

**TARTU ÜLIKOOL
BIOLOOGIA-GEOGRAAFIA TEADUSKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
EVOLUTSIOONILISE BIOLOOGIA ÕPPETOOL**

Jana Eensalu

Y kromosoomi haplogruppide E ja R1b fülogeograafia

Magistritöö

Juhendajad: vanemteadur Maere Reidla
vanemteadur Toomas Kivisild

Tartu 2005

SISUKORD

KASUTATUD LÜHENDID	3
MÕISTED	4
SISSEJUHATUS	5
1. KIRJANDUSE ÜLEVAADE	6
1.1 Y KROMOSOOM.....	6
1.2 Y KROMOSOOMI MARKERID	7
1.3 EUROOPA ASUSTAMINE ANATOOMILISELT KAASAEGSE INIMESE POOLT	9
1.4 HAPLOGRUPP R	12
1.5 HAPLOGRUPP E	15
2. MATERJAL JA METOODIKA.....	19
2.1 KATSEMATERJAL.....	19
2.1.1 Haplogrupp R1b3.....	19
2.1.2 Haplogrupp E.....	19
2.2 DNA AMPLIFIKATSIOON.....	20
2.3 AGAROOSGEEL-ELEKTROFOREES	23
2.3.1 YAP-i määramine	24
2.4 RFLP ANALÜÜS	24
2.5 SEKVENEERIMINE.....	25
2.5.1 PCR-i produkti puhastamine.....	25
2.5.2 Sekveneerimine.....	26
2.6 KASUTATUD MEETODID JA PROGRAMMID	27
3. TULEMUSED	29
3.1 HAPLOGRUPP R1B3	29
3.2 HAPLOGRUPP E	36
KOKKUVÕTE	42
SUMMARY	43
TÄNUSÕNAD	45
KASUTATUD KIRJANDUS	46
LISAD	54

KASUTATUD LÜHENDID

AZF	asoospermia faktor (<i>azoospermia factor</i>)
Bp	aluspaar (<i>base pair</i>)
ddNTP	didesoksünukleotiidtrifosfaat
DNA	desoksüribonukleiinhape
DYS	Y kromosoomi DNA lühike kordusjärjestus (<i>DNA Y-chromosome Short Tandem Repeat</i>)
dNTP	desoksünukleotiidtrifosfaat
EDTA	etüleendiamiintetraatsetaat
F praimer	juhtiva ahela praimer
Hg	haplogrupp
Ht	haplotüüp
LGM	viimase jääaja maksimum (<i>last glacial maximum</i>)
Mb	Mega aluspaar
mtDNA	mitokondriaalne DNA
np	nukleotiidi positsioon
NRY	Y kromosoomi mitterekombineeruv ala (<i>nonrecombining region of Y</i>)
PAR	pseudoautosomaalne regioon (<i>pseudoautosomal region</i>)
PCR	polümeraasne ahelreaktsioon (<i>polymerase chain reaction</i>)
RFLP	restriktsioonifragmentide pikkuse polümorfism (<i>restriction fragment lenght polymorphism</i>)
rpm	pööret minutis (<i>rotation per minute</i>)
R praimer	mahajääva ahela praimer
SNP	ühenukleotiidiline polümorfism (<i>single nucleotide polymorphism</i>)
SRY	Y kromosoomi sugu determineeriv regioon (<i>sex determinig region of Y</i>)
STR	lühike kordusjärjestus (<i>short tandem repeat</i>)
TMRCa	aeg viimase ühise eellaseni (<i>time to the most recent common ancestor</i>)
YAP	Y kromosoomi Alu polümorfism (<i>Y chromosome Alu polymorphism</i>)
YCC	Y Kromosoomi Konsortsium (<i>Y Chromosome Consortium</i>)
Yp	Y kromosoomi lühike õlg
Yq	Y kromosoomi pikk õlg

MÕISTED

Haplotüüp – haplogrupi-sisene alamliin, mis on defineeritud Y kromosoomi mitterekombineeruva ala STR-de alleelide kombinatsiooniga

Haplogrupp – bialleelsete markeritega (SNP-dega) defineeritud monofüleetiline üksus.

Klaad - monofüleetiline üksus

Geenitriiv - alleelisageduste juhuslik muutumine põlvkonnast põlvkonda

Rajaja-efekt - geenitriivi üks mehhanisme, mis seisneb selles, et emapopulatsioonist irdunud üksikud isendid panevad aluse tütarpopulatsioonile, mille geneetiline struktuur erineb oluliselt emapopulatsiooni omast

Geenivoog (migratsioon) – alleelisageduste muutumine inimeste liikumise tõttu populatsioonide vahel

Pudelikaelaefekt – ajutise ehk mööduva geneetilise triivi erijuht, kus populatsiooni ajutine arvukuse vähenemine põhjustab juhusliku alleelisageduste muutumise

Efektiivne populatsiooni suurus – idealiseeritud populatsiooni teoreetiline suurus, millel on samas suurusjärgus geenitriiv kui reaalsel populatsioonil

SISSEJUHATUS

Inimpopulatsioonide päritolu ja neis toimuvate evolutsiooniprotsesside uurimisel kasutatakse arheoloogiliste ja keeleuuringute kõrval üha rohkem populatsioonigeneetilisi uurimusi. Viimastes on kasutamist leidnud ka Y kromosomaalne DNA, mille abil saab iseloomustada isaliini ajalugu.

Molekulaarbioloogide käsutuses on inimese DNA koos sinna kogunenud mutatsioonidega. Mutatsioonide abil on võimalik grupeerida inimesi ning taastada nende ühist põlvnemislugu, samuti uurida populatsioonide päritolu ning saada teadmisi migratsioonide kohta.

Kõige rohkem on kasutatud molekulaargeneetilistes uuringutes mitokondriaalset DNA-d (mtDNA). Hiljem on ka alustatud Y kromosoomi uuringuid. Y kromosoom sisaldab võrreldes mtDNA-ga palju laialdasemat polümorfismide spektrit. Inimese isaliinide fülogeneesi uuringutes on põhiliselt kasutatud konservatiivseid punktmutatsioone, insertioone, deletsioone ja teiselt poolt väga varieeruvaid lühikesi kordusjärjestusi (peamiselt di-, tri- ja tetranukleotiidseid kordusi). Vähem on kasutatud penta- ja heksanukleotiidseid kordusjärjestusi. Need võivad olla aga täiendavaks võimaluseks fülogeneetilises analüüsis.

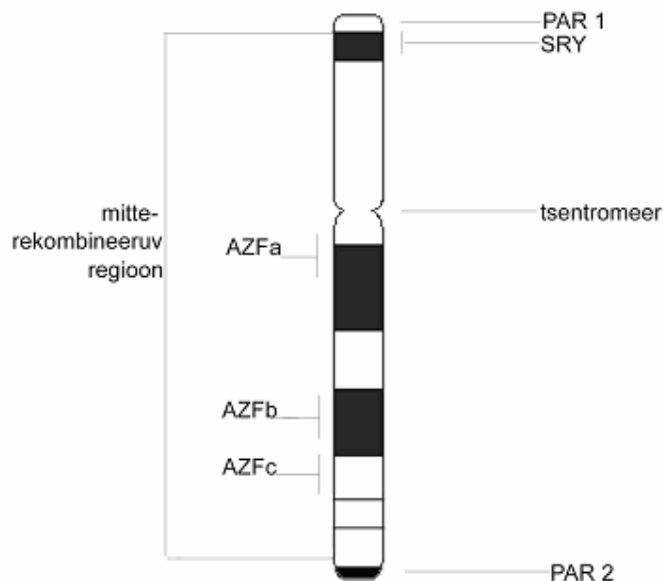
Käesoleva töö eesmärgiks on uurida võrdlevalt eurooplaste paleoliitiliseks pärandiks peetava Y kromosomaalse haplogrupi (hg) R1b ning neoliitilise immigrandi staatusesse kuuluva haplogrupi E fülogeograafiat Euroopas. Selleks määrasin Lähis-Idast, Euroopast, Loode-Aafrikast ning Kaukaasia piirkonnast pärit populatsioonides haplogruppi E ja tema alamhaplogruppe defineerivaid ühenukleotiidilisi polümorfisme (SNP-sid).

Teiseks eesmärgiks on Euroopas laialt ja suure sagedusega levinud Y kromosoomi haplogrupi R1b topoloogia täpsustamine penta- ning heksanukleotiidsete kordusjärjestuste abil, et määratleda paremini tema geograafilise leviku mustreid ja mitmekesisuse määra Euroopas.

1. KIRJANDUSE ÜLEVAADE

1.1 Y KROMOSOOM

Y kromosoom on ~60 Mb pikkune lineaarne kromosoom, mis paikneb isasorganismi rakkude tuumas. Y kromosoomi eukromatiin on ~30 Mb suurune. Y kromosoom on soo-spetsiifiline ning haploidne. Haploidsuse tõttu suurem osa kromosoomist ei rekombineeru meiosis. Rekombinatsioon toimub Y ja X kromosoomil ainult pseudoautosomaalsete terminaalsete regioonide vahel (PAR 1 ja PAR 2) (joonis 1). Rekombinatsioon ei toimu alal, mis jääb nende piirkondade vahele – Y kromosoomi mitte-rekombineeruv regioon (NRY – *nonrecombining region of Y chromosome*) (Jobling et al. 1997). Pärandumise viisist tulenevalt kandub Y kromosoom edasi isalt pojale ja muutused tekivad vaid järk-järgult kuhjuvate mutatsioonide tulemusel. Y kromosoomide erinevusi uurides saab teha järeldusi minevikus populatsioonides toimunud evolutsiooniliste ja ajalooliste sündmuste kohta.



Joonis 1. Y kromosoomi skeem (Jobling ja Tyler-Smith (2000) artiklist pärit joonise põhjal modifitseeritud). AZF - asoospermia faktor (*azoospermia factor*), SRY – Y kromosoomi sugu determineeriv regioon (*sex determining region of Y*).

Eeldades naissoo ja meessoos esindajate võrdset arvu, on Y kromosoomi arv populatsioonis $\frac{1}{4}$ iga autosoomi koopia arvust (Jobling ja Tyler-Smith 2000). Väikse

efektiivse populatsiooni suuruse tõttu on Y kromosoom eriti mõjutatud geenitriivi poolt (Jobling ja Tyler-Smith 2000).

Y kromosoomi puhul uuritakse põhiliselt punktmutatsioone (SNP - *Single nucleotide polymorphism* ehk ühenukleotiidne polümorfism), mikro- ja minisatelliite, insertsioone, deletsioone, inversioone ja duplikatsioone.

Y kromosoomi kasutatakse meditsiinigeneetikas (Jobling ja Tyler-Smith 2000), sugupuude uurimisel (Jobling 2001), evolutsioonilistes töodes isaliini uurimiseks (Kayser et al. 2001; Stumpf ja Goldstein 2001; Jobling ja Tyler-Smith 2003), kriminalistikas ning isaduse määramisel (Jobling et al. 1997; Kayser et al. 1997).

1.2 Y KROMOSOOMI MARKERID

Y kromosoom sisaldab mitmeid erinevaid polümorfisme, mida saab kasutada markeritena haplogrupeerimisel. Esimesed polümorfismid Y kromosoomi mitterekombineerivas regioonil avastati ~15 aastat tagasi (Casanova et al. 1985; Ngo et al. 1986). Erinevad lookused omavad erinevat muteerumise kiirust. Kombineerides erinevaid markereid saab uurida haplogrupisest mitmekesisust ning evolutsiooni erinevates ajaskaalades. Haplogruppide varieeruvus on enamasti seotud nende vanusega ning tavaliselt on nooremad haplogrupid vähem divergentsed.

Y kromosomaalsed markerid jagunevad kahte suurde kategooriasse: bialleelsed markerid ning multialleelsed markerid (Jobling ja Tyler-Smith 2000).

Bialleelsed markerid hõlmavad SNP-e ning insertsioone (nt. *Alu* elemendi insertsioon) ja deletsioone. Bialleelseid markereid iseloomustab madalam mutatsiooniline sagedus ning seetõttu saab nende abil uurida varaseid demograafilisi sündmusi inimajaloos. SNP-d esindavad unikaalseid või peaaegu unikaalseid sündmusi inimese evolutsioonis. SNP-d on kõige levinumad polümorfismide tüübid, moodustades üle 90% polümorfismidest (Collins et al. 1997). Uusi SNP mutatsioone tekib keskmiselt kiirusega 2×10^{-8} mutatsiooni aluspaari kohta põlvkonnas (Nachman ja Crowell 2000). Bialleelsete markerite alleelikombinatsioonide alusel defineeritakse Y kromosoomi haplogruppe (joonis 2) (Jobling et al. 1997).

Antud töös käsitletud *Alu* element kuulub lühikeste hajutatud kordusjärjestuste perekonda, mis on retrotranspositsiooniga levinud üle primaatide genoomi viimase 65 miljoni aasta jooksul. YAP (*Y chromosome Alu polymorphism*) on alguse saanud

stabiilsest 300 bp suurusest insertioonist inimese Y kromosoomi pika õla spetsiifilisse saiti (Yq11) (lookusesse *DYS 287*) kahe nukleotiidi vahele (512. ja 513. positsioonis) (Hammer 1994). Kuna šimpansidel ja gorilladel ei ole Y kromosoomis YAP elementi, võib järeldada, et YAP element inserterus *Homo erectus*'e Y kromosoomi peale inimese ja šimpansi lahknemist. Arvatavasti on YAP elemendi vanuseks 29 000–334 000 aastat (Hammer 1994).

Multialleelsete markeritena kasutatakse kordusjärjestustest minisatelliite ja mikrosatelliite e. STR-e (STR - *short tandem repeat*).

Lühikesi kordusjärjestusi iseloomustab kõrge mutatsioonikiirus ning alleelide arvu suur varieeruvus lookuses (3-49 alleelini) (de Knijff et al. 1997). STR-d koosnevad 1-6 bp kordustest ning korduste arv võib indiviidide vahel varieeruda. STR-de keskmine mutatsiooni määr on 0.2% põlvkonna kohta (Heyer et al. 1997; Jobling et al. 1999; Kayser et al. 2000), kuid mutatsioonikiirused võivad erinevatel STR-del erineda (Caglia` et al. 1997; Heyer et al. 1997; Jobling et al. 1999; Kayser et al. 2000). Trinukleotiidsed mikrosatelliidid muteeruvad suurema kiirusega kui tetranukleotiidsed mikrosatelliidid (Chakraborty et al. 1997). STR-de pikkusvariantide kombinatsioone Y kromosoomi mitterekombineerivas osas nimetatakse haplotüüpideks ja nad on heaks indikaatoriks haplogrupisisesele varieeruvusele (deKnijff et al. 1997). STR-d võimaldavad seega eristada suhteliselt hiljuti lahknenud fülogeneetilisi liine (Brion et al. 2003).

STR-de mutatsioonikiiruste kohta on saadud mitmeid eri tulemusi sõltuvalt valitud meetodist. Kayser et al. (2000) on saanud isa-poeg paarides keskmiseks mutatsiooni määraks $2,8 \times 10^{-3}$ asendust lookuse kohta põlvkonnas, kuid Forster et al. (2000) sai $2,6 \times 10^{-4} - 2,7 \times 10^{-5}$ mutatsiooni lookuse kohta põlvkonna kohta (põlvkonna pikkus on 20 aastat). Zhivotovsky et al. (2004) kohaselt saadi $6,9 \times 10^{-4}$ asendust lookuse kohta põlvkonnas (põlvkonna pikkus on 25 aastat). Mutatsiooni kiirust mõjutab ka kindlasti see, kas see on arvatatud di-, tri- või tetranukleotiidsete STR-de põhjal (Ellegren 2004).

