16260T T16278C C16311T T16362C A16482G C16519T	-
L57	C16294T	A16129G T16187C C16189T T16223C G16230A T16278C C16294T C16311T C16519T	-
L60	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L61	T16189C T16304C	A16129G T16187C T16223C G16230A T16278C T16304C C16311T C16519T	-
L62	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L63	A16293G T16311C	A16129G T16187C C16189T T16223C G16230A T16278C A16293G C16519T	-

L64	A16235G C16291T A16293G	A16129G T16187C C16189T T16223C G16230A A16235G T16278C C16291T A16293G C16311T C16519T	-
L65	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L66	T16304C	A16129G T16187C C16189T T16223C G16230A T16278C T16304C C16311T C16519T	-
L69	T16189C T16304C	A16129G T16187C T16223C G16230A T16278C T16304C C16311T C16519T	-
L74	CRS	A16129G T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-
L75	T16362C A16482G	A16129G T16187C C16189T T16223C G16230A T16278C C16311T T16362C A16482G C16519T	-
L83	T16263C	A16129G T16187C C16189T T16223C G16230A T16263C T16278C C16311T C16519T	-
L84	G16129A	T16187C C16189T T16223C G16230A T16278C C16311T C16519T	-

LISA 3. Hg-de H (a), J (b) ja T (c) mediaanvõrgustike koostamisel nukleotiidpositsioonidele antud kaalud.

a)

HVS-I positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal
16051	10	72g	10	4024	10	8448	10
16070	10	73	10	4095	10	8563	10
16092	10	93	2	4136	3	8578	10
16093	1	146	1	4216	3	8592	10
16104	10	150	1	4336	10	8634	3
16129	1	152	1	4512	10	8790	3
16140	10	186a	10	4646	10	8860	10
16145	2	195	1	4659	3	9055	10
16162	10	204	2	4688	10	9123	10
16168	3	239	10	4702	3	9163	3
16172	1	251	10	4727	10	9254	10
16176	10	263	10	4733	10	9299	10
16178	3	327	10	4745	10	9325	3
16186	3	340	3	4763	10	9380	10
16189	1	408a	10	4769	10	9554	3
16192	1	456	10	4793	10	10394	10
16193	10	477	10	4877	3	10463	10
16209	2	619	3	5004	10	11150	10
16219	10	709	1	5319	10	11253	10
16220c	10	750	10	5460	2	11719	10
16223	1	951	10	5471	10	12684	3
16235	10	961	10	5472	3	12864	10
16239	3	1438	10	5480	10	13056	10
16256	2	2181	3	5785	10	13194	10
16260	2	2259	10	5899.XC	3	13575	3
16261	10	2352	10	5979	10	13680	10
16263	3	2581	10	6152	3	13708	10
16274	2	2619	3	6216	10	13762g	10
16291	1	2626	10	6231	3	14208	3
16293	10	2706	10	6237a	10	14220	10
16294	2	2851	10	6365	10	14325	3
16301	3	3010	10	6380	3	14364	10
16304	10	3145	10	6620	3	14365	10
16311	1	3226	3	6716	10	14497	10
16316	10	3388a	10	6776	10	14582	10
16325	2	3397	3	7028	10	14766	10
16327	3	3434	3	7148	10	14872	10
16349	3	3525a	10	7337	10	15326	10
16354	10	3606	3	7630	3	15670	3
16356	10	3796	10	8020	10	15713g	10
16362	1	3861	10	8164	3	15781	10
16390	1	3915	10	8381	10	15833	10
16482	10	3992	10	8406	3	15884	2

b)

HVS-I positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal
16069	10	150	01	9632	03
16126	10	152	01	9635	03
16145	03	185	03	10172	03
16172	01	188	03	10398	10
16193	03	195	01	10499	03
16221	03	215	03	11251	10
16222	03	228	03	11377	03
16224	01	242	03	11623	02
16231	03	295	10	12007	03
16235	01	319	03	12083	03
16244	01	462	03	12127	03
16261	03	482	03	12612	10
16271	01	485	01	13656	03
16278	01	489	10	13681	03
16319	01	513	03	13708	10
16320	01	574	01	13722	03
16390	01	709	01	13879	03
		2158	03	13899	01
		3010	03	13934	03
		3308	01	14029	03
		3392	01	14133	03
		3394	03	14401	01
		5054	01	14569	01
		5460	01	14798	03
		5633	03	14854	03
		6345	03	14894	01
		6464	03	14905	01
		6554	03	15001	03
		7299	03	15013	03
		7302	03	15257	03
		7391	03	15452	10
		7476	03	15721	01
		7789	03	15812	03
		8269	03	15924	01
		8557	03		

c)

