

VILNIAUS UNIVERSITETAS

INGRIDA UKTVERYTĖ

LIETUVOS ETNOLINGVISTINIŲ GRUPIŲ GENETINĖS STRUKTŪROS  
ANALIZĖ REMIANTIS INFORMATYVIAIS GENOMO ŽYMENIMIS

Daktaro disertacija  
Biomedicinos mokslai, Medicina (06 B)

Vilnius, 2014 metai

Disertacija rengta 2010–2014 metais Vilniaus universiteto Medicinos fakulteto  
Žmogaus ir medicininės genetikos katedroje

**Mokslinis vadovas:**

Prof. habil. dr. Vaidutis Kučinskas (Vilniaus universitetas, biomedicinos  
mokslai, medicina – 06 B)



2.5.1. Egzomo sekos nustatymas ir žymenys.....	41
2.5.2. Populiacijų tyrimai naudojant egzomo duomenis.....	43
2.6. Populiacijų tyrimų problematika .....	45
3. DARBO METODAI.....	47
3.1. Tiriamieji.....	47
3.2. Genomo žymenų tyrimo metodai .....	49
3.3. Duomenų analizės metodai .....	50
3.3.1. Y chromosomos haplogrupių ir haplotipų analizė .....	50
3.3.2. mtDNR haplogrupių ir haplotipų analizė.....	51
3.3.3. Tolygiai po genomą pasiskirsčiusių autosomų VNP analizė .....	52
3.3.4. Egzomo autosomų žymenų analizė.....	54
4. TYRIMŲ REZULTATAI .....	55
4.1. Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis Y chromosomos haplogrupių ir haplotipų pasiskirstymu. ....	55
4.2. Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis mtDNR haplogrupių pasiskirstymu .....	77
4.3. Lietuvių populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis autosomų VNP pasiskirstymu.....	101
4.4. Lietuvių populiacijos genetinė struktūra, remiantis egzomo DNR sekos variantų pasiskirstymu .....	119
4.5. Apibendrinimas.....	129
5. IŠVADOS.....	134
6. LITERATŪROS SĄRAŠAS.....	136
PADĖKA .....	156
1 PRIEDAS. TYRIMO METODAI .....	157

2 PRIEDAS. PAPILDOMA MEDŽIAGA .....	243
APIE AUTOREŲ .....	258

## SANTRUMPU SAŽAŠAS

3' NTS – 3' netransliuojama sritis (angl. *3' UTR, 3' Untranslated Region*)

5' NTS – 5' netransliuojama sritis (angl. *5' UTR, 5' Untranslated Region*)

A – Aukštaitija

AMOVA – molekulinės įvairovės analizė (angl. *Analysis of Molecular Variance*)

ASW – Afrikos kilmės Šiaurės Rytų JAV gyventojai (angl. *African Ancestry in Southwest US*)

bp – bazių pora (angl. *base pair*)

CEU – Šiaurės ar Vakarų Europos kilmės Jutos gyventojai (angl. *Utah Residents with Northern and Western European Ancestry*)

CHB – kinų populiacija iš Pekino miesto, Kinija (angl. *Han Chinese in Beijing, China*)

CHD – kinų populiacija iš Denverio miesto, Koloradas (angl. *Chinese in Metropolitan Denver, Colorado*)

CHS – kinų populiacija iš Pietų Kinijos (angl. *Southern Han Chinese, China*)

CLM – kolumbiečių populiacija iš Kolumbijos (angl. *Colombian in Medellin, Colombia*)

CV – dažnas sekos variantas (angl. *Common variant*) MAF >5 %

db – duomenų bazė (angl. *database*)

DNR – deoksiribonukleorūgštis (angl. *DNA, Deoxyribonucleic Acid*)

FIN – suomių populiacija iš Suomijos (angl. *Finnish in Finland*)

GBR – britų populiacija iš Anglijos ir Škotijos (angl. *British in England and Scotland*)

GIH – indų kilmės Hiustono gyventojai (angl. *Gujarati Indians in Houston, Texas*)

HVR – mtDNR hipervariabilus regionas (angl. *Hypervariable Region*)

IBS – Iberijos regiono populiacija iš Ispanijos (angl. *Iberian populations in Spain*)

ISOGG – Tarptautinė genetinės genealogijos draugija (angl. *International Society of Genetic Genealogy*)

YRI – jorubų populiacija iš Nigerijos (angl. *Yoruba in Ibadan, Nigeria*)

JPT – japonų populiacija iš Japonijos (angl. *Japanese in Tokyo, Japan*)

kb – kilobazė, 1 000 nukleotidų (angl. *kilobase*)

LD – nepusiausvira sankiba, alelių asociacija (angl. *Linkage Disequilibrium*)

LFV – retesnis sekos variantas (angl. *Low-frequency Variant*) MAF 0,5–5 %

LWK – luchjų populiacija iš Kenijos (angl. *Luhya in Webuye, Kenya*)

MAF – retojo alelio dažnis (angl. *Minor Allele Frequency*)

Mb – megabazė, 1 000 000 nukleotidų (angl. *Megabase*)

MDS – daugiamačių skalių metodas (angl. *Multidimensional Scaling*)

MKK – *Kinyawa* masajų populiacija iš Kenijos (angl. *Maasai in Kinyawa, Kenya*)

mtDNR – mitochondrijų genomo DNR (angl. *mtDNA, Mitochondrial DNA*)

MXL/MEX – meksikiečių kilmės Los Andželo gyventojai (angl. *Mexican Ancestry in Los Angeles, California*)

$N_e$  – efektyvus populiacijos dydis (angl. *Effective Population Size*), idealios populiacijos, kuriai būdingas genų dreifas lygus tiriamos populiacijos genų dreifui, dydis

NJ – artimiausių kaimynų jungimu paremtas klasterizavimo metodas (angl. *Neighbor Joining*)

np – nukleotidų pora (angl. *Nucleotide Pair*)

NR1 – nerekombinuojanti Y chromosomos dalis (angl. *Non-recombining Region of the Y chromosome*)

nt – nukleotidas (angl. *Nucleotide*)

PA – Pietų Aukštaitija

PAR – pseudoautosominis regionas (angl. *Pseudoautosomal Region*)

PI – pasiklovimo intervalas (angl. *CI, Confidence Intervals*)

PK – pagrindinė komponentė (angl. *PC, Principal Component*)

PKA – pagrindinių komponentių analizė (angl. *PCA, Principal Component Analysis*)

PL – paskutinis ledynmetis, pasibaigęs prieš ~11 700 metų (angl. *LGM, Last Glacial Maximum*)

PUR – puerto rikičių populiacija iš Puerto Riko (angl. *Puerto Rican in Puerto Rico*)

PŽ – Pietų Žemaitija

RA – Rytų Aukštaitija

rCRS – Kembridžo referentinė mtDNR seka, NC\_012920.1, GI:251831106 (angl. *Revised Cambridge Reference Sequence*)

RNR – ribonukleorūgštis (angl. *RNA, Ribonucleic Acid*)

rRNR – ribosominė ribonukleorūgštis (angl. *rRNA, Ribosomal Ribonucleic Acid*)

RSRS – rekonstruota žmogaus referentinė mtDNR seka (angl. *Reconstructed Sapiens Reference Sequence*)

RV – retas sekos variantas (angl. *Rare Variant*) MAF <0,5 %

SN – standartinis nuokrypis (angl. *SD, Standard Deviation*)

ŠŽ – Šiaurės Žemaitija

TMRCAs – laikas iki artimiausio bendro protėvio (angl. *Time to Most Recent Common Ancestor*)

tRNR – transportinė ribonukleorūgštis (angl. *tRNA, Transport Ribonucleic Acid*)

TSI – Toskanos regiono populiacija iš Italijos (angl. *Toscans in Italy*)

TTKS – trumpos tandeminės kartotinės sekos (angl. *STR, Short Tandem Repeats*)

UEP – vienetinė DNR sekos pakaita (angl. *Unique Event Polymorphism*)

VA – Vakarų Aukštaitija

VNP – vieno nukleotido polimorfizmas (angl. *SNP, Single Nucleotide Polymorphism*)

VU MF ŽMGK – Vilniaus universiteto Medicinos fakulteto Žmogaus ir medicininės genetikos katedra

VŽ – Vakarų Žemaitija

Ž – Žemaitija



## 1. ĮVADAS

Genetinės įvairovės skirtumai nustatomi ne tik tarp populiacijų iš skirtingų žemynų (Henn *et al.* 2010) ar to paties žemyno (Nelis *et al.* 2009), bet ir populiacijos viduje (Jakkula *et al.* 2008). Nevienodą genetinę įvairovę tarp populiacijų ir jų viduje lemia kalbiniai, kultūriniai ar geografiniai skirtumai. Pastarasis, manoma, gali būti viena iš pagrindinių esamo skirtingo populiacijų genetinio fondo priežasčių. Tiriant populiacijų struktūrą naudojami skirtingi genetiniai žymenys: (1) Y chromosomos haplogrupės ir haplotipai; (2) mtDNR haplogrupės ir haplotipai; (3) tolygiai po visą genomą pasiskirstę dažni variantai; (4) koduojančios genomo dalies, egzomo, DNR sekos variantai. Tiriant daugiau genetinių žymenų ir pasirinkus tinkamus duomenų kokybės parametrus galima nustatyti tarp artimų populiacijų ar populiacijos viduje esamus mažus genetinės įvairovės skirtumus. Neįvertinti genetinės įvairovės skirtumai tarp populiacijų ar jų viduje gali lemti potencialiai klaidingai teigiamus rezultatus ir išvadas apie tiriamą populiaciją. Nustatyti tiriamos populiacijos genetinės įvairovės skirtumai yra svarbūs tiriant mažas imtis ar tyrimams, kurių tikslas rasti mažus genetinių variantų efektus. Pastaraisiais tyrimais nustatyta arba nenustatyta asociacija tarp genetinio veiksnio ir požymio gali būti dėl neįvertintų esamų genetinės įvairovės skirtumų tiriamoje populiacijoje.

Ankstesnių tyrimų analizės rezultatai parodė Pietų Aukštaitijos statistiškai reikšmingus skirtumus nuo likusių etnolingvistinių grupių, remiantis dviejų iš 12 tirtų kraujo grupių sistemų (P, LW) P1 ir LW<sup>b</sup> alelių dažnių pasiskirstymu. Taip pat statistiškai reikšmingi skirtumai nustatyti tarp Šiaurės Žemaitijos ir Pietų Aukštaitijos, remiantis *Alu* TPA25 pasiskirstymu tirtoje lietuvių populiacijoje (Kučinskas 2001). Remiantis mtDNR ar Y chromosomos haplogrupių ir haplotipų pasiskirstymu, statistiškai reikšmingų skirtumų tarp Lietuvos etnolingvistinių grupių nerasta (Kasperavičiūtė *et al.* 2004).

**Naujumas.** Šis darbas, remiantis detaliau nustatyta mtDNR, Y chromosomos haplogrupių ir haplotipų įvairove ir pirmą kartą tolygiai po genomą pasiskirsčiusių ir egzomo DNR sekos variantų pasiskirstymu, įvertina

šešių Lietuvos etnolingvistinių grupių genetinę struktūrą, taip pat lietuvių populiacijos padėtį kitų analizuotų populiacijų atžvilgiu. Pirmą kartą pateiktas darbo metu naudotų skirtingų genomo žymenų, pritaikytų populiacijos struktūrai nustatyti, palyginimas.

**Aktualumas.** Darbo metu gauti rezultatai papildo ankstesnių tyrimų rezultatus ir taps pagrindu tolesniems lietuvių populiacijos genetinės struktūros tyrimams. Lietuvių populiacijos genetinės struktūros nustatymas, analizuojant daugelio genomo žymenų alelių pasiskirstymą, bus aktualus kitiems planuojamiems ar vykdomiems lietuvių populiacijos tyrimams, kurių strategijos sudarymui ir rezultatams svarbus esamas netolygus genetinės įvairovės pasiskirstymas. Šiame darbe, naudojant visame pasaulyje plačiai taikomus naujos kartos metodus, gauti duomenys ir rezultatai suteikia galimybę palyginti lietuvių populiaciją su kitų populiacijų naujausiais duomenimis. Analizuojant išsamius duomenis, galima įvertinti mažus skirtumus tarp kaimyninių populiacijų, pastarųjų įtaką lietuvių populiacijos genetinio fondo formavimuisi ar reikšmingus lietuvių populiacijos istorijos įvykius. Turimi duomenys, taikant naujos kartos technologijas, skatintų bendradarbiauti su kitomis tarptautinėmis mokslininkų grupėmis atliekant pasaulinius tyrimus.

**Tęstinumas.** Tolygiai praplėstos tiriamųjų imties skirtingų genomo žymenų molekuliniai genetiniai tyrimai leistų patikslinti ir (ar) patvirtinti šio darbo gautus rezultatus. Y chromosomos filogenetinio medžio dažniausių haplogrupių (N1c1 ir R1a1a) detalesni tyrimai suteiktų galimybę geriau įvertinti šių grupių plitimo į dabartinę Lietuvos teritoriją ar jos viduje. Pirmieji bendrosios lietuvių populiacijos plataus masto ir egzomo sekoskaitos tyrimai ne tik suteikia naujos informacijos, bet ir iškelia daug naujų klausimų. Šiems klausimams spręsti reikalingi didesnės imties tyrimai ir nauji duomenų analizės metodai. Vilniaus universiteto Medicinos fakulteto Žmogaus ir medicininės genetikos katedroje įdiegti naujos kartos metodai suteikia galimybę kurti lietuvių populiacijos referentinio genomo duomenų bazę.

### ***Tikslas***

Įvertinti lietuvių populiacijos genetinę struktūrą ir padėtį kitų populiacijų atžvilgiu, naudojant informatyvius žmogaus genomo žymenis.

### ***Uždaviniai***

- (1) Nustatyti ir įvertinti lietuvių populiacijos vidinę genetinę struktūrą ir padėtį kitų populiacijų atžvilgiu pagal Y chromosomos genetinius žymenis.
- (2) Remiantis mitochondrijų DNR genetiniais žymenimis, nustatyti ir įvertinti lietuvių populiacijos vidinę genetinę struktūrą ir padėtį kitų populiacijų atžvilgiu.
- (3) Naudojant tolygiai po visą branduolio genomą pasiskirsčiusius genetinius žymenis, nustatyti ir įvertinti lietuvių populiacijos vidinę genetinę struktūrą ir padėtį kitų populiacijų atžvilgiu.
- (4) Nustatyti egzomo, koduojančios branduolio genomo dalies, sekos variantų pasiskirstymą tirtoje lietuvių populiacijoje.
- (5) Palyginti tarpusavyje lietuvių populiacijos genetinės struktūros analizei naudotus genomo žymenis.

### ***Ginamieji teiginiai***

- (1) Lietuvių populiacija pasižymi homogeniškumu – skirtumai tarp grupių, sudarytų remiantis geografinė Lietuvos etnolingvistinių grupių padėtimi, sudaro <2 %.
- (2) Y chromosomos haplogrupių ir haplotipų genetinė įvairovė pasiskirsto pagal šiaurės–pietų ašį, mtDNR haplogrupių ir haplotipų – vakarų–rytų ašį tirtoje lietuvių populiacijoje.
- (3) Tolygiai po genomą pasiskirsčiusių VNP alelių įvairovei būdinga šiaurės vakarų–pietryčių ašis tirtoje lietuvių populiacijoje.
- (4) Genetiniai atstumai tarp lietuvių ir kitų analizuotų populiacijų priklauso nuo geografinių atstumų, t. y. populiacijų tarpusavio geografinės padėties.
- (5) Skirtingų genomo žymenų alelių įvairovė pasiskirsto netolygiai tirtoje lietuvių populiacijoje.

### ***Darbo aprobacija***

Darbo rezultatai paskelbti dviejuose periodiniuose ir viename specialiaame mokslo Lietuvos recenzuojamuose leidiniuose, taip pat dvejuose užsienio ISI leidiniuose, pristatyti šešiose mokslinėse konferencijose.

### ***Disertacijos tema paskelbtų straipsnių sąrašas***

- (1) I. Lazaridis, N. Patterson, A. Mittnik, <...>, V. Kučinskas, <...>, **I. Uktverytė**, <...>. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 2014; 513: 409–13.
- (2) V. Kučinskas, **I. Uktverytė**, A. Molytė. Šiuolaikinių molekulinų genetinių ir biostatistinių metodų taikymas dabartinėse ir istorinėse populiacijose. *Metodai Lietuvos archeologijoje. Mokslas ir technologijos praeičiai pažinti*. Vilnius: Vilniaus universiteto leidykla, 2013, p. 22–64 (ISBN 978-609-459-278-2).
- (3) **I. Uktverytė**, A. Molytė, V. Kučinskas. Lietuvos populiacijos etnolingvistinių grupių genetinių ir geografinių atstumų analizė pagal Y chromosomos trumpas tandemines kartotines sekas. *Laboratorinė medicina*, 2013, t. 15, nr. 1(57), p. 3–8.
- (4) V. Kučinskas, **I. Uktverytė**. Genetic variation and genomic origin of Lithuanians. IFEH 12<sup>th</sup> World Congress on Environmental Health Vilnius, Lithuania 22–27 May 2012. *Medimond Publisher proceedings*, 2012, p. 7–11 (ISBN-978-88-7587-664-7).
- (5) **I. Uktverytė**, O. Balanovsky, E. Balanovska, V. Kučinskas. Genetinė įvairovė tarp Lietuvos etnolingvistinių grupių remiantis Y chromosomos DNR sekų tyrimais. *Laboratorinė medicina*, 2011, t. 13, nr. 2(50), p. 75–79.

### ***Stendinių pranešimų sąrašas***

- (1) **I. Uktverytė**, R. Meškienė, L. Ambrozaitytė, I. Domarkienė, A. Pranculis, N. Burokienė, A. Coj, A. Mažeikienė, V. Kasiulevičius, Z. A. Kučinskienė,

- V. Kučinskas. LITGEN – revealing genetic structure of the population of Lithuania. Europos žmogaus genetikos draugijos konferencija, Paryžius, Prancūzija, 2013. Vol. 20, suppl. 2, p. 394 (P16.067).
- (2) **I. Uktverytė**, M. Li, M. Stoneking, V. Kučinskas. mtDNA haplogroups in the population of Lithuania. Europos žmogaus genetikos draugijos konferencija, Niurnbergas, Vokietija, 2012. Vol. 20, suppl. 1, p. 257 (P10.41).
- (3) R. Meškienė, **I. Uktverytė**, J. Arasimavičius, L. Ambrozaitytė, L. Viniarskaitė, A. Irnius, V. Kučinskas. Research of the Nutritional Genomics Markers Specific in the Population of Lithuania. Konferencija „11<sup>th</sup> Baltic Congress of Laboratory Medicine”, Vilnius, Lietuva, 2012. Vol. 14, spec. suppl., p. 42 (ISSN 1392-6470).
- (4) **I. Uktverytė**, O. Balanovsky, S. Frolova, M. Kuznetsova, E. Balanovska, V. Kučinskas. The place of the population of Lithuania between Northern and Eastern Europe: Y chromosome analysis. Europos žmogaus genetikos draugijos konferencija, Amsterdamas, Olandija, 2011. Vol. 19, suppl. 2, p. 345 (P10.79).

### *Žodinių pranešimų sąrašas*

- (1) A. Molytė, **I. Uktverytė**, V. Kučinskas. Kritinė genetinių ir geografinių atstumų analizė pagal Y chromosomos trumpus tandeminius pasikartojimus žmonių populiacijose. Žodinis pranešimas jaunųjų mokslininkų konferencijoje „Bioateitis: gyvybės ir geomokslų perspektyvos“, Vilnius, Lietuva, 2012.
- (2) **I. Uktverytė**, O. Balanovsky, S. Frolova, M. Kuznetsova, E. Balanovska, V. Kučinskas. Application of Evolving Y chromosome Genetic Markers for Analysis of Lithuanian Population. Tarptautinė konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, Vilnius, Lietuva, 2012.
- (3) **I. Uktverytė**, O. Balanovsky, S. Frolova, M. Kuznetsova, E. Balanovska, V. Kučinskas. Analysis of the Population of Lithuania using

Y Chromosome Genetic Markers. Konferencija „11<sup>th</sup> Baltic Congress of Laboratory Medicine“, Vilnius, Lietuva, 2012.

- (4) I. Pepalytė, **I. Uktverytė**, V. Dirsė, V. Kučinskas. Characteristics of the Genomic Structural Variation in the Lithuanian Population. Tarptautinė konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, Vilnius, Lietuva, 2012.

## 2. TYRIMŲ APŽVALGA

### 2.1. Dabartinės Lietuvos teritorijos populiacijos formavimasis

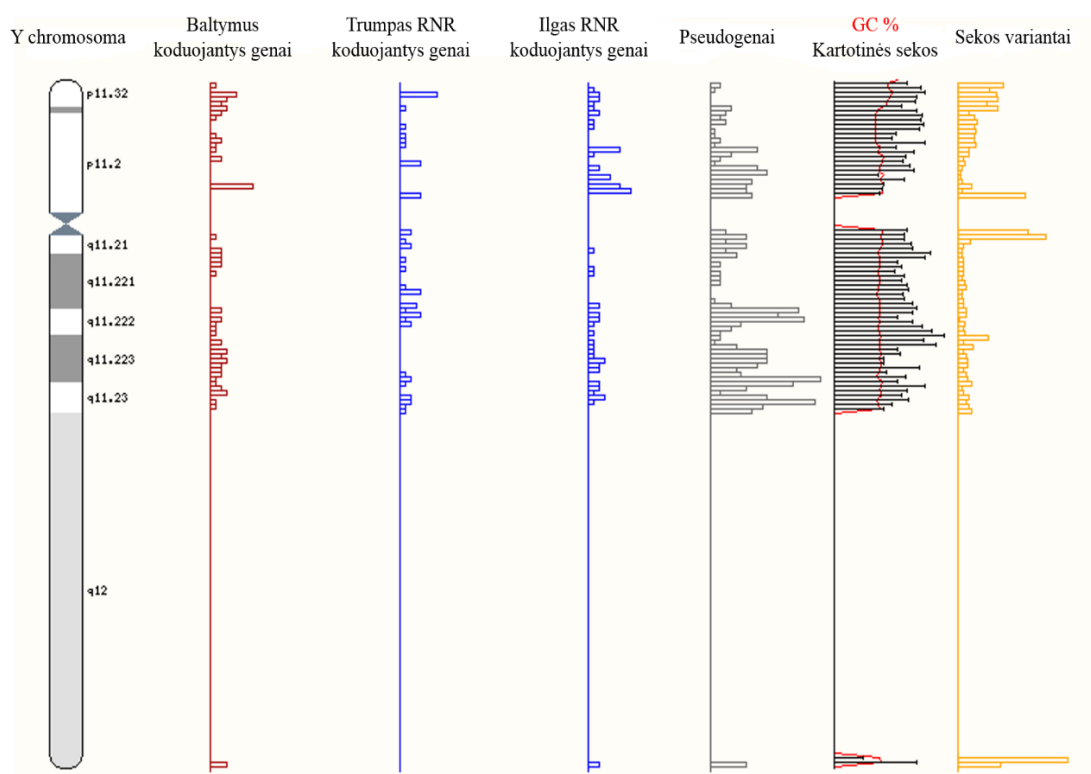
Galimybė apgyvendinti Lietuvos teritoriją atsirado tik po paskutiniojo ledynmečio (PL, angl. *LGM, Last Glacial Maximum*) pabaigos, t. y. prieš 11–12 tūkst. metų. Dabartinės Lietuvos teritorijoje po PL pirmieji buvo paleolito eros gyventojai. Keliamos dvi baltų genčių teritorijos pirmųjų gyventojų hipotezės: (1) pirmieji apgyvendino nežinomos kilmės gyventojai, t.y. ne finougrai ar indoeuropiečiai; (2) pirmieji apgyvendino Uralo kalbų grupei priklausę gyventojai (Laitinen *et al.* 2002). Aktyvus neolito kultūros plitimas Lietuvos teritorijoje prasidėjo prieš 3 500–6 500 metų (Rimantienė 1996). Mezolito ir neolito laikotarpiais vykęs baltų genčių formavimasis, indoeuropiečių paplitimas ir Neolito kultūros plitimas dabartinės Lietuvos teritorijoje lieka neaiškus, nes antropologiniai radiniai iš šių laikotarpių reti (Butrimas *et al.* 1985). Manoma, kad baltų genčių formavimasis, remiantis kalbos ir archeologiniais tyrimais, prasidėjo prieš ~5 000 metų. Įtaka ir ryšiai su kaimyninėmis populiacijomis nėra iki galo aiškūs. Archeologinių paminklų tyrimai rodo, kad finougrų įtaka baltų gentims buvo minimali (Rimantienė 1996). Slavų genčių invazija prasidėjo prieš 1 500 metų ir padalino baltų apgyvendintą teritoriją į vakarinę dalį – Prūsiją, Lietuvą, Latviją, ir rytinę, kurios gentys buvo asimiliuotos slavų populiacijos (Zinkevičius 1996). Artimų baltų genčių jungimasis, lėmęs ir dabartinės Lietuvos kalbos formavimąsi, prasidėjo prieš ~1 000 metų. Vėlesniu laikotarpiu dėl valdymo formos gyventojų minimali migracija lėmė Lietuvos etnolingvistinių grupių – šiaurės, pietų, vakarų žemaičių ir rytų, pietų, vakarų aukštaičių – susiformavimą.

### 2.2. Populiacijų tyrimai naudojant Y chromosomos DNR žymenis

#### 2.2.1. Y chromosoma

Y chromosoma, perduodama tik vyriškąja linija, t. y. iš tėvo sūnui, sudaro daugiau nei 59 milijonai bazių porų (Mb, angl. *Megabase*), t. y. 2–3 % haploidinio žmogaus branduolio genomo. Šiuo metu Y chromosomoje yra

nustatyti 72 polipeptidus koduojantys, 57 trumpų ir 70 ilgų RNR koduojančių genų, 346 pseudogenai ir 378 554 trumpi sekos variantai (2.1 pav.) (Flicek *et al.* 2014). Skiriamos dvi Y chromosomos dalys: pseudoautosominiai regionai (PAR, angl. *Pseudoautosomal Region*), esantys abiejuose Y chromosomos galuose (PAR1 ir PAR2), ir nerekombinuojantis regionas (NRY, angl. *Non-recombining Y*). PAR1, 26 Mb ilgio Y chromosomos trumpojo peties regionas, ir PAR2, 400 kilobazių (kb, angl. *kilobase*) ilgio Y chromosomos ilgojo peties regionas, turi homologinius PAR regionus X chromosomoje, tarp kurių vyksta rekombinacija apsiikeičiant genetinė informacija vyriškosios mejozės metu (Freije *et al.* 1992; Rappold 1993). NRY, nedalyvaujantis rekombinacijoje vyriškosios mejozės metu, sudaro 95 % Y chromosomos (Tilford *et al.* 2001).



2.1 pav. Y chromosomos genų, kartotinių sekų ir sekos variantų pasiskirstymas (Flicek *et al.* 2014)

Nevykstant rekombinacijai, NRY sukaupia genetinė variacija, t. y. vieno nukleotido polimorfizmai (VNP, angl. *SNP, Single Nucleotide Polymorphism*) ar evoliucijos metu įvykusios vienietinės DNR sekos pakaitos (UEP, angl. *Unique Event Polymorphism*), perduodama iš kartos į kartą. Molekuliniai



genetiniai NRY, didžiausio nerekombinuojančio žmogaus genomo fragmento ir informatyviausio haplotipo, tyrimai pritaikomi Y chromosomos filogenetinio medžio rekonstrukcijai ir populiacijų evoliucijos tyrimams (Underhill, Kivisild 2007). Abiejų lyčių bendroje populiacijoje Y chromosomos efektyvus populiacijos dydis ( $N_e$ ) atitinka 1/4 autosomų, 1/3 X chromosomos ir ~1 mitochondrijų genomo DNR (mtDNR, angl. *mtDNA*, *mitochondrial DNA*)  $N_e$  (Jobling, Tyler-Smith 2003).

### 2.2.2. Y chromosomos DNR žymenys

Populiacijų tyrimams ir Y chromosomos filogenetinio medžio rekonstrukcijai naudojami Y chromosomos UEP. Tiriant genealogijas nustatytas mutacijų dažnis Y chromosomoje siekia  $3 \cdot 10^{-8}$  (95 % PI  $8,9 \cdot 10^{-9} - 7 \cdot 10^{-8}$ ) mutacija/nukleotidui/kartai, kuris artimas filogenetinių linijų tyrimais nustatytam dažniui –  $2,3 \cdot 10^{-8} - 6,3 \cdot 10^{-8}$  mutacija/nukleotidui/kartai (Kuroki *et al.* 2006; Xue *et al.* 2009).

Trumpos tandeminės kartotinės sekos (TTKS, angl. *STR*, *Short Tandem Repeat*), kurių kartotinis motyvas 1–6 nt ilgio, yra Y chromosomos DNR žymuo, naudojamas populiacijų tyrimams. Žmogaus genome dažniausios yra dinukleotidinės kartotinės sekos, kurios sudaro iki 0,5 % genomo (Lander *et al.* 2001), tačiau tyrimams dažniau naudojamos 4–6 nt ilgio kartotinės sekos dėl mažesnio gebėjimo mutuoti. Mechanizmai, kuriantys TTKS žymenų variaciją: (1) netolygi perkryža mejozės metu; (2) retrotranspozicija; (3) grandinių netikslus suporavimas. Pastarasis yra pagrindinis mechanizmas, formuojantis TTKS variaciją. Grandinių netikslus suporavimas yra dažnesnis negu nustatomas TTKS variacijos lygis, kuris, manoma, yra pusiausvyra tarp grandinių netolygaus suporavimo ir šių klaidų taisymo mechanizmo (angl. *Mismatch Repair Mechanism*). TTKS variacija priklauso nuo kartotinio motyvo, ilgio, genotipo, nukleotidų sudėties, aplinkinių sekų, rekombinacijos lygio, lyties, amžiaus (Fan, Chu 2007). Tiriant TTKS pasiskirstymą Y chromosomos filogenetinio medžio haplogrupių viduje nustatytas mutacijų dažnis  $6,9 \cdot 10^{-4}$  ( $\pm 5,7 \cdot 10^{-4}$ ) mutacija/žymeniui/kartai, tiriant genealogijas

(tėvo ir vaiko poras) – mutacijų dažnis varijavo nuo  $3,78 \cdot 10^{-4}$  (95 % PI  $1,38 \cdot 10^{-5}$ – $2,02 \cdot 10^{-3}$ ) iki  $7,44 \cdot 10^{-2}$  (95 % PI  $6,51 \cdot 10^{-2}$ – $9,09 \cdot 10^{-2}$ ) mutacija/žymeniui/kartai (Ballantyne *et al.* 2010; Zhivotovsky *et al.* 2004).

TTKS neribojamo alelių skaičiaus modelis (angl. *IAM, Infinite Alleles Model*), kai mutacija sukuria populiacijoje naują TTKS alelį ir kiekviena mutacija turi vienodą tikimybę (Kimura, Crow 1964), neapibrėžė tinkamai stebimo TTKS mutacijų modelio. Vieno žingsnio mutacijų modelis (angl. *SMM, Stepwise Mutation Model*), turintis išankstines prielaidas: (1) mutacijos, apimančios nedidelį nukleotidų skaičių; (2) vienoda mutacijų, didinančių ir mažinančių kartotinio motyvo skaičių, tikimybė; (3) neribojamas alelių skaičius; (4) nuo kartotinio motyvo nepriklausantis mutacijų dažnis (Ota, Kimura 1973), yra artimesnis ir geriau apibrėžiantis stebimą TTKS mutacijų modelį. Kuriami nauji SMM variantai, kurių prielaidos skiriasi nuo klasikinio modelio, norint kuo tiksliau apibrėžti stebimą TTKS mutacijų modelį.