Kõige varieeruvamad Y kromosoomi markerid on minisatelliidid, mille pikkus on 10-100 aluspaari ja mutatsiooni määr ~6-11% põlvkonna kohta (Jobling et al. 1998). Minisatelliidid annavad teavet väga hiljutiste populatsiooni sündmuste kohta (Hurles et al. 1998). Seni on neid markereid kasutatud suhteliselt vähe (Ellegren 2004).

1.3 EUROOPA ASUSTAMINE ANATOOMILISELT KAASAEAGSE INIMESE POOLT

Üldiselt arvatakse, et anatoomiliselt kaasaegne inimene – *Homo sapiens sapiens* saabus Euroopasse esmalt paleoliitikumis ~40 000-50 000 aastat tagasi. Põlluharimine tekkis Lähis-Idas neoliitikumis ~10 000 aastat tagasi ning levis seejärel Euroopasse.

Siamaani ei ole aga ühist seisukohta Euroopa geenitiigi paleoliitilise ja neoliitilise komponendi osakaalu kohta. On arutletud paleoliitiliste küttide-korilaste ja Lähis-Ida neoliitiliste põlluharijate panuse üle.

Neoliitikumi osakaalu suhtes ollakse kahel eri arvamusel. Deemilise difusiooni mudeli (DDM, Ammerman ja Cavalli-Sforza 1984) kohaselt levisid tehnoloogiad koos inimestega. Selle teooria kohaselt võiks eeldada Lähis-Ida neoliitiliste põlluharijate märkimisväärsel panust praegusesse Euroopa geenitiiki (Barbujani ja Bertorelle 1998; Chikhi et al. 2002).

Kultuurilise difusiooni mudeli (KDM, Zvelebil 2000; Richards et al. 1996) järgi toimus põllumajandusele üleminek valdavalt ideede ja tehnoloogiate levikuga, mitte inimeste liikumisega (Richards et al. 2000; Semino et al. 2000; Richards et al. 2002; Underhill et al. 2000). Küsimus on selles, kas migreeruvaid põlluharijaid ja nende järglasi oli suhteliselt vähe (KDM) või palju (DDM).

mtDNA ja Y kromosoomi põhjal on pakutud neoliitiliste põlluharijate panuseks Euroopa geenitiiki umbes 25% (Richards 2003). See migratsioon, mis tõi põllumajanduse Euroopasse, ilmselt ei hõlmanud suurt hulka inimesi ning olulist rolli mängisid kontaktid ja tehnoloogiate ülevõtmine. Richards et al. (1996) leidsid, et Euroopas on ainult mõned mtDNA liinid neoliitilise päritoluga, kuna enamuse vanus jääb 15000 ja 50000 aasta vahele. Seega suurem osa tundub olevat pärit kesk- või hilisest ülem-paleoliitikumist.

Haplogruppide geograafiline levik ja alleelide vanuse hinnangud on kokkusobivad kahe paleoliitilise ja ühe neoliitilise migratsiooniga, mis on kujundanud tänapäevast Euroopa geenitiiki (Semino et al. 2000).

Anatoomiliselt kaasaegse inimese väljarändamise kohta Aafrikast on samuti pakutud mitu võimalust. Euraasiasse saamiseks on põhilised kolm teed. Lisaks Gibraltari väinale, mis ühendab Ibeeria poolsaart ja Marokot (Bosch et al. 2001), on Levanti koridor (Vahemere idarannik, mis hõlmab praegust Iisraeli, Liibanoni, Süüriat ja Jordaaniat) ja Bab el Mandebi väin, mis jääb Somaali ja Araabia poolsaarte vahele

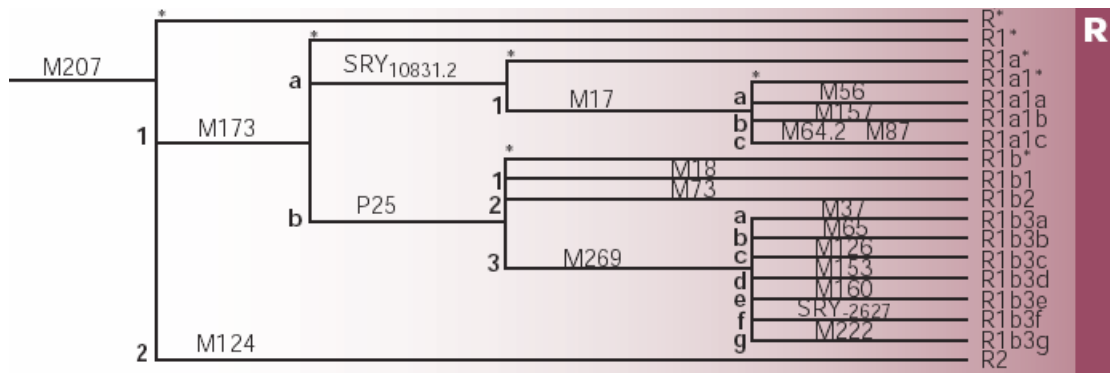
(Cavalli-Sforza et al. 1994; Kivisild et al. 1999; Quintana-Murci et al. 1999; Stringer 2000; Underhill et al. 2001; Bar-Yosef 1998; Nebel et al. 2001). Seda, et Levanti koridor on olnud oluline migratsioonitee Aafrikast välja, on näidatud mitmete haplogruppide põhjal. Põhjatee kulges piki Niilust üle Siinai poolsaare Lähis-Itta ja hargnes sealt edasi üle Euraasia (Maca-Meyer et al. 2001). Näiteks haplogrupid E3b1, E3b3 ning E3a näitavad geenivoogu Egiptuse ja Levanti vahel ülem-paleoliitikumis ja mesoliitikumis ning seda, et E3b3 ja E3b1 haplogruppe kandvad inimesed migreerusid Anatooliasse läbi Levanti koridori, toetab fakt, et Türgis esinevad mitmekesised E3b3 ja E3b1 (Cinnioğlu et al. 2004).

Tuleb siiski rõhutada, et paljud viimase aja populatsioonigeneetika-, arheoloogia- ja antropoloogiaalased tööd viitavad Lõuna Rannikuteele (Lõunatee) kui peamisele või ka ainsale Aafrikast väljunud migratsiooniteele (Quintana-Murci et al. 1999; Stringer 2000; Kivisild et al. 2003). Lõunatee kulges oletatavasti piki India ookeani rannikut, Ida-Aafrikast üle Bab el Mandebi väina, Araabia poolsaare ja Pärsia lahe Lõuna-Aasiasse ning sealt mööda rannikut Austraaliani.

Samuti on leitud kahesuunalist geenivoogu üle Gibraltari väina. Euroopa haplogruppide panust Loode-Aafrika geenitiiki on hinnatud 4-le ja Loode-Aafrika populatsioonide panust Ibeeria geenitiiki 7-le protsendile (Bosch et al. 2001).

1.4 HAPLOGRUPP R

Paleoliitiliseks pärandiks peetav haplogrupp R, mida defineerib mutatsioon M207, jaguneb alamhaplogruppideks R1 ja R2 ning parafüleetiliseks grupiks R* (joonis 3).



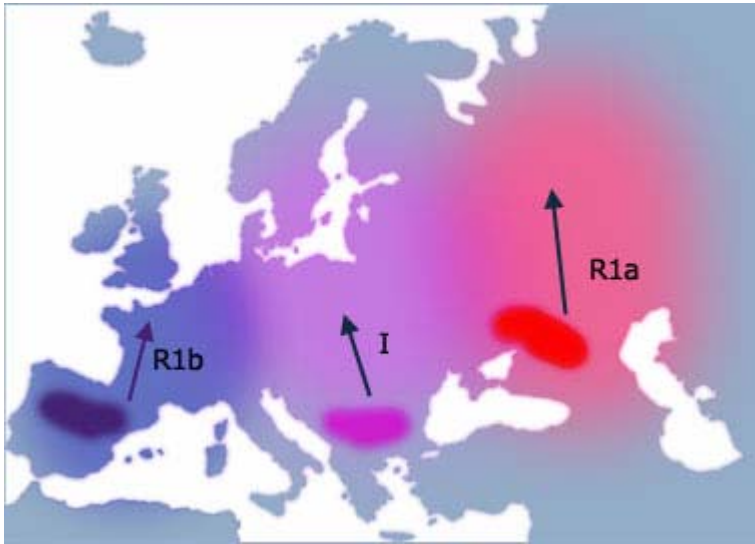
Joonis 3. Haplogrupi R fülogeneetiline puu (Jobling ja Tyler-Smith 2003)

Haplogrupp R1, mida defineerib mutatsioon M173, on Euroopas kõige sagedasem haplogrupp, omades sagedust ~50% (Semino et al. 2000). R1 päritolu Euroopas on seostatud kõige varasema Euroopa ja Aasia lääneosa asustamise faasiga anatoomiliselt kaasaegse inimese poolt (Semino et al. 2000). Kuna mutatsiooni M173 vanuseks on STR-de põhjal hinnatud ~30 000 aastat (Semino et al. 2000), siis on välja pakutud, et R1 kirjeldab varaseimat, ülem-paleoliitikumis aset leidnud ekspansiooni Kesk-Aasiast Euroopasse (Semino et al. 2000; Underhill et al. 2001; Wells et al. 2001).

Arvestades sõsarklaadide R1 ja R2 geograafilist levikut (R2 esineb Indias, Pakistanis, Iraanis ja Lõuna- ning Kesk-Aasias) ja STR diversiteeti on võimalik, et Aasia lääne- ning lõunaosa olid R1 ja R1a diferentseerumise kohaks (Kivisild et al. 2003).

Haplogrupid R1a ja R1b omavad vastupidist sageduste jaotust (Semino et al. 2000; Rosser et al. 2000): R1b esineb kõrge sagedusega Lääne-Euroopas (Underhill et al. 2000; Bosch et al. 2001), R1a on väga harv Lääne-Euroopas ning sagedus suureneb ida suunas ja on sage (30-60%) Ida-Euroopas, Kesk-Aasias ja Loode-Indias (Semino et al. 2000; Wells et al. 2001; Passarino et al. 2001; Kivisild et al. 2003). Haplogrupi R1b kõrge sagedus Lääne-Euroopas (eriti baskidel ja brittidel) võib olla triivi tulemus, kuna LGM (*Last glacial*

maximum – viimase jääaja maksimum) ajal oli populatsiooni arvukus väga väike (Semino et al. 2000).



Joonis 4. R1a ja R1b levik refuugiumidest ~12 000 aastat tagasi.

<http://www.dnaheritage.com/masterclass2.asp>

Haplogrupi R1a vanuseks on pakutud ~15 000 aastat (Semino et al. 2000; Wells et al. 2001). M17 võis olla seotud Kurgaani kultuuri levimisega, mis sai alguse Lõuna-Venemaa/Ukraina steppidest ja laienes ajavahemikul 3000 – 1000 eKr Lääne-Euroopasse, Kesk-Aasiasse ja Indiasse (Passarino et al. 2001; Quintana-Murci et al. 2001; Wells et al. 2001). Teine võimalus on, et R1a tekkis Aasias (Zerjal et al. 1997). R1a STR-de madal diversiteet Kesk-Aasias (Kivisild et al. 2003; Zerjal et al. 2002) viitab hiljutisele rajaja-efektile või triivile, mis võib olla põhjuseks, miks M17 omab Lääne- ja Kesk-Aasias kõrget sagedust. Kuna R1a sagedus on Türgi idaosas kõrgem, siis arvatakse, et haplogruppi R1a kandvad individid sisenesid Anatooliasse Iraani platoo kaudu, kus varieeruvus on märgatavalt kõrgem (Quintana-Murci et al. 2001).

Semino et al. (2000) on välja pakkunud, et R1a ja R1b erinev levik Euroopas on erinevatest refuugiumidest lähtuvate ekspansioonide tagajärg. Pärast jää taandumist umbes 12 000 aastat tagasi algas R1b ekspansioon Ibeeriast ja R1a ekspansioon praeguselt Ukraina alalt (joonis 4). Antud hüpoteesi toetab ka R1a mikrosatelliitide maksimaalne varieeruvus Ukrainas (Passarino et al. 2001).

Kõik siiani analüüsitud haplogruppi R1b kuuluvad individid Euroopast ning Lähis-Idast kannavad ka mutatsiooni M269, mis defineerib alamhaplogrupi R1b3. Cruciani et al. (2002) kirjeldasid mutatsiooni M269 esimest korda. M269 on transitsioon (T→C) geenis IF1AY (Y kromosoomis eukarüootne translatsiooni initsatsiooni faktor 1A). Haplogrupp R1b3 esineb Euroopas sagedusega 40-80%. STR varieeruvus viitab sellele, et viimasel jääajal liikus R1b3 refuugiumidesse Ibeerias ja Anatooolias, kust ta levis põhja poole hilis-ülem-paleoliitikumis ja holotseenis (Cinnioğlu et al. 2004).

R1b3 puhul on välja pakutud migratsiooni Türgist Kagu-Euroopa ja Kaukaasia suunas, kuna Türgis on haplogrupp R1b3 laialt levinud (Cinnioğlu et al. 2004; Arredi et al. 2004). Alamhaplogrupi R1b3 levik Euroopas (Cruciani et al. 2002; Semino et al. 2000; Wells et al. 2001; Kivisild et al. 2003) viitab sellele, et R1b3 levis Euroopas juba paleoliitikumis, saabudes Lääne-Aasiast samaaegselt Aurignac'i kultuuriga. Haplogrupi R1b päritolukoht ei ole täpselt teada, kuid on välja pakutud nii Kesk-Aasiat (Semino et al. 2000; Wells et al. 2001) kui ka Anatooliat (Cinnioğlu et al. 2004). Mutatsioon P25, mis defineerib R1b, on dateeritud ~23 000 aastat vanaks (Karafet et al. 1999).

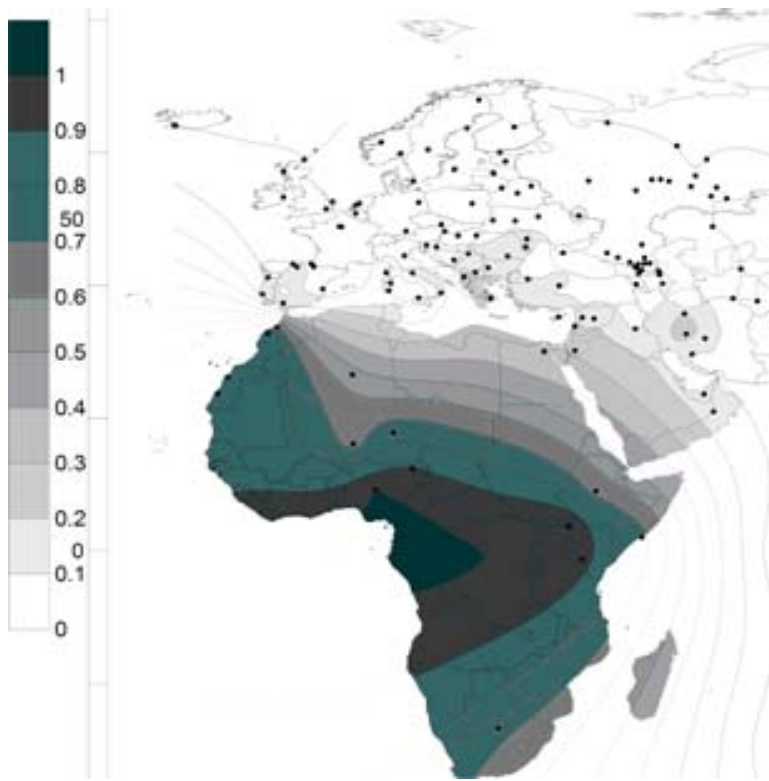
Kuigi R1b3 liine leitakse üle Euroopa märgatava sagedusega (Cruciani et al. 2002), ei ole sobivaid bialleelseid markereid, mis suurendaks haplogrupi R1b3 resolutsiooni, praeguseks näidatud. R1b3 alamhaplogruppidest R1b3a-R1b3g (Jobling ja Tyler-Smith 2003) on näidatud ainult R1b3f levik prantslastel ja baskidel (Rosser et al. 2000) ning R1b3d esinemine 16% Hispaania baskidel (Underhill et al. 2000). Alamhaplogruppi R1b3f defineeriv mutatsioon SRY₂₆₂₇ (transitsioon C→T geenis SRY positsioonis 2627) võis tekkida Ibeeria poolsaarel (Hurles et al. 1999, Brion et al. 2003) ning mutatsiooni vanuseks hinnati 1650-3452 aastat (95 % CI 1044-16001 aastat).