HVS-I positsioonid	Antud kaal	HVS-II ja kodeeriva ala positsioonid	Antud kaal
16104	03	152	01
16126	10	195	01
16163	03	204	01
16182.1C	03	207	01
16183.1C	03	709	10
16186	03	930	03
16189	01	1888	10
16192	01	2850	03
16256	01	4639	03
16261	01	4688	03
16278	01	4917	10
16294	10	5147	03
16296	03	5277	03
16298	03	5426	03
16304	03	5800	01
16324	03	6489a	03
16362	01	6671	03
		7022	03
		8270	03
		8281-8219d	03
		8697	10
		9254	03
		10463	10
		11242g	03
		11251	10
		12633a	03
		11812	03
		12171	01
		13368	10
		13965	03
		14053	01
		14178	03
		14233	03
		14687	03

LISA 4. Käesolevas töös uuritud waleslaste mtDNA HVS-I ja kodeeriva piirkonna andmed Cambridge referentsjärjestuse (rCRS) ja inimese rekonstreeritud referentsjärjestuse (RSRS) suhtes. Tabelid: a) hg H, b) hg U, c) hg-d J ja T, d) hg-d V, HV, I jne., e) täisjärjestused. Punasega on tabelites märgitud erinevused kahe referentsjärjestuste vahel.