### 2.2.3. Y chromosomos filogenetinis medis

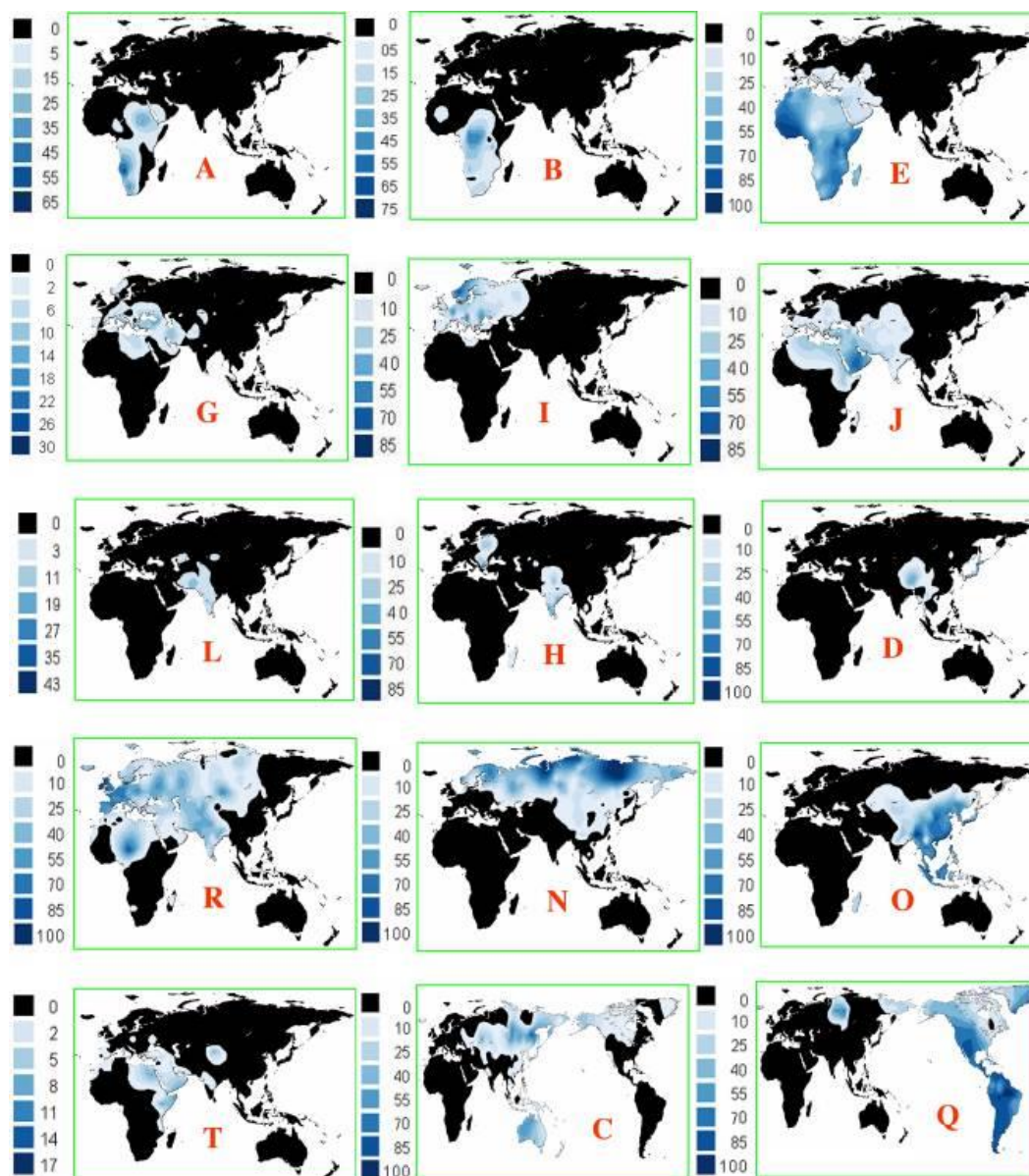
Nustačius UEP, Y chromosoma priskiriama haplogrupei. Haplogrupė – Y chromosomų, turinčių bendrą protėvinį alelį, grupė. Nustačius TTKS, Y chromosoma priskiriama haplotipui. Haplotipas – nehomologinių alelių derinys, perduodamas kartu kitai kartai. Nustačius haplogrupę ir haplotipą Y chromosoma priskiriama filogenetinei linijai (de Knijff 2000). Y chromosomos filogenetinis medis – diagrama, parodanti evoliucinius ryšius tarp atskirų haplogrupių ar filogenetinių linijų. Y chromosomos filogenetinis medis sudarytas hierarchiniu principu – Y chromosoma, priskirta haplogrupei N1c1, turi filogenetinės linijos N, N1, N1c, N1c1 protėvinius alelius. UEP tyrimai leidžia tirti ir vertinti pagrindines Y chromosomos filogenetinio medžio filogenetines linijas ir sąlyginai ilgai trunkančius populiacijos evoliucijos procesus, o TTKS tyrimas – patikslinti filogenetinę liniją ir vertinti sąlyginai trumpus populiacijos evoliucijos procesus. Pirmą Y chromosomos filogenetinį medį, naudodamas suvienodintą nomenklatūrą, paskelbė Y chromosomos konsorciumas (angl. *Y Chromosome Consortium*) (YCC 2002).

Y chromosomos filogenetini medį, kurį recenzuoja Tarptautinė genetinės genealogijos draugija (ISOGG, angl. *International Society of Genetic Genealogy*), sudaro 20 pagrindinių filogenetinių linijų A–T (ISOGG 2014). Daugelį Y filogenetinio medžio detalizavimo, nustatant naujus žymenis ir identifikuojant jų vietą filogenetiniame medyje, atliko dvi mokslinių grupės, kurioms vadovauja M. Hammer ir P. Underhill. Šios dvi grupės įvedė UEP nomenklatūrą: (1) nurodoma atitinkamai „M“ arba „P“ raidė; (2) mutacijos nustatymo eiliškumas (Karafet *et al.* 2008).

Dominuojančios filogenetinės linijos: (1) Afrikos žemyne – A, B, E; (2) Artimuosiuose Rytuose – G, J; (3) Europoje – E, G, I, J, N, R; (4) Azijoje – C, N, O; (5) Malaizijoje – M, O; (6) Australijoje – C, K; (7) Šiaurės Amerikos žemyne – Q; (8) Pietų Amerikos žemyne – P, Q (2.2 pav.) (Jobling, Tyler-Smith 2003). Europos Centrinėje ir Rytų Viduržemio regiono dalyje dominuojanti filogenetinė linija – E (Semino *et al.* 2004), Pietų Europoje – G (Rootsi *et al.* 2012), Šiaurės ir Pietų Europoje – I (Rootsi *et al.* 2004), Centrinėje ir Rytų Europoje ir Europos Centrinėje ir Rytų Viduržemio regiono dalyje – J (Di Giacomo *et al.* 2004), Šiaurės Rytų Europoje – N (Rootsi *et al.* 2007), Vakarų ir Centrinėje Europoje – R1b, Rytų Europoje – R1a (Jobling, Tyler-Smith 2003; Kayser *et al.* 2005). Y chromosomos, priklausančios filogenetinei linijai R, nustatomos >50 % Europos vyriškosios lyties atstovų (Myres *et al.* 2011). Beveik visi Europos vyriškosios lyties asmenys gali būti priskirti vienai iš 11 Europoje dominuojančių filogenetinių linijų – E3a(P46), E3b(M79, M123, V12, V13), G(M201), I1a(M253), I1b1(P37), I1b2(M223), J2(M172), N2(P43), N3(P105), R1a(M17), R1b(M269) (Wiik 2008).

Nustatytas Y filogenetinio medžio filogenetinių linijų A3–R amžius yra 101–115 tūkst. metų, filogenetinių linijų aptinkamų už Afrikos žemyno ribų amžius 57–74 tūkst. metų. Dažniausios Vakarų Europoje filogenetinės linijos R1b amžius 4,3–13 000 metų, remiantis 8,97 Mb ilgio Y chromosomos NRY fragmento sekoskaitos rezultatais. Nustatytas R1b amžius patvirtina šios filogenetinės linijos kilmę neolito, prasidėjusio prieš 12 000 metų Artimuosiuose Rytuose, laikotarpiu (Wei *et al.* 2013). Remiantis filogenetinės linijos A1b

analize, Y chromosomos filogenetinio medžio kilmė siejama su Šiaurės Vakarų Afrika (Scozzari *et al.* 2012). Filogenetinių linijų pasiskirstymas (dažnis), haplotipo įvairovė ir filogenetinių linijų amžius galbūt yra pakankama informacija daryti prielaidas apie filogenetinių linijų kilmę ir plitimą (Underhill *et al.* 2010).



2.2 pav. Y chromosomos filogenetinio medžio 15 pagrindinių filogenetinių linijų dažnių pasiskirstymas pasaulyje (Chiaroni *et al.* 2009)

#### 2.2.4. Y chromosomos filogeografija Europoje

Diskusijų keliantis Y chromosomos imigracijos ir plitimo Europoje klausimas sprendžiamas analizuojant detalesnių Y chromosomos haplogrupių pasiskirstymą Europoje. Galimos Y chromosomos kryptys į Europą: Artimieji

Rytai – Pietryčių Europa, Vakarų Azija – Centrinė Europa, Šiaurės Azija – Šiaurės Europa. Prasidėjus PL Šiaurės Europos gyventojai turėjo persikelti į mažiau atšiaurų prieglobstį Europoje: Iberijos pusiasalį, Centrinę Europą (dabartinė Ukrainos teritorija), Balkanus, Sibirą. Pirminis ir pakartotinis po PL Europos apgyvendinimas lieka detalesnių Y chromosomos filogeografinių tyrimų objektas (Wiik 2008).

Filogenetinė linija N Šiaurės Rytų Europą iš Šiaurės ar Vakarų Azijos pasiekė prieš 8–10 tūkst. metų (Shi *et al.* 2013), tai patvirtina anksčiau atliktų tyrimų analizuojant Šiaurės Rytų Europos populiacijų TTKS įvairovę haplogrupės N3 viduje. Apskaičiuotas laikas iki artimiausio filogenetinės linijos N3 bendro protėvio (TMRCA, angl. *Time to Most Recent Common Ancestor*) – 8 800 ( $\pm 1\ 500$ ) metų. Manoma, kad filogenetinės linijos R1a, kuriai apskaičiuotas TMRCA – 10 700 ( $\pm 1\ 400$ ) metų, plitimas į Pabaltijo regioną vyko iš Centrinės, o ne Rytų Europos, tačiau neatmestinas genetinio fondo maišymasis su vėlesne R1a migracija iš Rytų Europos (Lappalainen *et al.* 2008). Remiantis Y chromosomos sekos nustatymo duomenimis filogenetinės linijos R1a, pradėjusios formotis dabartinėje Irano ir Rytų Turkijos teritorijoje, atšakos R1a1a(M417), kurios TMRCA – 5 800 (95 % PI 4 800–6 800) metų, dvi grupės paplitusios Europoje ir Pietų Azijoje (Underhill *et al.* 2014). Filogenetinės linijos R1a plitimą iš Centrinės Europos patvirtina nustatytas haplogrupės R1a1a7 (R1a1a1b1a1 – (ISOGG 2014)) didžiausias dažnis (~30 %) Pietų Lenkijoje, kuris smarkiai mažėja visomis kryptimis. Tirtose Lenkijos populiacijoje apskaičiuotas R1a1a7 TMRCA – 11 000 metų, tirtose Europos populiacijose – 7 900 ( $\pm 2\ 600$ ) (Underhill *et al.* 2010). Filogenetinės linijos I1a, kurios didžiausias dažnis yra Skandinavijos pusiasalyje, migracija į Šiaurės Europą vyko iš Vakarų Europos pasibaigus PL (Rootsi *et al.* 2004). Filogenetinės linijos R1b(M412) (R1b1a2a1 (ISOGG 2014)) didžiausias dažnis matomas Vakarų Europoje su šios linijos vidinių haplogrupių išplitimu Europos pietvakarių ir šiaurės kryptimis. Apskaičiuotas filogenetinės linijos R1b(M412) TMRCA – 8 870 ( $\pm 1\ 708$ ) metų (Myres *et al.* 2011).

### 2.2.5. Europos populiacijų Y chromosomos filogenetinių linijų tyrimai

Daugelio Europos ir pasaulio populiacijų Y chromosomos filogenetinių linijų tyrimų rezultatai (2.1 lentelė) leidžia analizuoti šių populiacijų vidinę genetinę struktūrą ir (ar) padėti kitų populiacijų atžvilgiu. Y chromosomos filogenetinių linijų tyrimų apimtys keitėsi genotipuojant: (1) kelis UEP ar TTKS žymenis, leidžiančius nustatyti pagrindines haplogrupes (Kasperavičiūtė *et al.* 2004; Laitinen *et al.* 2002); (2) kelias dešimtis ar šimtus UEP ar TTKS žymenų (Karafet *et al.* 2010; Karafet *et al.* 2008; Varzari *et al.* 2013); (3) atliekant kelių Mb ilgio Y chromosomos NRY fragmento sekos nustatymą (Poznik *et al.* 2013; Rootsi *et al.* 2013; Underhill *et al.* 2014; Wei *et al.* 2013).

2.1 lentelė. Y chromosomos filogenetinių linijų tyrimai Europoje ir pasaulyje

Populiacija	Imtis	UEP	TTKS	Šaltinis
Estija	118	18	9	(Lappalainen <i>et al.</i> 2008)
Suomija	536	18	9	(Lappalainen <i>et al.</i> 2008)
Švedija	160	18	9	(Lappalainen <i>et al.</i> 2008)
Norvegija	1 766	5	9	(Dupuy <i>et al.</i> 2006)
Grenlandija	215	35	–	(Sanchez <i>et al.</i> 2004)
Danija	194	35	–	(Sanchez <i>et al.</i> 2004)
Latvija	113	18	9	(Lappalainen <i>et al.</i> 2008)
Lenkija	913	10	7	(Kayser <i>et al.</i> 2005)
Baltarusija	68	23	–	(Kharkov <i>et al.</i> 2005)
Ukraina	53	37	17	(Varzari <i>et al.</i> 2013)
Rusija	414	23	–	(Malyarchuk, Derenko 2008)
Čekija	257	21	–	(Luca <i>et al.</i> 2007)
Kroatija	109	16	8	(Barac <i>et al.</i> 2003)
Rumunija	54	37	17	(Varzari <i>et al.</i> 2013)
Moldova	125	37	17	(Varzari <i>et al.</i> 2013)
Britanija	110	13 <sup>1</sup>	17	(King, Jobling 2009)
Vokietija	1 215	10	7	(Kayser <i>et al.</i> 2005)
Italija <sup>2</sup>	3 325	–	–	(Brisighelli <i>et al.</i> 2012)
Graikija	171	29	37	(King <i>et al.</i> 2011)
Turkija	205	29	37	(King <i>et al.</i> 2011)

<sup>1</sup>Pagal (Karafet *et al.* 2008)

<sup>2</sup>Autoriaus tyrimo rezultatų ir apibendrintų literatūros šaltinių duomenys

Dalis tyrimų analizuoja populiacijų vidinę genetinę struktūrą, remiantis pagrindinių Y chromosomos haplogrupių pasiskirstymo duomenimis

(Brisighelli *et al.* 2012; Dupuy *et al.* 2006; Lappalainen *et al.* 2009; Raitio *et al.* 2001). Gauti pagrindinių Y chromosomos haplogrupių duomenys leidžia analizuoti populiacijų padėtį, kurią lėmė istoriniai įvykiai (Kayser *et al.* 2005; Lappalainen *et al.* 2008; Luca *et al.* 2007; Marjanovic *et al.* 2005), kalbiniai (Balanovsky *et al.* 2008), geografiniai (Balanovsky *et al.* 2013) skirtumai ar jų sąveika (Balanovsky *et al.* 2013; King *et al.* 2011; Varzari *et al.* 2013) kitų populiacijų atžvilgiu. Dalis nustatytų skirtumų tarp populiacijų, remiantis Y chromosomos haplogrupių pasiskirstymo duomenimis, lieka nepaaiškinti (Rebala *et al.* 2007). Tikslią evoliucijos veiksnių (migracija, genų dreifas, mutacijos, gamtinė atranka), formuojančių skirtumus tarp populiacijų, įtaką populiacijos genetinei struktūrai sunku įvertinti.

Dalis tyrimų tikslina Y chromosomos filogenetinio medžio atskirų stambių filogenetinių linijų vidinę struktūrą. Gauti rezultatai leidžia patikslinti Y chromosomos filogenetinės grupės amžių ir (ar) nustatyti naujus ryšius tarp tiriamų populiacijų. Nauji duomenys suteikia daugiau informacijos, padedančios įvertinti galimus evoliucijos veiksnius ar jų sąveikas, kurios galėjo turėti įtakos populiacijų genetinės struktūros susiformavimui (Battaglia *et al.* 2013; Larmuseau *et al.* 2012; Mendez *et al.* 2011; Scozzari *et al.* 2012; Trombetta *et al.* 2011).

#### 2.2.6. Lietuvos populiacijos Y chromosomos filogenetinių linijų tyrimai

T. Zerjal kartu su kitais mokslininkais ištyrė 363 vyriškosios lyties asmenis iš aštuonių Skandinavijos ir Pabaltijo regiono populiacijų, iš kurių 38 – iš Lietuvos populiacijos. Tyrimo metu, nustatčius 10 UEP (12f2, YAP, SRY-8299, sY81, M9, LLY22g/HindIII, Tat, 92R7, SRY-2627) genotipus, Y chromosomos priskirtos haplogrupėms (1–4, 9, 16, 21, 26). Tirtos Lietuvos populiacijos genetinė įvairovė, remiantis nustatytomis haplogrupėmis, artima tirtai Latvijos populiacijai ir siekė 0,660 ( $\pm 0,06$ ). Nustatyta, kad artimiausi, remiantis apskaičiuotais  $F_{ST}$  ir  $\Phi_{ST}$  atstumais pagal haplogrupių ir haplotipų pasiskirstymą, lietuviai ir latviai, kurie atsiskyrę nuo kitų tirtų populiacijų. Didžiausi molekulinės įvairovės analizės (AMOVA, angl. *Analysis of Molecular*

*Variance*) skirtumai nustatyti, populiacijas sugrupavus pagal geografinius ir kalbinius skirtumus. Manoma, kad genetinius skirtumus lėmė geografiniai atstumai, tačiau nustatyti skirtumai tarp baltų ir finougrų (estų) rodo, nors ir mažesnę, kalbos įtaką populiacijų genetinio fondo formavimuisi (Zerjal *et al.* 2001).

V. Laitinen kartu su kitais mokslininkais ištyrė 346 vyriškos lyties asmenis iš trijų Pabaltijo regiono populiacijų, iš kurių 114 – iš Lietuvos populiacijos. Tyrimo metu, nustačius penkių UEP (YAP, Tat, M9, 92R7, SRY-1532) genotipus, Y chromosomos priskirtos haplogrupėms (1–4, 16, 26). Tirtos Lietuvos populiacijos genetinė įvairovė, remiantis haplogrupių pasiskirstymu, artima tirtai Latvijos populiacijai ir siekė 0,663 ( $\pm 0,022$ ). Nustatyti trumpiausi, tačiau statistiškai nereikšmingi  $F_{ST}$  atstumai tarp lietuvių, latvių, estų ir marių. Kita artimiausia grupė tirtoms keturioms populiacijoms (lietuviai, latviai, estai ir maria) – rusai, lenkai, baltarusiai ir suomia. Naudojant artimiausių kaimynų jungimu (NJ, angl. *Neighbor Joining*) grindžiamą klasterizavimo metodą, lietuviai, latviai, estai ir maria sujungti į vieną klasterį. Remiantis gautais rezultatais iškelta hipotezė, kad finougrai (estai ir maria) ir baltai-indoeuropiečiai (lietuviai ir latviai) yra kilę iš bendro finougrų protėvio, nes nėra duomenų, kad estai būtų kalbėję kita nei finougrų kalba, o baltai (lietuviai ir latviai) galėjo protėvinę finougrų kalbą pakeisti indoeuropiečių kalba (Laitinen *et al.* 2002).

D. Kasperavičiūtė kartu su kitais mokslininkais ištyrė 196 vyriškosios lyties asmenis iš lietuvių populiacijos. Tyrimo metu, nustačius penkių UEP (YAP, Tat, M9, 92R7, SRY–1532) genotipus, Y chromosomos priskirtos haplogrupėms (Hg 1, 2, 3, 4, 8, 9, 16, 21, 26), nustačius devynių (DYS19 (arba DYS394), DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 ir DYS385) TTKS žymenų alelius, Y chromosomos priskirtos haplotipams. Y chromosomų, priklausančių haplogrupei N3, DYS385I ir DYS385II žymenys tirti atskirai. Nustatytos dažniausios haplogrupės – R1a (45 %) ir N3 (37 %). Lietuvių populiacija, remiantis N3 haplogrupės pasiskirstymu ir gautais rezultatais, taikant daugiamatės skalės metodą (MDS, angl. *Multidimensional Scaling*),



remiantis  $F_{ST}$  atstumais pagal haplogrupių pasiskirstymą, yra artimesnė finougrams nei indoeuropiečiams. Tirtos lietuvių populiacijos genetinė įvairovė, remiantis haplogrupių pasiskirstymu, siekė 0,653 ( $\pm 0,02$ ). Nustatyti 123 skirtingi haplotipai tirtiems 196 vyriškosios lyties asmenims ir apskaičiuota haplotipų įvairovė siekė 0,985 ( $\pm 0,004$ ). Tirtos lietuvių populiacijos homogeniškumą patvirtino nustatyti statistiškai nereikšmingi  $R_{ST}$  atstumai ir AMOVA rezultatai. Nustatytas tirtos lietuvių populiacijos augimas, priklausomai nuo skaičiavimui naudojamo mutacijų dažnio, prieš ~1 000 arba 7 000–8 000 metų. Apibendrinus nustatytas išskirtinis tirtos lietuvių populiacijos homogeniškumas, o susiformavę kalbiniai skirtumai nepastebimi genetinėje populiacijos struktūroje. Remiantis haplogrupių pasiskirstymu apskaičiuoti trumpiausi genetiniai atstumai tarp tirtos lietuvių populiacijos, lyginant su kitomis Europos populiacijomis, ir finougrų (estai, suomiai), tik šiek tiek toliau – slavai (rusai, lenkai). Aptikto didelio R1a haplogrupės dažnio viena iš hipotezių yra migracija iš Centrinės Europos. Nustatyti statistiškai reikšmingi skirtumai, remiantis haplogrupės N3 haplotipų įvairovės tyrimo rezultatais, tarp tirtų lietuvių, latvių ir estų populiacijų. Lietuvių ir estų populiacijų TTKS žymenų įvairovės haplogrupės N3 viduje tyrimų rezultatai patvirtina skirtingą estų populiacijos istoriją. Nustatyta skirtinga TTKS žymenų haplogrupių R1a ir N3 viduje genetinė įvairovė leidžia manyti apie jau susiformavusių populiacijų, haplogrupių R1a ir N3, susidūrimą. Nustatytą populiacijos augimą prieš ~1 000 metų papildė archeologiniai ir antropologiniai radiniai (Kasperavičiūtė *et al.* 2004).

L. Roewer mokslininkų grupės tyrimas parodė, kad lietuvių populiacija, remiantis  $F_{ST}$  atstumų matrica pagal septynių Y chromosomos TTKS (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS39, DYS392) žymenų alelių pasiskirstymą, artimiausia latvių populiacijai ir kartu su tirtomis slavų populiacijomis (Kroatija, Lenkija, Baltarusija, Rusija, Ukraina, Slovėnija) formuoja bendrą Rytų Europos grupę. Analizė parodė, kad Lietuvos ir Latvijos populiacijos skiriasi nuo Estijos populiacijos ir galimas paaiškinimas yra kalbiniai skirtumai (Roewer *et al.* 2005).

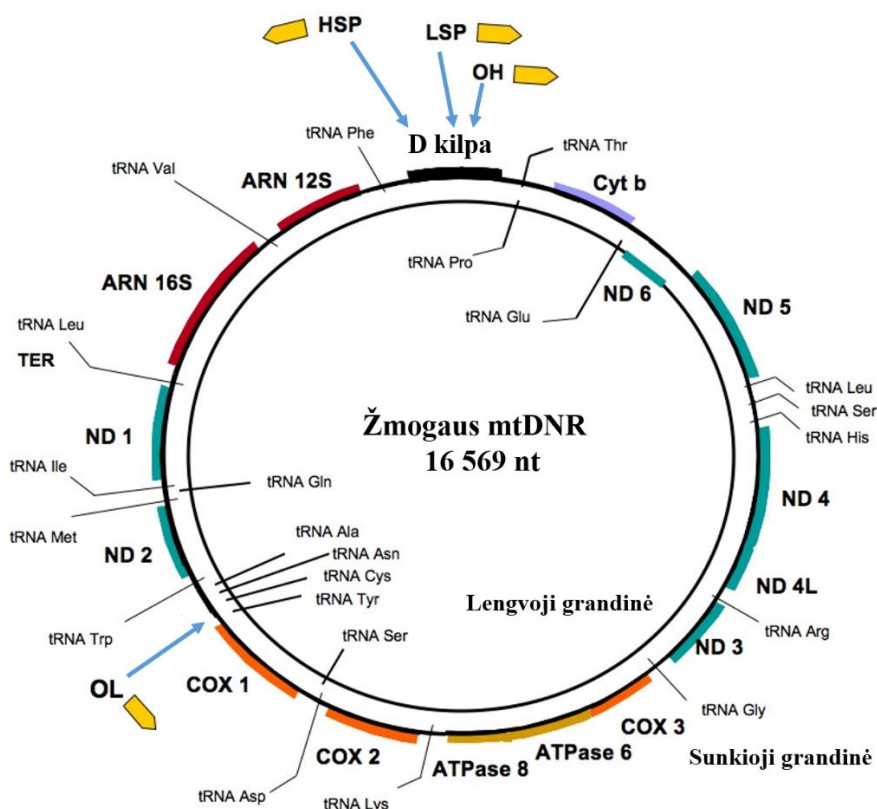
T. Lappalainen kartu su kitais mokslininkais ištyrė 1 223 vyriškosios lyties asmenis iš septynių Skandinavijos, Šiaurės Rytų Europos ir Pabaltijo regiono populiacijų, iš kurių 164 iš Lietuvos populiacijos. Nustačius 18 UEP (SRY-1532, M216, M203, P14, M170, M253, P37, M223, 12f2, M9, LLY22g, P43, Tat, M45, P36, M207, P25, M17) genotipus, Y chromosomos priskirtos haplogrupėms, nustačius devynių (DYS19, DHS385a/b, DHS388, DHS389I, DHS389II, DHS390, DHS391, DHS392, DHS393) TTKS žymenų genotipus, Y chromosomos priskirtos haplotipams. Atlikta TTKS analizė, sujungiant tirtas lietuvių ir latvių populiacijas, paaiškina >20 % variacijos haplogrupės N3 viduje. Nustatyti skirtumai tarp baltų (lietuvių ir latvių) ir finougrų (estų) tik haplogrupės N3 atveju atmeta susidariusio kalbinio barjero, kuris lemtų skirtumus visų haplogrupių atžvilgiu, priežastį. Nepastebėta sumažėjusi įvairovė haplogrupės N3 viduje atmeta galimą įkūrėjo efekto paaiškinimą. Galima nustatyti skirtumų priežastis – dviejų etapų migracija į Pabaltijo regioną, po kurios vyko baltų ir finougrų genų fondo maišymasis. Šis paaiškinimas atmeta iki šio tyrimo iškeltą hipotezę apie baltų (lietuvių ir latvių) galimos protėvinės finougrų kalbos asimiliaciją su indoeuropiečių migracijos banga (Lappalainen *et al.* 2008).

### 2.3. Populiacijų tyrimai naudojant mtDNR žymenis

#### 2.3.1. Mitochondrijų DNR (mtDNR)

Haploidinė mtDNR – ląstelės mitochondrijose esančios nuo kelių dešimčių iki kelių tūkstančių kopijų žiedinės dvigrandės 16 569 nukleotidų porų (np) ilgio DNR molekulės. Vienoje iš mtDNR grandinių yra palyginti gausu purinų (guanidinų) – tai sunkioji grandinė (angl. *Heavy chain*), kitoje gausu pirimidinų (citidinų) – tai lengvoji grandinė (angl. *Light chain*). Skiriamos dvi mtDNR sritys: (1) D kilpa arba reguliacinė sritis (angl. *D-loop*; *CR*, *control region*; 16 024–576 np), vienintelė mtDNR sritis, sudaryta iš trijų nukleotidų grandinių; (2) koduojanti sritis (CR, angl. *Coding region*; 577–16 023 np). D kilpoje išskiriami du hipervariabilūs regionai I ir II (HVRI ir HVRII). mtDNR pasižymi keliomis savybėmis: (1) genetinis kodas palyginti nedaug skiriasi nuo

universaliojo branduolio genetinio kodo; (2) stebimi sanklotiniai genai; (3) nėra introninių sekų; (4) neturi poli-A signalinių sekų, kurios suformuojamos RNR brendimo metu; (5) nuo visos mtDNR sintetinama viena RNR molekulė, kuri brendimo metu suskaidoma į atskirų genų RNR. Žinomi 37 mtDNR genai, koduojantys 2 rRNR, 22 tRNR ir 13 polipeptidų. mtDNR turi du, sunkiosios ir lengvosios grandinių, skaitymo rėmelius (angl. *HSP*, *LSP*) ir du, sunkiosios ir lengvosios grandinių, replikacijos taškus (angl. *OH*, *OL*), kas lemia nesimetrinę mtDNR replikaciją (2.3 pav.) (Bellance *et al.* 2009). Nerekombinuojanti mtDNR yra perduodama per citoplazmą išskirtinai tik motinine linija t.y. iš motinos palikuonims, tėvinės mtDNR kopijos perdavimas retas ir evoliucijos mechanizmų nepalaikomas reiškinys. Tik labai retais atvejais pasitaiko mtDNR paveldėjimas iš tėvo, kai sutrinka paveldėjimo mechanizmas (Schwartz, Vissing 2002), arba rekombinacija tarp motininės ir tėvinės mtDNR (Kraytsberg *et al.* 2004).



2.3 pav. Žmogaus mtDNR struktūra (Bellance *et al.* 2009).

Organizme didžioji dalis mtDNR molekulių yra identiškos – tai vadinamoji homoplazmija. Įvykus mutacijai lytinių ląstelių linijoje matomas organizmo,

organų, audinių, ląstelių ploidiškumo variantas – heteroplazmija (Jansen 2000). Heteroplazmijos paplitimo, kuris gali siekti iki 14 % tirtos populiacijos, nustatant iki 1 % kitos kopijos mtDNR, įvertinimas priklauso nuo tyrimo metodų jautrumo (Tully *et al.* 2000). Taikant naujos kartos DNR sekos nustatymo metodus, apskaičiuota mtDNR heteroplazmija siekė 45 % tirtos imties (Abecasis *et al.* 2010). mtDNR paveldėjimo būdas leidžia greitai, t. y. kelių kartų laikotarpiu, heteroplazmiją pakeisti į homoplazmiją (Jansen 2000).

### 2.3.2. mtDNR žymenys

mtDNR, neturinčioje didelių introninių sekų, dažnesni vieno ar kelių nukleotidų ilgio sekos variantai nei stambių, kelių šimtų ar tūkstančių np ilgio, fragmentų mutacijos, kurios dažniausiai susijusios su fenotipiniais pokyčiais (Grady *et al.* 2014). Ankstesniais tyrimais nustatyta, kad mtDNR būdingas didesnis mutacijų dažnis negu branduolio genomui (Ingman *et al.* 2000). Šiuos rezultatus patvirtina naujais tyrimais, naudojant naujos kartos DNR sekos nustatymo metodus, apskaičiuotas branduolio genomo mutacijų dažnis –  $1,0\text{--}1,2 \cdot 10^{-8}$  mutacija/nukleotidui/kartai (Abecasis *et al.* 2010). Galimo didesnio mtDNR mutacijų dažnio priežastys: (1) mažiau efektyvus klaidų taisymas; (2) mutageninės aplinkos sąlygos; (3) didesnis replikacijų skaičius (Baer *et al.* 2007).

Žinoma, kad mutacijų dažnis skirtingose mtDNR dalyse nevienodas. Reguliacinė mtDNR sritis (HVRI ir HVRII) pasižymi didesniu mutacijų dažniu negu koduojanti sritis (577–16 023 np). Pastebėta, kad tai pačiai mtDNR sričiai nustatomi skirtingi mutacijų dažniai, priklausomai nuo naudoto metodo: (1) tiriant genealogijas; (2) tiriant filogenetines linijas. Tiriant genealogijas, nustatytas mutacijų dažnis HVI regionui –  $1,24 \cdot 10^{-6}$  mutacija/nukleotidui/metams (Madrigal *et al.* 2012) ar reguliacinei sričiai (HVRI ir HVRII) –  $0,62 \cdot 10^{-6}$  mutacija/nukleotidui/metams (Heyer *et al.* 2001). Tiriant filogenetines linijas, nustatytas reguliacinės srities (HVRI ir HVRII) mutacijų dažnis varijuoja  $0,26 \cdot 10^{-6}$ – $3,0 \cdot 10^{-6}$  mutacija/nukleotidui/metams (Rosset *et al.* 2008). Skirtingų mtDNR pozicijų savybės gali turėti įtakos

mutacijų dažnių skaičiavimui. Nustatytos HVRI ir HVRII regionų pozicijos (146, 150, 152, 195, 16 189, 16 311, 16 362, 16 519), pasižyminčios didesniu variabilumu – mutacijų karštieji taškai (angl. *Hot Spot*), kitos pozicijos (477, 493, 16 108, 16 219), yra mažai varijuojančios (Galtier *et al.* 2006; van Oven, Kayser 2009). Imant daug mutacijų karštųjų taškų skaičiavimui genealogijų tyrimo metodu, gautas įvertis būtų didesnis nei realus. Manoma, kad mutacijų dažnis, tiriant filogenetines linijas ir atsižvelgiant į karštųjų taškų kartotines mutacijas, įvertinamas tiksliau, nes į skaičiavimą įtraukiama daug variabilių pozicijų ilgesniu laikotarpiu. Ankstesniais tyrimais aptiktas ir sunkiai paaiškinamas mtDNR mutacijų dažnių, nustatytų skirtingais tyrimo metodais, skirtumas mažėja, atliekant išsamesnius tyrimus ir taikant naujus analizės metodus. Nustatytas koduojančios srities (577–16 023 np) mutacijų dažnis –  $1,26 \cdot 10^{-8}$  mutacija/nukleotidui/metams (Mishmar *et al.* 2003) arba  $1,691 \cdot 10^{-8}$  mutacija/nukleotidui/metams (Atkinson *et al.* 2009). Manyta, kad naujos mutacijos koduojančioje mtDNR srityje neretai pašalinamos gamtinės atrankos dėl galimo neigiamo poveikio fenotipui, tačiau, kaip ir reguliacinėje mtDNR srityje, nustatytos koduojančios dalies pozicijos (709, 1 719, 3 010, 5 460, 10 398, 11 914, 13 105, 13 708, 15 884), pasižyminčios didesniu variabilumu – mutacijų karštieji taškai (van Oven, Kayser 2009). Koduojančios srities mutacijų karštieji taškai lemia atskirų filogenetinių linijų nepriklausomas kartotines mtDNR mutacijas (homoplazmija), palaikomas gamtinės atrankos. Išaiškėjo, kad filogenetinių linijų kartotinėms mutacijoms būdingas mažesnis patogeniškumas, t. y. funkcinė reikšmė nežymi (Pereira *et al.* 2011), tačiau pastebėtos kartotinių mutacijų pozicijos, turinčios funkcinę reikšmę ir susijusios su geresne adaptacija (Levin *et al.* 2013). Koduojančioje mtDNR srityje matomas ne tik skirtingas mutacijų dažnis, bet ir skirtingas tranzicijų ir transversijų santykis – vidutiniškai 21,5:1, sinoniminių mutacijų atveju 34,8:1 (Kivisild *et al.* 2006). Didžiausias variabilumas, ypač nesinoniminių mutacijų atveju, nustatytas mtDNR *ATP6* gene. Didžiausias *ATP6* geno variabilumas arktinio klimato, *MTCYB* geno – vidutinio klimato, *COI* geno – atogrąžų klimato populiacijose iškėlė hipotezę dėl galimo gamtinės atrankos poveikio prisitaikant

prie klimato (Mishmar *et al.* 2003). Kito tyrimo rezultatai parodė, kad nesinoniminių mutacijų dominavimas yra jaunų filogenetinių linijų požymis, nepriklausomas nuo klimato sąlygų (Kivisild *et al.* 2006).