Praeguseks on haplogrupi R1b3 puhul osutunud üheks informatiivsemaks polümorfismiks *TaqI* p49a,f. See polümorfism on *DYS1* lookuses (neli DAZ geeni AZF-c (Azoospermia faktor-c) regioonis) Y kromosoomi mitterekombineerivas osas (Ngo et al. 1986). *TaqI* p49a,f haplotüüp 15 on levinud Lääne-Euroopas, kuid Türgis esineb madala sagedusega (Semino et al. 1996; Jobling et al. 1996). Haplotüüp 35 on seevastu levinud Balkanil, Kaukaasias ja Türgis (Semino et al. 2000). Haplotüüpide 15 ja 35 vastanduv levik võib peegeldada Euroopa uuesti asustamist Ibeerias ja Anatooliast (Cinnioğlu et al. 2004).

1.5 HAPLOGRUPP E

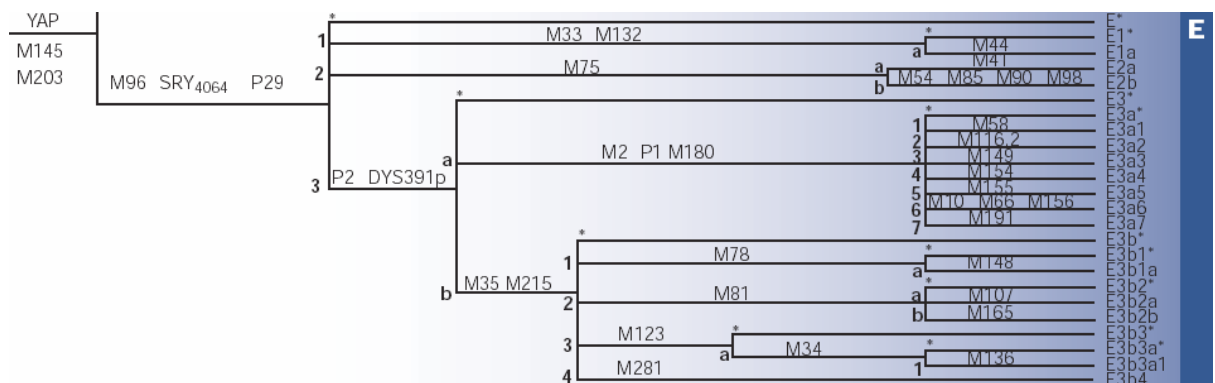
Aafrikas divergeerunud haplogrupp E (joonis 5), mida iseloomustavad mutatsioonid SRY₄₀₆₄, M96 ja P29 ning YAP, jõudis Lähis-Itta tõenäoliselt viimase jääaja järgselt ning levis edasi Euroopasse. Enim töid (Semino et al. 1996; Semino et al. 2004; Cruciani et al. 2004; Luis et al. 2004; Cinnioğlu et al. 2004) toetab arvamust, et haplogrupp E jõudis Lähis-Idast Euroopasse neoliitiliste põlluharijatega, kes mõjutasid oluliselt Vahemere piirkonda.

Haplogrupp E koosneb kolmest alamhaplogrupist E1, E2 ja E3 ning parafüleetilisest rühmast E* (joonis 6). Haplogrupid E1 ja E2 esinevad madala sagedusega üle terve Aafrika. Nende levik Euroopas on peamiselt piirdunud lõunaosaga (Underhill et al. 2000).



Joonis 5. Haplogrupi E (Aafrikas ja Lääne-Euraasias) levik vastavalt Helgason et al. (2000); Rosser et al. (2000); Semino et al. (2000); Bosch et al. (2001); Wells et al. (2001); Zerjal et al. (2001); Laitinen et al. (2002); Passarino et al. (2002); Karafet et al. (2002); Al-Zahery et al. (2003); Barac et al. (2003); Francalacci et al. (2003); Maca-Meyer et al. (2003); Nasidze et al. (2003); Cruciani et al. (2004); Tajima et al. (2004) andmete (Siiri Rootsi doktoritöö).

Haplogruppi E3, mida defineerib mutatsioon P2, on leitud lisaks Aafrikale ka Euroopas ja Lääne-Aasias, kus ta esineb sagedusega alla 25% (Cruciani et al. 2004). Haplogrupp E3 jaotub kaheks monofüleetiliseks haplogrupiks - E3a ja E3b (YCC 2002), mis Aafrikas on levinud suure sagedusega. E3b on kõige sagedasem haplogrupp Loode-Aafrikas (Underhill et al. 2000), kuid esineb ka Euroopas ja Lääne-Aasias (Cruciani et al. 2004). Parafüleetiline rühm E3* esineb peamiselt etiooplastel (Semino et al. 2004).



Joonis 6. Haplogrupi E fülogeneetiline puu (Jobling ja Tyler-Smith 2003)

Antud töös huvipakkuv haplogrupp E3b jaguneb alamklaadideks E3b1, E3b2, E3b3 ja E3b4 ning parafüleetiliseks grupiks E3b* (joonis 4).

E3b TMRCA-ks (Time to the most recent common ancestor – aeg kõige hiljutisema ühise esivanemani) on pakutud 25 600 aastat (95% CI 24 300–27 400 aastat) (Cruciani et al. 2004), mis on lähedane mutatsiooni M35 vanuse hinnangule, 30 000 ± 6000 aastat (Bosch et al. 2001). Semino et al. (2004) on pakkunud M35 tekkeajaks ~29 200 aastat tagasi.

E3 ja E3b fülogeograafia ja STR-de varieeruvus näitavad, et ilmselt on E3b pärit Ida-Aafrikast (Cruciani et al. 2004). Seda toetab ka parafüleetiliste gruppide E3* ja E3b* kõrgeim sagedus samas piirkonnas (Semino et al. 2004).

Haplogrupp E3b1, mida iseloomustab mutatsioon M78, esineb Lähis-Idas, Põhja- ja Ida-Aafrikas ning Euroopas, kus ta on E3b alamhaplogruppidest kõige sagedasem (Cruciani et al. 2004). M78 puhul on pakutud TMRCA 23 200 aastat tagasi (95% CI 21 100–25 400 aastat) (Cruciani et al. 2004).

E3b1 on Cruciani et al. (2004) poolt jagatud kolmeks klastriks, millel on kindel levikuala: klaster α (Balkani poolsaar, Lähis-Idas on piirdunud Türgiga), klaster β (Loode-Aafrikas) ning klaster γ (Ida-Aafrikas). Neljas klaster δ ei ole selge levikumustriga. Klaster α sageduste jaotus Euroopas näitab migratsioone neoliitikumis või neoliitikumi järgselt Balkani poolsaarelt läände kuni Ibeeria poolsaareni ning tõenäoliselt ka kagusse – Türgisse (Cruciani et al. 2004).

STR andmed viitavad E3b1 Ida-Aafrika päritolule (Semino et al. 2004). Hilisemaid populatsiooni ekspansioone, mis hõlmasid klastrit α Euroopas (TMRCA 7800 aastat; 95% CI 6300–9200 aastat), β Loode-Aafrikas (5200 aastat; 95% CI 3200–7500 aastat) ja γ Ida-Aafrikas (9600 aastat; 95% CI 7200–12900 aastat) peaks arvestama kui põhilisi panustajaid haplogrupi E3b1 suhteliselt kõrgele sagedusele.

Alamhaplogruppi E3b3 defineerib marker M123. E3b3 on levinud Lähis-Idas ning Ida-Aafrikas esinedes ka Põhja-Aafrikas ja Euroopas (Semino et al. 2004). E3b3* on leitud Etioopias, Lähis-Idas, Euroopas ja Põhja-Aafrikas (Cruciani et al. 2004). E3b3 levik viitab tema Lähis-Ida päritolule, kuna Ida-Aafrikas piirdub E3b3 leviala Etioopiaga ning teda ei ole seni leitud Etioopia naaberriikidest - Somaaliast ja Keeniast, kuid on leitud enamusel Lähis-Ida populatsioonidel. Samuti on haplogrupi E3b3 STR-de varieeruvus Lähis-Ida populatsioonidel suurem kui Ida-Aafrika populatsioonidel (Cruciani et al. 2004). E3b3 toodi Lähis-Idast Euroopasse tõenäoliselt neoliitiliste põlluharijate poolt. Euroopas võib E3b1 ja E3b3 liine pidada Lähis-Idast lähtuva neoliitilise ekspansiooni tunnusmärkeks (Cruciani et al. 2004; Underhill et al. 2001).

E3b2, mida defineerib mutatsioon M81, on kõige sagedasem haplogrupp Põhja-Aafrikas (Underhill et al. 2000; Cruciani et al. 2002). Väljaspool antud piirkonda on tema harv esinemine tõenäoliselt seletatav Põhja-Aafrikast lähtuva geenivooga (Cruciani et al. 2004; Semino et al. 2004). Haplogrupi E3b2 päritolukohaks on pakutud Põhja-Aafrikat (Semino et al. 2004). E3b2 suhteliselt väike vanus (TMRCA 5600 aastat, 95% CI 4600–6300 aastat) ja madal haplotüübiline varieeruvus Euroopa ja Aafrika E3b2 rajajaliinide vahel toetavad hiljutise geenivoo hüpoteesi Loode-Aafrika ja Ibeeria vahel (Cruciani et al. 2004) ajasügavuses 2300-8600 aastat (Kayser et al. 2000; Zhivotovsky et al. 2004; Semino et al. 2004). E3b2 levik langeb kokku berberite populatsioonide (asustavad Lääne ja Kesk-Sahaarat, Kirde-Aafrika lääneosa) praeguse levikualaga (Cruciani et al. 2004). E3b2

madalam varieeruvus on seletatav kas suhteliselt hiljutise päritoluga, millele järgnes ekspansioon, või hiljutise ekspansiooniga pärast pudelikaelaefekti (Semino et al. 2004). Arvatakse, et hiljutine geenivoog tõi E3b Loode-Aafrikast Ibeeriasse Islami okupatsiooni tulemusena (Cruciani et al. 2004). Mitte kõiki E3b indiviide Ibeerias ei saa vaadelda Aafrika geenivoo tulemusena, vaid geenivoog võis tulla ka Euroopast (Cruciani et al. 2004; Gonçalves et al. 2005).

Haplogrupp E3a, mida defineerivad mutatsioonid M2, P1 ning M180, on kõige sagedasem Lääne- ja Lõuna-Aafrikas (Underhill et al. 2001; Semino et al. 2000, 2002; Karafet et al. 1999). Tema praegune geograafiline jaotus võib olla Lääne-Aafrikast lähtunud Bantu ekspansiooni tulemus (Passarino et al. 1998; Underhill et al. 2001; Cruciani et al. 2002; Luis et al. 2004). Keenias on E3a sagedus ~52%, kuid Etioopias (Underhill et al. 2000; Semino et al. 2002) ja Sudaanis (Underhill et al. 2000) peaaegu null. E3a päritolukohaks on pakutud nii Lääne-Aafrikat, millele viitas sageduse muster ja mikrosatelliitne võrgustik (Cruciani et al. 2002), kui ka Sahaara kõrbest lõunapoole jäävat Aafrikat (Hammer et al. 1998), mida Bantu ekspansioon läbis, kuivõrd E3a omab just selles piirkonnas suurt sagedust (Rosser et al. 2000).

Paragrupp E3b* on levinud Ida- ja Lõuna-Aafrikas ning esineb ka madala sagedusega Põhja-Aafrikas ja Euroopas (Cruciani et al. 2004). Cruciani et al. (2004) leidsid E3b* bantudelt Keeniast, kuid ei leidnud seda haplogruppi Kesk-Aafrikast, kust Bantu ekspansioon alguse sai. E3b* esineb enamasti Sahaara kõrbest lõuna poole jäävatel populatsioonidel (Tansaania, Keenia) (Underhill et al. 2000).

2. MATERJAL JA METOODIKA

2.1 Katsematerjal

DNA proovid on kogutud indiviididelt, kes pole omavahel sugulased ning on andnud sellekohasteks uuringuteks oma nõusoleku.

2.1.1 Haplogrupp R1b3

Töös kasutati 20 DNA proovi, mis olid eelnevalt M269 SNP alusel määratud haplogruppi R1b3. Neist 5 pärinesid lõuna-rootslastelt Malmöst, 5 lõuna-prantslastelt, 5 horvaatidelt ja 5 türklastelt. Kõigil kahekümnel haplogruppi R1b3 kuuluval proovil kontrolliti üheteistkümne penta- ja ühe heksanukleotiidsed kordusjärjestuse korduste arvu. Täiendavalt STR-i P11 kordusjärjestuse korduste arvu kontrolliti 10 lõuna-prantslasel, 6 eestlasel, 2 rumeenlasel, 6 türklasel ja 8 portugallasel ning STR-i P24 kordusjärjestuse korduste arvu 13 türklasel, 14 šveitslasel ja 12 eestlasel.

2.1.2 Haplogrupp E

Töös kasutati DNA proove, mis olid pärit Marokost (97), Kuveidist (204), Kreekast (Rhodos (91), Kreeta (93), Kreeka mandriosa (75), Peloponnesose poolsaar (42)), Rumeeniast (338), Albaaniast (51), Šveitsist (77), Ukrainast (116), Sloveeniast (55), Bosniast (113), Horvaatiast (458), Küproselt (201), Lõuna-Prantsusmaalt (37), Moldaaviast (139) (moldaavlased ja gagauusid), Saudi-Araabiast (216), Kaukaasiast (706), Slovakkias (200). Kõigil 3309 proovil määrati YAP-i olemasolu ning YAP positiivsetel määrati subhaplogruppe defineerivad mutatsioonid M35, M78, M81, M123, M2, M33.

Ise määrasin YAP-i 1019 proovil (Sloveenia, Saudi-Araabia, Küpros, Maroko, Kuveit, Rhodos, Kreeka mandriosa, Kreeta, Patras) ning analüüsisin M35 (Albaania, Šveits, Bosnia, Horvaatia) ja teised subhaplogruppe defineerivad markerid (Ukraina, Moldaavia, Kaukaasia, Lõuna-Prantsusmaa, Slovakkia, Rumeenia).