Tabelid saab alla laadida siit: <http://www.evolutsoon.ebc.ee/wales/>

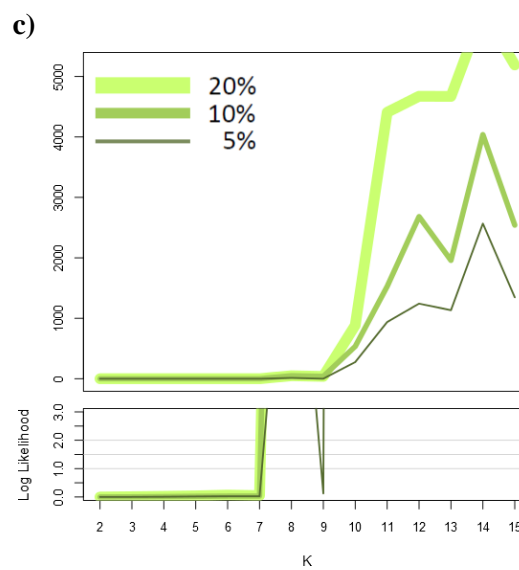
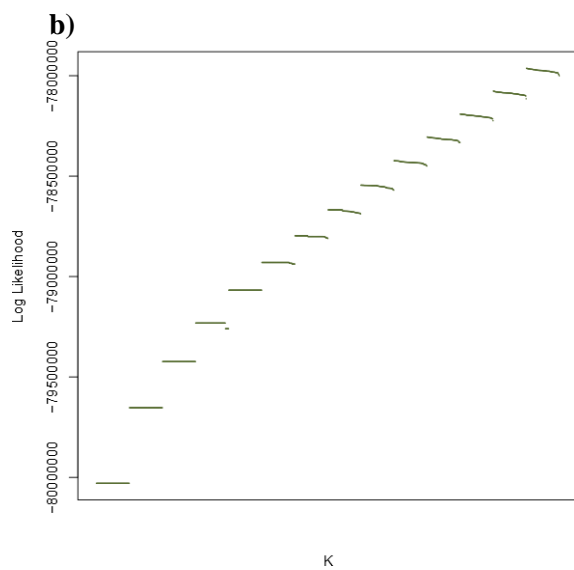
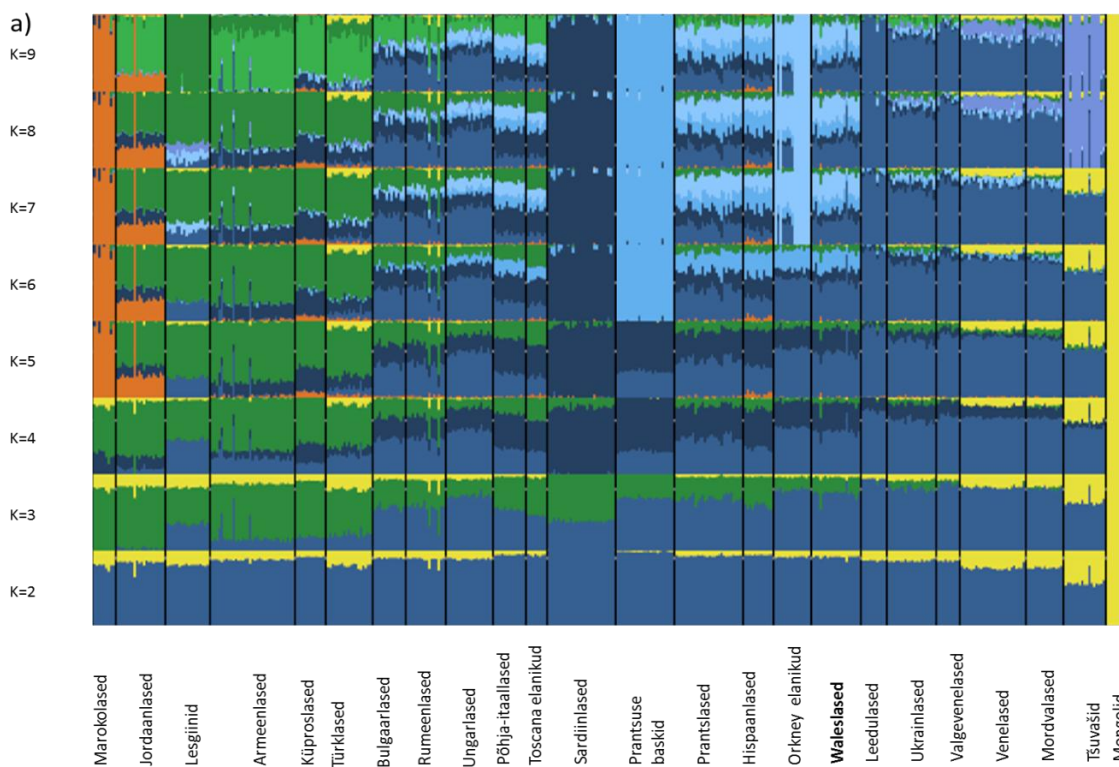
LISA 5. MtDNA põhikomponentanalüüsis (PCA) kasutatud populatsioonide haplogruppide sagedused ja 95% usalduspiirid. Helehalliga on esiletoodud need populatsioonid, kus waleslastega võrreldes oli erinevus kõige suurem ja statistiliselt usaldusväärne. Tabeli andmed pärinevad (Helgason et al., 2001; Richard et al., 2007; García et al., 2011 ja viited seal sees).

Parema resolutsiooniga tabeli saab alla laadida siit: <http://www.evolutsoon.ebc.ee/wales/>