Dalis mtDNR hipervariabilių sričių: vieno ar kelių citidinių įsiterpimas 309, 315, 16 193 srityse, tranzicija 16 519 srityje, transversija 16 182 ir 16 183 srityse, pašalinamos iš tolesnės tyrimų analizės (Behar *et al.* 2012).

### 2.3.3. mtDNR filogenetinis medis

Nustaćius koduojančios ir reguliacinės (HVRI ir HVRII) sričių UEP genotipus, mtDNR priskiriama haplogrupei – haplotipų, turinčių bendrą protėvinį alelį, grupei. mtDNR, kaip ir Y chromosomos, filogenetinis medis sudaromas hierarchiniu principu. Haplogrupės U5b mtDNR turi filogenetinės linijos U, U5 ir U5b protėvinius alelius. Pirminiai mtDNR filogenetinių linijų tyrimai naudojo tik reguliacinės srities (HVRI ir HVRII) ar jos dalies sekos informaciją. Pastebėta, kad ne visos filogenetinės linijos identifikuojamos tinkamai naudojant tik mtDNR reguliacinės srities informaciją (van Oven, Kayser 2009). Naujausi tyrimai mtDNR filogenetinio medžio tikslinimui naudoja visos mtDNR sekos informaciją (Chandrasekar *et al.* 2009; van Oven, Kayser 2009). mtDNR sekos variantų identifikavimui ar filogenetinio medžio sudarymui naudoja rekonstruotą žmogaus referentinę seką (RSRS, angl. *Reconstructed Sapiens Reference Sequence*) vietoj iki šiol naudotos atnaujintos Kembridžo referentinės sekos (rCRS, angl. *Revised Cambridge Reference Sequence*; NC\_012920.1; GI:251831106) (Behar *et al.* 2012).

Pirmas mtDNR filogenetinis medis, tiriant vietinius Šiaurės ir Pietų Amerikos žemynų gyventojus, turėjo keturias pagrindines filogenetines linijas A–D (Torroni *et al.* 1993). Paskelbti filogenetinio medžio hierarchinio grupavimo principai (Richards *et al.* 1998) vėliau buvo pritaikyti Y chromosomos filogenetiniam medžiui. mtDNR filogenetinio medžio pagrindines filogenetines linijas (A–Z, nėra O), išskyrus L0, L1, L2, L4, L5, L6, jungia už Afrikos žemyno ribų paplitusi viena pagrindinė filogenetinė linija L3, išsišakojanti į tris stambias filogenetines linijas M, N, R (van Oven, Kayser

2009). mtDNR sekos variantai 309.1C(C), 315.1C, nukleotidų (AC) išskritos/intarpai 515–522, 16 182C, 16 183C, 16 193.1C(C) ir 16 519 pozicijose neįtraukiami į mtDNR filogenetinį medį.

Dominuojančios filogenetinės linijos: (1) Afrikos žemyne – L (išskyrus iš L3 kilusias N, M) (Zheng *et al.* 2012); (2) Artimuosiuose Rytuose – H, I, J, K, T, U, X; (3) Europoje – H, I, J, K, T, U, V (Loogvali *et al.* 2004; Pala *et al.* 2012; Simoni *et al.* 2000); (4) Azijoje – A, B, C, D, F, G, M (Derenko *et al.* 2010; Derenko *et al.* 2003; Summerer *et al.* 2014); (5) Malaizijoje – B, M, P, Q; (6) Australijoje – M, P, O, S (Kayser 2010); (7) Šiaurės ir Pietų Amerikos žemyne – A, B, C, D (Brown *et al.* 1998). Pietryčių Europoje nustatomos mtDNR haplogrupės – H (41 %), J (8 %), K (5 %) T (6,5 %), U (15,6 %, kur U5 sudaro iš viso 8,2 %), HV, I, M, V, W, X (<5 %) (Cvjetan *et al.* 2004); Centrinėje ir Rytų Europoje H (41 %), U (21 %), T, J (po 8 %), HV, I, K, prie-V, V, W, X (<5 %) (Vidrova *et al.* 2008); Šiaurės Europoje H (40–50 %), Vakarų Europoje H (36–42 %), Pietvakarių Europoje H (41–50,2 %), U (22,6 %), J (8,9 %), T (6,9 %), V, HV, O, X (<5 %) (Dahmany *et al.* 2006; Pereira *et al.* 2005; Richards *et al.* 2000). Dominuojančios filogenetinės linijos Europoje H (30–41 %), U (12–20 %), T (10 %), J (8 %), K, V (<10 %) (Fu *et al.* 2012; Pala *et al.* 2012). Haplogrupės H, I, J, N1b, T, U, V, W, X nustatomos >98 % Europos populiacijos (Mishmar *et al.* 2003).

mtDNR filogenetinio medžio iki seniausio bendro protėvio (RSRS) amžius >170 000 metų ir viršija nustatytą Y chromosomos filogenetinio medžio amžių (101–115 000 metų). Pagrindinės filogenetinės linijos L3, paplitusios už Afrikos žemyno ribų, amžius 70 000 metų. mtDNA filogenetinio medžio išsišakojimo proveržis buvo prieš 40 000 metų (Behar *et al.* 2012). Filogenetinė linija L, jungianti L0, L1, L2, L4, L5, L6, nustatoma tik Afrikos žemyne (Mishmar *et al.* 2003). Genetiniai, archeologiniai duomenys šią pasaulio dalį sieja su mtDNR kilme. Keliamos dvi paplitimo iš Afrikos žemyno hipotezės: (1) Šiaurės Rytų Afrika – Artimieji Rytai; (2) Pietų Afrika – Pietų Azija, kurios iki šiol lieka nepatvirtintos (Beyin 2011; Quintana-Murci *et al.* 1999).

#### 2.3.4. mtDNR filogeografija Europoje

Neolito laikotarpio vykusį kultūrinę kaitą – pasisavinamąjį ūkį (rinkimą ir medžioklę) pakeitęs gamybinis ūkis (žemdirbystė ir gyvulininkystė), galėjo prasidėti Pietryčių Europoje, kurią pirmieji gamybinio ūkio atstovai pasiekė iš Artimųjų Rytų prieš 9 000 metų. Pirmykščiai Europos gyventojai įsisavino gamybinį ūkį pereinamuoju neolito laikotarpiu prieš 4 000–8 000 metų (Haak *et al.* 2010). Iškeltos dvi gamybinio ūkio plitimo teorijos: (1) iš gamybinio ūkio atstovų perimta tik kultūra (Lukes *et al.* 2004); (2) gamybinio ūkio atstovai iš Artimųjų Rytų pakeitė pirmykštes Europos populiacijas (Fu *et al.* 2012; Haak *et al.* 2010; Rasteiro, Chikhi 2013). Populiacijų migracija prieš, po PL į ir iš prieglobsčių Centrinėje, Pietvakarių ir Pietų Europoje turėjo įtakos dabartiniam Europos genetiniam fondui. Manoma, kad tam tikros mtDNR haplogrupės pakartotinai ar naujai paplito iš Europos prieglobsčių: (1) haplogrupės H1, H3, V, U5b iš dabartinės Šiaurės Ispanijos ir Pietų Prancūzijos teritorijos; (2) haplogrupė U5b3 iš dabartinės Italijos teritorijos; (3) haplogrupės U4, U5a iš Rytų Europos (Pala *et al.* 2012). Gamybinio ūkio plitimas galėjo būti atsitiktinis procesas, turėjęs nevienodą įtaką Europos prieglobsčio zonų genetiniam fondui (Hervella *et al.* 2012). Šio proceso tikslesnis vaizdas vis dar yra diskusijų objektas.

Tyrimai parodė, kad prieš agrarinės kultūros paplitimą, pasisavinamojo ūkio Europos populiacijose dominavo haplogrupė U (82 %), o haplogrupė H tuo laikotarpiu nedominavo (Bramanti *et al.* 2009). Tyrimų rezultatai rodo galimą vėlesnį haplogrupės H augimą, kuris prasidėjo prieš 9 000 metų ir vyksta iki šiol. Haplogrupės U galimas pirmas augimas vyko dviem etapais – pirmas prieš 10–20 000 metų ir antras prieš 4 000 metų, kuriuos skyrė laikinas mažėjimas, vykęs prieš 5–6 000 metų (Fu *et al.* 2012). Gauti duomenys leidžia iškelti hipotezę, kad haplogrupė H, galėjo plisti Europoje kartu su gamybinio ūkio plitimu. Ankstyvojo neolito (~7 500 metų) populiacijų genetinio fondo, kur haplogrupė H sudarė 20 %, įtaka dabartiniam genetiniam fondui, manoma, buvo menka. Dabartinis paplitimas, kurio židinytis galėjo būti Centrinėje Europoje, pasiektas viduriniame neolite prieš 6 000 metų. Manoma, kad haplogrupės H



populiacijų genetinis fondas vėliau buvo papildomas Europoje paplitusių kultūrų genetiniu fondu (Brotherton *et al.* 2013). Nustatytas pastovus haplogrupių U ir H augimas, prasidėjęs prieš 4 000 metų, kai išnyksta archeologiniai pasisavinamojo ūkio kultūros įrodymai, gali rodyti šių dviejų populiacijų susijungimą (Fu *et al.* 2012). Motininės ir tėvinės linijų tyrimų rezultatų skirtumai galimi dėl skirtingos istorijos, kuriai turėjo įtakos skirtingos kultūros (Rasteiro, Chikhi 2013). Atlikus 1000 genomų projekto tiriamųjų mtDNR analizę, pasiūlyta kita haplogrupių plitimo Europoje hipotezė. Nustatyta, kad filogenetinės linijos HV, H, H1, H3, J1c, T1, T2, U5a1, U5a, K1, V, W, U2'3'4'7'8'9 susiformavo dar prieš 10 000 metų. Dviejų filogenetinių linijų (J1c3 ir T2b) augimas vyko prieš ~10 000 metų. Daugelio tirtų filogenetinių linijų, išskyrus HV ir U2'3'4'7'8'9, augimas nustatytas pasibaigus PL. Filogenetinių linijų augimo pradžia vyko anksčiau nei gamybinio ūkio plitimas Europoje (Zheng *et al.* 2012).

Haplogrupės U4/H1b, U5 ir U5a, dominavusios tarp pasisavinamojo ūkio atstovų mezolito laikotarpiu Europoje ir dabartinėse Baltijos jūros rytinėse populiacijose (lietuviai, latviai), sudarė 74 % tirtose paskutinių pasisavinamojo ūkio atstovų grupėje Skandinavijoje. Gauti rezultatai leido iškelti hipotezę, kad duobelinės keramikos kultūros (angl. *Pitted Ware Culture*) atstovai galėjo būti Baltijos jūros rytinių teritorijų, kurios galbūt buvo kai kurių pasisavinamojo ūkio atstovų prieglobsčiu, populiacijų (lietuviai, latviai) protėviai (Malmstrom *et al.* 2009).

### 2.3.5. Europos populiacijų mtDNR filogenetinių linijų tyrimai

Populiacijų mtDNR tyrimai plėtėsi nuo HVRI ir (ar) HVRII regionų (Kasperavičiūtė *et al.* 2004; Salas *et al.* 2000) iki visos mtDNR sekos nustatymo (Derenko *et al.* 2012; Gunnarsdottir *et al.* 2011; Schonberg *et al.* 2011). Analizuojant populiacijos vidinę genetinę struktūrą, padėti kitų populiacijų atžvilgiu (Delfin *et al.* 2014; Grzybowski *et al.* 2007; Gunnarsdottir *et al.* 2011; Schlebusch *et al.* 2013; Schonberg *et al.* 2011), norima nustatyti skirtumų priežastį – geografinę ar kalbinę. Visos mtDNR sekos nustatymo metodas,

paskatino plėsti mtDNR filogenetinio medžio filogenetinių linijų detalizavimo, filogenetinių linijų ar populiacijų paplitimo, kilmės ir migracijos tyrimus (Cardoso *et al.* 2013; Chandrasekar *et al.* 2009; Malyarchuk *et al.* 2010; Pala *et al.* 2009; van Oven, Kayser 2009). Daugelio tirtų Europos, Artimųjų Rytų, Vakarų Azijos ir kitų populiacijų ar atskirų filogenetinių linijų visos mtDNR nukleotidų sekos duomenys prieinami kitoms tyrimų grupėms (2.2 lentelė).

2.2 lentelė. Visos mtDNR tyrimai Europoje ir pasaulyje

Populiacija	db <i>GenBank</i> numeris <sup>1</sup>	Šaltinis
Samai	AY882379, DQ902708	(Achilli <i>et al.</i> 2005; Ingman, Gyllensten 2007)
Baltarusija	KC867103	(Kushniarevich <i>et al.</i> 2013)
Kazanės totoriai	GU122975	(Malyarchuk <i>et al.</i> 2010)
Lenkija	JX307099, JX128041, JX266260	(Mielnik-Sikorska <i>et al.</i> 2013)
Ukraina	JX307099, JX128041, JX266260	(Mielnik-Sikorska <i>et al.</i> 2013)
Rusija	JX307099, JX128041, JX266260	(Mielnik-Sikorska <i>et al.</i> 2013)
Slovakija	JX307099, JX128041, JX266260	(Mielnik-Sikorska <i>et al.</i> 2013)
Čekija	JX307099, JX128041, JX266260	(Mielnik-Sikorska <i>et al.</i> 2013)
Gruzija	HM852756	(Schonberg <i>et al.</i> 2011)
Turkija	HM852756	(Schonberg <i>et al.</i> 2011)
Armėnija	HM852756	(Schonberg <i>et al.</i> 2011)
Azerbaidžanas	HM852756	(Schonberg <i>et al.</i> 2011)
Iranas	HM852756	(Schonberg <i>et al.</i> 2011)
Aškenaziai	KC878709, JX273243	(Costa <i>et al.</i> 2013)
Ispanija	AY882379	(Achilli <i>et al.</i> 2005)
Italija	AY882379	(Achilli <i>et al.</i> 2005)
Sardinija	GQ129143	(Pala <i>et al.</i> 2009)

<sup>1</sup>Duomenų bazė (db) *NCBI DNA & RNA PopSet*

Europos ar gretimų geografinių regionų populiacijų mtDNR tyrėjai sprendžia, priklausomai nuo geografinės padėties, tiriamai populiacijai svarbius klausimus. Rytų Europos genetinis fondas: slavų, finougrų, baltų populiacijų ryšiai, migracija (Grzybowski *et al.* 2007; Kushniarevich *et al.* 2013; Mielnik-Sikorska *et al.* 2013), Pietryčių Europos genetinis fondas: Balkanų populiacijų ryšiai, gamybinio ūkio plitimas (Bosch *et al.* 2006; Fu *et al.* 2012; Sarac *et al.* 2014), Pietvakarių Europos genetinis fondas: Pietų ir Vakarų Europos populiacijų ryšiai, filogenetinių linijų kilmė, geografinio prieglobsčio hipotezė, migracija (Cardoso *et al.* 2013; Garcia *et al.* 2011; Hernandez *et al.* 2014; Pereira

*et al.* 2005), Artimųjų Rytų genetinis fondas: populiacijų ryšiai, agrarinės kultūros plitimas, geografinio prieglobsčio hipotezė (Pala *et al.* 2012).

### 2.3.6. Lietuvos populiacijos mtDNR filogenetinių linijų tyrimai

D. Kasperavičiūtė kartu su kitais mokslininkais ištyrė lietuvių populiacijos 180 asmenų. Nustačius mtDNR HVRI regiono (16 024–16 400 np) nukleotidų seką ir skaidymo restrikcijos endonukleazėmis metodu įvertinus tris pozicijas (73, 7 028, 14 766), mtDNR priskirta haplogrupei (H, H1, H3, H4, H5, H8, V, HV, preV, U, K, U3, U4, U5a, U5a1, U5b, U5b1, U5, J, J1, J1b1, T, T1, I, W, X, Kita). Tolesnei analizei naudotas 356 np ilgio mtDNR fragmentas (16 028–16 383 np). Ištyrus 180 asmenų HVRI regioną, nustatytos 76 variabilios sritys ir 95 skirtingi haplotipai. Nustatyta genetinė įvairovė (0,947–0,984) ir vidutiniai poriniai skirtumai (3,61–4,98) artimi kitų Europos populiacijų įverčiams. mtDNR HVRI 356 np ilgio fragmento analizės ( $F_{ST}$  ir AMOVA) rezultatai parodė Lietuvos populiacijos homogeniškumą – nerasta skirtumų tarp šešių Lietuvos etnolingvistinių grupių ar dviejų pagrindinių grupių – aukštaičių ir žemaičių. MDS, remiantis  $F_{ST}$  atstumų matrica pagal mtDNR haplotipų (356 np) pasiskirstymą, parodė, kad tirta lietuvių populiacija artima slavų (Rusija, Lenkija) populiacijoms ir išsidėsto tarp Šiaurės Europos finougrių (Estija, Suomija) ir Vakarų Europos indoeuropiečių. Statistiškai reikšmingi  $F_{ST}$  skirtumai nustatyti tik nuo labiausiai nutolusių populiacijų (Italija, Ispanija, baskai, Islandija). Nustatytos Europos populiacijoms būdingos dažniausios haplogrupės sudarė 97 % tirtos lietuvių populiacijos. MDS, remiantis  $F_{ST}$  atstumais pagal haplogrupių pasiskirstymą, artima MDS, remiantis  $F_{ST}$  atstumais pagal mtDNR haplotipų (356 np) pasiskirstymą. Nenustatyta sumažėjusi mtDNR genetinio fondo ar haplogrupių įvairovė gali būti susijusi su skirtinga tėvinės ir motininės linijų istorija arba dėl panašaus populiacijų, formavusių baltus, mtDNR genetinio fondo (Kasperavičiūtė *et al.* 2004).

A. Kushniarevich kartu su kitais mokslininkais atlikta pagrindinių komponentų analizė (PKA, angl. *PCA, Principal Component Analysis*), remiantis mtDNR haplogrupių pasiskirstymu, parodė, kad lietuvių populiacija

kartu su rusų, lenkų, ukrainiečių ir baltarusių populiacijomis formuoja bendrą grupę. Estai, čekai, slovakai ir Balkanų regiono populiacijos nutolusios nuo suformuotos Rytų Europos populiacijų grupės (Kushniarevich *et al.* 2013).

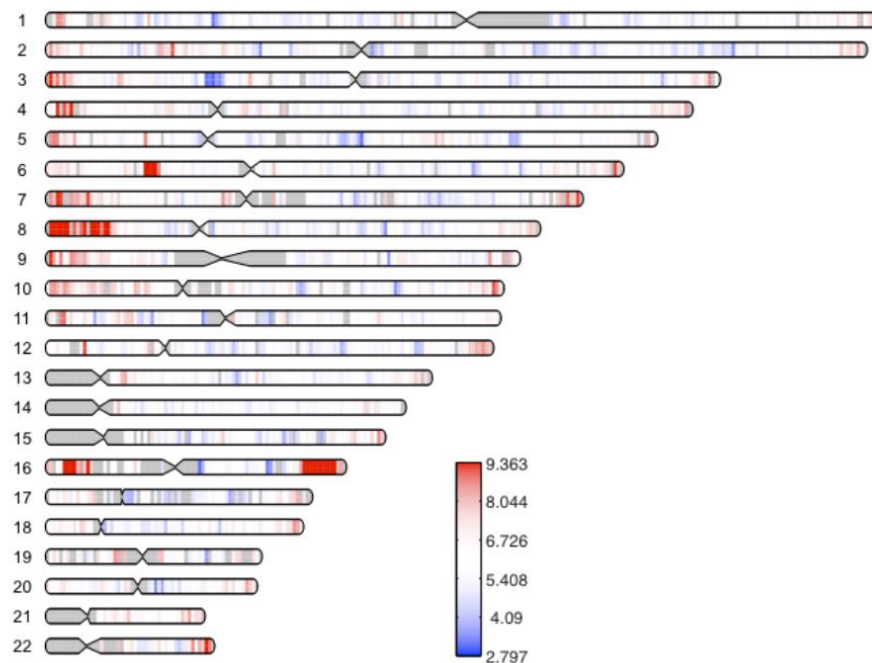
## 2.4. Populiacijų plataus masto tyrimai naudojant autosomų žymenis

### 2.4.1. Autosomų žymenų naudojimas populiacijų tyrimams

Žmogaus genomo projekto metu gauti rezultatai tapo kitų projektų – HapMap (3 etapai), 1000 genomų (3 etapai), atskirų populiacijų ar individualių genomų – pagrindu (Abecasis *et al.* 2010; Altshuler *et al.* 2010; Lander *et al.* 2001). Žmogaus genomo projektas ir jo rezultatai skatino naujų technologijų kūrimą ir genetinių duomenų prieinamumą mažesnėms tyrimų grupėms. Vykdomi fundamentiniai ir taikomieji tyrimai bei jų rezultatai leido atsakyti į daugelį klausimų, kartu išskėlė dar daugiau naujų klausimų ir problemų, kurios bus sprendžiamos artimiausiu metu. Nuolat pildoma žmogaus genomo variantų duomenų bazė yra iki šiol vykdomų tyrimų rezultatas. Šiuo metu žinoma daugiau nei 67 milijonai žmogaus genomo variantų (VNP; trumpos, 1–50 np, iškritos ar intarpai, angl. *Indel*) (Flicek *et al.* 2014).

Žmogaus autosomose vidutinis atstumas tarp dviejų VNP yra 8–900 nt, tačiau matomas netolygus VNP pasiskirstymas visame genome – padidėjusio ar sumažėjusio VNP tankio DNR fragmentai (2.4 pav.). Didesnis VNP tankis nustatytas chromosomų galuose, telomeroose, ir atskirose chromosomų dalyse. Netolygus VNP pasiskirstymas gali būti susijęs su rekombinacijos (Myers *et al.* 2008) ir mutacijų (Rogozin, Pavlov 2003) karštaisiais taškais. Taip pat matomas skirtingas VNP variabilumas, kurio mastui apibūdinti nebeužtenka tik polimorfizmo apibrėžimo t. y. kai  $MAF > 1\%$  populiacijoje. Žmogaus genome matomi DNR molekulių fragmentai su VNP alelių deriniu, t. y. haplotipu, kurie perduodami kitai kartai nepakitę, kol neįvyksta rekombinacija. Haplotipai gali būti sudaryti iš kelių dešimčių, šimtų ar tūkstančių VNP alelių derinio. Nustatyti raktiniai VNP (angl. *TagSNPs*), kurie yra specifiniai atskiram haplotipui. Tiriant raktinius VNP, greičiau ir efektyviau įvertinami individualaus žmogaus genomo haplotipai. Turima genetinė informacija leidžia nustatyti ir atrinkti polimorfinius

variantus, kurie būtų tolygiai pasiskirstę po tiriamą DNR fragmentą ar visą genomą. VNP tyrimui gali būti atrinkti pagal poreikį, atsižvelgiant į: (1) tyrimo tikslą, susijusį su tikslinių genomo fragmentų, požymių ar viso genomo analize; (2) tiriamą požymį, manoma, kad dažnesnius požymius lemia dažniau pasitaikančių variantų sąveika, retus požymius, reti ar vienetiniai variantai; (3) tiriamųjų imtį, skirtumai tarp mažų ar artimų imčių nustatomi, genotipuojant dešimtis, šimtus retų ar šimtus tūkstančių dažnų variantų; (4) turimas genotipavimo technologijas. Dažnai tyrimams naudojami komerciniai tiksliniai ar bendriniai genotipavimo lustai su atrinktais Europos populiacijose dažniais VNP (MAF > 5 %).



2.4 pav. VNP pasiskirstymas žmogaus autosomose. Raudona ar mėlyna spalva nurodo VNP tankį 1 Mb ilgio DNR fragmente (VNP skaičiaus santykis su nukleotidų skaičiumi 1 Mb ilgio fragmente, pilka spalva pažymėti regionai, kuriuose nustatytas nukleotidų skaičius mažesnis nei 75 % 1Mb) (Abecasis *et al.* 2010)

Populiacijų genų fondą evoliucijos metu formuojantys procesai (mutacijos, migracija, gamtinė atranka, butelio kaklelio efektas, įkūrėjo efektas, genetinis dreifas) gali būti vertinami tiriant VNP alelių pasiskirstymą. Haplotipų ir stebimo statistinio reiškinių – nepusiausviros sankibos (LD, angl. *Linkage Disequilibrium*), t. y. alelių asociacijos, struktūros, kuri priklauso nuo populiacijos amžiaus, t. y. įvykusių mejozių skaičiaus, rekombinacijos dažnio,

genų konversijos, genų dreifo, gamtinės atrankos, analizė gali suteikti informacijos apie populiacijų evoliucijos procesą. Nepusiausvira sankiba priklauso nuo genomo srities ir tiriamos populiacijos, todėl tik empiriniai daugelio žymenų ir populiacijų tyrimai gali suteikti informacijos apie nepusiausviroi sankibos struktūrą ir pritaikyti tiriamos populiacijos evoliucijos analizei (Zhao *et al.* 2003).

#### 2.4.2. Europos populiacijų tyrimai naudojant autosomų žymenis

Žinios, suteiktos fundamentinių žmogaus genomo tyrimų, ir sukurtos naujos technologijos, leidžiančios vienu metu genotipuoti šimtus tūkstančių genetinių žymenų, paskatino išsamius daugelio Europos, pasaulio populiacijų ar etninių grupių tyrimus (2.3 lentelė).

Tiriant Europos populiacijų genetinio fondo struktūrą, atlikta PKA, remiantis autosomų žymenų alelių įvairove, parodė Europos, ypač centrinės dalies, populiacijų homogeniškumą ( $PK1 < 10\%$ ). Analizuojant PK1, kuri aprašo pagrindinę Europos populiacijų variacijos dalį, matomas Europos šiaurės–pietų ar šiaurės vakarų–pietryčių ašies genetinės variacijos gradientas. Genetinei variacijai tarp Europos populiacijų įtakos turi populiacijų geografinė padėtis ir atstumai. Šiuos rezultatus patvirtina PKA, kurios didžiausią dalį aprašomos variacijos lemia tiriamų populiacijų geografinė padėtis. Taip pat nustatyti  $F_{ST}$  atstumai, kurie koreliuoja su geografiniais atstumais tarp tiriamų populiacijų. Naudojami plataus masto genotipavimo duomenų analizės metodai homogeniškoje Europos populiacijų grupėje nustato skirtumus ne tik tarp populiacijų, bet ir tiriamų populiacijų, neturinčių geografinių barjerų, viduje. Analizės metodai leidžia nustatyti protėvinius genetinius žymenis (angl. *AIM, ancestry-informative markers*), kurie pasižymi didžiausiu variabilumu tarp atskirų populiacijų ir pagal kuriuos galima konkrečiai populiacijai teisingai priskirti iki 90 % nežinomos kilmės asmenų. Gauti dažnų sekos variantų genotipavimo duomenys leidžia nustatyti uždaras populiacijas, pavyzdžiui, samiai, ar populiacijų evoliucijoje įvykusį staigų sumažėjimą (butelio kaklelio efektas), kuriems dėl intensyvios migracijos būdingi didesni LD fragmentai nei

heterogeniškomis populiacijoms. Pasirinktos tinkamos analizės metu skaičiuojamų įverčių slenkstinės reikšmės, neleidžia prarasti svarbios genetinės variacijos, lemiančios artimų populiacijų skirtumus (Huckins *et al.* 2014; Nelis *et al.* 2009).

2.3 lentelė. Plataus masto autosomų žymenų tyrimai Europoje ir pasaulyje

Populiacijos / etninės grupės	Imtis	DNR genotipavimo lustas	Šaltinis
Šiaurės Rytų Europa	1 564	Human370CNV <sup>1</sup> HumanHap300 <sup>1</sup>	(Nelis <i>et al.</i> 2009)
Šiaurės Europa	945	GeneChip <sup>®</sup> Human Mapping 250K Sty Array <sup>2</sup>	(Salmela <i>et al.</i> 2008)
Europa	2 514	GeneChip <sup>®</sup> Mapping 500K <sup>2</sup>	(Lao <i>et al.</i> 2008)
Europa	2 257	GeneChip <sup>®</sup> Mapping 500K <sup>2</sup>	(Ralph, Coop 2013)
Europa	2 907	Human660W-Quad <sup>1</sup>	(Huckins <i>et al.</i> 2014)
Europa	4 110	Genome-Wide Human SNP Array 5.0/6.0 <sup>2</sup>	(Moskvina <i>et al.</i> 2010)
Romai	152	Genome-Wide Human SNP Array 6.0 <sup>2</sup>	(Mendizabal <i>et al.</i> 2012)
Romai	27	Genome-Wide Human SNP Array 6.0 <sup>2</sup>	(Moorjani <i>et al.</i> 2013)
Žydai	121	Human610-Quad <sup>1</sup> Human660W-Quad <sup>1</sup>	(Behar <i>et al.</i> 2010)
Samai	344	GeneChip <sup>®</sup> Human Mapping 100K <sup>2</sup>	(Huyghe <i>et al.</i> 2011)
Olandija	4 441	Genome-Wide Human SNP Array 6.0 <sup>2</sup>	(Abdellaoui <i>et al.</i> 2013)
Airija, D. Britanija	3 367	Genome-Wide Human SNP Array 5.0/6.0 <sup>2</sup>	(O'Dushlaine <i>et al.</i> 2010)
Graikija	56	Custom BeadChip (1 813 SNP) <sup>1</sup>	(Stathias <i>et al.</i> 2012)
Islandija	885	HumanHap300 <sup>1</sup>	(Price <i>et al.</i> 2009)
Švedija	1 525	HumanHap550 <sup>1</sup>	(Salmela <i>et al.</i> 2011)

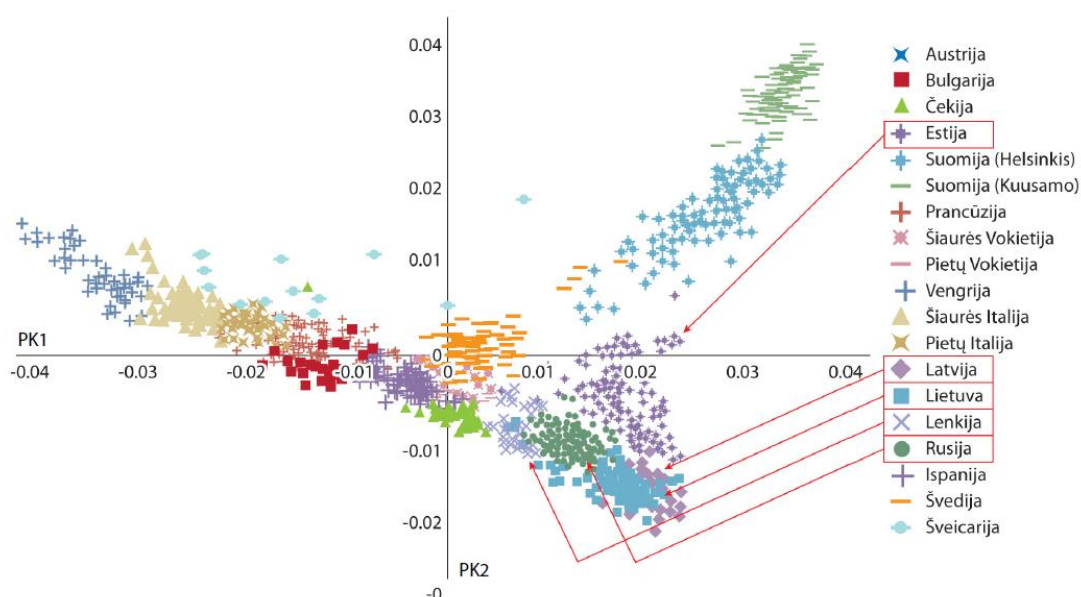
<sup>1</sup>*Illumina Inc.*, JAV

<sup>2</sup>*Affymetrix Inc.*, JAV

#### 2.4.3. Lietuvos populiacijos tyrimai naudojant autosomų žymenis

M. Nelis kartu su kitais mokslininkais ištyrė 3 378 asmenis (19 populiacijų), iš kurių 95 asmenys iš Lietuvos populiacijos. Šie asmenys buvo genotipuoti naudojant komercinius DNR lustus (*Illumina Human370CNV Beadchip*, *HumanHap300 Beadchip*). Atlikus genotipuotų mėginių kokybės kontrolę (3 112 tiriamieji) iš kiekvienos populiacijos atrinkta po 100 mėginių, jeigu po

kokybės kontrolės buvo daugiau nei 100 tiriamųjų iš vienos populiacijos. Galutinę tyrimo imtį sudarė 1 564 tiriamieji. Atlikus VNP kokybės kontrolę, tolesnei analizei naudoti 273 464 autosomų žymenų genotipų duomenys. Atlikus PKA, įtraukiant HapMap projekto Europos, Azijos ir Afrikos populiacijų duomenis, nustatyta, kad PK1 aprašė 36,6 %, PK2 23,8 % variacijos, ir vizualizavus gautus duomenis, matyti, kad tirtos populiacijos sudaro trikampį: kampuose (1) Suomija; (2) Baltijos regionas (Estija, Latvija, Lietuva), Lenkija ir Vakarų Rusija; (3) Italija; (4) trikampio centre Vidurio ir Vakarų Europos populiacijos (2.5 pav.).