Lisaks sellele kasutasin oma fülogeneetilistes ja statistilistes analüüsides Siiri Rootsi (Albaania/YAP), Hela Helpi (Lõuna-Prantsusmaa/YAP, M35), Ildus Kutuevi (Kaukaasia regiooni populatsioonid/YAP, M35), Daniel Dioni (Sveits/YAP), Marina Bermisheva (Ukraina, Moldaavia, Moldaavia gagauusid/YAP, M35), Marijana Peričići ja Lovorka Barać-Lauci (Bosnia-Hertzegoviina ja Horvaatia/YAP), Monica Grobei (Rumeenia/YAP, M35, M78), Marián Baldoviči (Slovakkia/YAP) poolt tehtud restriksioonanalüüsi ning sekveneerimise seni publitseerimata andmeid.

2.2 DNA amplifikatsioon

Y kromosoomi DNA amplifitseerimiseks kasutatud praimerite järjestused on ära toodud tabelites 1 ja 2. Tabelis 1 on ära toodud ka sekveneerimispraimerid.

Tabel 1. Penta- ja heksanukleotiidsete kordusjärjestuste määramisel Y kromosoomi DNA lõikude amplifitseerimiseks kasutatud praimerite järjestused

Y STR		Praimeri järjestus 5'→3' suunas	Kordusühik	Praimeri seond. temp. (°C)	Produkt (bp)	Sekv. praimer
P11 ²	F	GATAAAGTGAAGGCCTGTATTGTAT	TGTAT	54	208	F
	R	TTGTTTCAGGTGAGGAATATGCTA				
P10 ²	F	TTGCCCAGTGAGCCAAGTTC	ATTGC	54	233	R
	R	AACTGCATGCCACCTATTGG				
DYS592 ¹	F	TGGCAGGTTTCTGTAATTCCT	AAGAA	54	262	F
	R	TTTTCGTTTCGTTTCATTTCG				
P24 ²	F	CGGGTAATGGAGTGGAATGA	TTCCA	60	284	R
	R	TTCCATTCTTTCTTCTGACA				
P36 ²	F	CCATGGGCTCTGCTAAGCTA	AAAAC	62	287	F
	R	TCAATCTTCAACCCACAGACC				

Tabel 1 jätkub

Y STR		Praimeri järjestus 5'→3' suunas	Kordusühik	Praimeri seond. temp. (°C)	Produkt (bp)	Sekv. praimer
P28 ²	F	TGAGCTCACATTTTCATTCACA	CATTC	62	285	F
	R	AATGGAGTGGAATTGGATGG				
DYS643 ¹	F	GCAGGATGGTGAGGGATAAAAG	CTTTT	54	308	R
	R	ATGCATAGGAATGGTGGGTC				
DYS594 ¹	F	AGAGACAGCGGAAAACACCC	AAATA	60	322	R
	R	ATTCTGAGGATGCAGAGCAGTG				
H3 ²	F	CAAGCCCAAAGTTCTTAACTTCC	AGAAGG	54	321	R
	R	TTCAGATAACCGTGCCCTTT				
P19 ²	F	GCCATCAGATAACAAAAGACGGGC	ATACA	54	355	F
	R	GGCCAAAGAGTGAAAGAGAAGGCA				
P21 ²	F	CCTGCCCACTCATCCAAAGTTTGA	AAAGG	54	357	R
	R	TTGCATTCTGGATGTGCTGACTCC				
P13 ²	F	CCAGATAAACTGAAGGGATTG	AAGGG	56	214	R
	R	AGAGGGAGGCTTCTTCTCAGAA				

P – pentanukleotiidne STR

H – heksanukleotiidne STR

¹ Kayser et al. (2004)

² Meie labori andmed

PCR-i segu komponendid ning DNA amplifikatsiooni tingimused on ära toodud tabelites 3 ja 4.

Tabel 2. Haplogruppide E ja R1b3 bialleelsete markerite määramisel Y kromosoomi DNA lõikude amplifitseerimiseks kasutatud praimerite järjestused

Marker		Praimeri järjestus 5'→3' suunas	Praimerite seond (°C)	Allikas
YAP	F	CAG GGG AAG ATA AAG AAA TA	54	Hammer ja Horai (1995)
	R	ACT GCT AAA AGG GGA TGG AT		
M35	F	TAA GCC TAA AGA GCA GTC AGA G	58	Underhill et al. (2000)
	R	AGA GGG AGC AAT GAG GAC A		
M78	F	CTT CAG GCA TTA TTT TTT TTG GT	51	Underhill et al. (2000)
	R	ATA GTG TTC CTT CAC CTT TCC TT		
M81	F	ACT TAA TTT ATA GTT TCA ATC CCT CA	52	Underhill et al. (2000)
	R	TTC ATG GAG ATG TCT GTA TCT GG		
M2	F	AGG CAC TGG TCA GAA TGA AG	57	Underhill et al. (2000)
	R	ATT GGA AAA TAC AGC TCC CC		
M123	F	TGG TAA ACT CTA CTT AGT TGC CTT T	52	Underhill et al. (2000)
	R	CAG CGA ATT AGA TTT TCT TGC		
M33	F	TTG AAA AAA TAC AGT GGA AC	51	Underhill et al. (2000)
	R	CAA GTG TTT AAG GAT ACA GA		
M269	F	TTTTTGATCTTAGAAAGGAAACG	53	Cruciani et al. (2002)
	R	TTCTGAGGCACATATGATAAAAA		
M281	F	TGG TAA ACT CTA CTT AGT TGC CTT T	52	Semino et al. (2004)
	R	CAG CGA ATT AGA TTT TCT TGC		

Tabel 3. PCR-i segu komponendid 15µl-s

Komponent	Tüvilahuse kogus	Tüvilahuse kontsentratsioon ja koostis	Kontsentratsioon PCR-i segus
BD puhver	1,5 µl	(pH 9.4-9.5 800mM Tris-HCl, 200 mM (NH ₄) ₂ SO ₄) 10x	1/10
Tartrasiin puhver	1,5 µl	(750mM Tris HCl pH 8,8; 200mM (NH ₄) ₂ SO ₄ ; 0.1% Triton X100; 5% Ficoll 400; 10mM Tartrasiin) 10x	1/10
MgCl ₂	1,5/1,8 µl	25 mM	2,5 mM/3mM
dNTP	0,15 µl	dATP,dCTP, dGTP, dTTP 10 mM	0,1 mM

Tabel 3 jätkub

Komponent	Tüvilahuse kogus	Tüvilahuse kontsentratsioon ja koostis	Kontsentratsioon PCR-i segus
Taq DNA/hot-start Taq DNA polümeraas	0,15 µl	5 U/µl	0,05U/µl
F praimer	0,3 µl	10 pmol/µl	0,2 pmol/µl
R praimer	0,3 µl	10 pmol/µl	0,2 pmol/µl
deioniseeritud vesi	9,23/8.93 µl		
genoomne DNA	2 µl		ca10-50ng

Halvemini amplifitseeruvatel DNA-del tõsteti $MgCl_2$ kontsentratsiooni 3mM-ni. DNA-del, kus Taq DNA polümeraasi FIREPol (Solis BioDyne) kasutades PCR-i produkti ei saadud, kasutati nõ. Hot-start Taq DNA polümeraasi Hot FIREPol (Solis BioDyne). Sellisel juhul oli esmase denaturatsiooni kestvuseks 25 minutit.

Kui tartrasiin puhvriga produkti ei saadud, kasutati BD puhvrit (Solis BioDyne).

Tabel 4. PCR-i tingimused

Etapp	°C	aeg
Esmase DNA denaturatsioon (Hot FIREPol Taq polümeraasi puhul)	94	25 min
DNA denaturatsioon	94	30 sek
Praimerite seondumine (vt tabel 1, 2)		25-60 sek
Ekstensioon	72	45 sek
Lõplik ekstensioon	72	3 min

DNA amplifitseeriti 44 tsükliga. Kõik antud töös tehtud PCR-i reaktsioonid tehti kaanesoojendusega termoblokis “Biometra UNO II”.

2.3 Agaroosgeel-elektroforees

Saadud PCR-i produkti ja restriksioonfragmentide pikkusi kontrolliti agaroosgeel-elektroforeesil. DNA fragmendid lahutati 2%-lises agaroosgeelis (0,5 kordses TBE puhvril (45mM Tris-boraat, 0,05M EDTA)). Kasutatud pinge oli 140V, elektroodide distants 32 cm. DNA visualiseerimiseks ultraviolettkiirguses lisati geelile etiidumbromiidi lõpp-

kontsentratsiooniga 0,5 µg/ml. PCR-i produkti kontrolliks kanti geelile 4,5 µl amplifitseeritud DNA-d ja 2µl foreesivärvi (broomfenoolsinine ja 20% fycoll 400). Tulemused jäädvustati UV valguses.

2.3.1 YAP-i määramine

Markeri DYS287 (YAP) varieeruvuse määramisel piisas ainult DNA amplifitseerimisest vastavate praimeritega ning seejärel PCR-i produktide pikkuse kontrollist agarogeelil. Kui insertioon esineb, siis amplifitseerub produkt pikkusega 455bp, kui insertiooni ei ole, on produkti pikkus 150bp.

2.4 RFLP analüüs

Restriktaasid lõikavad DNA-d teatud järjestuse motiivi (äratundmissaidi) esinedes. Mutatsioonid järjestuses võivad põhjustada restriksioonisaidi kao või tekke. Restriksioonifragmendid lahutatakse agarogeel-elektroforeesil. RFLP analüüsiga määratud markerid (M35, M78, M81, M2 ja M269) ning restriksiooniks kasutatud ensüümid on toodud tabelis 5.

Restriksiooni reaktsiooniks lisati 10µl PCR produktile 0,3U restriктаasi ja 1 µl restriksiooni puhvrit. Restriksioon toimus 37° C juures väljaarvatud *TaiI* puhul, millega lõikamine toimus 65 C juures, minimaalselt kaks tundi.

Tabel 5. Markerid, mille analüüsiks kasutati RFLP analüüsi

Marker	np	Alleeli-variant	Restriktaas	Restriktsiooni-sait	Produktide pikkused	Puhver	Allikas
M35	168	¹ G	<i>Bse</i> NI	5' ACTGGN↓3'	169, 344 bp	B ⁺	Underhill <i>et al.</i> (2000)
		C		3' TGAAC↑CN 5'	513 bp		
M78	197	¹ C	<i>Ac</i> II	5' C↓CGC3'	196, 105 bp	O ⁺	Underhill <i>et al.</i> (2000)
		T		3' GGC↑G 5'	301 bp		
M81	147	¹ C	<i>Tai</i> I	5' ACGT↓	276, 146 bp	R ⁺	Underhill <i>et al.</i> (2000)
		T		3' ↑TGCA 5'	422 bp		
M2	168	¹ A	<i>Nla</i> III	5' CATG↓3'	42, 65, 102 bp	NEB4, BSA	Underhill <i>et al.</i> (2000)
		G		3' ↑GTAC 5'	65, 144 bp		
M269	358	¹ T	<i>Mva</i> I	5' CC↓AGG3'	427 bp	R ⁺	Underhill <i>et al.</i> (2000)
		C		3' GGT↑CC 5'	154, 44 bp		

¹ eellasalleel.

Restriktaasid *Bse*NI, *Ac*II, *Tai*I ning *Mva*I ja nende restriktsioonipuhvrid on firmalt MBI Fermentas.

Restriktaas *Nla*III ja reaktsioonipuhver NEBuffer 4 on saadud firmalt New England BioLabs Inc.

2.5 Sekveneerimine

Penta- ja heksanukleotiidsete kordusjärjestuste korduste arvu ja markerite M123, M281 ja M33 (praimerid, vt tabel 3) alleelivariandi määramiseks kasutati Y kromosoomi DNA sekveneerimist.

2.5.1 PCR-i produkti puhastamine

PCR-i produkti sekveneerimise puhul on oluline PCR-i produkti puhastamine praimeritest ning ahelasse lülitamata vabadest nukleotiididest. Selleks lisati 10 µl PCR-i produktile 1 µl puhastussegu, mis koosnes:

eksonukleaas I (10U/µl) (<i>Exo</i> I <i>E. coli</i> 'st, Amersham Pharmacia Biotech)	1 osa
kreveti aluseline fosfataas (1U/µl) (SAP, Amersham Pharmacia Biotech)	9 osa

Töötlemine toimus termobloki "Biometra UNO II" järgmistel tingimustel:

* vabade nukleotiidide defosforüleerimine ja üheaahelalise

DNA (praimerid) lõhkumine 37⁰C 20 min

* ensüümide inaktiveerimine 80⁰C 15 min

2.5.2 Sekveneerimine

Y kromosoomi DNA primaarjärjestuse määramiseks kasutati Sangeri didesoksüterminatsiooni meetodit. DNA sünteesitakse *in vitro* DNA polümeraasi poolt, mis kasutab matriitsina üheaahelalist DNA-d. Reaktsioonisegus sisalduvad kindlas kontsentratsioonis didesoksünukleotiidtrifosfaadid (ddNTP-d), mille kaasamisel sünteesireaktsiooni lõpeb konkreetse DNA ahela süntees. Et eri pikkusega DNA fragmente saaks detekteerida, siis on ddNTP-dega seotud neli erinevat fluorestseeruvat värvi.

Sekveneerimisreaktsioon DYE

* sekveneerimispraimer (5 pmol/μl) 1 μl

* puhastatud PCR-i produkt 5 μl

* sekveneerimisegu 4 μl

Sekveneerimisegu koosneb 1 μl DYE Sequencing Premix (Amersham DyeamicTM ET Terminator Cycle Sequencing Kit; Amersham Pharmacia Biotech, USA) ja 3 μl lahjenduspuhvrist (100μl puhvrit B; MgCl₂ 30μl (lõpp-kontsentratsiooniga 1,875mM) ja 270μl deioniseeritud vett)

Reaktsiooni tingimused on ära toodud tabelis 6.

Tabel 6. Sekveneerimisreaktsiooni tingimused

Etapp	Temperatuur (°C)	Aeg (sek)
DNA denaturatsioon	94	20
praimerid seondumine	54	15
ekstensioon	60	60

Tehti 50 tsükli.

Sadestamine

10 µl produktile lisati:

2 µl lahust, mis koosneb 1,5 M CH₃COONH₄ (ammooniumatsetaat), [pH>8]/EDTA (250mM) ja 20mg/ml punane dekstraan suhtes 1:1

30 µl külma (-20⁰C) 96% etanooli

hoiti -20⁰C juures ~30 min

tsentrifuugiti 15 min 13 000 rpm, eemaldati supernatant

pesti 200 µl 70% etanooliga

tsentrifuugiti 8 min 13 000 rpm, eemaldati supernatant

pesti 200 µl 70% etanooliga

tsentrifuugiti 8 min 13 000 rpm, eemaldati supernatant

lasti sademel kuivada +37⁰C 10 min

suspendeeriti 10 µl MegaBACE geelile kandmise lahuses (70% formamiid, 1 mmol/l EDTA) (Amersham Pharmacia Biotech, USA)

Reaktsiooni produktid analüüsiti sekvenaatoriga MegaBACETM 1000 (Amersham Biosciences), milles kasutatakse fluorestseeruvat märgist, mis on kovalentselt seotud reaktsiooni produktiga. Iga ddNTP oli märgitud kahe fluorestseeruva märgisega, proovid kanti geelis ühele rajale, fluorokroomid aktiveeriti ja detekteeriti skaneeriva CCD kaameraga.

Sekveneeritud DNA järjestuste analüüsil kasutati GCG programmipaketti, versioon 10,0 (Genetics Computer Group, Madison, Wisconsin) ja selle alamprogrammi SeqLab.