Populatsioon	Haplogruppid															
	HHV	VHV0	U3	U4	U5	U8	U**	K	T1	T2	J	I	W	X	MU0**	
Loode hispaanlased (baskideta)	N=533	254	65	6	9	57	1	8	29	3	40	30	13	3	2	11
	Sagedus	0,477	0,122	0,011	0,016	0,108	0,002	0,016	0,055	0,006	0,075	0,057	0,025	0,005	0,004	0,021
	Usalduspiirid	(.435 - .519)	(.097 - .153)	(.005 - .023)	(.009 - .031)	(.084 - .137)	(.001 - .011)	(.008 - .003)	(.039 - .078)	(.002 - .017)	(.056 - .101)	(.004 - .008)	(.015 - .042)	(.002 - .016)	(.001 - .014)	(.012 - .038)
Hispaania baskid	N=462	269	42	2	2	40	6	6	25	4	25	22	1	4	11	2
	Sagedus	0,583	0,091	0,005	0,005	0,088	0,013	0,014	0,054	0,009	0,054	0,047	0,003	0,010	0,024	0,005
	Usalduspiirid	(.537 - .627)	(.068 - .012)	(.002 - .017)	(.001 - .016)	(.065 - .117)	(.006 - .027)	(.007 - .029)	(.037 - .079)	(.004 - .022)	(.037 - .079)	(.031 - .007)	(.001 - .013)	(.004 - .023)	(.014 - .043)	(.001 - .016)
Kirde hispaanlased	N=363	146	20	5	10	32	1	10	24	4	41	33	5	10	8	20
	Sagedus	0,403	0,054	0,014	0,026	0,088	0,002	0,028	0,066	0,012	0,113	0,092	0,014	0,027	0,023	0,054
	Usalduspiirid	(.354 - .454)	(.035 - .082)	(.006 - .033)	(.014 - .048)	(.063 - .122)	(0 - .014)	(.016 - .051)	(.045 - .096)	(.005 - .03)	(.084 - .015)	(.066 - .126)	(.006 - .032)	(.015 - .049)	(.012 - .044)	(.035 - .083)
Prantsuse baskid	N=81	48	4	0	5	2	0	0	3	3	0	14	0	0	2	0
	Sagedus	0,593	0,049	0,000	0,062	0,025	0,000	0,000	0,037	0,037	0,000	0,172	0,000	0,000	0,025	0,000
	Usalduspiirid	(.484 - .694)	(.02 - .012)	(0 - .044)	(.027 - .137)	(.008 - .086)	(0 - .044)	(0 - .044)	(.013 - .103)	(.013 - .103)	(0 - .044)	(.106 - .269)	(0 - .044)	(0 - .044)	(.008 - .086)	(0 - .044)
Prantslased	N=825	372	53	6	18	61	7	15	80	13	57	63	18	10	9	45
	Sagedus	0,452	0,065	0,007	0,021	0,074	0,008	0,018	0,097	0,016	0,069	0,076	0,022	0,012	0,011	0,055
	Usalduspiirid	(.418 - .486)	(.05 - .083)	(.003 - .015)	(.014 - .034)	(.058 - .093)	(.004 - .017)	(.011 - .003)	(.078 - .119)	(.009 - .027)	(.054 - .089)	(.006 - .096)	(.014 - .034)	(.007 - .023)	(.006 - .021)	(.041 - .073)
Bretoonid	N=298	142	10	6	6	25	2	6	35	9	21	18	6	7	1	1
	Sagedus	0,476	0,034	0,022	0,020	0,084	0,007	0,021	0,119	0,030	0,069	0,061	0,019	0,025	0,004	0,005
	Usalduspiirid	(.42 - .532)	(.019 - .061)	(.01 - .045)	(.009 - .042)	(.058 - .121)	(.002 - .024)	(.01 - .044)	(.087 - .161)	(.016 - .056)	(.046 - .104)	(.039 - .094)	(.009 - .042)	(.013 - .05)	(.001 - .02)	(.001 - .021)
Irlased	N=302	136	20	2	4	26	0	3	35	5	25	31	7	6	2	2
	Sagedus	0,449	0,067	0,008	0,012	0,087	0,000	0,010	0,117	0,015	0,081	0,102	0,024	0,018	0,005	0,005
	Usalduspiirid	(.394 - .505)	(.044 - .101)	(.003 - .026)	(.005 - .032)	(.06 - .124)	(0 - .012)	(.004 - .029)	(.085 - .158)	(.006 - .036)	(.056 - .118)	(.073 - .141)	(.012 - .048)	(.008 - .041)	(.002 - .022)	(.002 - .022)
Šotlased	N=891	396	38	13	23	61	2	7	58	21	68	130	38	10	17	7
	Sagedus	0,445	0,043	0,015	0,026	0,069	0,002	0,008	0,065	0,024	0,076	0,146	0,043	0,011	0,019	0,008
	Usalduspiirid	(.413 - .478)	(.032 - .058)	(.009 - .025)	(.017 - .039)	(.054 - .088)	(.001 - .008)	(.004 - .016)	(.051 - .083)	(.016 - .036)	(.06 - .095)	(.124 - .171)	(.032 - .058)	(.006 - .02)	(.012 - .03)	(.004 - .016)
Inglased	N=142	68	6	4	1	12	0	4	7	3	13	16	3	2	1	2
	Sagedus	0,480	0,042	0,028	0,007	0,085	0,000	0,028	0,049	0,021	0,091	0,113	0,021	0,014	0,007	0,014
	Usalduspiirid	(.399 - .562)	(.02 - .089)	(.011 - .07)	(.002 - .038)	(.05 - .143)	(0 - .025)	(.011 - .07)	(.024 - .098)	(.008 - .06)	(.054 - .015)	(.071 - .176)	(.008 - .06)	(.004 - .049)	(.002 - .038)	(.004 - .049)
Waleslased*	N=268	150	11	0	1	16	1	0	21	4	12	36	7	3	2	4
	Sagedus	0,559	0,042	0,000	0,003	0,059	0,003	0,000	0,078	0,016	0,045	0,134	0,025	0,012	0,009	0,017
	Usalduspiirid	(.499 - .617)	(.024 - .073)	(0 - .014)	(.001 - .019)	(.037 - .094)	(.001 - .019)	(0 - .014)	(.052 - .116)	(.007 - .039)	(.026 - .077)	(.098 - .018)	(.012 - .051)	(.004 - .033)	(.003 - .028)	(.007 - .004)