2.5 pav. Europos populiacijų PKA, remiantis 273 454 autosomų VNP alelių pasiskirstymu, vizualizuoti rezultatai (Nelis *et al.* 2009)

Pašalinus HapMap Azijos ir Afrikos populiacijų duomenis, PK1 aprašė 8,7 %, PK2 4,9 % tirtos variacijos, kas rodo Europos genetinio fondo vientisumą. Apskaičiuoti poriniai  $F_{ST}$  atstumai tarp HapMap projekto metu tirtų Šiaurės ar Vakarų Europos kilmės Jutos gyventojų (CEU, angl. *Utah Residents with Northern and Western European Ancestry*) ir kitų populiacijų statistiškai reikšmingai koreliavo su geografiniais atstumais ( $r^2 = 0,382$ ,  $p < 0,01$ ). Apskaičiuoti  $F_{ST}$  atstumai tarp CEU ir Estijos (0,004), Latvijos (0,006), Lietuvos (0,005), Vakarų Rusijos (0,004) populiacijų buvo tarpiniai iš visų nustatytų. Paaiškėjo trečias barjeras, remiantis geografiniais atstumais tarp tirtų populiacijų



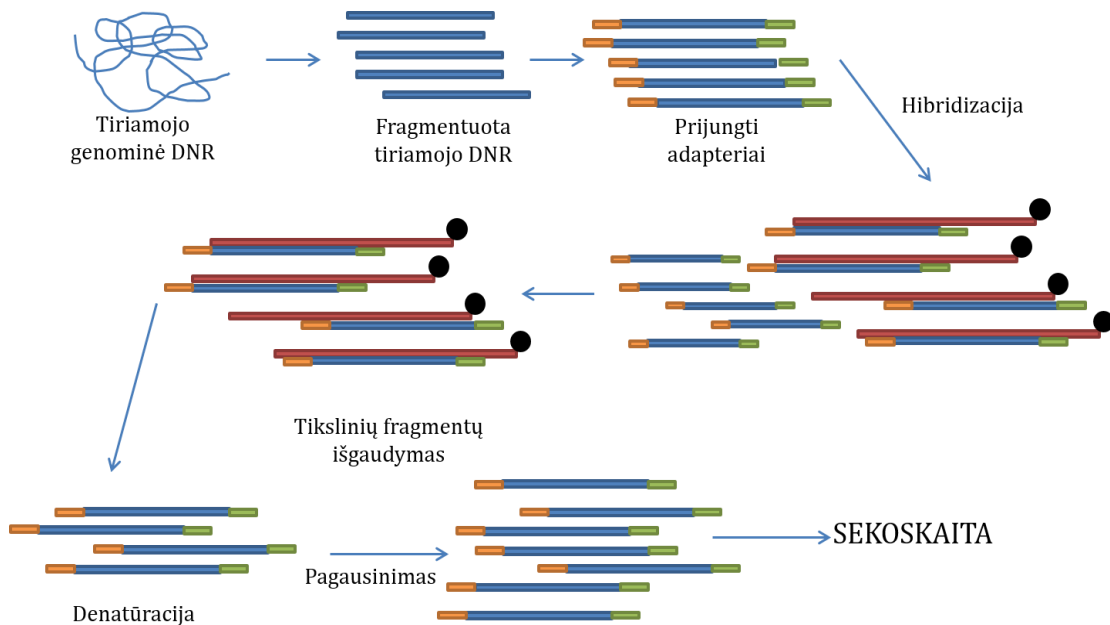
sostinių ir nustatytais  $F_{ST}$  atstumais tarp Vakarų Rusijos, Lenkijos, Lietuvos ir Bulgarijos populiacijų, bei ketvirtas barjeras tarp Pabaltijo regiono, Lenkijos ir Švedijos (Nelis *et al.* 2009). Daugėjant plataus masto atskirų etninių grupių ar Europos populiacijų autosomų žymenų tyrimų, įtraukiamos mažos imtys iš daugelio Europos populiacijų, tarp kurių ir Lietuvos populiacija (Behar *et al.* 2010; Mendizabal *et al.* 2012).

## 2.5. Populiacijų tyrimai naudojant egzomo žymenis

### 2.5.1. Egzomo sekos nustatymas ir žymenys

Egzomo, koduojančios genomo dalies, sudarančios iki 1,5 % genomo, sekos nustatymo fundamentalieji ir taikomieji tyrimai pradėti pasibaigus Žmogaus genomo projektui, kuris paskatino naujų technologijų kūrimą (Lander *et al.* 2001). Pavienių genų ar DNR fragmentų tyrimus pakeitė visos žmogaus genomo koduojančios dalies sekos nustatymo tyrimai. Nustatytų žmogaus genomo variantų pasiskirstymas paskatino naują variantų pagal dažnius skirstymą, kuris sudarytas iš penkių kategorijų: (1) dažnas sekos variantas (CV, angl. *Common variant*), kai  $MAF > 5\%$  populiacijoje; (2) retesnis sekos variantas (LFV, angl. *Low-frequency variant*), kai  $MAF 0,5-5\%$  populiacijoje; (3) retas sekos variantas (RV, angl. *Rare variant*), kai  $MAF < 0,5\%$  populiacijoje; (4) vienetinis sekos variantas (angl. *Singleton*) nustatomas vienam iš visų tirtų asmenų; (5) privatus sekos variantas (angl. *Private Variant*), kai sekos variantas nustatytas tik vienoje genealogijoje (Abecasis *et al.* 2010; Abecasis *et al.* 2012). Žmogaus genomo ir egzomo tyrimais nustatyti dažni sekos variantai ( $MAF 5\%$ ) sudaro  $>99\%$ , retesni ( $MAF 1\%$ )  $>98\%$ , reti ( $MAF \sim 0,1\%$ )  $\sim 50\%$  variacijos artimose projektų metu tirtoms (Europos CEU, TSI, GBR, FIN, IBS; Afrikos YRI, LWK, ASW; Rytų Azijos CHB, JPT, CHS; Amerikos žemynų MXL, CLM, PUR) populiacijose, pavyzdžiui, UK 10K projekto tiriamieji. Genomo variantai iš projektų sukauptos duomenų bazės nustatomi rečiau tolimesnių ar uždarų populiacijų, pavyzdžiui, SardiNIA projekto tiriamųjų, genetiniame fonde. Dauguma likusių nenustatytų variantų yra vienetiniai ar privatūs DNR sekos variantai, kurių duomenų bazę išplės tik atskirų populiacijų išsamūs

genomo ar egzomo sekos nustatymo tyrimai, nes 17 % LFV (MAF 0,5–5 %) nustatyta bendros kilmės populiacijose, 53 % RV (MAF < 0,5 %) nustatyta atskirose populiacijose, o beveik visi dažni variantai (MAF > 10 %) nustatyti tarp visų tirtų populiacijų (Abecasis *et al.* 2010; Abecasis *et al.* 2012).



2.6 pav. Egzomo sekos nustatymo schema.

Egzomo sekos nustatymo iš skaidytos genomines DNR mišinio hibridizacijos metodu išrenkami koduojančios DNR dalies fragmentai. Atsitiktinis genomines DNR skaidymo procesas ir naudojamas genomo koduojančios dalies fragmentų mišinys hibridizacijos metu leidžia nustatyti ne tik genomo koduojančią dalį, bet ir aplinkines sekas, kaip 5' ar 3' netransliuojamos sritys, promotoriai, intronai (2.6 pav.). Šiuo metu žinoma daugiau kaip 20 000 genų (Flicek *et al.* 2014), koduojančių polipeptidus, tačiau esamos DNR sekos nustatymo technologijos leidžia nustatyti ne visus žinomus genus ar jų fragmentus. Guanidino ir citidino gausių sekų (angl. *GC rich*), esančių šalia pirmojo egzono ar heterochromatino, nustatymas lieka iššūkiu, skatinantis kurti naujas technologijas. Nustatytų DNR sekos variantų patvirtinimas nepriklausomais metodais leidžia nustatyti naudojamų technologijų, kurioms būdingos skirtingos sisteminės klaidos, klaidingai teigiamų rezultatų įvertį (MacArthur *et al.* 2012).

Egzomo sekos nustatymo taikomųjų tyrimų rezultatai išplėtė iki tol žinotų patogeninių variantų, lemiančių pagal Mendelio dėsnius paveldimas ligas, sąrašą (Gilissen *et al.* 2011). Fundamentinių egzomo sekos nustatymo tyrimų objektai: (1) migracija ir populiacijų maišymasis, vykęs geografinio regiono viduje; (2) populiacijos augimas ar staigus sumažėjimas; (3) neutralūs, žalingi ar pranašumą suteikiantys sekos variantai. Tačiau plataus masto VNP genotipavimas ar mažos apimties DNR sekų nustatymas, pasižymintys mažesne statistine galia ar tikslumu, vis dar plačiai naudojami fundamentiniams klausimams spręsti. Egzomo sekos nustatymo metodu gauti duomenys sujungia genomo koduojančios dalies sekos variabilias ir nevariabilias sritis ir išvengiama galimo gautų duomenų kryptingumo dėl išankstinio tikslinių variabilių sričių parinkimo, kuo remiasi plataus masto genotipavimo metodas. Egzomo sekos nustatymo metodas, kurio vienas iš privalumų yra didesnis sekos padengimas (angl. *Coverage Depth*) negu viso genomo sekos nustatymo metodo, leidžia aptikti koduojančios dalies naujus, dar neaprašytus ar kitais tyrimais nenustatytus, retus sinoniminius ar nesinoniminius sekos variantus. Taip pat šiuo tyrimo metodu galima nustatyti vienetinius ar privačius koduojančios sekos variantus, kurie vis dar yra nenustatyti ir neaprašyti daugelio projektų, tiriančių sekos variantų pasiskirstymą atskirose populiacijose. Tiriant populiacijų genetinę struktūrą, kitaip nei tiriant fenotipą lemiančius variantus, siekiant išvengti klaidingai identifikuotos populiacijos vidinės struktūros ar populiacijų grupavimosi dėl trūkstamos informacijos (angl. *Missing data*), stengiamasi pašalinti kiek įmanoma daugiau koduojančios sekos sričių, kurios nustatytos ne visiems tiriamiesiems. Egzomo sekos nustatymo metodas apima tik mažą dalį nekoduojančios dalies, esančios aplink koduojančią dalį, variantų ir lieka neįvertintas daugelio nekoduojančios dalies variantų, atsakingų už reguliavimą, pasiskirstymas (Tennesen *et al.* 2011).

#### 2.5.2. Populiacijų tyrimai naudojant egzomo duomenis

Egzomo sekos nustatymo tyrimų etapas keičia vykdomus populiacijų evoliucijos tyrimus, analizuojančius vienos iš tėvinių linijų, Y chromosomos ar

mtDNR, žymenis ar plataus masto dažnus autosomų žymenis, kurių rezultatai priklauso nuo turimų technologijų ir taikomų metodų. Y chromosomos ir mtDNR dėl paveldėjimo būdo ir nevykstančios rekombinacijos tyrimai vis dar lieka populiacijų evoliucijos tyrimų objektas.

Pirmas individualus genomas paskelbtas praėjus septyneriems metams po Žmogaus genomo projekto (Wheeler *et al.* 2008), pirmieji skirtingų populiacijų genomo ir egzomo sekos nustatymo tyrimų rezultatai paskelbti po 9 metų nuo žmogaus genomo projekto pabaigos (2.4 lentelė). Remiantis visose tirtose populiacijose nustatytais DNR sekos variabiliomis pozicijomis, rasti skirtumai ( $F_{ST} = 0,071$ ) tarp pagrindinių grupių (Afrikos, Azijos, Europos). Didžiausi dažnių skirtumai (>80 %) tarp populiacijų būdingi nesinoniminiams variantams ( $n = 139$ ), kurie gali būti svarbūs adaptacijai. Manoma, kad gamtinė atranka veikia ne tik naujus, bet ir esamas sekos variabilias pozicijas, kurių aleliai tampa fiksuoti tam tikrose populiacijose (Abecasis *et al.* 2010; Abecasis *et al.* 2012). Tiriant genomo koduojančios seką (3 205 genus) daugiau retų variantų ( $MAF < 5\%$ ) nustatyta Afrikos, mažiau Europos populiacijose (Raska, Zhu 2011). Norint aprašyti vienodą variacijos dydį, naudojant retus sekos variantus ( $MAF < 5\%$ ) reikia daugiau komponentų negu naudojant dažnus sekos variantus. Pastebėta, kad naudojant retus sekos variantus galima atskirti populiacijas ne tik pagal geografinę padėtį, bet ir pagal bendrą kilmę (Baye *et al.* 2011).

2.4 lentelė. Europos ir pasaulio populiacijų tyrimai naudojant egzomo žymenis

Populiacija	Imtis	Metodas	Šaltinis
Europa, Azija, Afrika	742	WGS <sup>1</sup> , WES <sup>2</sup>	(Abecasis <i>et al.</i> 2010)
Europa, Azija, Afrika	697	WES	(Raska, Zhu 2011)
Europa, Azija, Afrika	697	WES	(Baye <i>et al.</i> 2011)
Prancūzija, Kanados prancūzai	139	WES	(Casals <i>et al.</i> 2013)
Europa	6 636	WES	(Zawistowski <i>et al.</i> 2014)
Šiaurės ir Pietų Amerika	181	WGS, WES	(Gravel <i>et al.</i> 2013)

<sup>1</sup>WGS – genomo sekos nustatymo metodas (angl. *WGS, Whole Genome Sequencing*)

<sup>2</sup>WES – egzomo sekos nustatymo metodas (angl. *WES, Whole Exome Sequencing*)

Analizuojant ir interpretuojant žmogaus egzomo sekos duomenis, svarbu įvertinti metodo trūkumus: (1) egzonų skaičių; (2) paralogus; (3) sekos padengimą; (4) skirtingas technologijas ar reagentus; (5) klaidingai teigiamus

duomenis; (6) nekoduojančią (reguliacinę) seką (Tennesen *et al.* 2011). Žmogaus egzomo sekos duomenų analizės metodai tebėra daugelio tyrėjų grupių sprendžiamas uždavinys.

## 2.6. Populiacijų tyrimų problematika

Pirmieji žmogaus genomo tyrimai parodė, kad asmenys vienas nuo kito skiriasi tik 0,1 % genomo. Didžiąją dalį (~99 %) visos variacijos sudaro skirtumai tarp individų. Tolesni tyrimai parodė, kad galima nustatyti skirtumas tarp populiacijų iš skirtingų žemynų, taip pat tarp populiacijų iš to paties žemyno ar net populiacijos viduje. Norint nustatyti mažus skirtumus tarp artimų populiacijų tyrimui reikalinga reprezentatyvi imtis, kokybiški polimorfinių genetinių žymenų duomenys ir adekvatūs analizės metodai.

Skirtumams tarp populiacijų įvertinti reikalingi informatyvūs genomo žymenys. Genetinių žymenų paieška kartu su imties parinkimu tapo vienu iš svarbiausių planuojamo tyrimo užduočių. Genetinis žymuo turėtų tenkinti bent kelis reikalavimus: jis turi pasižymėti (1) variabilumu/polimorfiškumu, (2) būti nustatomas turimais molekuliniais genetiniais metodais, (3) pats tyrimas palyginti nebrangus. Technologijų plėtra ir įgytos žinios leido plėsti populiacijų tyrimams galimų naudoti genetinių žymenų skaičių. Nuolat pildomas genetinių žymenų sąrašas paskatino kurti vieningą nomenklatūrą, kas leistų gautus tyrimų rezultatus ir duomenis pakartoti ar palyginti su kitų mokslininkų grupių gautais rezultatais.

Nustatytų genetinių žymenų gausa taip pat paskatino kurti naujas technologijas, leidžiančias vienu metu nustatyti daugelį genetinių žymenų. Naujoms technologijoms, kuriomis galima atlikti didesnės apimties tyrimus, būdingi trūkumai. Siekiant jų išvengti ar sumažinti įtaką gautiems rezultatams, reikia atlikti tinkamą gautų duomenų kokybės kontrolę.

Tyrimo imties sudarymas tapo kitu svarbiu uždaviniu vykdant populiacijų tyrimus. Reprezentatyvios imties įtraukimas neretai leidžia nustatyti net ir mažus skirtumus populiacijos viduje. Dažnu atveju dalis populiacijų dėl kultūrinių, kalbinių, socialinių, ekonominių priežasčių yra neprieinamos populiacijų tyrimų

vykdytojams. Tai skatina tarptautinį bendradarbiavimą tyrėjų grupių, kurios taiko naujos kartos technologijas, ir mokslininkų, kurie gali į tyrimą įtraukti uždarų ar mažesnių populiacijų mėginius. Mėginių iš daugelio skirtingų populiacijų ištyrimas ir duomenų dalinimasis leidžia mokslininkų grupėms į savo vykdomo tyrimo duomenų analizę įtraukti kuo daugiau reprezentatyvių imčių iš kitų populiacijų ir išvengti kryptingo rezultatų publikavimo.

Analizuojant atskiras genetinių žymenų grupes, t. y. Y chromosomos, mtDNA ar tolygiai po visą genomą pasiskirstę autosomų žymenys, gauti rezultatai ir išvados aprašo tik dalį tiriamos populiacijos, t. y. tėvinę ar motininę linijas. Neretai šie populiacijų tyrimų rezultatai nėra papildomų kitų genetinių žymenų tyrimų rezultatais ir išvadomis.

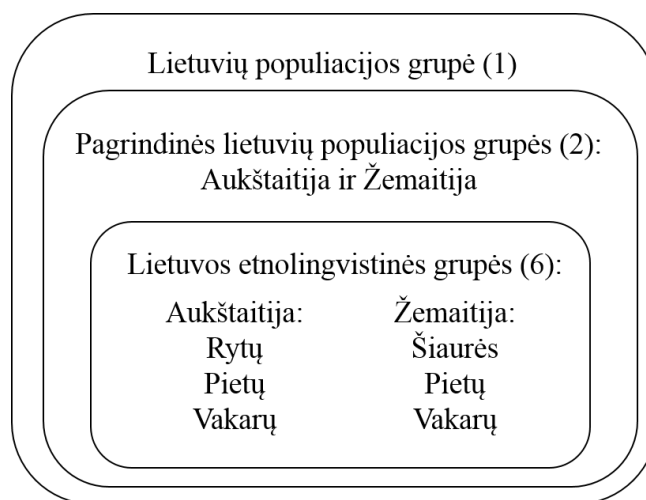
Didelės tiriamųjų imtys ir naujos kartos technologijomis gauti rezultatai neretai tampa iššūkių duomenų saugojimui ir analizei. Reikalingi prieinami analizės metodai, leidžiantys vienu metu analizuoti daugelio asmenų didelės apimties genetinę informaciją.

Populiacijų genetinių tyrimų rezultatai neretai suteikia unikalios informacijos. Šiems rezultatams ar iškeltoms hipotezėms patvirtinti reikalingas tarp atskirų mokslo šakų bendradarbiavimas. Kitų disciplinų, t. y. archeologijos, antropologijos, istorijos, tyrimai gali patvirtinti ir paaiškinti dalį populiacijų tyrimų rezultatų.

### 3. DARBO METODAI

#### 3.1. Tiriamieji

Tiriamųjų grupes sudaro atsitiktinai atrinkti negiminingi asmenys, kurie nurodė bent tris lietuvių tautybės kartas iš bendros lietuvių populiacijos. Bendra lietuvių populiacija (lietuvių populiacijos grupė) skirstoma į dvi pagrindines grupes (pagal tarmių būrius) – Aukštaitija ir Žemaitija, kiekviena jų dalinama į tris etnolingvistines grupes (pagal patarmes): Pietų, Vakarų, Rytų Aukštaitija ir Šiaurės, Pietų ir Vakarų Žemaitija (Girdenis, Zinkevičius 1966) (3.1 pav.).



3.1 pav. Darbe taikytas tiriamųjų grupavimas

Tiriamųjų genominė DNR iš veninio kraujo, surinkto 1994–1995 m. ir LITGEN projekto (VP1-3.1-ŠMM-07-K-01-013) metu 2011–2013 m., išskirta VU MF ŽMGK darbuotojų.



3.2 pav. Šešių Lietuvos etnolingvistinių grupių apimamos teritorijos

Darbo metu tirtų asmenų skaičiaus pasiskirstymas pateikiamas žemiau 3.3–3.6 paveiksluose.

<b>Aukštaitija (A)</b>		<b>Žemaitija (Ž)</b>	
Rytų (RA)	57	Šiaurės (ŠŽ)	63
Pietų (PA)	38	Pietų (PŽ)	47
Vakarų (VA)	50	Vakarų (VŽ)	38
Iš viso:	145	Iš viso:	148
293			

3.3 pav. Negiminingų asmenų, kuriems nustatyta Y chromosomos haplogrupės ir haplotipai, skaičiaus pasiskirstymas tarp šešių Lietuvos etnolingvistinių grupių

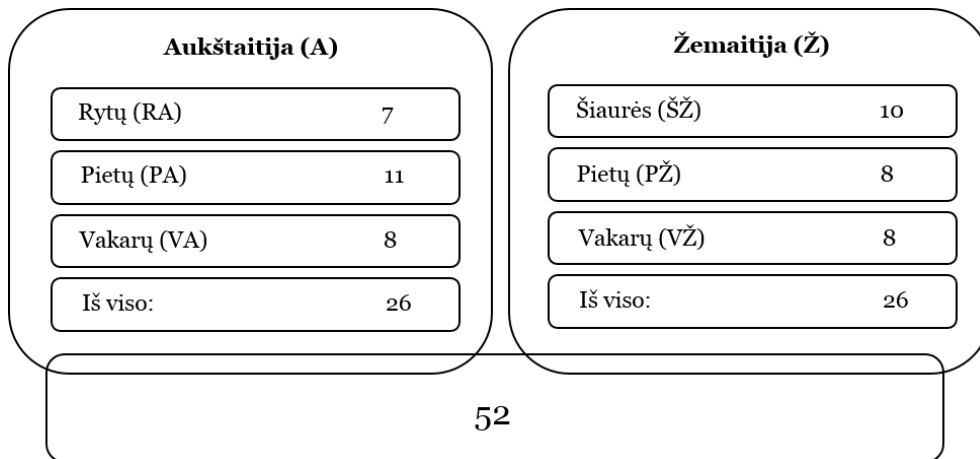
<b>Aukštaitija (A)</b>		<b>Žemaitija (Ž)</b>	
Rytų (RA)	55	Šiaurės (ŠŽ)	62
Pietų (PA)	37	Pietų (PŽ)	43
Vakarų (VA)	37	Vakarų (VŽ)	34
Iš viso:	137	Iš viso:	139
267			

3.4 pav. Negiminingų asmenų, kuriems nustatyta visa mtDNR seka, skaičiaus pasiskirstymas tarp šešių Lietuvos etnolingvistinių grupių

<b>Aukštaitija (A)</b>		<b>Žemaitija (Ž)</b>	
Rytų (RA)	46 (34)	Šiaurės (ŠŽ)	48 (36)
Pietų (PA)	48 (34)	Pietų (PŽ)	45 (34)
Vakarų (VA)	46 (35)	Vakarų (VŽ)	20 (15)
Iš viso:	140 (103)	Iš viso:	113 (85)
253 (188)			

3.5 pav. Negiminingų asmenų, kurie genotipuoti pagal 719 666 VNP, skaičiaus pasiskirstymas tarp šešių Lietuvos etnolingvistinių grupių. Skliaustuose pateiktas po pirminės ir antrinės kokybės kontrolės likusių asmenų skaičiaus pasiskirstymas





3.6 pav. Negiminingų asmenų, kuriems nustatyta egzomo DNR seka, skaičiaus pasiskirstymas tarp šešių Lietuvos etnolingvistinių grupių

Genetiniams ir genominiams tyrimams vykdyti buvo gauti Vilniaus regioninio biomedicininų tyrimų etikos komiteto leidimai: (1) „Lietuvos populiacijos genetinė įvairovė ir sandaros kitimai, susiję su evoliucija ir dažniausiai paplitusiomis ligomis“, Nr. 158200-05-329-79, data: 2011-05-03; (2) vietinis leidimas VU MF ŽMGK vykdomiems tyrimams.

### 3.2. Genomo žymenų tyrimo metodai

Darbo metu genomine DNR išskirta fenolio ir chloroformo arba automatizuotu, naudojant magnetines daleles, metodais. Atliktas išskirtos genomines DNR koncentracijos ir švarumo matavimas spektrofotometru. Y chromosomos VNP genotipuoti buvo naudojami TaqMan<sup>®</sup> zondai, Y chromosomos TTKS – *Applied Biosystems<sup>®</sup> AmpFlSTR<sup>®</sup> Yfiler<sup>™</sup>* rinkinys vykdant fragmentų kapiliarinę elektroforezę. Visos mtDNR sekoskaita taikant tikslių fragmentų išgaudymą atlikta *Illumina GAI, GAI* genetiniu analizatoriumi (vidutinis sekos padengimas >350). Tolygiai po visą genomą pasiskirstę autosomų žymenys genotipuoti *Illumina Infinium<sup>®</sup> HD* technologijos lustais (719 666 VNP). Egzomo, koduojančios genomo dalies, sekoskaita taikant tikslių fragmentų išgaudymą atlikta *5500 SOLiD<sup>™</sup> Sequencer* genetiniu analizatoriumi (vidutinis nustatytų pakaitų sekos padengimas >40). Detalūs naudotų tyrimo metodų protokolai pateikti 1 priede.

### 3.3. Duomenų analizės metodai

Analizės metu nustatytos Y chromosomos arba mtDNR haplogrupės, esant poreikiui, jungiamos hierarchiniu principu į stambesnes grupes, t.y. pavyzdžiui, Y chromosomos haplogrupės G2a ir G2ab3 jungiamos į vieną haplogrupę G2 arba G2a, kurios dažnis lygus dviejų sujungtų haplogrupių dažnių sumai.

Iš Y chromosomos haplotipų analizės pašalinti DYS385a ir DYS385b žymenys, kurie susidarė įvykus Y chromosomos fragmento, apimančio DYS385, duplikacijai. DYS385a ir DYS385b žymenys nustatomi naudojant bendrą pradmenų porą, todėl neįmanoma įvertinti tiksliai nustatytų fragmentų kilmės, t.y. DYS385a ar DYS385b. Vietoj nustatyto Y chromosomos žymens DYS389II, susidedančio iš DYS389I ir DYS389b, alelio įverčio, naudojamas DYS389II ir DYS389I žymenų alelių skirtumas, t.y. DYS389b žymens alelio reikšmė.

Programinė įranga, naudota Y chromosomos, mtDNA, tolygiai po visą genomą ir egzomo žymenų analizei, ir atliktų darbų trumpas aprašymas pateikti 3.1–3.4 lentelėse.

#### 3.3.1. Y chromosomos haplogrupių ir haplotipų analizė

Pirminė duomenų analizė atlikta Rusijos medicinos mokslų akademijos Medicininės genetikos tyrimų centro RAMN Populiacinės genetikos laboratorijos (Maskva, Rusijos Federacija) darbuotojų.

3.1 lentelė. Programinė įranga, naudota pirminei ir antrinei Y chromosomos haplotipų ir haplogrupių analizei

<b>Duomenų analizė</b>	<b>Programinė įranga</b>	<b>Analizė</b>
<b>Pirminė</b>	<i>GeneMapper® Software v4.0 (GMSv4.0)</i>	TTKS žymenų fragmentų ilgių ir alelių nustatymas
<b>Antrinė</b>	<i>Haplogroup Predictor</i> <sup>1</sup> (Athey 2005)	Haplogrupių ir haplotipų atitikimo su tam tikra tikimybe tikrinimas
	<i>Arlequin v3.5.1.2</i> <sup>1</sup> (Excoffier, Lischer 2010)	Haplogrupių, haplotipų įvairovės įverčių skaičiavimas
		F <sub>ST</sub> atstumų, remiantis haplogrupių pasiskirstymu, ir R <sub>ST</sub> atstumų, remiantis haplotipų pasiskirstymu, skaičiavimas (Belle <i>et al.</i> 2010)
		AMOVA analizė

	<b>Phylip v3.69<sup>1</sup></b> (Felsenstein 2004)	Nei (D) genetinių atstumų (Nei 1972) skaičiavimas (Zoosmann-Diskin 2010) Hierarchinio klasterizavimo UPGMA (angl. <i>UPGMA, Unweighted Pair Group Method with Arithmetic Mean</i> ) arba NJ metodais dendrogramos sudarymas
	<b>SPSS Statistics Desktop v22.0.0</b> (bandomoji versija) <sup>1</sup>	MDS, remiantis Nei (D) ir F <sub>ST</sub> atstumais pagal haplogrupių pasiskirstymą
	<b>Microsoft® Office Excel</b>	Grafikų braižymas Duomenų analizė
	<b>R v3.0.3 (Ade4, Rcmdr)<sup>1</sup></b>	Dviejų atstumų matricų panašumo įvertinimas (Mantelio kriterijus) PKA
	<b>Barrier v2.2<sup>1</sup></b> (Manni <i>et al.</i> 2004)	Haplogrupių įvairovės barjerų nustatymas
	<b>Network v4.6.1.2<sup>1</sup></b> (www.fluxus-engineering.com)	Filogenetinių medžių, remiantis Y chromosomos 15 TTKS žymenų pasiskirstymu haplogrupių viduje, sudarymas
	<b>BATWING<sup>1</sup></b> (Wilson <i>et al.</i> 2003)	TMRCA, remiantis 15 TTKS žymenų pasiskirstymu haplogrupės viduje, skaičiavimas

<sup>1</sup>Laisvos prieigos

### 3.3.2. mtDNR haplogrupių ir haplotipų analizė

Pirminė duomenų analizė atlikta Makso Planko Evoliucinės antropologijos instituto (Leipcigas, Vokietija) darbuotojų: (1) nustatytų fragmentų lygiavimas; (2) mtDNR rekonstravimas pagal referentinį genomą, leidžiant du nukleotidų neatitikimo (angl. *Mismatch*) atvejus ir vieną tarpo (angl. *Gap*) atvejį; (3) kokybės kontrolė: pozicijos, kurių padengimas siekė vieną kartą ir (ar) kuriose MAF >30 %, pažymėtos N, t. y. nenustatyta.

3.2 lentelė. Programinė įranga, naudota pirminei ir antrinei mtDNR haplotipų ir haplogrupių analizei

Duomenų analizė	Programinė įranga	Analizė
Pirminė	<b>Haplogrep<sup>1</sup></b> (Kloss-Brandstatter <i>et al.</i> 2011)	Haplogrupių nustatymas
Antrinė	<b>Muscle<sup>1</sup></b> (Edgar 2004)	Dauginis lygiavimas su rCRS
	<b>BioEdit v7.1.3.0<sup>1</sup></b> (Hall 1999) <b>DnaSP v5.10.01</b>	Sekos redagavimas, tarpų ar kitų pozicijų, nurodytų fragmentų ar tiriamųjų šalinimas

(Librado, Rozas 2009)	Sekos variantų analizė
<b>Arlequin v3.5.1.2</b> <sup>1</sup> (Excoffier, Lischer 2010)	Haplogrupių, haplotipų įvairovės įverčių, Tajima D įverčių skaičiavimas F <sub>ST</sub> atstumų pagal haplogrupes ir porinių skirtumų ar Φ <sub>ST</sub> atstumų pagal haplotipus skaičiavimas (Barbieri <i>et al.</i> 2012; Kasperavičiūtė <i>et al.</i> 2004)
	AMOVA analizė
<b>SPSS Statistics Desktop v22.0.0</b> (bandomoji versija) <sup>1</sup>	MDS
<b>Microsoft® Office Excel</b>	Grafikų braižymas Duomenų analizė
<b>R v3.0.3 (Ade4, Rcmdr)</b> <sup>1</sup>	Dviejų atstumų matricų panašumo įvertinimas (Mantelio kriterijus) PKA
<b>Barrier v2.2</b> <sup>1</sup> (Manni <i>et al.</i> 2004)	Haplogrupių ir haplotipų įvairovės barjerų nustatymas
<b>Network v4.6.1.2</b> <sup>1</sup> (www.fluxus-engineering.com)	Filogenetiniai medžių, remiantis haplotipų (577–16 023 np) pasiskirstymu haplogrupių viduje, sudarymas Vidutinio mutacijų skaičiaus (ρ) filogenetinėje linijoje, naudojant haplotipus (16 564 np), skaičiavimas
<b>Filogenetinių linijų amžiaus skaičiuoklė</b> (Costa <i>et al.</i> 2013)	Filogenetinių linijų amžiaus ir PI pagal haplotipų (16 564 np) įvairovę nustatymas
<b>Programavimo kalba Python v2.7.4</b>	Duomenų tvarkymas ir apdorojimas, naudojant automatinius scenarijus (angl. <i>Script</i> ) <sup>2</sup>

<sup>1</sup>Laisvos prieigos

<sup>2</sup>Atlikta bendradarbiaujant su specialistu

### 3.3.3. Tolygiai po genomą pasiskirsčiusių autosomų VNP analizė

Pirminė duomenų analizė atlikta VU MF ŽMGK jaunesniosios mokslo darbuotojos ir doktorantės I. Domarkienės.

Atliktas protėvinių alelių pagal PK (angl. *PCA*IMS, *Principal Components Ancestral Informative Markers*) nustatymas. Iš PK1–100 pasirinkta po 200 VNP, kuriems suteikti didžiausi svorio koeficientai (*EIGENSOFT v5.0.1*). Kiekvienos PK kiekvienam iš 200 VNP suteikti santykiniai svoriai, didžiausias svoris prilygintas 1, visiems kitiems VNP suteikti proporcingi santykiniai svorio koeficientai. Galutinis VNP koeficientas gautas VNP santykinį svorio

koeficientą dauginant iš konkretaus VNP pasitaikymo skaičiaus visose PK. Tolesnei analizei atrinkti VNP, kurių galutinis koeficientas >1.

3.3 lentelė. Programinė įranga, naudota pirminei ir antrinei tolygiai po visą genomą pasiskirsčiusių VNP analizei

Duomenų analizė	Programinė įranga	Analizė
Pirminė	<i>GenomeStudio v2011.1</i> (Illumina® <i>GenomeStudio 2011</i> , gamintojas <i>Illumina, Inc. 2003-2011</i> )	Parametrai (2 priedo 13 lentelė, 2 priedo 14 lentelė)
Antrinė	<i>PLINK v1.9</i> <sup>1</sup> (Purcell <i>et al.</i> 2007)	Autosomų VNP atrinkimas
		Asmenų, kuriems remiantis identiškumo pagal kilmę (angl. <i>IBD, Identical by Descent</i> ) PI HAT < 0,4, pašalinimas atsitiktine tvarka
		Inbrydingo koeficiento (F) įvertinimas, galimam mėginių užterštumui tyrimo metu nustatyti
		VNP, neatitinkančių Hardžio ir Vainbergo dėsnio, kai $p < 10^{-5}$ , pašalinimas
		VNP su trūkstamais genotipais (>1 %) pašalinimas
		VNP, kurių MAF < 0,01, pašalinimas
		Alelių asociacijos ( $r^2 < 0,2$ ) pašalinimas
		MDS, naudojant VNP ( $r^2 < 0,2$ )
		LD struktūros nustatymas
	<i>EIGENSOFT v5.0.1</i> <sup>1</sup> (Patterson <i>et al.</i> 2006)	PKA (VNP $r^2 < 0,2$ )
		PK statistinio reikšmingumo įvertinimas (Tracy-Widom kriterijus)
		VNP, kurių genetinė variacija tam tikroje PK didžiausia tarp populiacijų, nustatymas
	<i>Microsoft® Office Excel</i>	Grafikų braižymas
		Duomenų analizė
		PCAIMS nustatymas
<i>R v3.0.3 (Ade4, Rcmdr)</i> <sup>1</sup>	Dviejų atstumų matricių panašumo įvertinimas (Mantelio kriterijus).	
	Dviejų nepriklausomų imčių vidurkių palyginimas (nepriklausomų imčių t kriterijus)	
<i>Barrier v2.2</i> <sup>1</sup> (Manni <i>et al.</i> 2004)	Genetinės įvairovės barjerų nustatymas	
<i>Phylip v3.69</i> <sup>1</sup> (Felsenstein 2004)	Hierarchinio klasterizavimo UPGMA arba NJ metodais dendrogramos sudarymas	

	<b>Structure v.2.3.4<sup>1</sup></b> (Pritchard <i>et al.</i> 2000)	Modelio, paremto skirtingų grupių skaičiumi (K), kuris aprašo esamą genetinių duomenų variaciją tarp populiacijų, sudarymas
	<b>Programavimo kalba Python v2.7.4<sup>1</sup></b>	Duomenų tvarkymas ir apdorojimas, naudojant automatinius scenarijus <sup>2</sup>

<sup>1</sup>Laisvos prieigos

<sup>2</sup>Atlikta bendradarbiaujant su specialistu

### 3.3.4. Egzomo autosomų žymenų analizė

Pirminė, antrinė, tretinė sekoskaitos duomenų analizė atlikta kartu su VU MF ŽMGK mokslo darbuotoja dr. L. Ambrozaityte.