2.6 Kasutatud meetodid ja programmid

STR-de korduste arvu põhjal konstrueeriti mediaanvõrgustikud programmi Network 4.1.1.0 (Fluxus technology ltd; <http://www.fluxus-engineering.com/>) abil, kasutades mediaanvektorite ühendamise (*median joining*) (Bandelt et al. 1999) algoritmi.

Mediaanvõrgustik lubab üheaegselt esitada mitmeid alternatiivseid kõige säästlikumaid fülogeneetilisi puid. Võrgustike esitamiseks kasutati programmi netViz 3D 6.50.00.

Y kromosoomi haplogruppide sagedused koos 95%-liste usalduspiiridega arvutati kasutades programmi Sampling (Vincent Macaulay Software, modifitseeritud Mait Metspalu poolt).

Y kromosoomi STR-de korduste arvu varieeruvus arvutati Microsoft Excelis.

3. TULEMUSED

3.1 HAPLOGRUPP R1b3

Antud töö üheks põhiliseks eesmärgiks oli Y kromosoomi haplogrupi R1b3 fülogeograafia täpsustamine kasutades uusi penta- ja heksanukleotiidseid kordusjärjestusi, et määratleda paremini selle Euroopa ühe põhilise haplogrupi geograafilise leviku mustreid ja mitmekesisuse määra. Selleks uuritud populatsioonid valiti sellel tingimusel, et kaetud oleks võimalikult suur ala ja populatsioonid asuksid üksteisest kaugel.

Tabel 7 põhjal joonistati mediaanvektorite ühendamise (*median joining*) meetodil mediaanvõrgustik (joonis 7). Juurhaplotüüp on ära toodud tabelis 7. Välisgrupiks on kaks haplogruppi R1a kuuluvat indiviidi, mida kasutati puule juure panemiseks. Kõige olulisem erinevus ilmnas P24 markeri puhul, mis jaotab R1b3 haplotüübid vastavalt nende levikule kaheks. Kõik uuritud Türgi proovid omasid 10 kordusega alleeli, mis kattub R1a välisrühmaga. Valdav enamus Euroopa proovidest aga omavad antud lookuses 11 kordust (tabel 7). Sellega võiks paralleele tõmmata R1b3 TaqI p49a,f haplotüüpide (ht) 15 ja 35 levikuga. Enamus Türgi proovidest grupeerus Balkani ja Kaukaasia ht35-de kuuluvate proovidega ja Lääne-Euroopa proovid ht15-de kuuluvatega (Cinnioğlu et al. 2004). Ht35 varieeruvus, mis arvutati STR-de põhjal, oli Balkanil, Kaukaasias ja Iraagis madalam kui Türgis. Selle põhjal järeldati, et ht15 ja ht35 levik peegeldab Euroopa uuestiasustamist vastavalt Ibeerias ja Anatooliast.

Kuna P24 tundus olevat informatiivne, siis uuriti P24 lisaks veel 39-l R1b3 indiviidil Türgist, Eestist ja Šveitsist (tabel 8).

P24 grupisisene varieeruvus on ära toodud tabelis 9. Türgi R1b3 proovid osutusid endiselt homogeenseks 10 kordusega alleeli osas, samas kõige rohkem varieeruvaks osutusid Šveitsi proovid. Enamus Šveitsis Rhône jõe orus jõe ülemjooksul omasid 10 kordust, enamus alamjooksul aga 11 kordust. Üllatavalt leiti minu kasutuses olevate Šveitsi proovidel (Daniel Dion, publitseerimata andmed) ka haplogrupi G kõrge sagedus, mis muidu esineb äärmiselt harva Euroopas, olles aga tüüpiliselt sage haplogrupp Anatoolias ja Kaukaasias. Võimalik seletus R1b3 ja G haplogrupi kõrgele sagedusele võiks olla neoliitiline rajajaefekt.

Tabel 7. STR-de korduste arv ning korduste arvu varieerumine uuritud populatsioonides.

Populatsioon	Proovi nr.	H3	DYS592	P24	DYS594	P10	P19	P21	P36	P11	P28	P13	DYS643	GKV
Türgi	18	10	12	10	10	12	12	11	12	10	11	11	10	
Türgi	23	10	12	10	10	12	12	12	11	10	14	11	10	
Türgi	39	10	11	10	10	12	12	11	11	9	12	11	10	
Türgi	42	10	12	10	10	12	12	11	11	8	11	MA	MA	
Türgi	80	10	12	10	10	12	12	11	11	11	12	11	10	
Grupisisene varieeruvus		0	0.2	0	0	0	0	0.2	0.2	1.3	1.5	0	0	0,68
Horvaatia	5064	10	12	11	10	12	12	11	11	10	12	11	10	
Horvaatia	1749	10	12	12	10	12	12	11	11	11	11	11	10	
Horvaatia	4200	10	12	11	10	12	12	11	11	10	11	9	10	
Horvaatia	1946	10	12	11	10	12	12	11	11	10	11	MA	10	
Horvaatia	1916	10	12	10	10	12	12	11	MA	10	11	11	10	
Grupisisene varieeruvus		0	0	0.5	0	0	0	0	0	0.2	0.2	1	0	0,5
Lõuna-Rootsi	209	10	12	11	10	12	12	MA	11	10	11	MA	MA	
Lõuna-Rootsi	213	10	12	11	10	12	12	11	11	10	11	MA	10	
Lõuna-Rootsi	106	10	12	10	10	12	12	MA	11	10	11	MA	10	
Lõuna-Rootsi	139	9	12	11	10	12	12	11	11	10	11	MA	10	
Lõuna-Rootsi	232	10	12	11	10	12	12	11	12	10	MA	MA	MA	
Grupisisene varieeruvus		0.2	0	0.2	0	0	0	0	0.2	0	0	-	0	0,12
Lõuna-Prantsusmaa	50	10	12	11	10	13	12	11	11	10	11	11	10	
Lõuna-Prantsusmaa	54	10	12	11	10	12	12	11	MA	10	11	11	10	
Lõuna-Prantsusmaa	57	10	12	11	10	12	13	11	11	10	11	11	10	
Lõuna-Prantsusmaa	59	10	12	11	10	12	12	11	MA	10	13	11	10	
Lõuna-Prantsusmaa	78	9	12	12	10	12	MA	11	MA	10	12	MA	10	
Grupisisene varieeruvus		0.2	0	0.2	0	0.2	0.25	0	0	0	0.8	0	0	0,33
Kogu varieeruvus		0.09	0.05	0.41	0	0.05	0.05	0.06	0.17	0.37	0.71	0.33	0	
SSH		10	12	<u>11</u>	10	12	12	11	11	10	11	11	10	
JH		10	12	<u>10</u>	10	12	12	11	11	10	11	11	10	

MA (mitte amplifitseeruv), SSH (suurima sagedusega haplotüüp), JH (juurhaplotüüp), GKV (grupisisene korduste arvu keskmine varieeruvus populatsiooniti)

Tabel 8. P24 korduste arv uuritud populatsioonides.

Populatsioon	Proovi nr.	P24
Türgi	12	10
Türgi	18	10
Türgi	23	10
Türgi	39	10
Türgi	42	10
Türgi	46	10
Türgi	48	10
Türgi	52	10
Türgi	71	10
Türgi	80	10
Türgi	97	10
Türgi	100	10
Türgi	104	10
Šveits	6171	10
Šveits	6181	10
Šveits	6271	10
Šveits	6279	10
Šveits	6283	10
Šveits	6293	10
Šveits	6303	10
Šveits	7001	11
Šveits	7003	11
Šveits	7005	10
Šveits	7007	11
Šveits	7011	11
Šveits	10'027	11
Šveits	12'055	11
Eesti	333	11
Eesti	70	11
Eesti	388	11
Eesti	80	11
Eesti	122	11
Eesti	3	10
Eesti	50	11
Eesti	84	11
Eesti	315	11
Eesti	194	11
Eesti	1	11
Eesti	19	11

Tabel 9. P24 korduste arvu varieeruvus

Populatsioon	P24 varieeruvus
Türgi	0
Šveits	0.26
Eesti	0.08

Täiendav uuring tehti ka STR P11 puhul, sest tema polümorfsus algses valimis (tabel 7) lubas oletada, et P11 puhul võib olla abi R1b3 resolutsiooni suurendamisel (tabel 10). Lõuna-Prantsusmaa, Eesti ja Portugal olid homogeensed, samas Türgile geograafiliselt lähemas Rumeenias leiti varieeruvus. Sellest võib järeldada, et kõige suurem erinevate alleelide arv Türgis, võib peegeldada R1b3 varasemat esinemist antud piirkonnas.

Tabel 10. P11 korduste arvu tabel uuritud populatsioonides.

Populatsioon	Proovi nr.	P11
Lõuna-Prantsusmaa	60	10
	65	10
	67	10
	74	10
	77	10
	85	10
	87	10
	64	10
	63	10
	68	10
Eesti	331	10
	333	10
	388	10
	122	10
	103	10
	194	10
Rumeenia	348	10
	506	8
Türgi	12	10
	46	10
	48	10
	70	9
	72	10
	104	10
Portugal	1023	10

Tabel 10 jätkub

Populatsioon	Proovi nr.	P11
Portugal	1030	10
	109	10
	146	10
	232	10
	80	10
	84	10
	131	10

Kayser et al. (2004) arvavad, et alleelne varieeruvus on informatiivsem diversiteedist, kuna ta võtab arvesse mitte ainult seda kas erinevus eksisteerib alleelide vahel või mitte, vaid ka iga erinevuse ulatust.

Antud töös kasutati varieeruvuse hindamisel pentanukleotiidsete korduste ning heksanukleotiidsete korduste arvu muutlikkust nii populatsioonide vahel kui ka populatsioonisiselt. Grupisisene STR-ide korduste arvu keskmine varieeruvus leiti kõige suurem olevat Türgi populatsioonis (tabel 7). Tundub, et Türgi on olnud R1b3 divergeerumise kohalt oluline paik. Mitmekesisuse poolest oli Türgi järel kõige suurem muutlikkus täheldatav Horvaatia populatsioonis. Kõige väiksem grupisisene varieeruvus oli Lõuna-Rootsis. Cinnioğlu et al. (2004) leidsid ka Türgis kõige suurema STR-de varieeruvuse (0.31). Suur varieeruvus võib tekkida mitmel eri põhjusel. Ühest küljest on Türgi erinevate populatsioonide segunemise koht tema geograafilise asukoha tõttu. Teise võimalusena võib Anatoolia olla haplogrupi R1b3 algkoduks, vähemalt Euroopa populatsioonide seisukohalt, keda antud töös vaadeldud on. Minu andmete põhjal tundub tõenäolisem viimane variant, kuna Türgis on kõige suurem varieeruvus, mis järjest väheneb lääne suunas.

Kõigi nelja populatsiooni kogu varieeruvuse põhjal osutusid kõige varieeruvamateks STR-deks antud valimis P28 (tabel 7), P24 ning P11, mis võivad osutada potentsiaalselt informatiivseteks markeriteks R1b3 fülogeneetilise puu resolutsiooni suurendamisel.

3.2 HAPLOGRUPP E

Antud töös uuriti Y kromosomaalse haplogrupi E ja tema alamhaplogruppide levikut erinevates Euroopa, Lähis-Ida ja Aafrika populatsioonides.

Selleks tehti alamhaplogruppide sageduste kaardid (joonis 8, 9, 10), kuhu kaasati lisaks antud töö tulemustele Hela Helpi bakalaureusetöö ning kirjanduse andmed (vt lisas tabel 12). Tabelis 11 on esitatud antud töö tulemused haplogrupi E subhaplogrupeerimise kohta.

Antud töös on kõige rohkem uuritud Euroopa populatsioone. Uuritud populatsioonides on haplogrupi E sagedus varieeruv (2,7-39,2%), isegi geograafiliselt suhteliselt lähedaste piirkondade vahel. Näiteks võiks tuua Kreeka mandriosa sisesed erinevused, mille puhul Peloponnesose poolsaare ja ülejäänud Kreeka mandriosa vahel on märgata haplogruppi E kuuluvate indiviidide sageduse suur erinevus, vastavalt 30,95% ja 9,33%.

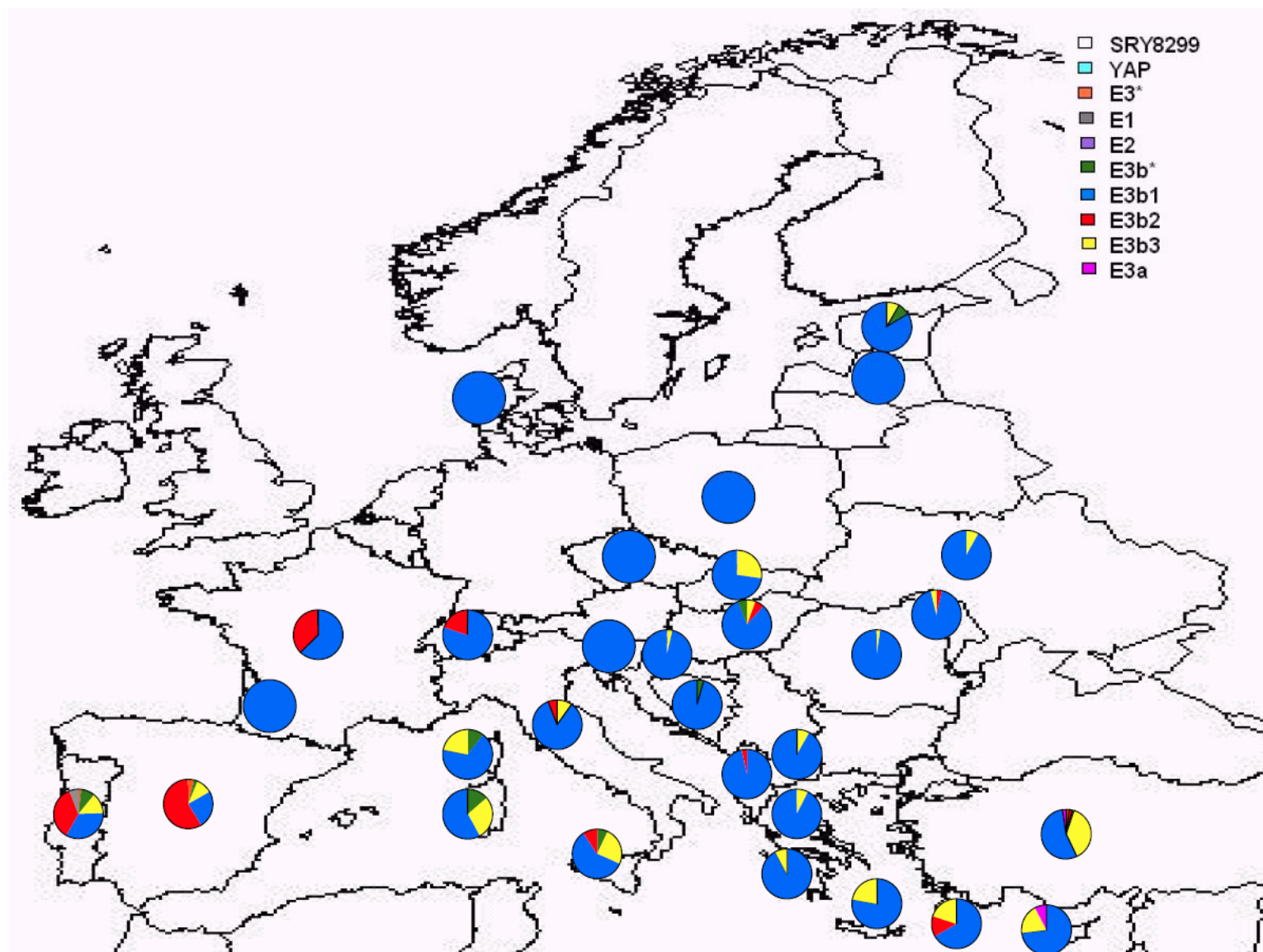
Euroopas on hg E3b valdavateks subhaplogruppideks E3b1 ja E3b3, mida on seostatud Euroopasse neoliitiliste põlluharijatega saabumisega Lähis-Idast (Cruciani et al. 2004). Selle arvamusega kooskõlas väheneb haplogrupi E sagedus Euroopas lõunast põhja suunas (joonis 5). Siinkohal tuleb aga rõhutada, et ainult sagedustel põhinev jaotus ei luba hinnata leviku ajalist sügavust: igasugune Lähis-Idast lähtuv migratsioon jätnuks samasuguse jälje. E3b liinide seostamine just põlluharijatega rajaneb seetõttu peamiselt STR lookustel põhinevatel ajahinnangutel (Cruciani et al. 2004). Oletades, et E3b liinid tõepoolest saabusid Euroopasse neoliitikumis, ilmneb minu töö ja kirjanduse andmetele tuginedes, et nende levik piirdus põhiliselt Vahemere rannikualadega, kus haplogrupi E sagedus on kõige suurem. Ilmselt hiljem on hg E indiviidid levinud põhja ning ida poole, kus enamasti esineb alamhaplogrupp E3b1 (joonis 8). E3b1 võis levida Balkanimaadest üle kogu Euroopa neoliitikumis, mida toetavad STR-de varieeruvuse andmed (Cruciani et al. 2004). Antud töö Euroopa indiviididel leitud E3b3 sagedus on suhteliselt madal (0,3-2,49%). Euroopas on E3b3 kõige rohkem levinud Ibeeria populatsioonidel ning Vahemeresaalartel.