Orkney elanikud	N=152	73	2	0	1	18	0	0	10	5	4	20	4	3	11	1
	Sagedus	0,480	0,013	0,000	0,007	0,118	0,000	0,000	0,066	0,033	0,026	0,132	0,026	0,020	0,072	0,007
	Usalduspiirid	(.402 - .559)	(.004 - .046)	(0 - .024)	(.002 - .036)	(.076 - .018)	(0 - .024)	(0 - .024)	(.036 - .117)	(.015 - .075)	(.011 - .066)	(.087 - .195)	(.011 - .066)	(.007 - .056)	(.041 - .125)	(.002 - .036)
Kornid	N=92	46	2	0	5	6	0	0	4	2	4	18	4	0	0	1
	Sagedus	0,500	0,022	0,000	0,054	0,065	0,000	0,000	0,043	0,022	0,044	0,196	0,044	0,000	0,000	0,010
	Usalduspiirid	(.4 - .6)	(.007 - .076)	(0 - .039)	(.024 - .121)	(.031 - .135)	(0 - .039)	(0 - .039)	(.017 - .106)	(.007 - .076)	(.018 - .107)	(.128 - .289)	(.018 - .107)	(0 - .039)	(0 - .039)	(.002 - .057)
Šotimaa Läänesaarte ja Skye elanikud	N=230	76	5	3	2	14	1	6	31	8	22	31	15	5	5	6
	Sagedus	0,317	0,022	0,026	0,009	0,078	0,009	0,026	0,135	0,035	0,096	0,135	0,065	0,022	0,017	0,009
	Usalduspiirid	(.273 - .394)	(.01 - .005)	(.005 - .037)	(.003 - .031)	(.037 - .01)	(.001 - .024)	(.012 - .056)	(.097 - .185)	(.018 - .067)	(.064 - .141)	(.097 - .185)	(.004 - .105)	(.001 - .005)	(.001 - .005)	(.012 - .056)
Sakslased	N=778	400	38	6	32	54	2	11	40	15	64	63	21	12	14	9
	Sagedus	0,514	0,049	0,008	0,041	0,070	0,002	0,014	0,051	0,020	0,082	0,081	0,027	0,016	0,018	0,011
	Usalduspiirid	(.479 - .549)	(.036 - .067)	(.004 - .017)	(.029 - .057)	(.054 - .09)	(.001 - .008)	(.008 - .025)	(.038 - .069)	(.012 - .032)	(.065 - .103)	(.064 - .102)	(.018 - .041)	(.009 - .027)	(.011 - .03)	(.006 - .021)
Šveitslased	N=224	114	11	1	8	17	2	3	12	4	18	24	3	4	1	2
	Sagedus	0,509	0,051	0,007	0,034	0,074	0,007	0,013	0,055	0,020	0,079	0,108	0,014	0,017	0,003	0,010
	Usalduspiirid	(.444 - .574)	(.029 - .088)	(.002 - .028)	(.017 - .067)	(.047 - .116)	(.002 - .029)	(.005 - .038)	(.032 - .092)	(.008 - .047)	(.005 - .121)	(.074 - .155)	(.005 - .039)	(.007 - .043)	(.001 - .022)	(.003 - .033)
Poolakad	N=436	205	20	2	22	38	2	5	15	9	41	34	8	16	8	11
	Sagedus	0,471	0,046	0,005	0,050	0,087	0,005	0,011	0,034	0,021	0,094	0,078	0,018	0,037	0,018	0,025
	Usalduspiirid	(.425 - .518)	(.03 - .007)	(.002 - .017)	(.033 - .075)	(.064 - .117)	(.002 - .017)	(.005 - .026)	(.021 - .055)	(.011 - .039)	(.007 - .125)	(.056 - .107)	(.009 - .035)	(.023 - .059)	(.009 - .035)	(.014 - .044)
Austrialased	N=99	48	1	1	4	8	1	1	7	2	7	11	2	1	1	3
	Sagedus	0,485	0,010	0,010	0,040	0,081	0,010	0,010	0,071	0,020	0,070	0,111	0,020	0,010	0,010	0,030
	Usalduspiirid	(.389 - .582)	(.002 - .054)	(.002 - .054)	(.016 - .099)	(.042 - .152)	(.002 - .054)	(.002 - .054)	(.035 - .139)	(.006 - .07)	(.035 - .138)	(.063 - .188)	(.006 - .07)	(.002 - .054)	(.002 - .054)	(.011 - .085)
Norralased	N=639	300	26	10	17	74	9	2	35	6	50	67	13	12	3	4
	Sagedus	0,470	0,041	0,016	0,027	0,116	0,014	0,003	0,055	0,009	0,078	0,105	0,020	0,019	0,005	0,006
	Usalduspiirid	(.432 - .509)	(.028 - .059)	(.009 - .029)	(.017 - .042)	(.093 - .143)	(.007 - .026)	(.001 - .011)	(.004 - .075)	(.004 - .02)	(.06 - .102)	(.084 - .131)	(.012 - .034)	(.011 - .033)	(.002 - .014)	(.002 - .016)
Soomlased	N=613	248	40	0	6	136	0	8	19	7	15	34	25	52	6	12
	Sagedus	0,405	0,065	0,000	0,010	0,222	0,000	0,013	0,031	0,011	0,024	0,055	0,041	0,085	0,010	0,020
	Usalduspiirid	(.367 - .444)	(.048 - .088)	(0 - .006)	(.005 - .021)	(.191 - .256)	(0 - .006)	(.007 - .026)	(.02 - .048)	(.006 - .023)	(.015 - .04)	(.04 - .077)	(.028 - .006)	(.065 - .011)	(.005 - .021)	(.012 - .034)
Islandlased	N=448	209	7	10	10	34	0	0	32	2	45	63	23	1	7	2
	Sagedus	0,467	0,016	0,022	0,022	0,076	0,000	0,000	0,071	0,004	0,100	0,140	0,051	0,002	0,016	0,005
	Usalduspiirid	(.421 - .513)	(.008 - .032)	(.012 - .041)	(.012 - .041)	(.055 - .104)	(0 - .008)	(0 - .008)	(.051 - .099)	(.001 - .016)	(.076 - .132)	(.111 - .175)	(.035 - .078)	(.001 - .012)	(.008 - .032)	(.002 - .017)