3.4 lentelė. Programinė įranga, naudota pirminei, antrinei, tretinei ir ketvirtinei egzomo DNR sekos variantų analizei

Duomenų analizė	Programinė įranga	Analizė
Pirminė Antrinė Tretinė	<b>LifeScope™ Genomic Analysis Software v2.5.</b>	Parametrai (2 priedo 16 lentelė)
Ketvirtinė	<b>Arlequin v3.5.1.2<sup>1</sup></b> (Excoffier, Lischer 2010)	Porinių skirtumų pagal 6 939 autosomų vieno nukleotido pakaitų alelių pasiskirstymą skaičiavimas
	<b>Microsoft® Office Excel</b>	Grafikų braižymas
		Duomenų analizė
	<b>R v3.0.3 (Ade4, Rcmdr)<sup>1</sup></b>	Venno diagramos sudarymas
		Tiesinės regresijos modelio sudarymas
		PKA
	<b>Programavimo kalba Python v2.7.4<sup>1</sup></b>	Duomenų tvarkymas ir apdorojimas, naudojant automatinius scenarijus <sup>2</sup>
	<b>Atviro kodo reliacinių duomenų bazių valdymo sistema PostgreSQL v.9.2.3<sup>1</sup></b>	Sukurta egzomo anototų pakaitų reliacinė duomenų bazė <sup>2</sup>
		Parašytos SQL užklauskos egzomo anototų pakaitų duomenų analizei <sup>2</sup>

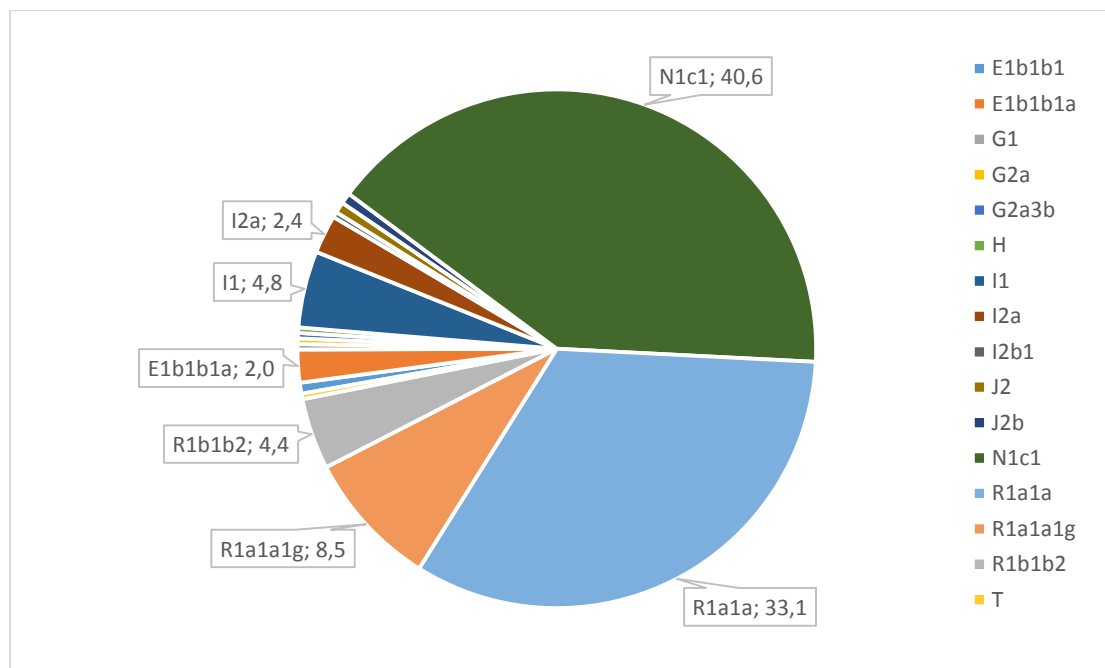
<sup>1</sup>Laisvos prieigos

<sup>2</sup>Atlikta bendradarbiaujant su specialistu

## 4. TYRIMŲ REZULTATAI

### 4.1. Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis Y chromosomos haplogrupių ir haplotipų pasiskirstymu

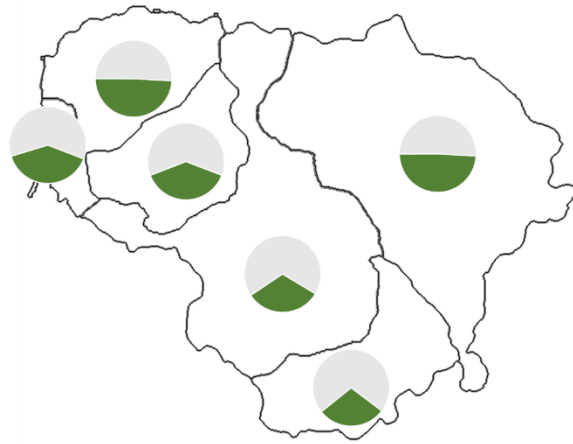
Tirtoje lietuvių populiacijos grupėje (3.2 pav.) nustatyta 16 skirtingų Y chromosomos haplogrupių (4.1 pav.). Šios haplogrupės priklauso aštuonioms skirtingoms filogenetinėms linijoms (E, G, H, I, J, N, R, T). Šešios (E, G, I, J, N, R), kurioms priklauso >99 % tirtos lietuvių populiacijos, yra dažniausios filogenetinės linijos Europoje (Wiik 2008).



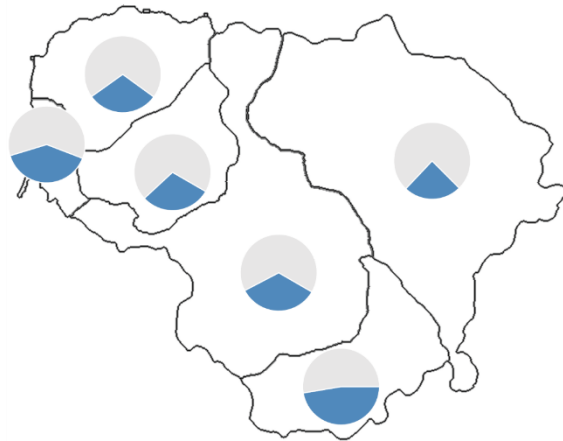
4.1 pav. Y chromosomos haplogrupių pasiskirstymas (dažnis, %) tirtoje Lietuvos populiacijos grupėje

Dažniausios haplogrupės tirtoje lietuvių populiacijos grupėje yra N1c1(M178) – 40,6 %, R1a1a(M198) – 33,1 %, kuri kartu su R1a1a1g(M458) sudaro 41,6 %. Trys dažniausios haplogrupės sudaro ~82 %, kitos 13 haplogrupių sudaro ~18 % tirtos lietuvių populiacijos.

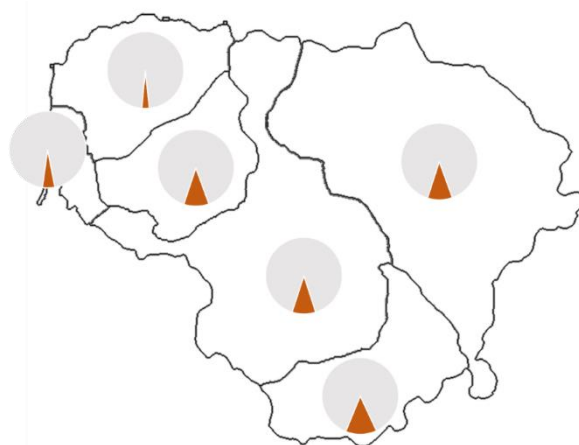
Matomas nustatytų dažniausių haplogrupių pasiskirstymo šiaurės–pietų ašimi gradientas. Haplogrupės N1c1 didžiausias dažnis nustatomas Šiaurės Lietuvoje, jis mažėja pietų kryptimi (4.2 pav.). Haplogrupės R1a1a ir R1a1a1g didžiausias dažnis nustatomas Pietų Lietuvoje, jis mažėja šiaurės kryptimi (4.3 pav., 4.4 pav.).



4.2 pav. Y chromosomos haplogrupės N1c1 pasiskirstymas (žalia spalva) tarp tirtų Lietuvos etnolingvistinių grupių



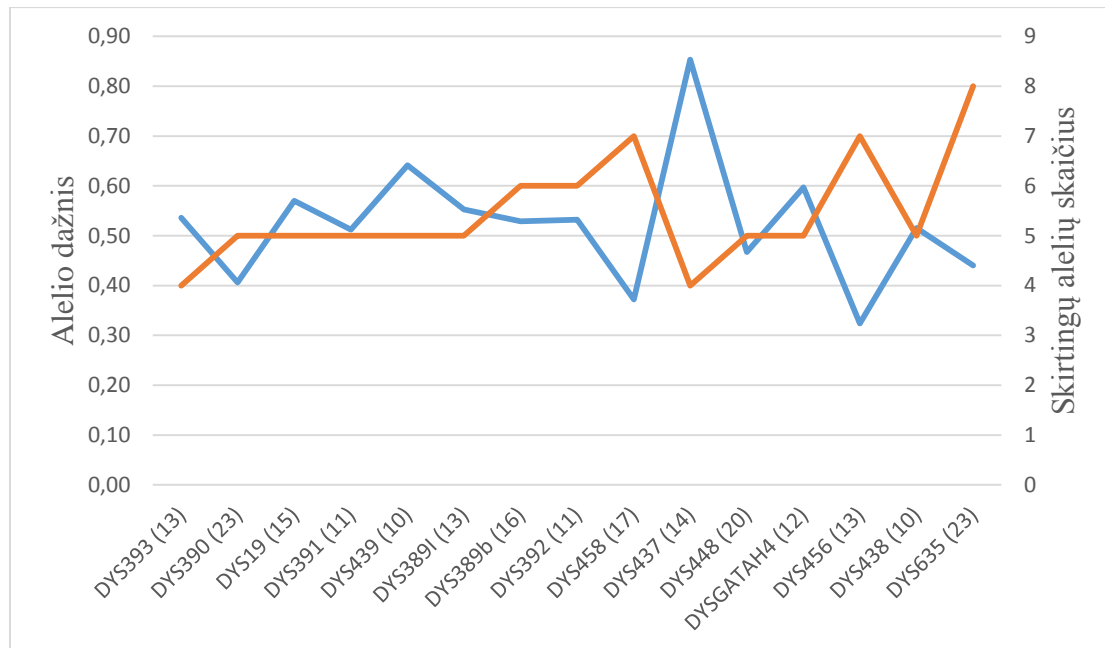
4.3 pav. Y chromosomos haplogrupės R1a1a pasiskirstymas (mėlyna spalva) tarp tirtų Lietuvos etnolingvistinių grupių



4.4 pav. Y chromosomos haplogrupės R1a1a1g pasiskirstymas (raudona spalva) tarp tirtų Lietuvos etnolingvistinių grupių



Tiriant 15 TTKS pasiskirstymą nustatyta, kad DYS458, DYS456, DYS635 yra variabiliausi, DYS437 mažiausiai variabilus žymuo tirtoje lietuvių populiacijos grupėje (4.5 pav.).



4.5 pav. Y chromosomos 15 TTKS žymenų dažniausių ir skirtingų alelių pasiskirstymas tirtoje Lietuvos populiacijos grupėje. Kiekvieno žymens pavadinime skliaustuose pateikti dažniausi aleliai, kurių dažnis populiacijoje pateikiamas mėlynos spalvos tiese (ašis kairėje). Kiekvieno žymens nustatytas skirtingų alelių skaičius pateikimas oranžinės spalvos tiese (ašis dešinėje)

DYS458 matomi septyni skirtingi aleliai, iš kurių dažniausias (17) nustatytas 37 % tirtos populiacijos; DYS456 – septyni skirtingi aleliai, iš kurių dažniausias (13) nustatytas 32 % tirtos populiacijos; DYS635 – aštuoni skirtingi aleliai, iš kurių dažniausias (23) nustatytas 44 % tirtos populiacijos; DYS437 – keturi skirtingi aleliai, iš kurių dažniausias (14) nustatytas 85 % tirtos populiacijos.

Tirtoje lietuvių populiacijos grupėje nustatyta 250 skirtingų haplotipų. Aukštaitijos grupėje skirtingi 137 (iš tirtų 145 asmenų), Žemaitijos grupėje skirtingi 129 (iš tirtų 148 asmenų) haplotipai. Tarp šešių Lietuvos etnolingvistinių grupių daugiausia skirtingų haplotipų nustatyta Pietų Aukštaitijoje, mažiausiai – Pietų Žemaitijoje. Tirtoje lietuvių populiacijos grupėje nustatytas dažniausias haplotipas, yra N1c1 haplogrupės viduje ir sudaro 1,7 % tirtos lietuvių populiacijos. Taip pat aptikti retesni haplotipai:

- (1) N1c1 haplogrupės viduje du haplotipai, kurių kiekvieno dažnis – 1,3 %;  
 (2) N1c1 haplogrupės viduje penki haplotipai, kurių kiekvieno dažnis – 1,0 %;  
 (3) 23 haplotipai, kurių kiekvieno dažnis – 0,7 %; (4) 219 haplotipų, kurie yra pavieniai tirtoje lietuvių populiacijos grupėje.

Apskaičiuoti Y chromosomos haplogrupių (16) ir haplotipų (250 remiantis 15 TTKS) įvairovės įverčiai, t. y. tikimybė, kad atsitiktinai parinkti du asmenys turės skirtingas haplogrupes arba haplotipus, šešioms Lietuvos etnolingvistinėms grupėms, dviem pagrindinėms grupėms (Aukštaitija ir Žemaitija) ir tirtai lietuvių populiacijos grupei. Taip pat apskaičiuoti Y chromosomos haplotipų, sudarytų iš 15 TTKS žymenų, įvairovės įverčiai dažniausių Y chromosomos haplogrupių viduje (4.1 lentelė).

Didžiausia Y chromosomos haplogrupių įvairovė tarp šešių Lietuvos etnolingvistinių grupių yra Vakarų Aukštaitijoje (0,7739 ( $\pm 0,0370$ )), mažiausia – Šiaurės Žemaitijoje (0,6667 ( $\pm 0,0439$ )). Lyginant dvi pagrindines Lietuvos grupes, Aukštaitijoje matoma didesnė haplogrupių įvairovė (0,7300 ( $\pm 0,0232$ )) negu Žemaitijoje (0,7023 ( $\pm 0,0258$ )).

4.1 lentelė. Haplogrupių (16) ir haplotipų (15 TTKS) įvairovės įverčiai tirtoje lietuvių populiacijos grupėje

Populiacijos grupė	Imtis	Haplogrupių įvairovė (SN)	Haplotipų skaičius	Haplotipų įvairovė (SN)
P. Aukštaitija	38	0,6899 ( $\pm 0,0537$ )	38	1,0000 ( $\pm 0,0060$ )
R. Aukštaitija	57	0,6898 ( $\pm 0,0480$ )	54	0,9981 ( $\pm 0,0037$ )
V. Aukštaitija	50	0,7739 ( $\pm 0,0370$ )	49	0,9992 ( $\pm 0,0042$ )
Aukštaitija	145	0,7300 ( $\pm 0,0232$ )	137	0,9991 ( $\pm 0,0010$ )
Š. Žemaitija	63	0,6667 ( $\pm 0,0439$ )	59	0,9980 ( $\pm 0,0033$ )
V. Žemaitija	38	0,6970 ( $\pm 0,0488$ )	37	0,9986 ( $\pm 0,0065$ )
P. Žemaitija	47	0,7595 ( $\pm 0,0414$ )	40	0,9926 ( $\pm 0,0062$ )
Žemaitija	148	0,7023 ( $\pm 0,0258$ )	129	0,9980 ( $\pm 0,0011$ )
Lietuva	293	0,7152 ( $\pm 0,0173$ )	250	0,9986 ( $\pm 0,0005$ )
Haplogrupė	Imtis	Haplogrupių įvairovė (SN)	Haplotipų skaičius	Haplotipų įvairovė (SN)
N1c1	119	–	85	0,9927 ( $\pm 0,0023$ )
R1a1a	97	–	91	0,9987 ( $\pm 0,0017$ )
R1a1a1g	25	–	24	0,9967 ( $\pm 0,0125$ )

Didžiausia Y chromosomos haplotipų įvairovė tarp šešių Lietuvos etnolingvistinių grupių yra Pietų Aukštaitijoje (1,0000 ( $\pm 0,006$ )), mažiausia – Pietų Žemaitijoje (0,9926 ( $\pm 0,0062$ )). Tarp dviejų pagrindinių Lietuvos grupių

Aukštaitijoje matoma didesnė haplotipų įvairovė (0,9991 ( $\pm 0,001$ )) negu Žemaitijoje (0,9980 ( $\pm 0,0011$ )). Tiroje lietuvių populiacijos grupėje apskaičiuotas vidutinis porinis haplotipų skirtumas – 8,8 ( $\pm 4,1$ ), t. y. atsitiktinai parinkti du haplotipai vidutiniškai skiriasi ~9 ( $\pm 4$ ) žymenų aleliais. Tiroje lietuvių populiacijos grupėje haplotipo žymens vidutinis įvairovės įvertis yra 0,5893 ( $\pm 0,3015$ ), t. y. ~59 % ( $\pm 30$  %) tikimybė, kad atsitiktinai parinktų haplotipų konkretaus žymens aleliai bus skirtingi.

Tiroje lietuvių populiacijos grupėje didžiausia Y chromosomos haplotipų įvairovė dažniausių haplogrupių (N1c1, R1a1a, R1a1a1g) viduje, nustatyta haplogrupei R1a1a – 0,9987 ( $\pm 0,0017$ ), 91 skirtingas haplotipas iš 97 tirtų asmenų. Mažiausia Y chromosomos haplotipų įvairovė nustatyta haplogrupei N1c1 – 0,9927 ( $\pm 0,0023$ ), 85 skirtingi haplotipai iš 119 tirtų asmenų. Haplogrupei R1a1a1g nustatyta haplotipų įvairovė – 0,9967 ( $\pm 0,0125$ ), 24 skirtingi haplotipai iš 25 tirtų asmenų.

Apskaičiuoti šešių Lietuvos etnolingvistinių grupių Nei (D) ir  $F_{ST}$  atstumai pagal Y chromosomos haplogrupių (E1b1b1 (+E1b1b1a), I1, I2a, N1c1, R1a1a, R1a1a1g, R1b1b2, kitos (+G1, G2a, G2a3b, H, I2b1, J2, J2b, T)) pasiskirstymą tiroje Lietuvos populiacijos grupėje (4.2 lentelė).

4.2 lentelė. Nei (D) ir  $F_{ST}$  atstumų tarp šešių Lietuvos etnolingvistinių grupių matrica. Virš įstrižainės Nei (D) atstumai, žemiau –  $F_{ST}$  atstumai. Paryškinti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių (angl. *Permutation*)) įverčiai. Trumpiniai: ŠŽ – Šiaurės Žemaitija, PŽ – Pietų Žemaitija, VŽ – Vakarų Žemaitija, VA – Vakarų Aukštaitija, PA – Pietų Aukštaitija, RA – Rytų Aukštaitija

	ŠŽ	PŽ	VŽ	VA	PA	RA
ŠŽ	0	0,003416	0,003673	0,005699	0,013052	0,002597
PŽ	-0,00224	0	0,002536	0,001633	0,006654	0,003187
VŽ	-0,00305	-0,01234	0	0,002300	0,004174	0,005575
VA	0,00875	-0,01391	-0,01282	0	0,003824	0,007274
PA	<b>0,04120</b>	0,00640	-0,00638	-0,00581	0	0,015460
RA	-0,00394	-0,00469	0,00471	0,01399	<b>0,04945</b>	0

Nustatyti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių)  $F_{ST}$  atstumai tarp Pietų Aukštaitijos ir Rytų Aukštaitijos, Pietų Aukštaitijos ir Šiaurės Žemaitijos. Pastarosioms etnolingvistinių grupių poroms apskaičiuoti didžiausi Nei (D)

atstumai. Rasta statistiškai reikšminga koreliacija tarp Nei (D) ir geografinių atstumų (4.2 lentelė) (Mantelio kriterijus,  $r = 0,47$ ;  $p = 0,0454$ ; 10 000 kėlinių).

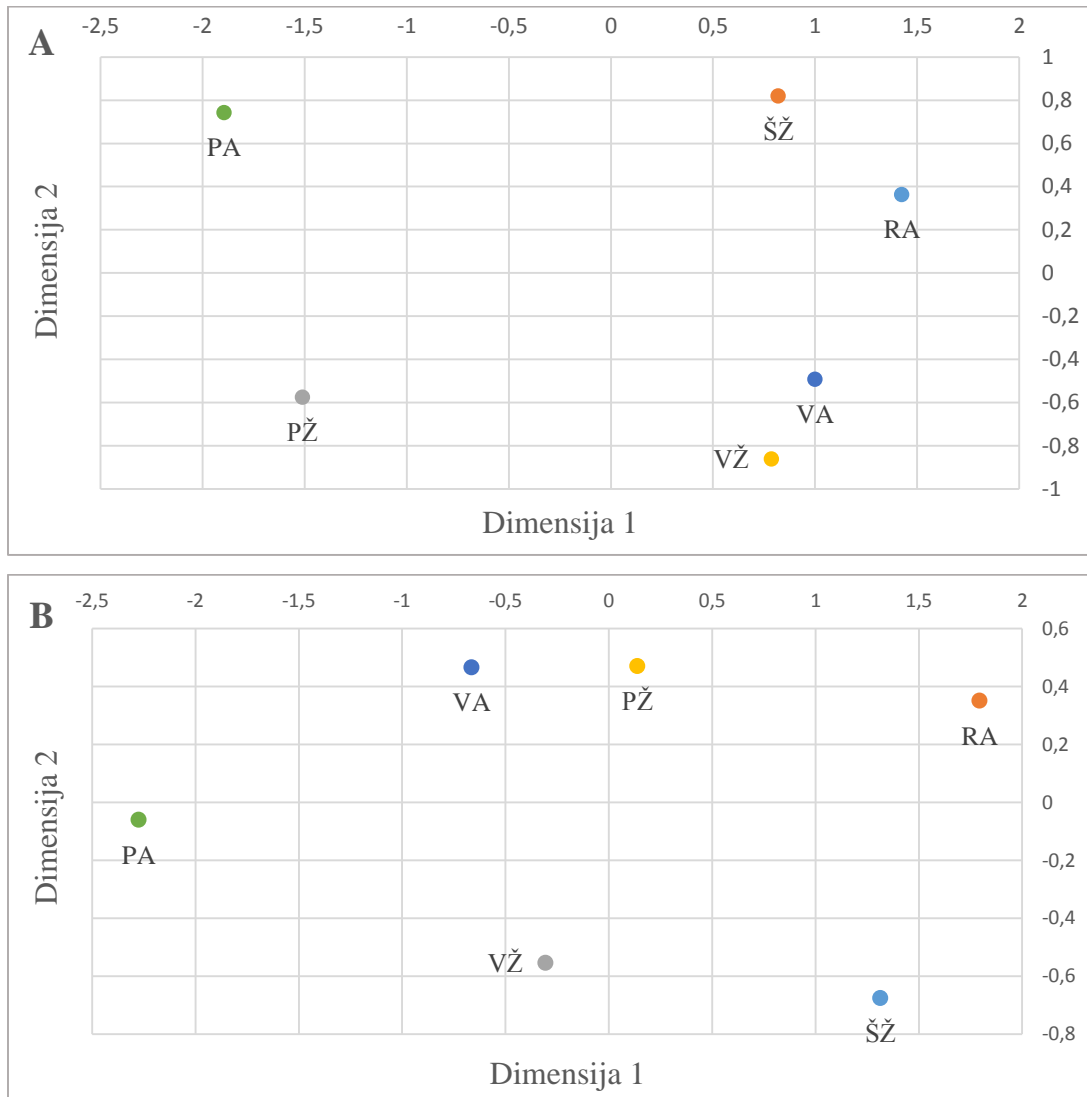
AMOVA analizė, remiantis haplogrupių  $F_{ST}$  ir haplotipų  $R_{ST}$  atstumais, atlikta jungiant tirtas šešias Lietuvos etnolingvistines grupes pagal geografinę padėtį ar kalbinius, kultūrinius skirtumus. Didžiausia haplogrupių ir haplotipų įvairovė, siekianti ~99 %, yra tarp tiriamų asmenų etnolingvistinių grupių viduje. Nustatyti haplogrupių ir haplotipų statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) įvairovės įverčiai tarp grupių, sudarytų pagal geografinę padėtį: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Vakarų, Pietų Žemaitija, Vakarų Aukštaitija; (3) Pietų Aukštaitija (4.3 lentelė).

4.3 lentelė. AMOVA analizės, pagal Y chromosomos haplogrupių (16) ir haplotipų (15 TTKS) pasiskirstymą tirtoje Lietuvos populiacijoje, rezultatai. Paryškinti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) įverčiai. Tirtos šešios Lietuvos etnolingvistinės grupės sujungtos pagal geografinę padėtį: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Vakarų Žemaitija, Vakarų Aukštaitija, Pietų Žemaitija; (3) Pietų Aukštaitija

Haplogrupės ( $F_{ST}$ atstumai)			
Genetinės įvairovės šaltinis	df <sup>1</sup>	Fiksacijos indeksas (F)	Įvairovė (%)
(1) tarp grupių	2	<b>0,0189</b> ( $p = 0,01683$ )	1,89
(2) tarp populiacijų grupių viduje	3	-0,00915 <sup>2</sup>	-0,9
(3) tarp tiriamųjų populiacijų viduje	287	0,00992 <sup>2</sup>	99,01
Haplotipai ( $R_{ST}$ atstumai)			
Genetinės įvairovės šaltinis	df <sup>1</sup>	Fiksacijos indeksas (F)	Įvairovė (%)
(1) tarp grupių	2	<b>0,02023</b> ( $p = 0,01653$ )	2,02
(2) tarp populiacijų grupių viduje	3	-0,00988 <sup>2</sup>	-0,97
(3) tarp tiriamųjų populiacijų viduje	287	0,01055 <sup>2</sup>	98,95

<sup>1</sup>df – laisvės laipsnis

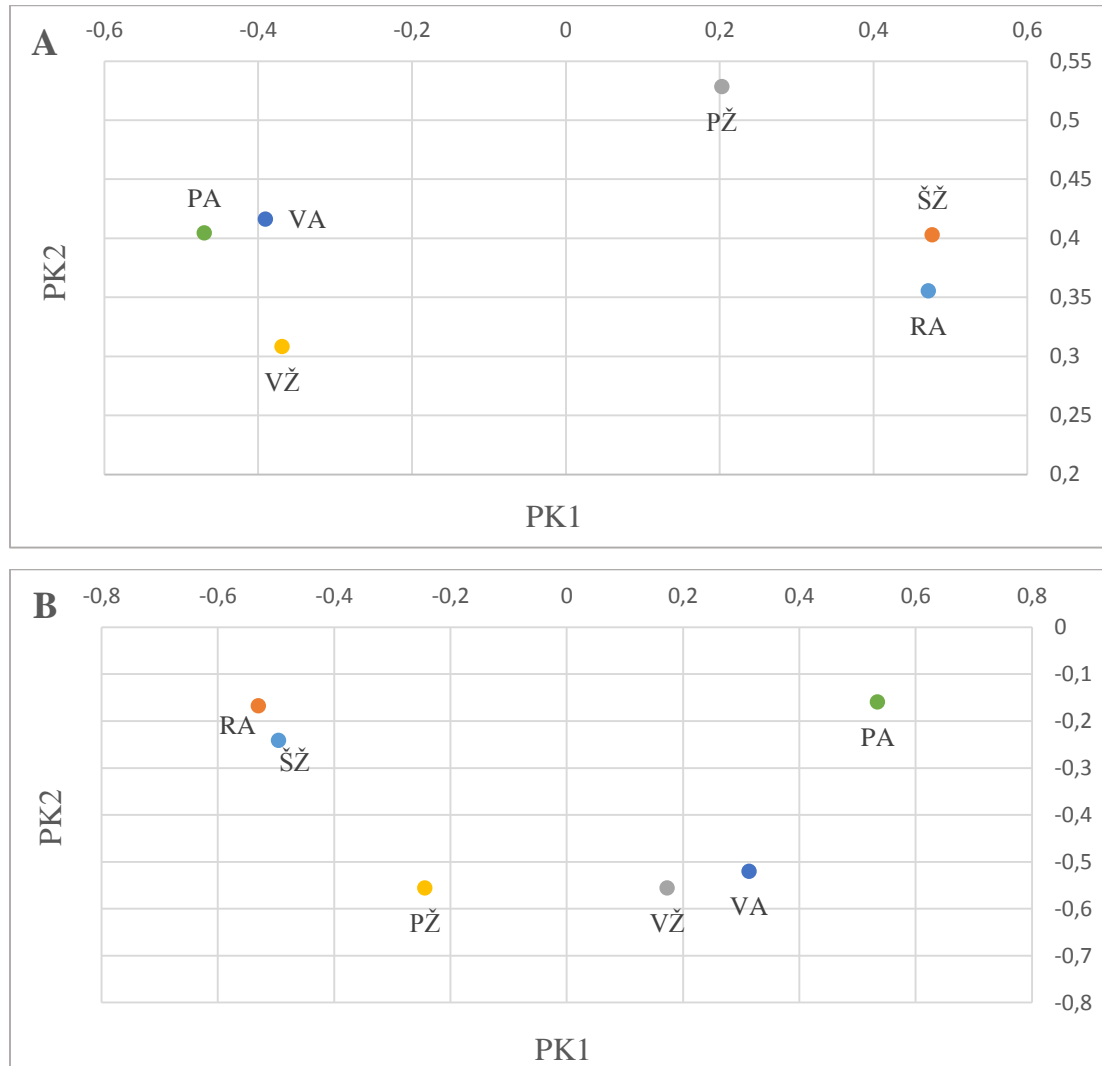
<sup>2</sup>Statistiškai nereikšmingi įverčiai ( $p > 0,05$ )



4.6 pav. MDS, remiantis (A)  $F_{ST}$  ir (B) Nei (D) atstumais pagal Y chromosomos haplogrupių (E1b1b1 (+E1b1b1a), I1, I2a, N1c1, R1a1a, R1a1a1g, R1b1b2, kitos (+G1, G2a, G2a3b, H, I2b1, J2, J2b, T)) pasiskirstymą. MDS paklaidos funkcija (angl. *Stress*) ir koreliacijos koeficientas ( $r^2$ ): (1)  $F_{ST}$  atstumų matrica,  $stress = 0,0025$ ;  $r^2 = 0,99996$ ; (2) Nei (D) atstumų matrica,  $stress = 0,0031$ ;  $r^2 = 0,99994$

Vizualizavus gautus rezultatus MDS metodu, remiantis  $F_{ST}$  atstumais pagal Y chromosomos haplogrupių pasiskirstymą, pirmoje ir antroje dimensijoje matomos keturios grupės: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Vakarų Žemaitija, Vakarų Aukštaitija; (3) Pietų Žemaitija; (4) Pietų Aukštaitija. Vizualizavus gautus rezultatus MDS metodu, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių pasiskirstymą, pirmoje ir antroje dimensijoje atsiduria trys grupės: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Vakarų Žemaitija, Vakarų Aukštaitija, Pietų Žemaitija; (3) Pietų Aukštaitija. Remiantis

pastarąjį MDS, tirta Lietuvos populiacijos grupė dalijama į Šiaurės, Vidurio ir Pietų Lietuvos populiacijų grupes (4.6 pav.).