Tabel 11. Haplogrupi E ja tema alamhaplogruppide sagedused protsentides

Populatsioon	n	Hg E	CI	E3b1	CI	E3b2	CI	E3b3	CI	E3a	CI	E1	CI	E3b*	CI
Maroko	97	61,86	51.9-70.9	10.3	5.7-18.0	41.24	31.9-51.2	1.03	0.2-5.6	6.19	2.9-12.9	1.03	0.2-5.6	2.06	0.6-7.2
Kuveit	204	11,27	6.8-15.2	2.94	1.4-6.3			1.96	0.8-4.9	4.41	2.4-8.2			0.98	0.3-3.5
Saudi-Araabia	216	9,26	6.5-14.4	7.4	4.6-11.7	0.46	0.1-2.5	0.93	0.3-3.3	0.93	0.3-3.3				
Kreeka mandriosa	75	9,33	4.7-18.1	8.0	3.8-16.4			1.33	0.3-7.1						
Rhodos (Kreeka)	91	16,48	10.3-25.5	10.99	6.1-19.1	2.2	0.7-7.6	3.3	1.2-9.2						
Kreeta (Kreeka)	93	9,68	5.2-17.4	7.53	3.7-14.7			2.15	0.7-7.5						
Peloponnesose poolsaar (Kreeka)	42	30,95	19.1-46.1	28.57	17.2-43.7			2.38	0.6-12.3						
Küpros	201	12,94	9.0-18.3	9.45	6.2-14.3			2.49	1.1-5.7	0.99	0.3-3.5				
Albaania	51	39,2	27.0-53.0	37.25	25.3-51.0	1.96	0.5-10.3								
Moldaavia (gagauusid)	79	22,78	14.9-33.2	21.52	13.9-31.8			1.27	0.3-6.8						
Moldaavia	60	15,0	8.2-26.2	13.33	7.0-24.2	1.67	0.4-8.8								
Sloveenia	55	3,64	1.1-12.3	3.64	1.1-12.3										
Bosnia ja Hertsegoviina	113	6,19	3.1-12.2	6.19	3.1-12.2										
Horvaatia	458	4,8	3.2-7.2	4.8	3.2-7.2										
Slovakkia	200	3,5	1.7-7.0	2.5	1.1-5.7			1.0	0.3-3.5						
Rumeenia	338	11,83	8.1-14.7	10.65	7.8-14.4			0.3	0.1-1.6						
Ukraina	116	3,45	1.4-8.5	3.45	1.4-8.5										
Šveits	77	6,49	2.9-14.3	5.19	2.1-12.6	1.3	0.3-6.9								
Lõuna-Prantsusmaa	37	2,7	0.6-13.8	2.7	0.6-13.8										
Kaukaasia	706	1,42	0.8-2.6	1.42	0.8-2.6										

n – valimi suurus

CI (confidence interval) – usalduspiirid 95% tõenäosusega



Joonis 8. Haplogrupi E alamhaplogruppide levik Euroopas.

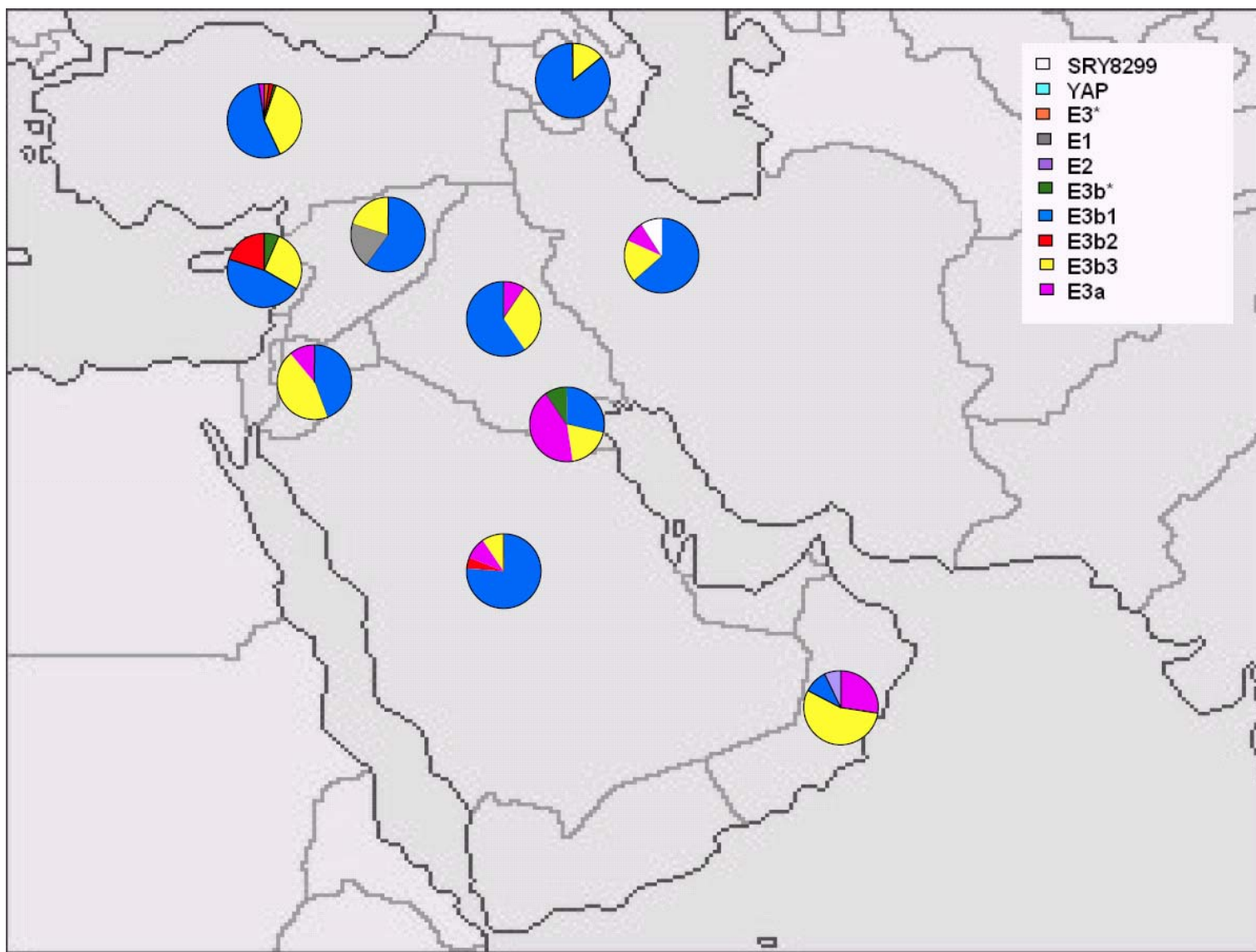
Kreeka mandriosa haplogrupi E sagedus (9,33%) on peaaegu sama Kreetal esineva (9,68%) sagedusega. Samas Peloponnesose poolsaarelt (30,95%) ja Rhodose (16,48%) saartelt saadud proovides erinevad selle haplogrupi sagedused oluliselt Kreeka mandriosa sagedusest, mida võib seletada geneetilise triivi ja ehk asutajaefektiga.

Euroopast tuleks eraldi välja tuua Vahemere idaosas asuv Kreeka saar Rhodos, mille kohta pole varem kirjanduses andmeid avaldatud. Haplogrupi E3b varieeruvus Rhodosel osutus suuremaks kui Kreekas ja Kreetal. Lisaks haplogruppidele E3b1 ja E3b3, esineb Rhodosel ka E3b2. Rhodose populatsiooni E3b liinide suur mitmekesisus on väga sarnane temale lähedal asuva Küprose populatsiooni mitmekesisusega. Ilmselt võib suuremat varieeruvust põhjendada sellega, et Anatoolia ja Rhodose vahel toimus kaubavahetus, mis ilmselt hõlmas ka inimeste rändeid, võimalik ka, et orjade vedu Aafrikast, mida näitab haplogrupi E3b2 esinemine. Anatoolias esinevad kõik need haplogrupid, mis leiti Rhodose saarelt.

Sarnaselt Euroopa põhja osaga on Kaukaasias E sagedus madal (1,42%). Antud töös uuritud adõgeedel, abaziinidel, kumõkkidel, karanogaidel, avaaridel, andidel ja lesgiinidel leiti ainult E3b1 alamhaplogruppi. Aseritel on varem leitud ka E3b3 (Cruciani et al. 2004).

Võrreldes Euroopaga on Lähis-Idas haplogrupi E sagedus kõrgem ja alamhaplogrupiline mitmekesisus suurem (joonis 9). Saudi-Araabias on esindatud kõik E3b alamhaplogrupid. Sagedaseim on E3b1, mis on esindatud enamikel Lähis-Ida populatsioonidel. E3b1-le järgneb sageduselt E3b3 (Underhill et al. 2000; Cinnioğlu et al. 2004; Semino et al. 2004; Cruciani et al. 2004), mida Kuveidis ja Saudi-Araabias leidis vastavalt 2,0% ja 0,9%.

Aafrikas on alamhaplogruppide levik selgelt piiritletud (joonis 10). Haplogrupi E3b2 levik Aafrikas langeb kokku berberite asualaga, s.o Põhja-Aafrika rannikuala Marokost Egiptuseni (Cruciani et al. 2004). Suurimaid sagedusi (~80%) on leitud Loode-Aafrikas Maroko berberitel (Bosch et al. 2001; Cruciani et al. 2002; Cann et al. 2002; Cruciani et al. 2004), kus ta on kõige levinum (joonis 10). E3b2 sagedus väheneb järsult ida suunas. Berbereid elab palju ka Maroko kõrval asetsevas Alžeerias ning seal esineb ka E3b2 kõrge sagedusega - 65%–80% (Bosch et al. 2001; Cruciani et al. 2002). Egiptuses, kus asub ainult mõni berberi küla, on E3b2 harv (1,9%). Haplogrupp E3b2 kõige lõunapoolsem leid asub Nigeris (Cruciani et al. 2004). Antud töös uuritud sagedus Maroko berberitel ja araablattel oli 41,2% (46 araablattel/14 berberit). Ilmselt antud juhul mõjutab sagedust araablattel suurem arv valimises, sest just neilt on leitud hg E3b2 madalamaid sagedusi (~31-52%) (Bosch et al. 2001; Cruciani et al.



Joonis 9. Haplogrupi E alamhaplogruppide levik Lähis-Idas.

2002). Marokost leiti ka üks indiid, kes kuulus haplogruppi E1. Haplogruppi E1 on Marokost leidnud ka Semino et al. (2004). Teda on leitud ka Malist, Sudaanist ja Lähis-Idast (Underhill et al. 2000). Üldiselt on E1 iseloomulik Lääne-Aafrikale (Cruciani et al. 2004) ning ilmselt võib E1 esinemist Marokos seletada rännetega Lääne-Aafrikast loodesse.

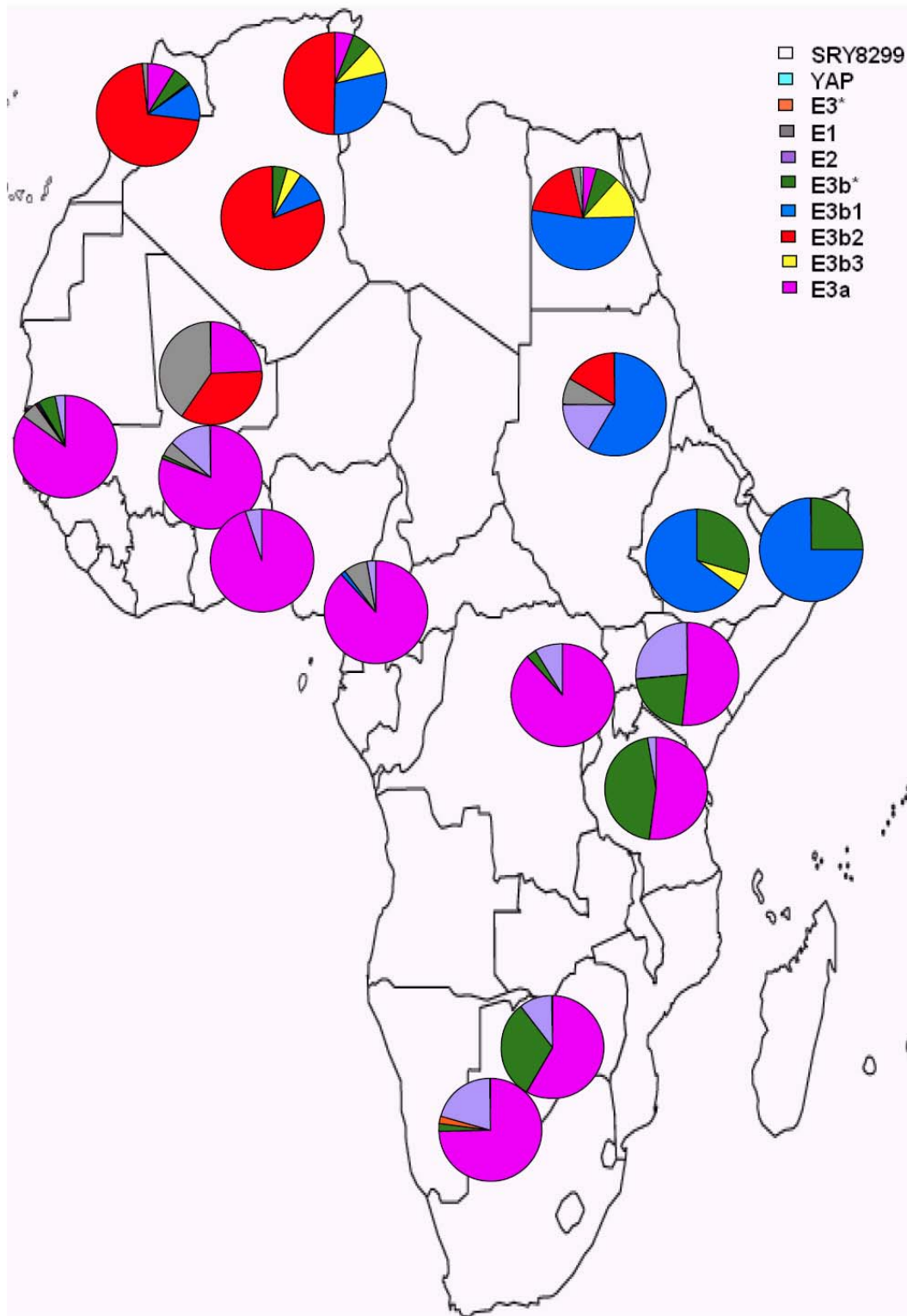
Kui E3b2 Euroopas üldiselt ei esine, siis Ibeeria populatsioonidel on leitud seda kuni 12% (Underhill et al. 2000; Bosch et al. 2001; Cruciani et al. 2004; Semino et al. 2004). Pooled Portugali E3b2 haplotüüpidest on jagatud berberite ja araablase populatsioonidega Marokost (Quintana-Murci et al. 2004). See viitab Marokost lähtunud geenivoole portugallaste populatsiooni (Gonçalves et al. 2005). Samas mtDNA ei näita olulist geenivoogu Marokost Ibeeria poolsaare populatsioonidesse.