* sh käesoleva töö 176 waleslast
 ** U1,U2,U6,U7
 *** L1,L2,L3,D/G,M,C,A,N1,F,B,R0,N*M,L3*

LISA 6. a) ADMIXTURE analüüsi abil ja 660 000 autosoomse SNP põhjal tuletatud Walesi populatsiooni struktuur Lääne-Euraasia populatsioonide kontekstis kahe kuni üheksa eellas komponendi puhul ($K=2$ kuni $K=9$). Mudelid $K>9$ jäeti ebausaldusväärsetena kõrvale. Hüpoteetilistest eellaspopulatsioonidest pärit komponendid on tähistatud eri värvidega ja esitatud vertikaalselt paigutatud tulpadena, millest igüks vastab ühele kindlale uuritud indiviidile. Erinevatesse populatsioonidesse kuuluvad invidiidid on eraldatud musta vertikaalse joonega. **b)** ja **c)** ADMIXTURE analüüsi *log-likelihood* skoorid (LLs) erinevate K väärtuste juures. Madala K väärtuse juures saavutasid kõik sada analüüsi/ arvutust (*run*) samasuguse või väga sarnased LLs-d, kuid kõrgemate K väärtuste juures LLs-d varieeruvad. Joonisel c on lisaks näidatud K -de LLs-de hajuvus vastavalt 5%, 10% ja 20% kõrgeimate LLs-ide fraktsioonide seas.



LIHTLITSENTS

Lihlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina **Anu Solnik** (sünnikuupäev: 04.05.1981)

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

Wales'i populatsiooni geneetiline analüüs emaliinide ja autosoomsete markerite varieeruvuse alusel Lääne-Euraasia kontekstis,

mille juhendajad on **Ene Metspalu ja Kristiina Tambets,**

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

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

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

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

Tartus, 21.08.2013