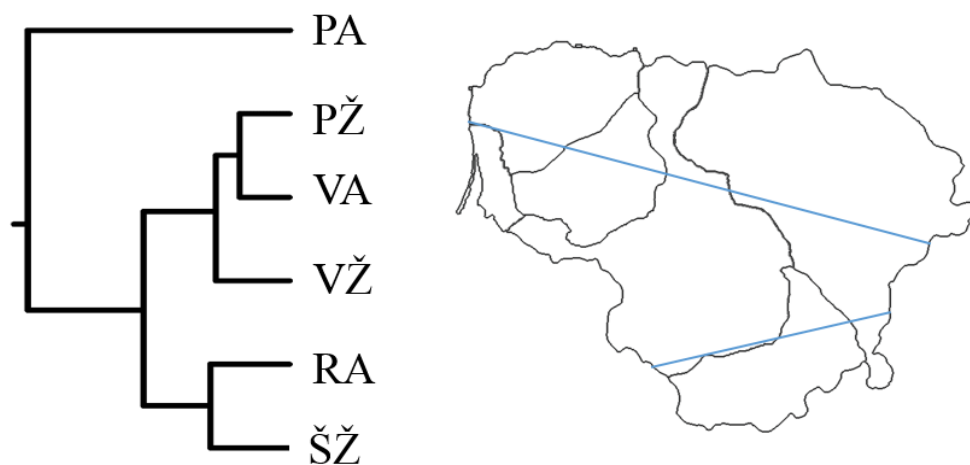


4.7 pav. PKA, remiantis (A)  $F_{ST}$  ir (B) Nei (D) atstumais pagal Y chromosomos haplogrupių (E1b1b1 (+E1b1b1a), I1, I2a, N1c1, R1a1a, R1a1a1g, R1b1b2, kitos (+G1, G2a, G2a3b, H, I2b1, J2, J2b, T)) pasiskirstymą, rezultatai

PKA, remiantis  $F_{ST}$  atstumais pagal Y chromosomos haplogrupių pasiskirstymą, rezultatai rodo, kad PK1 aprašo 48,47 %, PK2 – 30,71 % variacijos tarp tirtų šešių Lietuvos etnolingvistinių grupių. PK1 aprašo įvairovę tarp trijų klasterių (šiaurės–pietų ašis), gautų vizualizavus rezultatus: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Pietų Žemaitija; (3) Vakarų Žemaitija, Vakarų, Pietų Aukštaitija. PKA, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių pasiskirstymą, rezultatai rodo, kad PK1 aprašo 54,82 %, PK2 –

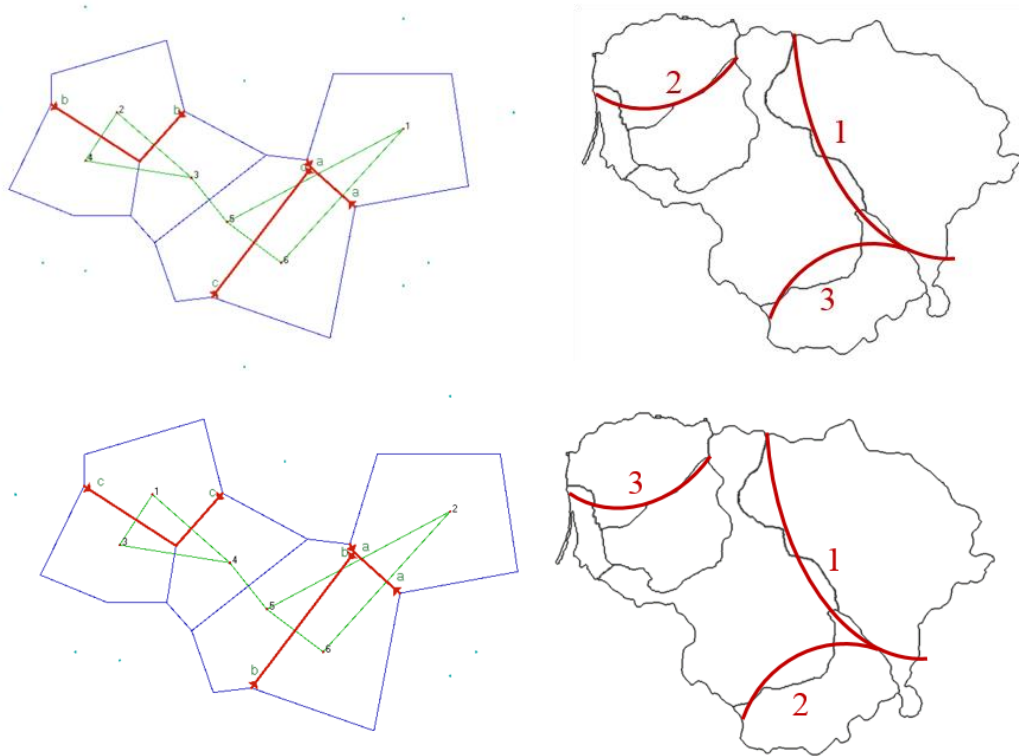
37,49 % variacijos tarp tirtų šešių etnolingvistinių grupių. PK1 aprašo variacijos dalį tarp tų pačių klasterių (4.7 pav.).

Hierarchinio klasterizavimo UPGMA metodu, remiantis Nei (D) atstumais, sudaryta tirtų šešių Lietuvos etnolingvistinių grupių dendrograma. Medyje matomos trys grupės: (1) Šiaurės Žemaitija, Rytų Aukštaitija; (2) Vakarų Žemaitija, Vakarų Aukštaitija, Pietų Žemaitija; (3) Pietų Aukštaitija, pastaroji etnolingvistinė grupė (Pietų Aukštaitija) nutolusi didžiausiu atstumu nuo kitų dviejų klasterių (4.8 pav.).



4.8 pav. Hierarchinio klasterizavimo UPGMA metodu, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių pasiskirstymą, sudaryta šešių Lietuvos etnolingvistinių grupių dendrograma. Paveikslo dešinėje linijomis atskirti dendrogramoje stebimi etnolingvistinių grupių klasteriai

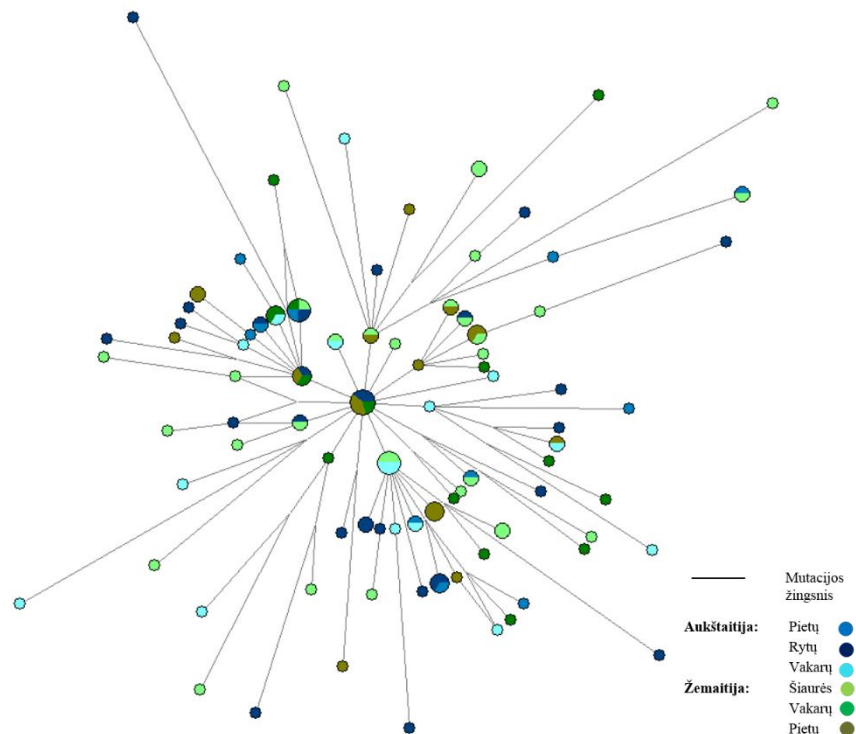
Nustatyti Y chromosomos haplogrupių įvairovės pasiskirstymo barjerai tarp tirtų šešių Lietuvos etnolingvistinių grupių. Pirmas barjeras atskiria Rytų Aukštaitiją nuo likusių etnolingvistinių grupių. Antras ir trečias barjerai matomi, priklausomai nuo naudotos atstumų matricos: (1) tarp Pietų Aukštaitijos ir Vakarų Aukštaitijos; (2) tarp Šiaurės Žemaitijos ir likusių Žemaitijos grupių (4.9 pav.).



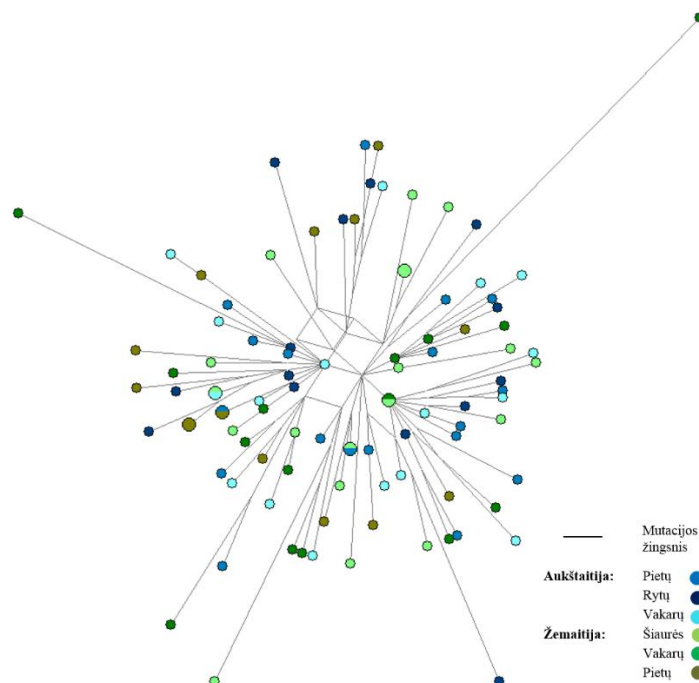
4.9 pav. Y chromosomos haplogrupių įvairovės barjerai tarp tirtų Lietuvos etnolingvistinių grupių. Kairėje šešių Lietuvos etnolingvistinių grupių schema pagal pateiktas koordinatas (2 priedo 1 lentelė) ir matomi barjerai (raudona linija, barjerų eiliškumo tvarka nurodyta skaičiais). Analizei naudota  $F_{ST}$  (viršutinis pav.) ir  $N_e$  (D) (apatinis pav.) atstumai

Sudaryti filogenetiniai medžiai, remiantis Y chromosomos haplotipų (15 TTKS) įvairove, dažniausių Y chromosomos haplogrupių N1c1, R1a1a, R1a1a1g viduje. Haplogrupės N1c1 viduje – artimas žvaigždės struktūrai filogenetinis medis, turintis centrinių haplotipą, kurio dažnis 1,7 %. Nuo centrinio haplotipo vienos mutacijos žingsniu atsišakoja devyni haplotipai, iš kurių penki nustatyti didesniu negu 0,7 % dažniu. Filogenetiniame medyje matomi septyni haplotipai, nutolę 1–2 mutacijomis daugiau nuo centrinio haplotipo nei kiti haplotipai (4.10 pav.). Haplogrupės R1a1a filogenetinis medis neturi centrinio haplotipo, kuris būtų nustatytas tirtoje lietuvių populiacijoje. Matomi penki haplotipai nutolę didesniu mutacijų skaičiumi palyginti su kitais haplotipais (4.11 pav.).

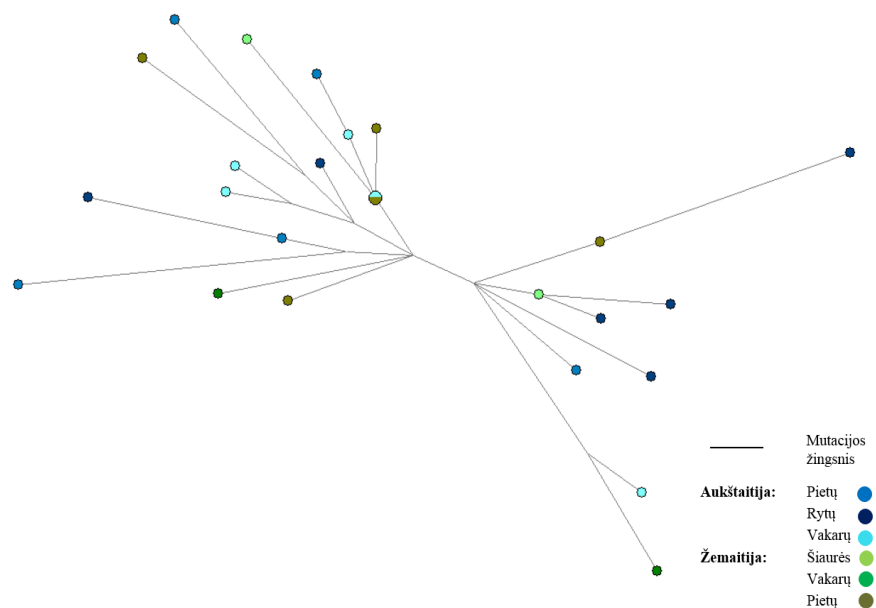




4.10 pav. Filogenetinis medis pagal Y chromosomos 15 TTKS pasiskirstymą haplogrupės N1c1 viduje. Filogenetinis medis sudarytas naudojant sumažintos medianos jungties metodą (angl. *Reduced Median*) (redukcijos laipsnis 1). Kiekvienam TTKS priskirtas svartinis koeficientas, atsižvelgiant į TTKS mutacijos dažnį (Ballantyne *et al.* 2010). (2 priedo 3 lentelė)



4.11 pav. Filogenetinis medis pagal Y chromosomos 15 TTKS pasiskirstymą haplogrupės R1a1a viduje. Filogenetinis medis sudarytas naudojant sumažintos medianos jungties metodą (redukcijos laipsnis 1). Kiekvienam TTKS priskirtas svartinis koeficientas, atsižvelgiant į TTKS mutacijos dažnį (Ballantyne *et al.* 2010)



4.12 pav. Filogenetinis medis pagal Y chromosomos 15 TTKS pasiskirstymą haplogrupės R1a1a1g viduje. Filogenetinis medis sudarytas naudojant sumažintos medianos jungties metodą (redukcijos laipsnis 1). Kiekvienam TTKS priskirtas svartinis koeficientas, atsižvelgiant į TTKS mutacijos dažnį (Ballantyne *et al.* 2010)

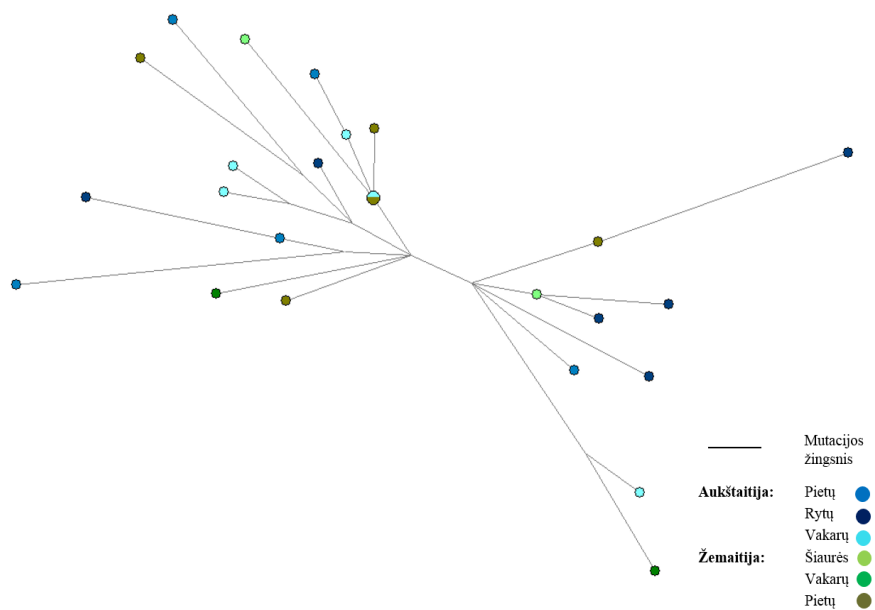
4.4 lentelė. Dažniausių Y chromosomos haplogrupių TMRCA ir populiacijos augimas, remiantis 15 TTKS pasiskirstymu

Mutacijų dažnis – $8,5 \cdot 10^{-3}$ mutacija/žymuo/karta (Ballantyne <i>et al.</i> 2010)			
Haplogrupė	Imtis	TMRCA (metai)	Populiacijos augimo pradžia (metai)
N1c1	119	8 312,3	4 456,1
R1a1a	97	15 512,3	4 580,5
R1a1a1g	25	8 606,6	4 206,7
Mutacijų dažnis – $6,9 \cdot 10^{-4}$ mutacija/žymuo/karta (Zhivotovsky <i>et al.</i> 2004)			
Haplogrupė	Imtis	TMRCA (metai)	Populiacijos augimo pradžia (metai)
N1c1	119	20 382,0	10 531,6
R1a1a	97	49 726,3	11 743,3
R1a1a1g	25	22 666,7	9 799,9
Mutacijų dažnis – $2 \cdot 10^{-3}$ mutacija/žymuo/karta (Luca <i>et al.</i> 2007)			
Haplogrupė	Imtis	TMRCA (metai)	Populiacijos augimo pradžia (metai)
N1c1	119	10 387,0	5 430,1
R1a1a	97	24 638,5	7 027,2
R1a1a1g	25	15 935,1	7 300,8

Remiantis Y chromosomos haplotipų (15 TTKS) įvairove tirtoje lietuvių populiacijoje, dažniausioms Y chromosomos haplogrupėms N1c1, R1a1a, R1a1a1g apskaičiuotas TMRCA. Modelio *a priori* parametrai: (1) pradinės pastovaus dydžio populiacijos su vėlesniu populiacijos augimu modelis; (2) pradinis populiacijos dydis, haplogrupei R1a1a1g 500, haplogrupėms N1c1,

R1a1a – 1 000; (3) populacijos augimo greitis 0–4 % kartai; (4) populacijos augimo pradžia; (5) mutacijos dažnis (2 priedo 4 lentelė) (Ballantyne *et al.* 2010; Kasperavičiūtė *et al.* 2004; Luca *et al.* 2007).

Apskaičiuotas haplogrupių TMRCA ir populacijos augimas labiausiai priklausė nuo *a priori* parametro – mutacijų dažnio. Haplogrupės N1c1 populacijos augimas varijavo ~4 456–10 531 metų ribose, haplogrupės R1a1a populacijos 4 580–11 743 metų ribose, haplogrupės R1a1a1g populacijos 4 207–9 800 metų ribose (

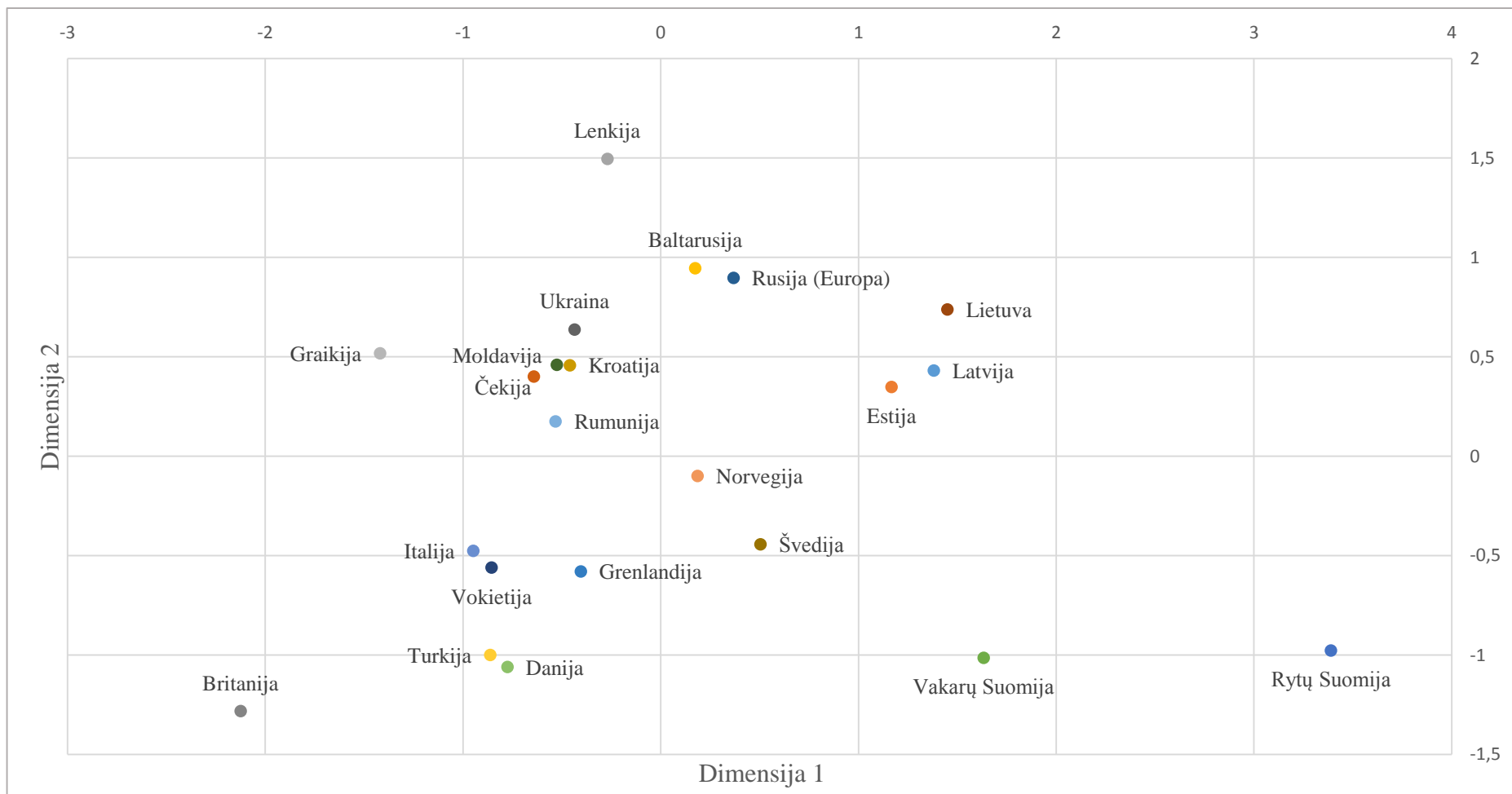


4.12 pav. Filogenetinis medis pagal Y chromosomos 15 TTKS pasiskirstymą haplogrupės R1a1a1g viduje. Filogenetinis medis sudarytas naudojant sumažintos medianos jungties metodą (redukcijos laipsnis 1). Kiekvienam TTKS priskirtas svertinis koeficientas, atsižvelgiant į TTKS mutacijos dažnį (Ballantyne *et al.* 2010)

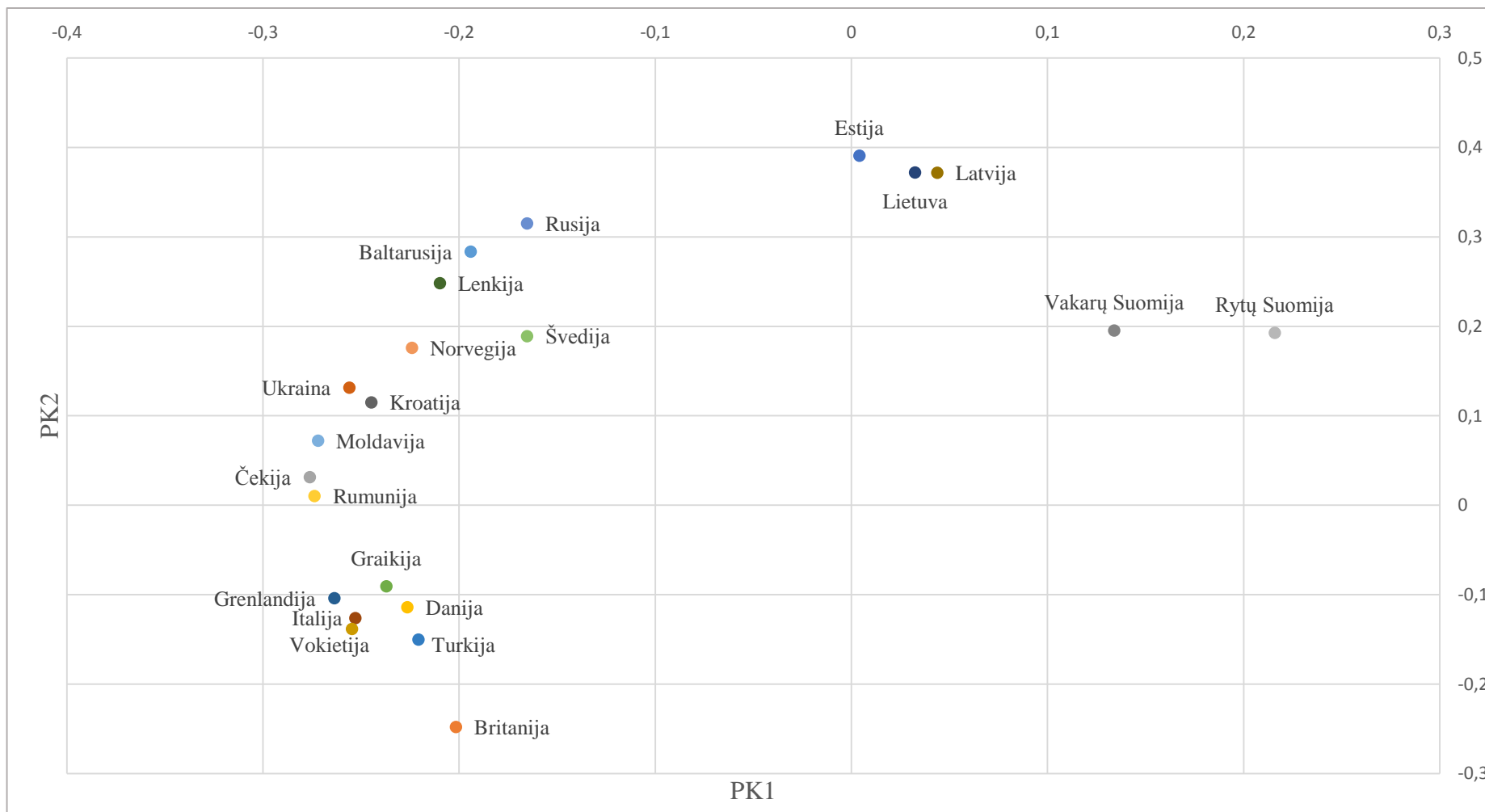
4.4 lentelė).

Taikant MDS, remiantis Nei (D) atstumais (2 priedo 6 lentelė ir 7 lentelė) pagal Y chromosomos haplogrupių (E1, G, I1, J2, N3, R1a, R1b) pasiskirstymą analizuotose Europos populacijose (2 priedo 5 lentelė), matomos penkios grupės: (1) Pabaltijo populacijos (Lietuva, Latvija, Estija); (2) Skandinavijos ir Šiaurės Europos populacijos (Norvegija, Švedija, Vakarų ir Rytų Suomija); (3) Rytų ir Centrinės Europos populacijos (Rusijos Europinė dalis, Baltarusija, Lenkija, Ukraina); (4) Centrinė ir Pietryčių Europos populacijos (Moldavija,

Čekija, Rumunija, Graikija); (5) Vakarų Europos populiacijos (Vokietija, Danija, Britanija, Grenlandija). Šios grupės atitinka populiacijų geografinę padėtį Europoje, išskyrus Italijos ir Turkijos populiacijas (4.13 pav.). Atlikus PKA, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių pasiskirstymą analizuotose Europos populiacijose, PK1 aprašė 55,45 %, PK2 – 29,36 % Y chromosomos haplogrupių įvairovės tarp Europos populiacijų (4.14 pav.). PK1 aprašoma variacija dėl Y chromosomos įvairovės skirtumų tarp Rytų ir Vakarų Suomijos, Pabaltijo (Lietuva, Latvija, Estija) ir likusių analizuotų Europos populiacijų. PK2 aprašo variaciją tarp Pabaltijo, Rytų slavų, Šiaurės, Centrinės ir Balkanų, Pietų ir Vakarų Europos populiacijų grupių.

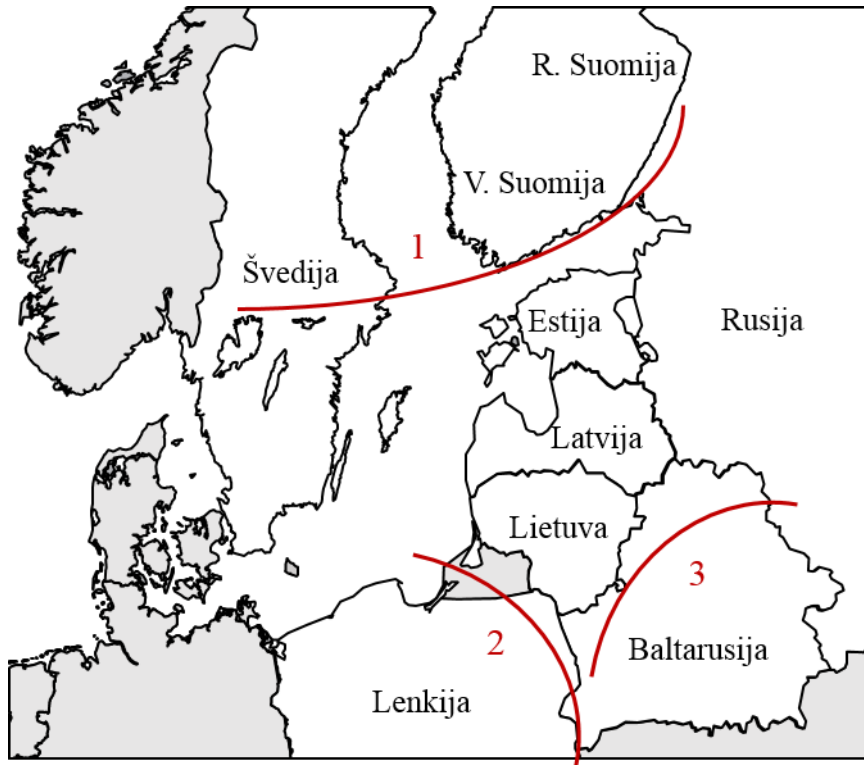


4.13 pav. MDS, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių (E1, G, I1, J2, N3, R1a, R1b) pasiskirstymą Europos populiacijose, vizualizuoti rezultatai. MDS paklaidos funkcija (angl. *Stress*) ir koreliacijos koeficientas ( $r^2$ ): (1) Nei (D) atstumų matrica,  $stress = 0,16853$ ,  $r^2 = 0,88982$



4.14 pav. PKA, remiantis Nei (D) atstumais pagal Y chromosomos haplogrupių (E1, G, I1, J2, N3, R1a, R1b) pasiskirstymą Europos populiacijose. PK1 aprašo 55,45 %, PK2 – 29,36 % Y chromosomos haplogrupių įvairovės tarp Europos populiacijų

Nustatyti Y chromosomos haplogrupių įvairovės barjerai tarp analizuotų kaimyninių Europos populiacijų (Lietuva, Latvija, Estija, Švedija, Rytų ir Vakarų Suomija, Rusija, Baltarusija, Lenkija). Pirmas barjeras atskiria Švedijos, Rytų ir Vakarų Suomijos populiacijas nuo kitų tirtų Europos populiacijų. Antras barjeras matomas tarp Lenkijos ir kitų analizuotų Europos populiacijų, trečias – tarp Baltarusijos ir kitų analizuotų Europos populiacijų (4.15 pav.).



4.15 pav. Y chromosomos haplogrupių įvairovės barjerai tarp analizuotų kaimyninių Europos populiacijų. Barjerai pažymėti raudona linija, eiliškumo tvarka pažymėta skaičiais. Analizei naudota Nei (D) atstumai

## APTARIMAS

D. Kasperavičiūtė kartu su kitais mokslininkais aprašė Lietuvos etnolingvistinių grupių struktūrą, remdamiesi Y chromosomos šešių haplogrupių ir haplotipų, sudarytų iš 9 TTKS žymenų, duomenimis. Taip pat lietuvių populiacija palyginta su kitų Pabaltijo populiacijų (Latvija, Estija) gautais duomenimis (Kasperavičiūtė *et al.* 2004). Kitų mokslininkų grupių atliktų tyrimų, kuriuose įtraukta ir Lietuvos populiacija, pateikti analizės rezultatai,

naudojant tik pagrindinių haplogrupių ar haplotipų (<10 TTKS) duomenis (Kushniarevich *et al.* 2013; Laitinen *et al.* 2002; Lappalainen *et al.* 2008). Dėl mokslininkų grupių kryptingai į tyrimus įtraukiamų mėginių, nenurodant jų tikslaus pasiskirstymo analizuojamoje populiacijoje, gauti rezultatai gali neatspindėti tikros populiacijos genetinės įvairovės.

Darbo metu nustatytas Y chromosomos haplogrupių (16) ir haplotipų (15 TTKS) pasiskirstymas, ištyrus didžiausią, palyginti su ankstesniais tyrimais, imtį iš šešių Lietuvos etnolingvistinių grupių. Y chromosomos haplogrupių (N1c1 ir R1a1a) dažnis artimas iki šio darbo nustatytam, tačiau šių haplogrupių pasiskirstymas skirtingas. Haplogrupės N1c1 dažnis siekia 40,3 % ir yra kiek mažesnis palyginti su kita dažniausia haplogrupu R1a1a (R1a1a1g – 8,5 %) – 41,6 %. Ankstesniais tyrimais nustatytas didesnis šių haplogrupių dažnių skirtumas – atitinkamai 37 % ir 45 % (Kasperavičiūtė *et al.* 2004). Nustatytas haplogrupės R1a1a atšakos R1a1a1g dažnis – 8,5 %, jos židiny, manoma, buvo dabartinėje Pietų ir Centrinėje Lenkijos teritorijoje (Underhill *et al.* 2010).

Tirtoje lietuvių populiacijos grupėje matomas haplogrupių N1c1, R1a1a ir R1a1a1g dažnių gradientas, kuris ankstesniais tyrimais nebuvo nustatytas. Haplogrupės N1c1 didžiausias dažnis yra Šiaurės Lietuvoje (Šiaurės Žemaitija ir Rytų Aukštaitija po ~49 %) ir mažėja pietų kryptimi (Pietų Aukštaitija ~29 %). Haplogrupių R1a1a ir R1a1a1g didžiausias dažnis yra Pietų Lietuvoje (Pietų Aukštaitija, atitinkamai ~47 % ir ~13 %) ir mažėja šiaurės kryptimi (Šiaurės Žemaitija, atitinkamai ~30 % ir ~3 % ir Rytų Aukštaitija, atitinkamai ~25 % ir ~11 %). Susiformavęs gradientas galimas dėl Y chromosomos haplogrupių (N1c1, R1a1a, R1a1a1g) plitimo dabartinėje Lietuvos teritorijoje krypties: (1) haplogrupė N1c1 pasiekė iš Šiaurės Europos, kur nustatomas didžiausias šios haplogrupės dažnis (Lappalainen *et al.* 2008); (2) haplogrupės R1a1a(M198) ir R1a1a1g(M458) pasiekė iš Centrinės ir (ar) Rytų Europos. Haplogrupės R1a1a1g didžiausias dažnis nustatomas dabartinėje Pietų ir Centrinėje Lenkijos teritorijoje. Haplogrupės R1a1a(M198) tikslesnei migracijos krypties nustatymui reikalinga detalesnė šios haplogrupės



diferenciacija. Šios plitimo hipotezės neatskleidžia tėvinės linijos populiacijos formavimosi proceso dabartinėje Lietuvos teritorijoje sudėtingumo.