E3a on valdavalt Aafrikas levinud haplogrupp, olles kõige suurema sagedusega esindatud Lääne- ja Lõuna-Aafrikas. Selline levik tänapäeval peegeldab tõenäoliselt Bantu migratsioonide mõju (Cavalli-Sforza et al. 1994; Passarino et al. 1998). Antud töös leiti hg E3a ka Marokos, ehkki suhteliselt madala sagedusega. E3a esinemist Marokos on kirjeldatud ka teistes töödes (Cruciani et al. 2002; Bosch et al. 2001; Cann et al. 2002).

Väljaspool Aafrikat on E3a levik piirdunud Lähis-Idaga, kus ta suurim sagedus on Omaanis (~6%; Cruciani et al. 2002; Luis et al. 2004). Siiski leiti antud töös kaks E3a indiid Küprosel.

Paragrapi E3b* esinemine Loode-Aafrikas viitab ilmselt Ida-Aafrika mõjudele, kus antud haplogrupi sagedus on kõige suurem (Cruciani et al. 2002; Semino et al. 2002). Antud töös leiti haplogruppi E3b* Marokost ning Kuveidist. Ilmselt on see paragrupp sattunud sinna rännete tulemusena.

Väljaspool Aafrikat on näha kõige suuremat mitmekesisust nii Türgis kui Ibeeria poolsaarel, mis ilmselt on olnud ajaloo jooksul mitmete migratsioonide kohaks. Kirjanduse ülevaates käsitletud neoliitiliste farmerite panus Euroopa geenitiiki näib olevat pigem piki Vahemerde kui läbi Kesk-Euroopa, seega mõjutades rohkem Lõuna-Euroopat kui Kesk-Euroopat (Semino et al. 2000).



Joonis 10. Haplogrupi E alamhaplogruppide levik Aafrikas.

KOKKUVÕTE

Käesolevas töös käsitleti Y kromosoomi haplogruppide E ja R1b3 varieeruvust Lääne-Euraasia ja Põhja-Aafrika populatsioonidel. Aafrikas, kus haplogrupi E sagedus ning mitmekesisus on kõige suurem, on alamhaplogruppide levik seotud kindlate regioonidega. Marokos esines kõige sagedamini haplogrupp E3b2, mille leviala piirnebki peamiselt Loode-Aafrikaga. Haplogrupi E sagedus Euroopas on suurim piki Vahemere rannikut ning peegeldab tõenäoliselt vastavalt kirjandusel tuginevatel STR andmetel neoliitiliste põlluharijate rändeid. Haplogrupi E sagedus väheneb Põhja-Euroopa suunas, kus esineb vaid hg E3b1. Põhja-Euroopaga sarnane on siiaaani suhteliselt vähe uuritud Kaukaasia piirkond, kus on samuti, kuigi väga väikese sagedusega, esindatud vaid hg E3b1. Käesolevas töös käsitletud Euroopa populatsioonidest oli haplogrupi E mitmekesisus suurim Vahemere saartel, kaasaarvatud Rhodosel, mille kohta varasemad andmed puudusid.

Haplogrupi R1b3 topoloogia täpsustamine on oluline, kuna siiaaani pole leitud sobivaid markereid, mis lahutaksid R1b3 alamharudeks ning oleksid iseloomulikud kindlatele geograafilistele regioonidele. Haplogrupi R1b3 sisest kordusjärjestuste (STR-id) varieeruvust analüüsiti populatsiooniti vastavalt nende geograafiale, et teha järeldusi minevikus toimunud populatsioonide liikumise kohta. Haplogrupi R1b3 STR-de varieeruvuse põhjal joonistatud fülogeneetilisel võrgustikul on näha, et STR P24 jaotab haplogrupi R1b3 haplotüübid vastavalt nende levikule kaheks. Türgis esindatud eellasalleel antud lookuses on harv Euroopas ja vastupidi, Euroopas valdavalt levinud tütaralleeli ei leidunud meie Türgi valimis.

Türgis on haplogrupi R1b3 haplogrupi-sisene lühikeste kordusjärjestuste varieeruvus kõrgem kui uuritud Euroopa populatsioonidel. Seega on tõenäoline, et hoolimata R1b3 kõrgest sagedusest Lääne-Euroopas, on selle haplogrupi praegune fülogeograafia seletatav tema päritoluga Lääne-Aasiast, pigem kui et Türgi R1b3 varieeruvus oleks tulenev Euroopast lähtunud tagasimigratsioonist.

Vastavalt tehtud tööle võib järeldada, et on kolm STR-i, mis võivad osutada potentsiaalselt informatiivseks Y kromosoomi haplogrupi R1b3 topoloogia lahutamisel. Nendeks on P24, P28 ning P11, mis näitavad kõige suuremat varieeruvust antud haplogrupi siseselt.

SUMMARY

Uniparentally inherited mtDNA variation in humans has been extensively studied for more than a decade while extensive studies on Y chromosome started more recently. Compared to mtDNA, Y chromosome has much wider spectrum of different genetic polymorphisms. For the purpose of reconstruction of the phylogenetic skeleton of paternal lineages conservative and nearly unique SNPs, insertions, deletions are the markers of choice while variable short tandem repeats (mostly bi-, tri- and tetranucleotide repeats) are widely used in forensic purposes and for the study of molecular variation with SNP defined haplogroups. Less attention has been paid on STR loci with longer repeat units, such as penta- and hexanucleotide repeats, which, however, can have potential use in phylogenetic analysis as having intermediate mutation rate to SNP markers and fastly evolving short tandem repeats.

The aim of this thesis was to study the phylogeography of two different Y chromosome haplogroups, one of which, haplogroup R1b, is considered to have Paleolithic legacy in Europe, while the second, haplogroup E3b, has been taken as the representative of a more recent Neolithic phase of immigration.

According to the present study, in order to explore the distribution of haplogroup E in Europe, Near East, Caucasus and Africa, information about this work and published work was combined and analyzed together. The total of 1019 individuals belonging to haplogroup E and its subhaplogroups were considered from the following populations: Slovenia, Saudi Arabia, Cyprus, Morocco, Kuwait, Rhodes, Greece mainland, Crete, Peloponnesos peninsula, Albania, Switzerland, Bosnia and Hertzegovina, Croatia, Ukraine, Moldova, area of Caucasus, Southern France, Slovakia and Romania. The results of these analyses show that certain sub-haplogroups of E are bound to specific geographic regions which is in concordance with previous work. In Morocco, the haplogroup E3b2 had the highest frequency as is usually restricted to North Western Africa. In Europe the frequency and variance of haplogroup E was much lower compared to Near East. The dominant subhaplogroup in Europe was E3b1. Higher frequencies are observed in the coast of Mediterranean and decrease towards north. Caucasus, which has been poorly explored so far, shows similarities with Northern Europe, having only subhaplogroup E3b1 which is found at low frequency there. The highest variance of haplogroup

E subclades was observed in Mediterranean islands including Rhodes, for which no data has been reported so far.

Another purpose of this thesis was to explore the topology of Y chromosomal haplogroup R1b, which is widespread and has a high frequency in Europe. I used penta- and hexanucleotide STRs to make conclusions about their geographic patterns and variation in Europe. Haplogroup R1b has been considered to have migrated to Europe probably in Late Pleistocene and to have survived the last ice age in different refugial areas. The total of 20 individuals from Southern Sweden, Turkey, Croatia and Southern France were analyzed for 11 pentanucleotide repeats. This analysis revealed that 3 markers, P11, P24, and P28 showed a high level of differentiation between the populations under examination. Therefore, an additional set of samples were examined for two of them. Most importantly, STR P24 variation divided haplogroup R1b3 haplotypes into two clusters with significant geographic association.

TÄNUSÕNAD

Esimesena ma tahaksin tänada oma juhendajaid Maere Reidlat ja Toomas Kivisilda kannatliku juhendamise eest. Samuti sooviksin tänada Siiri Rootsit ning Jüri Parikut kasulike nõuannete eest. Suured tänud Ille Hilpusele DNA proovide ettevalmistamise ning Jaan Linnule tehnilise abi eest.

Olen tänulik Ildus Kutuevile, Daniel Dionile, Marina Bermishevale, Marijana Peričićile, Lovorka Barać-Laucile, Monica Grobeile ning Marián Baldovičile võimaluse eest kasutada nende varasemate tööde andmeid.

Tänan Erika Tamme, Urmas Roostalut, Hela Helpi ja Monika Karminit ning ülejäänud laborikollektiivi igasuguse abi ning meeldiva töökeskkonna eest.

KASUTATUD KIRJANDUS

- Al-Zahery N, Semino O, Benuzzi G, Magri C, Passarino G, Torroni A, Santachiara-Benerecetti AS. (2003) Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. *Mol Phylogenet Evol.* 28:458-72.
- Ammerman AJ, Cavalli-Sforza LL (1984) The Neolithic transition and the genetics of populations in Europe. Princeton University Press, Princeton
- Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C. (2004) A predominantly neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet.* 2004 Aug;75(2):338-45. Epub 2004 Jun 16.
- Barac L, Pericic M, Klaric IM, Rootsi S, Janicijevic B, Kivisild T, Parik J, Rudan I, Villems R, Rudan P. (2003) Y chromosomal heritage of Croatian population and its island isolates. *Eur J Hum Genet.* 11:535-42.
- Bandelt HJ, Forster P, Rohl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 16:37-48.
- Barbujani G, Bertorelle G, Chikhi L (1998) Evidence for Paleolithic and Neolithic gene flow in Europe. *Am J Hum Genet* 62:488-92
- Bar-Yosef O (1998) The Natufian Culture in the Levant, Threshold to the Origins of Agriculture. *Evolutionary Anthropology* 6:159-177
- Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J (2001) High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 68:1019-29.
- Brion M, Salas A, Gonzalez-Neira A, Lareu MV, Carracedo A. (2003) Insights into Iberian population origins through the construction of highly informative Y-chromosome haplotypes using biallelic markers, STRs, and the MSY1 minisatellite. *Am J Phys Anthropol.* 122:147-61.
- Caglia A, Novelletto A, Dobosz M, Malaspina P, Ciminelli BM, Pascali VL (1997) Y-chromosome STR loci in Sardinia and continental Italy reveal islander-specific haplotypes. *Eur J Hum Genet* 5:288-92
- Cann HM, de Toma C, Cazes L, Legrand M-F, Morel V, Piouffre L, Bodmer J, et al (2002) A human genome diversity cell line panel. *Science* 296:261–262

- Casanova M, Leroy P, Boucekkine C, Weissenbach J, Bishop C, Fellous M, Purrello M, et al (1985) A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* 230:1403-6.
- Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The History and Geography of Human Genes*. Princeton University Press, Princeton
- Chakraborty R, Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc Natl Acad Sci U S A* 94:1041-6
- Chikhi L, Nichols RA, Barbujani G, Beaumont MA (2002) Y genetic data support the Neolithic demic diffusion model. *Proc Natl Acad Sci U S A* 99:11008-13.
- Cinnioğlu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, et al (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet.* 114:127-148
- Collins FS, Guyer MS, Charkravarti A (1997) Variations on a theme: cataloging human DNA sequence variation. *Science* 278:1580-1.
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA. (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet.* 70:1197-214.
- Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cali F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R. (2004) Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet.* 74:1014-22.
- de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* 110:134-49
- Ellegren, H. (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet.* 5:435-45.
- Flores C, Maca-Meyer N, Gonzalez AM, Oefner PJ, Shen P, Perez JA, Rojas A, Larruga JM, Underhill PA. (2004) Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. *Eur J Hum Genet.* 12:855-63.
- Forster P, Rohl A, Lunnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, Brinkmann B (2000) A short tandem repeat-based phylogeny for the human Y chromosome. *Am J Hum Genet* 67:182-96.

- Francaletti P, Morelli L, Underhill PA, Lillie AS, Passarino G, Useli A, Madeddu R, Paoli G, Tofanelli S, Calo CM, Ghiani ME, Varesi L, Memmi M, Vona G, Lin AA, Oefner P, Cavalli-Sforza LL. (2003) Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am J Phys Anthropol.* 121:270-9.
- Gonçalves R, Freitas A, Branco M, Rosa A, Fernandes AT, Zhivotovsky LA, Underhill PA, Kivisild T, Brehm A (2005) Y-chromosome lineages from Portugal, Madeira and Açores record elements of sephardim and berber ancestry *Ann of Hum Genet* 69:1-12
- Hammer MF. (1994) A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol.* 11:749-61.
- Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. (1998) Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. *Mol Biol Evol.* 15:427-41.
- Hammer MF, Horai S. (1995) Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet.* 56(4):951-62.
- Helgason A, Sigurdsson S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, Stefansson K. (2000) Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet.* 67:697-717.
- Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* 6:799-803.
- Hurles ME, Irvén C, Nicholson J, Taylor PG, Santos FR, Loughlin J, Jobling MA, et al (1998) European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. *Am J Hum Genet* 63:1793-806
- Hurles ME, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Perez-Lezaun A, Bosch E, Shlumukova M, Cambon-Thomsen A, McElreavey K, Lopez De Munain A, Rohl A, Wilson IJ, Singh L, Pandya A, Santos FR, Tyler-Smith C, Jobling MA. (1999) Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. *Am J Hum Genet.* 65:1437-48.
- Jobling MA (2001) In the name of the father: surnames and genetics. *Trends Genet* 17:353-7.
- Jobling MA, Heyer E, Dieltjes P, de Knijff P. (1999) Y-chromosome-specific microsatellite mutation rates re-examined using a minisatellite, MSY1. *Hum Mol Genet.* 8:2117-20.
- Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L, Crawford MH, Tyler-Smith C. (1996) Recurrent duplication and deletion

- polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet.* 5:1767-75.
- Jobling MA, Bouzekri N, Taylor PG (1998) Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). *Hum Mol Genet* 7:643-53.
- Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* 110:118-24
- Jobling MA, Tyler-Smith C (2000) New uses for new haplotypes the human Y chromosome, disease and selection [In Process Citation]. *Trends Genet* 16:356-62
- Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 4:598-612.
- Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W, Harihara S, de Knijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF. (1999) Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet.* 64:817-31.
- Karafet TM, Osipova LP, Gubina MA, Posukh OL, Zegura SL, Hammer MF. (2002) High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum Biol.* 74:761-89.
- Kayser M, Brauer S, Weiss G, Schiefenhover W, Underhill PA, Stoneking M (2001) Independent histories of human Y chromosomes from Melanesia and Australia. *Am J Hum Genet* 68:173-190.
- Kayser M, de Knijff P, Dieltjes P, Krawczak M, Nagy M, Zerjal T, Pandya A, Tyler-Smith C, Roewer L. (1997) Applications of microsatellite-based Y chromosome haplotyping. *Electrophoresis.* 18:1602-7.
- Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Kruger C, et al (2000) Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet* 66:1580-8.
- Kayser M, Kittler R, Erler A, Hedman M, Lee AC, Mohyuddin A, Mehdi SQ, Rosser Z, Stoneking M, Jobling MA, Sajantila A, Tyler-Smith C. (2004) A comprehensive survey of human Y-chromosomal microsatellites. *Am J Hum Genet.* 74:1183-97.
- Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, et al (1999) Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr Biol* 9:1331-1334

- Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, Metspalu E, et al (2003) The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet* 72:313-332
- Laitinen V, Lahermo P, Sistonen P, Savontaus ML (2002) Y-chromosomal diversity suggests that Baltic males share common Finno-Ugric-speaking forefathers. *Hum Hered* 53:68-78.
- Luis JR, Rowold DJ, Regueiro M, Caeiro B, Cinnioglu C, Roseman C, Underhill PA, et al (2004) The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am J Hum Genet* 74:532-544
- Maca-Meyer N, Gonzalez AM, Larruga JM, Flores C, Cabrera VM. (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* 2:13.
- Maca-Meyer N, Sanchez-Velasco P, Flores C, Larruga JM, Gonzales AM, Oterino A, Leyva-Cobian F (2003) Y Chromosome and Mitochondrial DNA Characterization of Pasiegos, A Human Isolate from Cantabria (Spain). *Annals of human Genetics* 67:327-339
- Marjanovic D, Bakal N, Pojskic N, Kapur L, Drobic K, Primorac D, Bajrovic K, Hadziselimovic R. (2005) Population data for the twelve Y-chromosome short tandem repeat loci from the sample of multinational population in Bosnia and Herzegovina. *J Forensic Sci.* 50:223-5.
- Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics* 156:297-304.
- Nasidze I, Sarkisian T, Kerimov A, Stoneking M (2003) Testing hypotheses of language replacement in the Caucasus: evidence from the Y-chromosome. *Hum Genet* 112:255-61.
- Nebel A, Filon D, Brinkmann B, Majumder PP, Faerman M, Oppenheim A (2001) The Y chromosome pool of Jews as part of the genetic landscape of the middle east. *Am J Hum Genet* 69:1095-112.
- Ngo KY, Vergnaud G, Johnsson C, Lucotte G, Weissenbach J. (1986) A DNA probe detecting multiple haplotypes of the human Y chromosome. *Am J Hum Genet.* 38:407-18.
- Passarino G, Semino O, Magri C, Al-Zahery N, Benuzzi G, Quintana-Murci L, Andellnovic S, et al (2001) The 49a,f haplotype 11 is a new marker of the EU19 lineage that traces migrations from northern regions of the Black Sea. *Hum Immunol* 62:922-32.
- Passarino G, Cavalleri GL, Lin AA, Cavalli-Sforza LL, Borresen-Dale AL, Underhill PA (2002) Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. *Eur J Hum Genet* 10:521-9.

- Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara-Benerecetti AS. (1998) Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *Am J Hum Genet.* 62:420-34.
- Quintana-Murci L, Semino O, Bandelt H-J, Passarino G, McElreavey K, Santachiara-Benerecetti AS (1999) Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nat Genet* 23:437-441
- Quintana-Murci L, Krausz C, Zerjal T, Sayar SH, Hammer MF, Mehdi SQ, Ayub Q, et al (2001) Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *Am J Hum Genet* 68:537-42.
- Quintana-Murci L, Bigham A, Rouba H, Barakat A, McElreavey K, Hammer M. (2004) Y-chromosomal STR haplotypes in Berber and Arabic-speaking populations from Morocco. *Forensic Sci Int.* 140:113-5.
- Richards M (2003) The Neolithic invasion of Europe. *Annu Rev Anthropol* 32:135-162
- Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, et al (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am J Hum Genet* 59:185-203
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, et al (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* 67:1251-1276
- Richards M, Macaulay V, Torroni A, Bandelt HJ (2002) In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet* 71:1168-74.
- Rootsi S (2004) Human Y-chromosomal variation in European populations. PhD thesis.
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, et al (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526-43.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, et al (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290:1155-9.
- Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, et al (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023-34.

- Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS (1996) A view of the neolithic demic diffusion in Europe through two Y chromosome-specific markers. *Am J Hum Genet* 59:964-8
- Semino O, Santachiara-Benerecetti A, Falaschi F, Cavalli-Sforza L, Underhill P (2002) Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am J Hum Genet* 70:265-8
- Stringer C (2000) Coasting out of Africa. *Nature* 405:24-5, 27
- Stumpf MP, Goldstein DB. (2001) Genealogical and evolutionary inference with the human Y chromosome. *Science*. 291:1738-42.
- Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, Scozzari R, et al (2004) The effective mutation rate at y chromosome short tandem repeats, with application to human population-divergence time. *Am J Hum Genet* 74:50-61.
- Zvelebil M (2000) The social context of the agricultural transition in Europe. In: Renfrew C, Boyle K (eds) *Archaeogenetics: DNA and the population prehistory of Europe*. McDonald Institute for Archaeological Research Monograph Series, Cambridge University, Cambridge, pp 57-79
- Zerjal T, Beckman L, Beckman G, Mikelsaar AV, Krumina A, Kucinskas V, Hurles ME, et al (2001) Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in Northern European populations. *Mol Biol Evol* 18:1077-87.
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, et al (1997) Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 60:1174-83
- Zerjal T, Wells R, Yuldasheva N, Ruzibakiev R, Tyler-Smith C (2002) A genetic landscape reshaped by recent events: y-chromosomal insights into central Asia. *Am J Hum Genet* 71:466-82
- Tajima A, Hayami M, Tokunaga K, Juji T, Matsuo M, Marzuki S, Omoto K, et al (2004) Genetic origins of the Ainu inferred from combined DNA analyses of maternal and paternal lineages. *J Hum Genet* 49:187-193
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley R, Oefner PJ, et al (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet*. 65:43-62
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, et al (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358-361

Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, et al (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc Natl Acad Sci U S A 98:10244-9.

YCC (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339-48

<http://www.dnaheritage.com/masterclass2.asp>

LISAD

Tabel 12. Haplogrupi E ja subhaplogruppide sagedused

	n	nE	E3a	E3*	E3b*	E3b3	E3b1	E3b2	E1	E2	YAP	SRY ₈₂₉₉	Allikas
AAFRIKA													
Maroko	244	182	16		11	1	21	130	3				
Maroko (araablast)	44	32	3		1		5	23					Bosch et al. 2001
Põhja-/Kesk-Maroko (berberid)	63	55	6		5		1	41	2				Bosch et al. 2001
Lõuna-Maroko (berberid)	40	35	1		3		5	26					Bosch et al. 2001
Maroko	97	60	6		2	1	10	40	1				antud töö
Kamerun	166	70	61				1		5	2			
Põhja-Kamerun	69	23	14				1		5	2			Cruciani et al.2002
Lõuna-Kamerun	83	36	36										Cruciani et al.2002
Lõuna-Kamerun (bantu)	14	11	11										Luis et al. 2004
Egiptus	263	113	5		8	15	59	21	3		1		
Egiptus (araablast)	147	56	3		4	10	26	12					Luis et al. 2004
Põhja-Egiptus	21	8				1	6	1					Cruciani et al. 2004
Lõuna-Egiptus	34	6					6						Cruciani et al. 2004
Egiptus	116	43	2		4	4	21	8	3		1		Hela Help
Algeeria	32	21			1	1	2	17					Semino et al. 2004
Tuneesia	58	32	2		2	3	9	16					Semino et al. 2004
Mali	44	37	9					13	15				Underhill et al. 2000
Burkina Faso	106	89	72		1				4	12			Cruciani et al.2002
Tanzaania (wairak)	43	35	18		16					1			Luis et al. 2004
Benin (fon)	100	100	95							5			Luis et al. 2004
Rwanda (hutu)	69	65	57		2					6			Luis et al. 2004
Somaalia	23	16			4		12						Cruciani et al. 2004
Senegal	139	132	112		7		1	1	7	4			Semino et al. 2002
Sudaan	40	12					7	2	1	2			Underhill et al. 2000
Lõuna-Aafika (bantu)	44	32	24	1	1					7			Underhill et al. 2000
Lõuna-Aafika (khoisian)	59	32	18		10					3			Cruciani et al.2002
Keenia (bantu)	29	19	10		4					5			Luis et al. 2004
Etiopia (amhara)	48	17			5	1	11						Semino et al. 2002

	n	nE	E3a	E3*	E3b*	E3b3	E3b1	E3b2	E1	E2	YAP	SRY ₈₂₉₉	Allikas
LÄHIS-IDA													
Liibanon	72	15			1	4	7	3					
Liibanon	42	8				2	5	1					Semino et al. 2004
Liibanon	30	7			1	2	2	2					Hela Help
Omaan	134	30	8			16	3			2			
Omaan (araablast)	121	28	8			15	2			2			Luis et al. 2004
Omaan	13	2				1	1						Cruciani et al.2002
Türgi	870	95	2	2	1	36	52	2					
Türgi (Istanbuli türklased)	46	6				1	4	1					Semino et al. 2004
Türgi (Konya türklased)	117	17				2	15						Semino et al. 2004
Kesk-Türgi	61	6				2	4						Cruciani et al.2002
Türgi	523	59	1	2		29	26	1					Cinnioğlu et al. 2004
Türgi	93	5	1		1	1	2						Hela Help
Türgi (kurdid)	30	2				1	1						Hela Help
Kaukaasia	803	14				2	12						
Kaukaasia	706	10					10						antud töö
Kaukaasia (aserid)	97	4				2	2						Cruciani et al.2002
Saudi-Araabia	216	21	2			2	16	1					antud töö
Iraak	218	20	2			6	12						Semino et al. 2004
Iraan	83	11	1			2	7					1	Hela Help
Jordania	31	9	1			4	4						Hela Help
Süüria	91	5				1	3		1				Hela Help
Kuveit	204	21	9		2	4	6						antud töö
EUROOPA													
Sitsiilia	191	44			3	11	26	4					
Sitsiilia	136	29				9	19	1					Cruciani et al.2002
Sitsiilia	55	15			3	2	7	3					Semino et al. 2004
Kreeka	159	27				3	24						
Kreeka	84	20				2	18						Semino et al. 2004

	n	nE	E3a	E3*	E3b*	E3b3	E3b1	E3b2	E1	E2	YAP	SRY ₈₂₉₉	Allikas
Kreeka mandriosa	75	7				1	6						antud töö
Albaania	95	31					30	1					
Albaania	44	11					11						Semino et al. 2004
Albaania	51	20					19	1					antud töö
Horvaatia	515	27				1	26						
Horvaatia	57	5				1	4						Semino et al. 2004
Horvaatia	458	22					22						antud töö
Ungari	236	17			1	1	14	1					
Ungari	53	5				1	4						Semino et al. 2004
Ungari (mustlased)	70	5					5						Hela Help
Ungari	113	7			1		5	1					Hela Help
Ukraina	209	12				1	11						
Ukraina	93	8				1	7						Semino et al. 2004
Ukraina	116	4					4						antud töö
Hispaania	427	66		2	2	6	18	38					
Hispaania/Kataloonia	33	2					1	1					Semino et al. 2004
Hispaania/Andaluusia	76	7					3	4					Bosch et al. 2001
Hispaania/Leon	60	6				1	2	3					Flores et al. 2004
Hispaania/Galicia	19	6		1		2	1	2					Flores et al. 2004
Hispaania/Valencia	31	5		1		2		2					Flores et al. 2004
Hispaania/Kantaabria	56	24			1			23					Cruciani et al.2002
Hispaania/Astuuria	90	12				1	9	2					Cruciani et al.2002
Lõuna-Hispaania	62	4			1		2	1					Cruciani et al.2002
Hispaania/Sevilla	155	11		1		2	1	7					Flores et al. 2004
Rumeenia	352	40				1	39						
Rumeenia	338	37				1	36						antud töö
Rumeenia	14	3					3						Cruciani et al. 2004
Moldaavia	139	27				1	25	1					
Moldaavia (gagauusid)	79	18				1	17						antud töö
Moldaavia	60	9					8	1					antud töö
Slovakkia	267	11				3	8						

	n	nE	E3a	E3*	E3b*	E3b3	E3b1	E3b2	E1	E2	YAP	SRY ₈₂₉₉	Allikas
Slovakkia	200	7				2	5						antud töö
Slovakkia	67	4				1	3						Hela Help
Portugal													
Põhja-Portugal	101	24			5	2	8	6	3				Gonçalves et al. 2005
Kesk-Portugal	102	16				4	6	5	1				Gonçalves et al. 2005
Lõuna-Portugal	100	9			1		1	6	1				Gonçalves et al. 2005
Põhja-Portugal	109	17		1		3	7	6					Flores et al. 2004
Prantusmaa	122	8					5	3					
Lõuna-Prantsusmaa	37	1					1						antud töö
Prantsusmaa	85	7					4	3					Cruciani et al.2002
Itaalia	385	50				5	42	3					
Põhja-Itaalia	67	8				1	6	1					Cruciani et al.2002
Kesk-Itaalia	89	12					10	2					Cruciani et al.2002
Lõuna-Itaalia	87	12				2	10						Cruciani et al.2002
Kesk-Itaalia	56	6					6						Semino et al. 2004
Kagu-Itaalia	86	12				2	10						Semino et al. 2004
Eesti	291	12			1	1	10						
Eesti	74	4			1		3						Cruciani et al.2002
Eesti	217	8				1	7						Hela Help
Bosnia ja Hertsegoviina	369	44			2		42						
Bosnia ja Hertsegoviina	256	37			2		35						Marjanovic et al.2005
Bosnia ja Hertsegoviina	113	7					7						antud töö
Makedoonia	59	12				1	11						Semino et al. 2004
Poola	99	4					4						Semino et al. 2004
Sardiinia	139	7			1	2	4						Semino et al. 2004
Kreeka/Rhodos	91	15				3	10	2					antud töö
Šveits	77	5					4	1					antud töö
Sloveenia	55	2					2						antud töö
Kreeka/Kreeta	93	9				2	7						antud töö

	n	nE	E3a	E3*	E3b*	E3b3	E3b1	E3b2	E1	E2	YAP	SRY ₈₂₉₉	Allikas
Küpros	201	26	2			5	19						antud töö
Kreeka/Peloponnesose poolsaar	42	13				1	12						antud töö
Prantsusmaa (baskid)	16	1					1						Cann et al. 2002
Prantsusmaa/Korsika	140	9			1	2	6						Cruciani et al.2002
Taani	35	1					1						Cruciani et al.2002
Läti	86	1					1						Hela Help
Tšehhi	53	4					4						Hela Help

n - indiviidide arv valimis

nE - haplogruppi E kuuluvate indiviide arv

Arvud tabelis tähistavad vastavasse haplogruppi kuuluvaid indiviide.