Apskaičiuoti statistiškai reikšmingi  $F_{ST}$  atstumai ( $p < 0,05$ ) tarp Pietų Lietuvos (Pietų Aukštaitijos) ir Šiaurės Lietuvos (Šiaurės Žemaitijos ir Rytų Aukštaitijos), kuriuos lemia nevienodas dažniausių haplogrupių N1c1, R1a1a ir R1a1a1g pasiskirstymas tarp tirtų šešių Lietuvos etnolingvistinių grupių. Ankstesniais tyrimais statistiškai reikšmingų skirtumų tarp Lietuvos etnolingvistinių grupių nenustatyta. D. Kasperavičiūtė kartu su kolegomis pateikė galimas priežastis: (1) pradinių populiacijų, iš kurių susiformavo dabartinės grupės (Aukštaitija ir Žemaitija) ar etnolingvistinės grupės, genetinis fondas galėjo būti homogeniškas; (2) tyrimui naudota imtis (30–40 asmenų iš kiekvienos etnolingvistinės grupės) galėjo būti per maža, norint nustatyti genetinės įvairovės skirtumus. Šio tyrimo metu, naudojant didesnės imties (iš viso 293 asmenys) duomenis, gauti rezultatai geriau paaiškina antrąją priežastį (Kasperavičiūtė *et al.* 2004).

Atlikta AMOVA parodė, kad didžiausia (~99 %) haplogrupių ir haplotipų įvairovė yra vidinė, t. y. tarp tiriamų asmenų etnolingvistinių grupių viduje. Gauti rezultatai rodo tirtos Lietuvos populiacijos didelį homogeniškumą. Šio darbo rezultatai patvirtina iki šiol gautus AMOVA rezultatus, kurie taip pat parodė didžiausią įvairovę (>99,9 %) tarp tirtų asmenų grupių viduje (Kasperavičiūtė *et al.* 2004). Kalbiniai ar kultūriniai skirtumai tirmoje lietuvių populiacijos grupėje nėra Y chromosomos haplogrupių ir haplotipų įvairovės šaltinis. Nustatyta, kad tirtų Lietuvos etnolingvistinių grupių geografinė padėtis Lietuvos teritorijoje (Šiaurės, Vidurio, Pietų Lietuva) yra galimas Y chromosomos haplogrupių ir haplotipų įvairovės šaltinis ( $p < 0,05$ ).

Vizualizavus gautus rezultatus MDS metodu pagal  $F_{ST}$  ir Nei (D) atstumus, nustatyti trys arba keturi tirtos lietuvių populiacijos grupės klasteriai. Vizualizavus MDS metodu pagal Nei (D) atstumus gautus rezultatus, pirmoje dimensijoje tirtos šešios Lietuvos etnolingvistinės grupės išsidėsto pagal geografinę padėtį šiaurės–pietų ašimi Lietuvos teritorijoje. Antroje dimensijoje – pagal geografinę padėtį vakarų–rytų ašimi Lietuvos teritorijoje.

Vizualizavus MDS metodu pagal  $F_{ST}$  atstumus gautus rezultatus, pirmoje ir antroje dimensijoje tirtos šešios Lietuvos etnolingvistinės grupės suformuoja keturis klasterius, kurie sunkiai paaiškinami kalbiniais, kultūriniais ar geografiniais skirtumais.

PKA, remiantis  $F_{ST}$  ir Nei (D) atstumais, rezultatai rodo įvairovės tarp etnolingvistinių grupių variaciją ir nevertinta vidinės įvairovės (tarp asmenų etnolingvistinių grupių viduje), kuri sudaro didžiąsą dalį (~99 %), variacijos. PK1 ir PK2 kartu aprašo ~79 % esamos haplogrupių įvairovės tarp tirtų šešių Lietuvos etnolingvistinių grupių. Vizualizavus PKA rezultatus matyti, kad PK1 aprašo Y chromosomos įvairovę tarp dviejų klasterių, kuriuose po tris etnolingvistines grupes (šiaurės rytų–pietvakarių ašis): (1) Šiaurės, Pietų Žemaitija ir Rytų Aukštaitija; (2) Vakarų Žemaitija, Pietų, Vakarų Aukštaitija.

Tarp tirtų šešių Lietuvos etnolingvistinių grupių nustatyti haplogrupių įvairovės barjerai atskiria Pietų Aukštaitiją nuo kitų penkių Lietuvos etnolingvistinių grupių ir Šiaurės Žemaitiją nuo kitų Žemaitijos etnolingvistinių grupių. Pastarieji rezultatai patvirtina MDS matomus etnolingvistinių grupių klasterius.

Haplotipų įvairovės sumažėjimas haplogrupės N1c1 viduje (0,9927 ( $\pm 0,0023$ )) matomas ir filogenetinio medžio pagal haplotipų (15 TTKS) pasiskirstymą struktūroje. Haplogrupės N1c1 haplotipų filogenetinis medis pasižymi glaustumu ir keliais žvaigždės struktūrai artimais (angl. *Star-like*) židiniai. Filogenetiniam medžiui būdingas centrinis haplotipas ir aštuoni, iš kurių tik trys pavieniai haplotipai, nutolę vienu mutacijos žingsniu nuo centrinio haplotipo. Žvaigždės struktūros filogenetinis medis matomas esant staigiam populiacijos sumažėjimui (butelio kaklelio efektas) arba mažam pradiniam populiacijos dydžiui (įkūrėjo efektas), po kurių vyksta populiacijos augimas. Galimas N1c1 haplogrupės augimas nustatytas pasibaigus PL, t. y. prieš 4 456–10 531 metų. Pastarasis laikotarpis taip pat apima baltų genčių galimą formavimosi pradžią, t. y. prieš ~5 000 metų, remiantis archeologiniais ir kalbos tyrimais. Haplogrupės R1a1a viduje nustatyta didesnė haplotipų įvairovė (0,9987 ( $\pm 0,0017$ )) negu N1c1 haplogrupėje. Šie rezultatai matomi ir

filogenetinio medžio pagal haplotipų (15 TTKS) pasiskirstymą struktūroje. Haplogrupės R1a1a filogenetiniam medžiui būdinga laisvesnė struktūra, ilgos šakos tarp haplotipų, t. y. haplotipus skiria daugiau mutacijos žingsnių, ir žvaigždės struktūrai artimų židinių yra mažai. Nustatytas haplogrupės R1a1a galimas augimas vykęs taip pat po PL, t.y. prieš 4 580–11 743 metų. Haplogrupės R1a1a augimo metu, atsižvelgiant į sudarytą filogenetinio medžio struktūrą (t. y. daugeliu mutacijos žingsnių nutolę haplotipai), galbūt vyko nevienkartinis genų fondo papildymas. Migracija galėjo vykti iš Centrinės ar Rytų Europos pagal nustatytą haplogrupių R1a1a ir R1a1a1g pasiskirstymo gradientą. Y chromosomos populiacijos genų fondo judėjimo suformuotas Y chromosomos įvairovės pasiskirstymas buvo nuolat veikiamas ir kitų evoliucijos veiksnių. Y chromosomos genetinės įvairovės skirtumai tarp šešių Lietuvos etnolingvistinių grupių yra daugelio evoliucijos procesų – migracijos, mutacijų, gamtinės atrankos, genų dreifo – sąveikos tiriamoje populiacijoje rezultatas. Darbe gauti rezultatai rodo skirtingas dažniausių haplogrupių N1c1 ir R1a1a formavimosi istorijas dabartinėje Lietuvos teritorijoje.

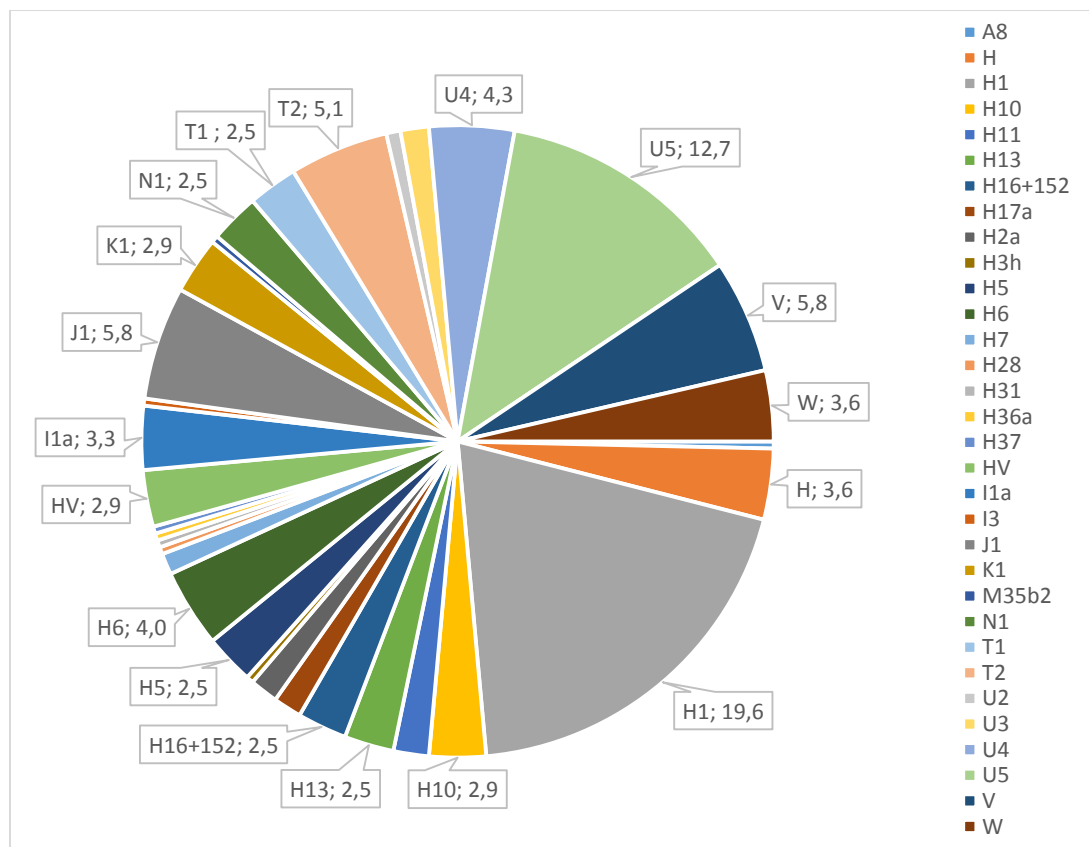
Vizualizavus gautus rezultatus MDS metodu, artimiausios lietuvių populiacijai iš analizuotų Europos populiacijų yra Latvija ir Estija, t. y. populiacijos nutolusios trumpu geografiniu atstumu šiaurės kryptimi. Kiek toliau išsidėsto slavų populiacijos (Rusija, Baltarusija, Lenkija). Gauti rezultatai patvirtina ankstesnių tyrimų duomenis (Kasperavičiūtė *et al.* 2004; Kushniarevich *et al.* 2013; Lappalainen *et al.* 2008).

PK1 (55,45 %) aprašo didžiąsą dalį Europos Y chromosomos įvairovės tarp Rytų, Vakarų Suomijos, Pabaltijo (Lietuva, Latvija, Estija) ir kitų analizuotų Europos populiacijų. PK2 (29,36 %) aprašo variaciją tarp Pabaltijo (Lietuva, Latvija, Estija), Rytų slavų (Rusija, Baltarusija, Lenkija), Šiaurės Europos (Norvegija, Švedija, Rytų ir Vakarų Suomija), Centrinės Europos ir Balkanų (Ukraina, Kroatija, Moldavija, Čekija, Rumunija), Pietų ir Vakarų Europos (Turkija, Graikija, Italija, Danija, Vokietija, Britanija, Grenlandija) populiacijų grupių.

Analizuojant kaimynines populiacijas, Y chromosomos haplogrupių įvairovės barjerai nustatyti tarp Skandinavijos kartu su Rytų ir Vakarų Suomijos populiacijomis, Lenkijos ir Baltarusijos populiacijų ir likusių analizuotų Europos populiacijų. Gauti rezultatai leidžia manyti, kad kalbos įtaka Y chromosomos genetinei įvairovei yra mažesnė už geografinius atstumus. Lenkijos populiacijos atskirtis galima dėl didžiausio haplogrupės R1a, kurios atšakos R1a1a1g formavimosi židinyje ir didžiausias dažnis yra Centrinėje ir Pietų Lenkijos teritorijoje, dažnio; taip pat dėl didesnės Vakarų Europos genetinio fondo įtakos (R1b). Nustatomas Y chromosomos haplogrupių įvairovės barjeras tarp Baltarusijos ir kitų analizuotų Europos populiacijų susidarė dėl Pietų Europos genų fondo įtakos (J2, E1), o tarp Švedijos, Rytų ir Vakarų Suomijos ir kitų analizuotų populiacijų – dėl atitinkamai didesnio haplogrupės I1 dažnio ir I1 bei didžiausio N1c1 dažnio. Įtaką gautiems Y chromosomos variacijoms tarp Europos populiacijų analizės rezultatams gali turėti straipsniuose pateikiamas skirtingo tikslumo Y chromosomos haplogrupių įvairovės pasiskirstymas skirtingose populiacijose.

#### 4.2. Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis mtDNR haplogrupių pasiskirstymu

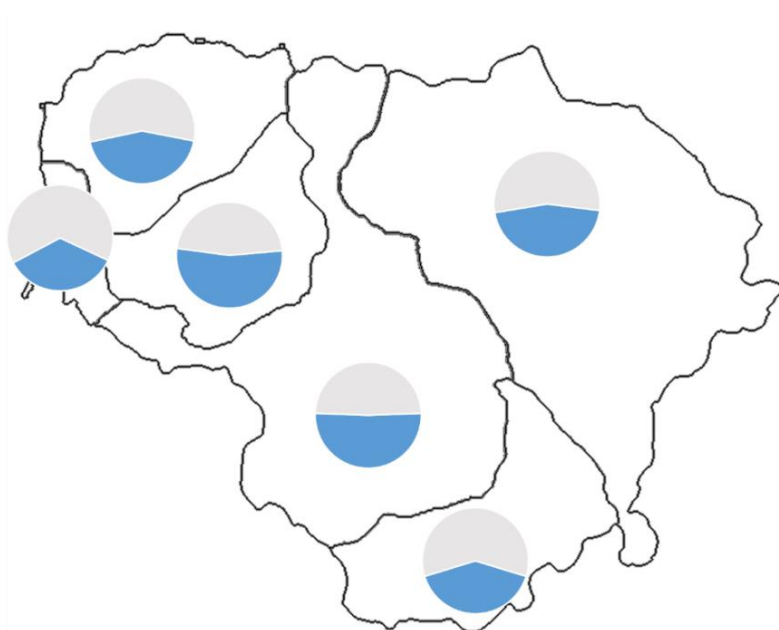
Tirtoje lietuvių populiacijos grupėje (3.4 pav.) nustatyta visos mtDNR seka su vidutiniu 352,6 ( $\pm 203,66$ ) kartų padengimu. Mažiausias padengimas – 23 kartai, didžiausias – 1 753 kartai. Tirtoje lietuvių populiacijos grupėje nustatyta 116 skirtingų haplogrupių, kurios hierarchiniu principu sujungtos į stambesnes 23 grupes arba smulkesnes 69 grupes. Nustatytos haplogrupės priklauso 11 skirtingų filogenetinių linijų (A, H, HV, I, J, K, M, N, T, U, V, W), iš kurių septynios (H, I, J, K, T, U, V) yra dažniausios Europoje.



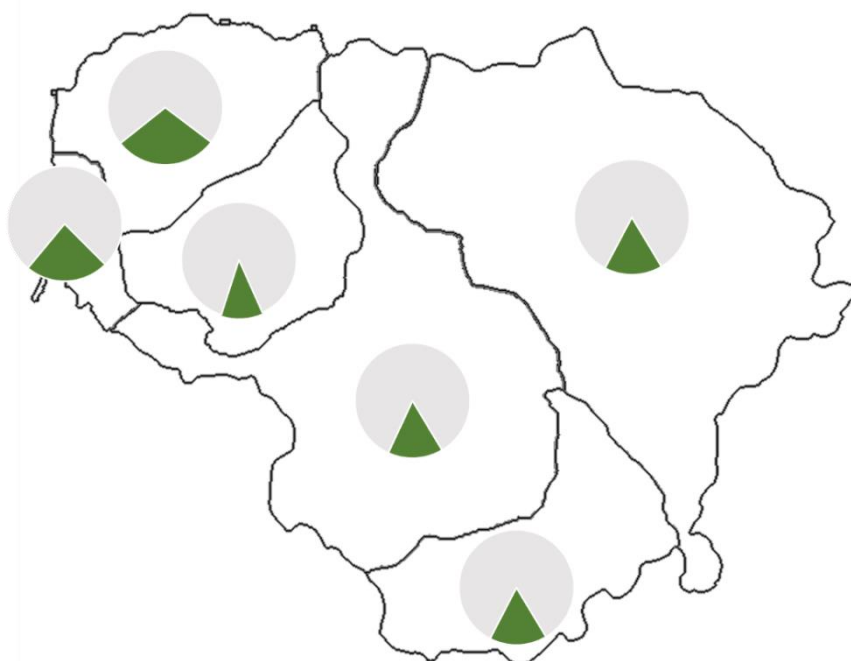
4.16 pav. mtDNR haplogrupių pasiskirstymas (dažnis, %) tirtoje lietuvių populiacijos grupėje

Dažniausios filogenetinės linijos tirtoje lietuvių populiacijos grupėje: (1) filogenetinė linija H sudaro 45,3 % (H – 3,6 %; H1 – 19,6 %, H5 – 2,5 %, H6 – 4,0 %, H10 – 2,9 %, H13 – 2,5 %, H16+152 – 2,5 %); (2) filogenetinė linija U sudaro 19,2 % (U4 – 4,3 %; U5 – 12,7 %); (3) filogenetinė linija T sudaro 7,6 % (T1 – 2,5 %; T2 – 5,1 %); (4) filogenetinė linija J1 sudaro 5,8 %; (5) filogenetinė linija V sudaro 5,8 %; (6) filogenetinės linijos A, HV, I, K, M,

N, W sudaro <5 % tirtos lietuvių populiacijos. Penkios dažniausios filogenetinės linijos (H, J1, T, U, V) sudaro ~84 %, likusios šešios filogenetinės linijos sudaro ~16 % tirtos lietuvių populiacijos (4.16 pav.).



4.17 pav. mtDNR filogenetinės linijos H pasiskirstymas (mėlyna spalva) tarp tirtų šešių Lietuvos etnolingvistinių grupių



4.18 pav. mtDNR filogenetinės linijos U pasiskirstymas (žalia spalva) tarp tirtų šešių Lietuvos etnolingvistinių grupių

Nustatytų dažniausių filogenetinių linijų pasiskirstymas neturi aiškios geografinės, kalbinės, kultūrinės priežasties. Filogenetinės linijos H mažiausias

dažnis nustatomas Vakarų Žemaitijoje (~35 %). Šiek tiek didesniu dažniu filogenetinė linija H nustatyta Šiaurės Rytų Lietuvoje – Šiaurės Žemaitijoje (~43 %) ir Rytų Aukštaitijoje (~45 %). Vidurio Lietuvoje – Pietų Žemaitijoje (~53 %) ir Vakarų Aukštaitijoje (~51 %) stebimas didžiausias dažnis, jis šiek tiek sumažėja Pietų Lietuvoje, t. y. Pietų Aukštaitijoje (~40 %) (4.17 pav.). Filogenetinės linijos U didžiausias dažnis nustatomas Šiaurės Vakarų Lietuvoje: Šiaurės Žemaitijoje (~29 %) ir Vakarų Žemaitijoje (~23 %). Daug mažesniu, bet tolygiu, dažniu filogenetinė linija U nustatyta likusiose keturiose Lietuvos etnolingvistinėse grupėse – Rytų Aukštaitijoje (~16 %), Pietų Aukštaitijoje (~16 %), Vakarų Aukštaitijoje (~15 %), ir mažiausiu dažniu – Pietų Žemaitijoje (~11 %) (4.18 pav.).

Matoma tolygi mtDNR haplotipų įvairovė tarp šešių Lietuvos etnolingvistinių grupių ( $0,999 (\pm 0,0004)$ ), t. y. >99 % tikimybė, kad atsitiktinai parinkti du haplotipai (mtDNR) skirsis. Nenustatyta haplotipų įvairovės skirtumų tarp Aukštaitijos ir Žemaitijos. Rasta 730 segreguojančių mtDNR pozicijų tirtoje lietuvių populiacijos grupėje. Aukštaitijoje ir Žemaitijoje atitinkamai – 519 ir 494 pozicijų. Tirtoje lietuvių populiacijos grupėje vidutinis porinis haplotipų skirtumas  $27,21 (\pm 12,0)$ , t. y. atsitiktinai parinkti du haplotipai (mtDNR) vidutiniškai skiriasi  $\sim 27 (\pm 12)$  nukleotidais. Tirtoje lietuvių populiacijos grupėje vidutinis nukleotidų įvairovės įvertis  $1,6 \cdot 10^{-3} (\pm 0,80 \cdot 10^{-3})$ , t. y. tikimybė, kad atsitiktinai parinktų dviejų haplotipų (mtDNR) konkrečios pozicijos aleliai bus skirtingi. Pietų Žemaitijoje palyginti su kitomis etnolingvistinėmis grupėmis, vidutinis porinių skirtumų skaičius ir nukleotidų įvairovė mažesni – atitinkamai  $22,55 (\pm 10,1)$  ir  $1,363 \cdot 10^{-3} (\pm 0,68 \cdot 10^{-3})$ . Vakarų Žemaitijoje matomas didesnis nei kitose etnolingvistinėse grupėse, vidutinis porinių skirtumų skaičius ir nukleotidų įvairovė – atitinkamai  $31,69 (\pm 14,2)$  ir  $1,915 \cdot 10^{-3} (\pm 0,95 \cdot 10^{-3})$ .

4.5 lentelė. mtDNR\* ir HVRI regiono įvairovės įverčiai tirtoje lietuvių populiacijoje

<b>mtDNR (16 548 np*)</b>							
<b>Etnolingvistinė grupė</b>	<b>N<sup>1</sup></b>	<b>HD (±SN)<sup>2</sup></b>	<b>S<sup>3</sup></b>	<b>π (±SN)<sup>4</sup> n·10<sup>-3</sup></b>	<b>k (±SN)<sup>5</sup></b>	<b>H<sup>6</sup></b>	<b>Tajima D<sup>7</sup></b>
Rytų Aukštaitija	55	0,998 (±0,004)	281	1,645 (±0,81)	27,2 (±12,1)	52	-2,00 (p = 0,005)
Pietų Aukštaitija	37	0,999 (±0,007)	268	1,706 (±0,85)	28,24 (±12,6)	36	-2,11 (p = 0,003)
Vakarų Aukštaitija	45	0,999 (±0,005)	291	1,677 (±0,83)	27,7 (±12,4)	44	-2,14 (p = 0,003)
Aukštaitija	137	0,999 (±0,001)	519	1,668 (±0,81)	27,61 (±12,2)	129	-2,34 (p < 0,001)
Šiaurės Žemaitija	62	0,998 (±0,003)	330	1,624 (±0,80)	26,87 (±11,9)	59	-2,19 (p = 0,002)
Pietų Žemaitija	43	0,998 (±0,006)	246	1,363 (±0,68)	22,55 (±10,1)	41	-2,23 (p < 0,001)
Vakarų Žemaitija	34	0,996 (±0,008)	279	1,915 (±0,95)	31,69 (±14,2)	32	-2,04 (p = 0,003)
Žemaitija	139	0,999 (±0,001)	494	1,619 (±0,79)	26,79 (±11,8)	128	-2,31 (p < 0,001)
Lietuva	276	0,999 (±0,0004)	730	1,644 (±0,80)	27,21 (±12,0)	252	-2,42 (p < 0,001)
<b>HVRI regionas (16 001–16 568 np)</b>							
<b>Etnolingvistinė grupė</b>	<b>N<sup>1</sup></b>	<b>HD (±SN)<sup>2</sup></b>	<b>S<sup>3</sup></b>	<b>π (±SN)<sup>4</sup> n·10<sup>-3</sup></b>	<b>k (±SN)<sup>5</sup></b>	<b>H<sup>6</sup></b>	<b>Tajima D<sup>7</sup></b>
Rytų Aukštaitija	55	0,983 (±0,009)	53	10,47 (±5,6)	5,92 (±2,87)	42	-1,67 (p = 0,0016)
Pietų Aukštaitija	37	0,970 (±0,018)	46	8,92 (±4,93)	5,05 (±2,51)	27	-1,94 (p = 0,01)
Vakarų Aukštaitija	45	0,982 (±0,013)	52	9,52 (±5,19)	5,39 (±2,65)	37	-1,92 (p = 0,008)
Aukštaitija	137	0,981 (±0,007)	84	9,76 (±5,22)	5,52 (±2,67)	94	-2,03 (p = 0,004)
Šiaurės Žemaitija	62	0,986 (±0,007)	58	10,51 (±5,63)	5,95 (±2,88)	47	-1,76 (p = 0,009)
Pietų Žemaitija	43	0,971 (±0,014)	45	8,03 (±4,47)	4,54 (±2,28)	29	-1,97 (p = 0,009)
Vakarų Žemaitija	34	0,991 (±0,009)	51	10,9 (±5,91)	6,17 (±3,01)	29	-1,85 (p = 0,014)
Žemaitija	139	0,985 (±0,004)	76	9,85 (±5,27)	5,57 (±2,69)	84	-1,88 (p = 0,015)
Lietuva	276	0,984 (±0,004)	101	9,81 (±5,23)	5,55 (±2,68)	151	-1,99 (p = 0,001)

\*Pašalinus visas pozicijas, kuriose bent vienam tirtam asmeniui nustatytas tarpas

<sup>1</sup>Imties dydis

<sup>2</sup>Haplotipų įvairovė ir standartinis nuokrypis (haplotipas – visa mtDNR seka)

<sup>3</sup>Segreguojančių pozicijų skaičius

<sup>4</sup>Nukleotidų įvairovė ir standartinis nuokrypis ( $n \cdot 10^{-3}$ )

<sup>5</sup>Vidutinis porinis skirtumų skaičius ir standartinis nuokrypis

<sup>6</sup>Haplotipų skaičius

<sup>7</sup>Tajima D įvertis ir p reikšmė



Lietuvių populiacijos grupei, pagrindinėms grupėms (Aukštaitija ir Žemaitija) ir šešioms Lietuvos etnolingvistinėms grupėms nustatytas statistiškai reikšmingas ( $p < 0,05$ ) Tajima D įvertis, kuris visoms minėtoms grupėms yra didesnis nei  $-2,0$ . Neigiamas Tajima D įvertis nustatomas, esant retų alelių pertekliui, kuris gali būti dėl kryptingo gamtinės atrankos veikimo arba populiacijos augimo. mtDNR reguliacinės srities HVRI regiono (16 001–16 568 np) nukleotidų įvairovė ( $\pi$ ) yra didesnė ( $\sim 8$  karus), palyginti su koduojančia sritimi (577–16 023 np) – atitinkamai  $9,81 \cdot 10^{-3}$  ( $\pm 5,23 \cdot 10^{-3}$ ) ir  $1,206 \cdot 10^{-3}$  ( $\pm 0,593 \cdot 10^{-3}$ ) (4.5 lentelė).

Reguliacinėje srityje ir *MT-ATP6* gene nustatytų pakaitų skaičius statistiškai reikšmingai didesnis, *MT-RNR1* (12S), *MT-RNR2* (16S), tRNR, *MT-CO1*, *MT-ND4L* genuose – mažesnis ( $\chi^2$  kriterijus,  $p < 0,05$ ) nei pagal šių sričių ilgį tikėtinas pakaitų skaičius. Reguliacinėje srityje nustatytas pakaitų skaičius  $\sim 3$  kartus didesnis nei tikėtinas, atsižvelgiant į reguliacinės srities ilgį. Visose mtDNR srityse tranzicijų skaičius yra didesnis negu transversijų skaičius (santykis  $\sim 16:1$ ). Reguliacinėje srityje ir *MT-ATP6* gene nustatytų tranzicijų skaičius statistiškai reikšmingai didesnis, *MT-RNR1* (12S), *MT-RNR2* (16S), tRNR, *MT-CO1*, *MT-ND4L* genuose – mažesnis ( $\chi^2$  kriterijus,  $p < 0,05$ ) nei pagal šių sričių ilgį tikėtinas tranzicijų skaičius. Reguliacinėje srityje nustatytų transversijų skaičius statistiškai reikšmingai didesnis, *MT-RNR2* (16S) gene – mažesnis ( $\chi^2$  kriterijus,  $p < 0,05$ ) nei pagal šių sričių ilgį tikėtinas transversijų skaičius. Nustatyta daugiau sinoniminių pakaitų, palyginti su nesinoniminėmis (santykis  $\sim 2:1$ ). Santykis tarp nesinoniminių pakaitų dalies nesinoniminėse pozicijose ir sinoniminių pakaitų dalies sinoniminėse pozicijose ( $p_N/p_S$ ) statistiškai reikšmingai varijavo tarp genų ( $\chi^2$  kriterijus,  $p = 0,0035$ ). Variaciją lėmė *MT-ATP6* ir *MT-ND6* genai ( $\chi^2$  kriterijus,  $p > 0,05$ ) (4.6 lentelė).

4.6 lentelė. mtDNR sričių segreguojančių pozicijų, tranzicijų, transversijų, sinoniminių, nesinoniminių pakaitų skaičius

Sritis	S <sup>1</sup>	TS <sup>2</sup>	TV <sup>3</sup>	Sinoniminės pakaitos	Nesinoniminės pakaitos	p <sub>N</sub> /p <sub>S</sub> <sup>4</sup>
Reguliacinė s. <sup>5</sup>	141	132	9	–	–	–
<i>MT-RNR1 (12S)</i>	24	22	2	–	–	–
<i>MT-RNR2 (16S)</i>	36	36	0	–	–	–
tRNR	37	35	2	–	–	–
<i>MT-ATP6</i>	42	40	2	15	27	0,640
<i>MT-ATP8</i>	11	11	0	6	5	0,269
<i>MT-COI</i>	43	40	3	36	7	0,065
<i>MT-CO2</i>	28	25	3	19	9	0,157
<i>MT-CO3</i>	35	35	0	22	13	0,197
<i>MT-CYTB</i>	48	45	3	29	19	0,219
<i>MT-ND1</i>	39	36	3	28	11	0,143
<i>MT-ND2</i>	48	44	4	34	14	0,140
<i>MT-ND3</i>	15	15	0	11	4	0,118
<i>MT-ND4L</i>	5	5	0	3	2	0,255
<i>MT-ND4</i>	52	48	4	45	8	0,063
<i>MT-ND5</i>	91	87	4	66	25	0,128
<i>MT-ND6</i>	18	15	3	16	2	0,042
Iš viso	713	671	42	330	146	–

<sup>1</sup>Segreguojančių pozicijų skaičius

<sup>2</sup>Tranzicijų skaičius

<sup>3</sup>Transversijų skaičius

<sup>4</sup>Santykis tarp nesinoniminių pakaitų dalies nesinoniminėse pozicijose ir sinoniminių pakaitų dalies sinoniminėse pozicijose

<sup>5</sup>Reguliacinė sritis: HVRI regionas (16 001–16 568 np) ir HVRII regionas (1–576 np)

Apskaičiuoti šešių Lietuvos etnolingvistinių grupių poriniai skirtumai pagal haplotipus (577–16 023 np) (2 priedo 9 lentelė) ir  $F_{ST}$ , Nei (D) (2 priedo 8 lentelė) atstumai pagal haplogrupes: (1) haplogrupių hierarchinis grupavimas į 23 grupes, kurių kiekvienos dažnis populiacijoje yra >1 %; (2) haplogrupių hierarchinis grupavimas į 69 grupes.

Nustatyti statistiškai reikšmingi (69 grupės,  $p < 0,05$ ; 10 000 kėlinių)  $F_{ST}$  atstumai tarp Pietų Žemaitijos ir Rytų Aukštaitijos, Pietų ir Šiaurės Žemaitijos. Statistiškai reikšmingas (23 grupės,  $p < 0,05$ ; 10 000 kėlinių)  $F_{ST}$  atstumas išlieka tarp Pietų ir Šiaurės Žemaitijos (4.7 lentelė). Nustatyta statistiškai nereikšminga koreliacija tarp Nei (D) ir geografinių atstumų (Mantelio kriterijus,  $r = 0,242$ ;  $p = 0,238$ ; 10 000 kėlinių).

4.7 lentelė.  $F_{ST}$  atstumų matrica. Virš įstrižainės  $F_{ST}$  atstumai (23 grupės). Žemiau įstrižainės –  $F_{ST}$  atstumai (69 grupės). Paryškinti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) įverčiai. Kursyvu pažymėtas įvertis, kurio  $p = 0,0693$

	PA	RA	VA	PŽ	VŽ	ŠŽ
PA	0	0,00848	-0,00698	-0,00104	-0,01049	0,00991
RA	0,00449	0	-0,0019	<i>0,01215</i>	-0,01344	-0,00294
VA	0,00242	-0,00038	0	-0,00069	-0,01035	0,00714
PŽ	0,00329	<b>0,00974</b>	0,00392	0	0,0011	<b>0,02343</b>
VŽ	0,00259	-0,00343	-0,00214	0,00392	0	-0,00985
ŠŽ	0,00388	-0,00488	0,00276	<b>0,00964</b>	-0,0067	0

AMOVA analizė, remiantis  $F_{ST}$  atstumais pagal haplogrupių pasiskirstymą, atlikta jungiant tirtas šešias Lietuvos populiacijos etnolingvistines grupes pagal geografinę padėtį ar kalbinius, kultūrinius skirtumus. Didžiausia haplogrupių įvairovė, siekianti ~99 %, yra vidinė, t. y. tarp tiriamų asmenų etnolingvistinių grupių viduje. Nustatyti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) haplogrupių įvairovės įverčiai tarp grupių, sudarytų pagal geografinę padėtį (vakarų–rytų ašis): (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų Aukštaitija; (4) Rytų Aukštaitija. Atlikta AMOVA analizė, remiantis poriniais skirtumais tarp mtDNR koduojančios dalies (577–16 023 np), kuriai būdingas mažesnis mutacijų dažnis ir sukaupta genetinė variacija ilgesniu laikotarpiu lyginant su reguliacine dalimi, haplotipų. Taip pat atlikta AMOVA analizė, remiantis poriniais skirtumais tarp visos mtDNR (16 569 np) haplotipų. Didžiausia haplotipų įvairovė, siekianti ~99 %, matoma tarp tiriamų asmenų etnolingvistinių grupių viduje. Nustatyti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) haplotipų (577–16 023 np) įvairovės įverčiai tarp grupių, sudarytų pagal geografinę padėtį (vakarų–rytų ašis): (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų, Rytų Aukštaitija. mtDNR haplotipų (16 569 np) įvairovės įverčiai statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) jungiant į tris grupes pagal geografinę padėtį (vakarų–rytų ašis): (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų, Rytų Aukštaitija (4.8 lentelė).

4.8 lentelė. AMOVA analizės, pagal mtDNR haplogrupių ir haplotipų pasiskirstymą tirtoje Lietuvos populiacijoje, rezultatai. Tirtos šešios Lietuvos etnolingvistinės grupės sujungtos pagal geografinį išsidėstymą: (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų Aukštaitija ir (ar) (4) Rytų Aukštaitija

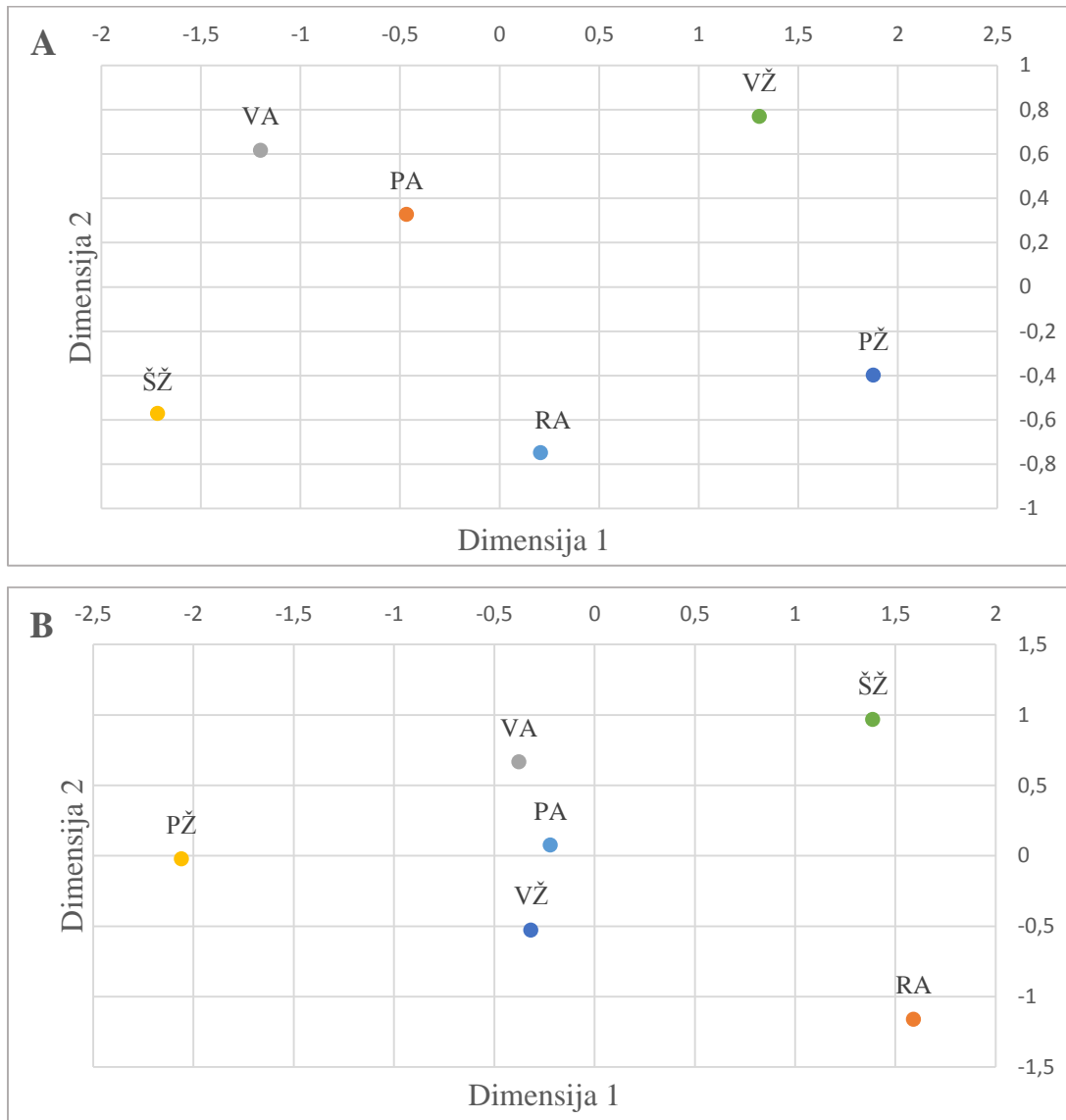
Haplogrupės (23 grupės) <sup>1</sup>			
Genetinės įvairovės šaltinis	df	Fiksacijos indeksas (F)	Įvairovė (%)
tarp grupių (4)	3	<b>0,01127</b> (p = 0,042)	1,13
tarp populiacijų grupių viduje	2	-0,00845 <sup>3</sup>	-0,84
tarp tiriamųjų populiacijų viduje	270	0,00291 <sup>3</sup>	99,71
Haplogrupės (69 grupės) <sup>1</sup>			
Genetinės įvairovės šaltinis	df	Fiksacijos indeksas (F)	Įvairovė (%)
tarp grupių (4)	3	0,0047 (p = 0,071)	0,47
tarp populiacijų grupių viduje	2	-0,00225 <sup>3</sup>	-0,22
tarp tiriamųjų populiacijų viduje	270	0,00246 <sup>3</sup>	99,75
Haplotipai (577–16 023 np) <sup>2</sup>			
Genetinės įvairovės šaltinis	df	Fiksacijos indeksas (F)	Įvairovė (%)
tarp grupių (3)	3	<b>0,00555</b> (p = 0,048)	0,55
tarp populiacijų grupių viduje	2	-0,00299 <sup>3</sup>	-0,30
tarp tiriamųjų populiacijų viduje	270	0,00257 <sup>3</sup>	99,74
Haplotipai (16 569 np) <sup>2</sup>			
Genetinės įvairovės šaltinis	df	Fiksacijos indeksas (F)	Įvairovė (%)
tarp grupių (3)	3	<b>0,00480</b> (p = 0,031)	0,48
tarp populiacijų grupių viduje	2	-0,00225 <sup>3</sup>	-0,22
tarp tiriamųjų populiacijų viduje	270	0,00257 <sup>3</sup>	99,74

<sup>1</sup>F<sub>ST</sub> atstumai pagal haplogrupių pasiskirstymą

<sup>2</sup>Poriniai skirtumai pagal haplotipų (16 569 np; 577–16 023 np) pasiskirstymą

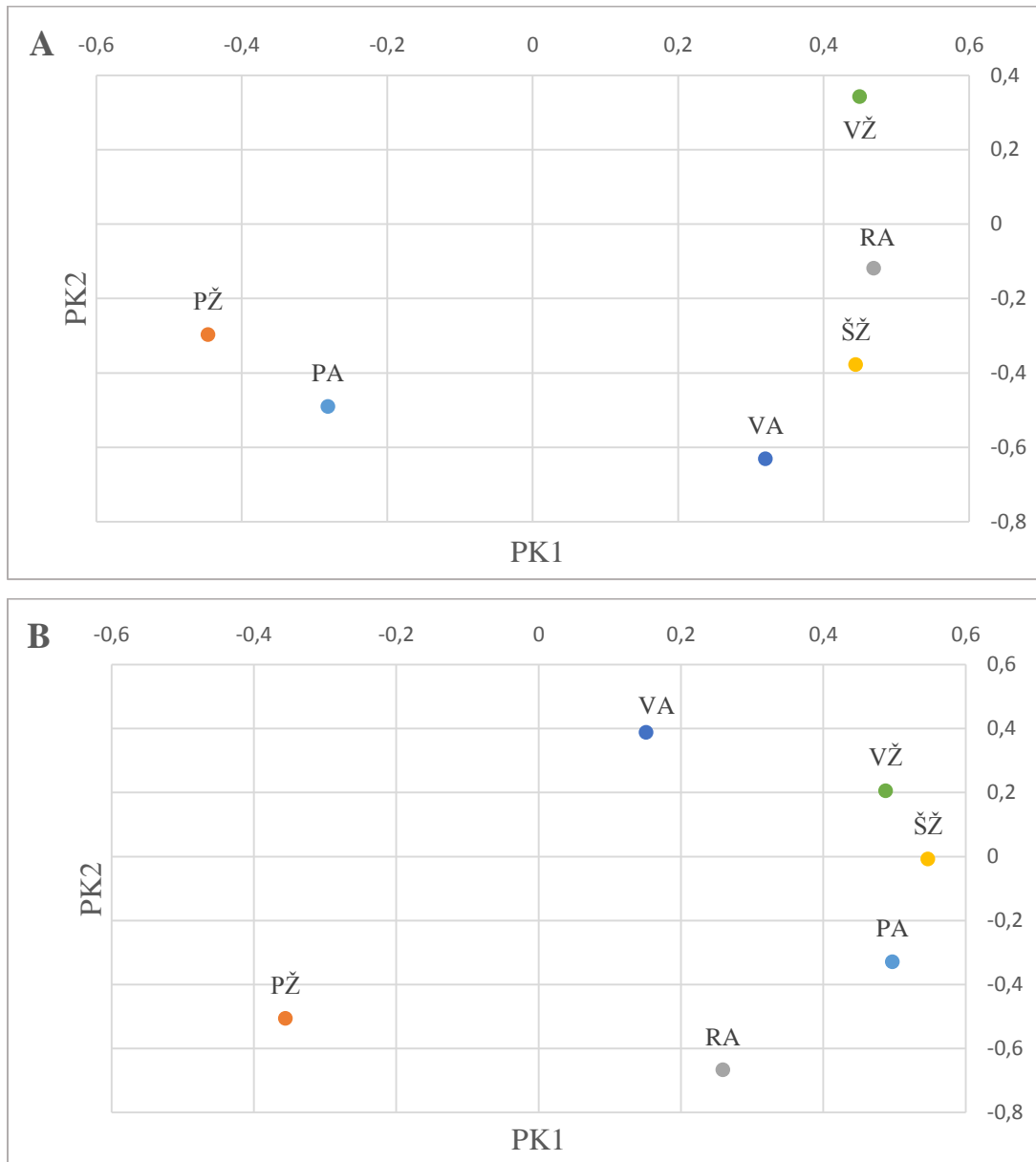
<sup>3</sup>Statistiškai nereikšmingi įverčiai (p > 0,05)

Vizualizavus gautus rezultatus MDS metodu, remiantis F<sub>ST</sub> atstumais pagal mtDNR haplogrupių (23 grupės) pasiskirstymą, pirmoje ir antroje dimensijoje nėra aiškių klasterių, t. y. tirtos šešios Lietuvos etnolingvistinės grupės yra išsidėčiusios atskirai. Vizualizavus MDS metodu, remiantis F<sub>ST</sub> atstumais pagal mtDNR haplogrupių (69 grupės) pasiskirstymą, gautus rezultatus, pirmoje ir antroje dimensijoje matomos keturios grupės: (1) Pietų Žemaitija; (2) Vakarų, Pietų Aukštaitija, Vakarų Žemaitija; (3) Šiaurės Žemaitija; (4) Rytų Aukštaitija. Pietų Žemaitija nutolusi didžiausiu atstumu nuo Šiaurės Žemaitijos ir Rytų Aukštaitijos, tarp kurių nustatyti statistiškai reikšmingi F<sub>ST</sub> atstumai (4.19 pav.).



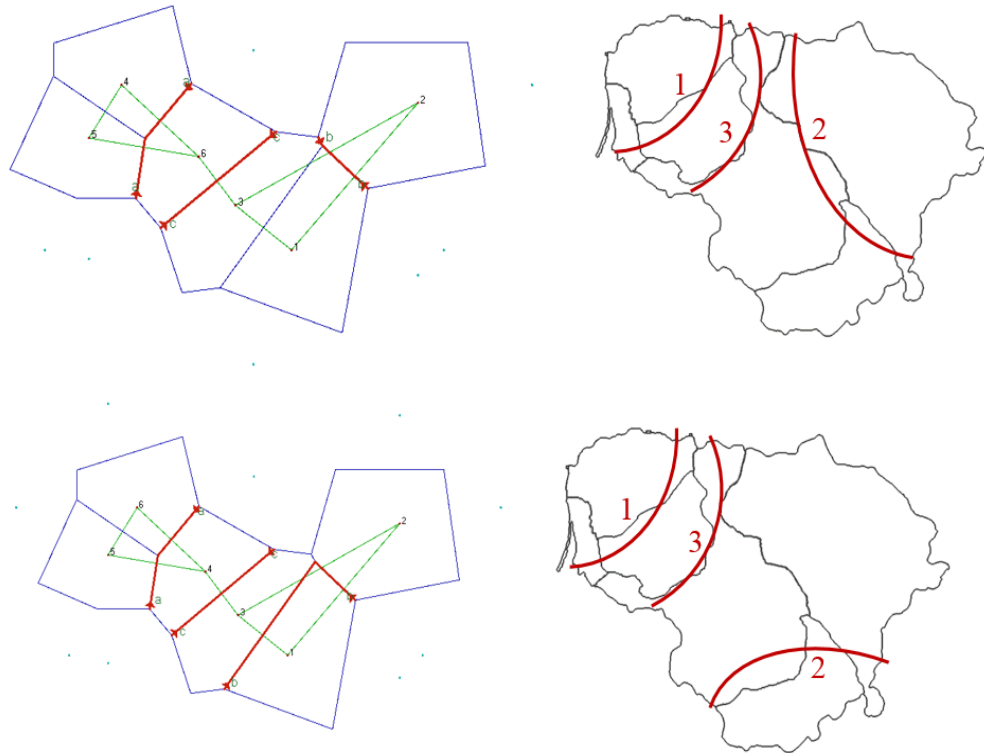
4.19 pav. MDS, remiantis  $F_{ST}$  atstumais pagal mtDNR haplogrupių (grupavimas į (A) 23 ir (B) 69 grupės) pasiskirstymą, vizualizuoti rezultatai. MDS paklaidos funkcija (angl. *Stress*) ir koreliacijos koeficientas ( $r^2$ ): (1)  $F_{ST}$  atstumų matrica (23 grupės),  $stress = 0,00286$ ,  $r^2 = 0,99994$ ; (2)  $F_{ST}$  atstumų matrica (69 grupės),  $stress = 0,00281$ ,  $r^2 = 0,99995$ .

PKA, remiantis  $F_{ST}$  atstumais pagal mtDNR haplogrupių (69 grupės) pasiskirstymą, rezultatai rodo, kad PK1 aprašo 63,63 %, PK2 – 21,49 % variacijos tarp tirtų šešių Lietuvos etnolingvistinių grupių. PKA, remiantis mtDNR koduojančios dalies (577–16 023 np) haplotipų įvairove, rezultatai rodo, kad PK1 aprašo 50,0 %, PK2 – 27,96 % variacijos tarp tirtų šešių Lietuvos etnolingvistinių grupių. Vizualizavus PKA rezultatus, matyti, kad PK1 aprašo mtDNR haplotipų įvairovę tarp Pietų Žemaitijos ir kitų penkių etnolingvistinių grupių (4.20 pav.).



4.20 pav. PKA, remiantis (A)  $F_{ST}$  atstumais pagal mtDNR haplogrupių pasiskirstymą ir (B) poriniais skirtumais tarp haplotipų (577–16 023 np), rezultatai

Nustatyti mtDNR haplogrupių įvairovės barjerai tarp tirtų šešių Lietuvos etnolingvistinių grupių. Pirmas barjeras atskiria Vakarų ir Šiaurės Žemaitiją nuo kitų etnolingvistinių grupių. Antras barjeras yra tarp Rytų Aukštaitijos ir kitų etnolingvistinių grupių, trečias barjeras tarp Žemaitijos ir Aukštaitijos etnolingvistinių grupių (4.21 pav.).

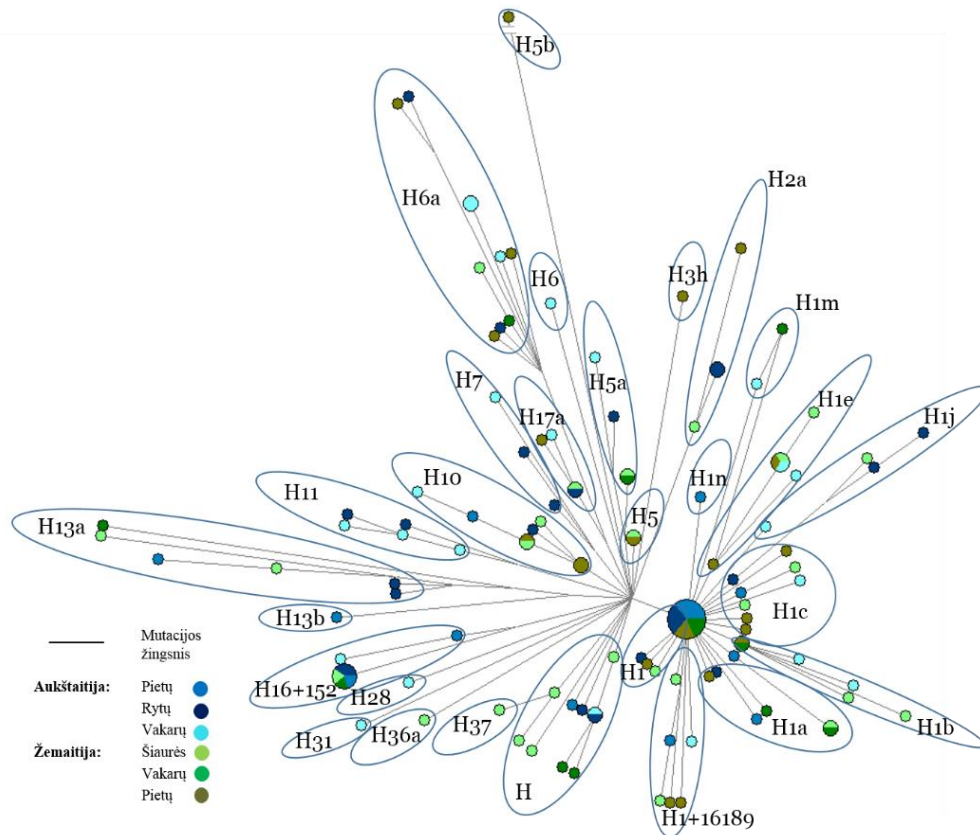


4.21 pav. mtDNR haplotipų (577–16 023 np) (viršuje) ir mtDNR haplogrupių (69 grupės) (apačioje) įvairovės barjerai tarp tirtų Lietuvos etnolingvistinių grupių. Kairėje šešių Lietuvos etnolingvistinių grupių schema pagal pateiktas koordinatas (2 priedo 1 lentelė) ir matomi barjerai (raudona linija, eiliškumo tvarka nurodyta skaičiais). Analizei naudoti poriniai skirtumai pagal haplotipų įvairovę ir  $F_{ST}$  pagal haplogrupių pasiskirstymą

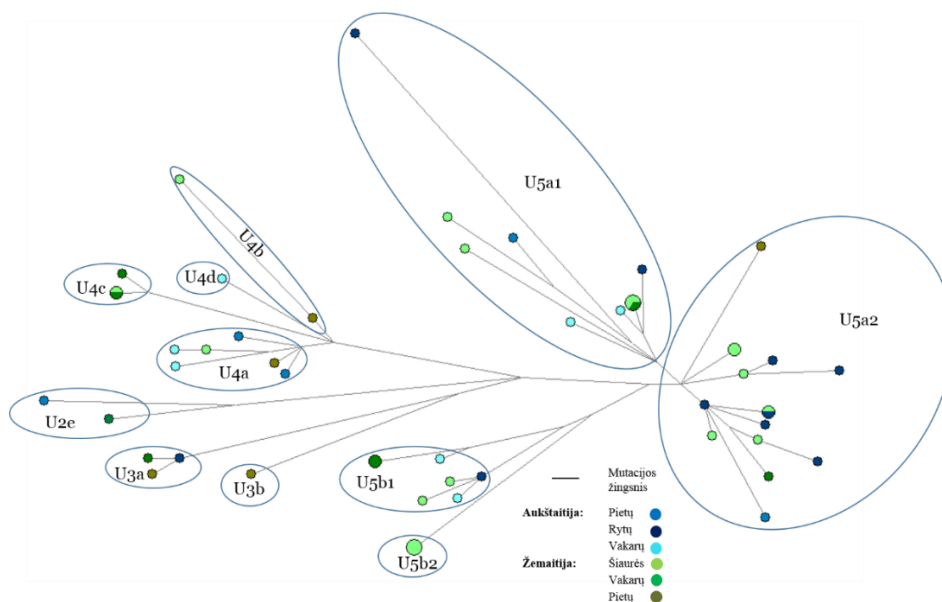
Sudaryti filogenetiniai medžiai, remiantis mtDNR haplotipų (577–16 023 np) pasiskirstymu, dažniausių tirtoje lietuvių populiacijoje mtDNR filogenetinių linijų H (45,3 %), U (19,2 %) viduje. Filogenetinės linijos H viduje dažniausiai haplogrupei H1 (19,6 %) būdingas artimas žvaigždės struktūrai filogenetinis medis, turintis centrinį haplotipą, kurio dažnis ~4,0 % tirtoje lietuvių populiacijos grupėje. Filogenetinio medžio struktūra rodo haplogrupės augimą, galbūt vykusį neolito laikotarpiu. Haplogrupės H1 klasteris atskirtas nuo kitų filogenetinio medžio haplogrupių G3010A pakaita, t. y. pakaita, apibūdinančia haplogrupę H1 mtDNR filogenetiniame medyje (4.22 pav.).

Filogenetinės linijos U viduje dažniausia haplogrupė U5a (9,0 %). Filogenetiniam medžiui būdinga laisvesnė struktūra, ilgos haplotipų šakos, t. y. haplotipus skiria daug mutacijos žingsnių. Nematyti haplogrupių specifinio pasiskirstymo tarp visų etnolingvistinių grupių ar tarp Šiaurės ir

Vakarų Žemaitijos, kurioms nustatytas didžiausias filogenetinės linijos U dažnis (4.23 pav.).



4.22 pav. Filogenetinis mtDNR filogenetinės linijos H medis pagal haplotipų (577–16 023 np) pasiskirstymą tirtoje Lietuvos populiacijoje. Filogenetinis medis sudarytas naudojant medianos jungties (angl. *median joining*) jungimo metodą ( $\epsilon = 1$ )



4.23 pav. Filogenetinis mtDNR filogenetinės linijos U medis pagal haplotipų (577–16 023 np) pasiskirstymą tirtoje Lietuvos populiacijoje. Filogenetinis medis sudarytas naudojant medianos jungties jungimo metodą ( $\epsilon = 1$ )



Nustatytas filogenetinių linijų H ir U haplogrupių, remiantis haplotipų (16 568 np, pašalinus keturias variabilias pozicijas 16 182, 16 183, 16 194, 16 519) pasiskirstymu, vidutinis mutacijų skaičius, amžius ir PI 95 % (Costa *et al.* 2013; Saillard *et al.* 2000; Soares *et al.* 2009).

Filogenetinės linijos H haplogrupių amžius varijuoja tarp ~1 000–16 000 metų, išskyrus H13, kurios nustatytas amžius 23 550 metų. Daugelio haplogrupių (H, H1a, H1b, H1c, H1e, H2a, H6a, H7, H10, H11, H16+152, H17a) nustatytas amžius siekia iki PL pabaigos. Dalies haplogrupių (H, H1a, H1c, H2a, H16+152, H17a) amžius siekia iki ~5 000 metų. Filogenetinės linijos U nustatytas haplogrupių amžius varijuoja tarp ~15 000–26 000 metų (4.9 lentelė).

4.9 lentelė. mtDNR filogenetinių linijų H ir U haplogrupių amžius

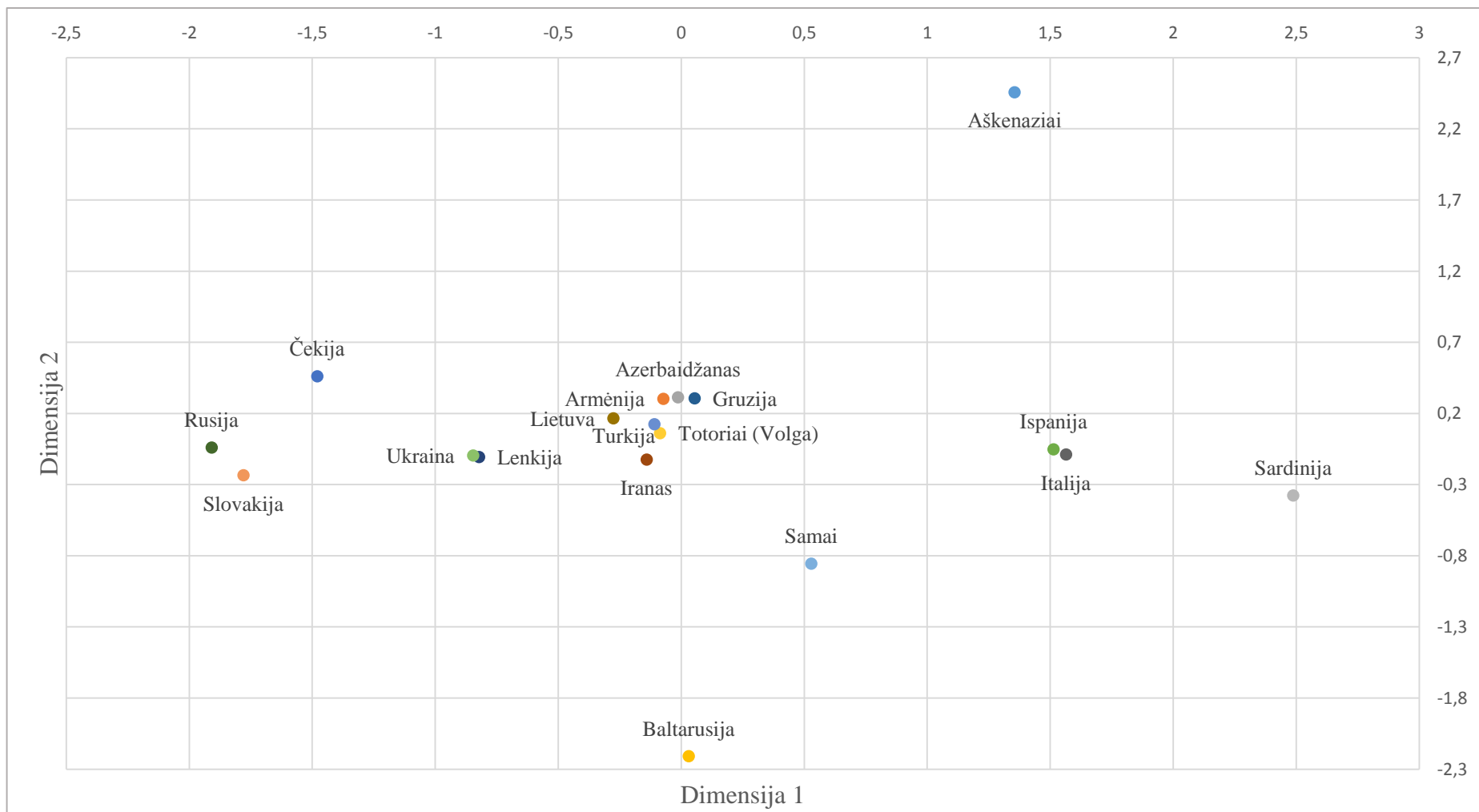
Haplogrupė	$\rho^1$	( $\pm$ SN)	Amžius	
			metai	PI 95 %
H	3,80	0,440	1 063	7 723–12 434
H1a	2,14	0,430	5 588	3 363–7 845
H1b	3,66	0,611	9 681	6 448–12 975
H1c	2,08	0,173	5 429	4 530–6 332
H1e	3,66	0,611	9 681	6 448–12 975
H1j	6,00	1,500	16 172	8 054–24 633
H2a	1,75	0,560	4 554	1 682–7 481
H5	6,14	0,878	16 567	11 762–21 489
H6a	3,50	0,450	9 246	6 865–11 660
H7	3,33	1,110	8 784	2 989–14 782
H10	2,38	0,390	6 214	4 187–8 268
H11	5,00	1,000	13 370	800–18 898
H13	8,57	1,220	23 550	16 677–30 633
H16+152	3,43	2,120	955	1 846–20 693
H17a	1,50	0,375	3 895	1 974–5 841
U4	6,66	0,722	18 042	14 053–22 110
U4a	5,83	1,138	15 701	9 524–22 076
U5a1	7,60	0,975	20 909	15 465–26 490
U5a2	5,60	0,453	15 046	12 570–17 554
U5b	9,50	1,330	26 281	18 707–34 097

<sup>1</sup>Vidutinis filogenetinės linijos mutacijų skaičius (angl. *rho factor*)

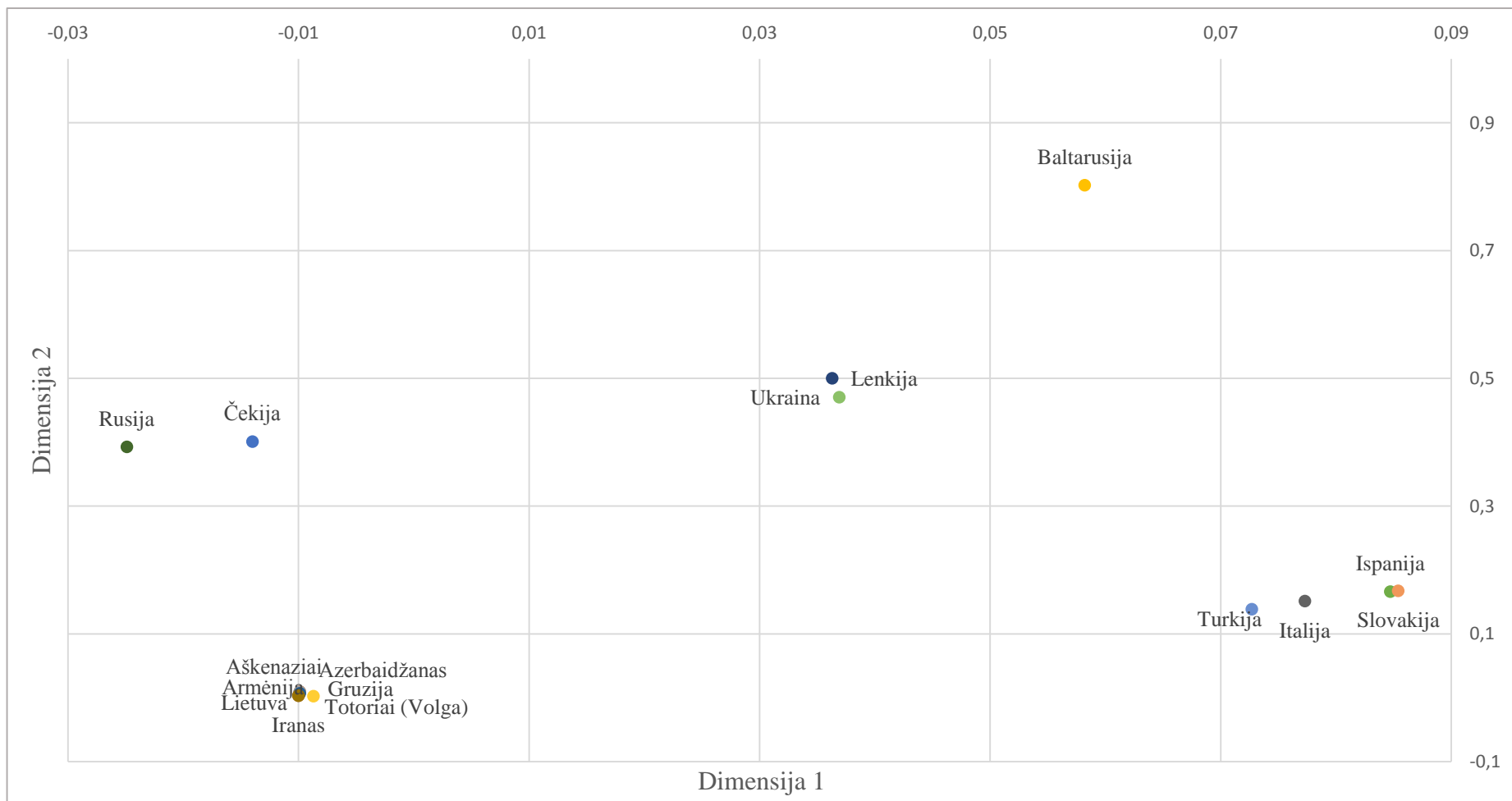
Vizualizavus gautus rezultatus MDS metodu, remiantis poriniais skirtumais (2 priedo 11 lentelė ir 12 lentelė) tarp mtDNR haplotipų (577–16 023 np) Europos populiacijose (2 priedo 10 lentelė), matomi populiacijų klasteriai: (1) slavų populiacijos (Lenkija, Ukraina, Rusijos europinė dalis, Čekija,

Slovakija); (2) Pietų Europos populiacijos (Italija, Ispanija, Sardinija); (3) aškenaziai; (4) lietuviai; (5) samiai; (6) Artimųjų Rytų, Vidurio Azijos populiacijos (Armėnija, Iranas, Azerbaidžanas, Gruzija, Turkija). Baltarusijos padėtis atokesnė nuo slavų populiacijų grupės (4.24 pav.). Vizualizavus gautus rezultatus MDS metodu, remiantis  $\Phi_{ST}$  (2 priedo 11 lentelė ir 12 lentelė) pagal mtDNR haplotipų (577–16 023 np) pasiskirstymą Europos populiacijose (2 priedo 10 lentelė), matomi klasteriai: (1) Rusija, Čekija; (2) Lenkija, Ukraina; (3) Baltarusija; (4) Vidurio Azijos populiacijos (Armėnija, Iranas, Azerbaidžanas, Gruzija) ir lietuviai; (5) Pietų Europos populiacijos (Italija, Ispanija, Sardinija) ir Slovakija (4.25 pav.).

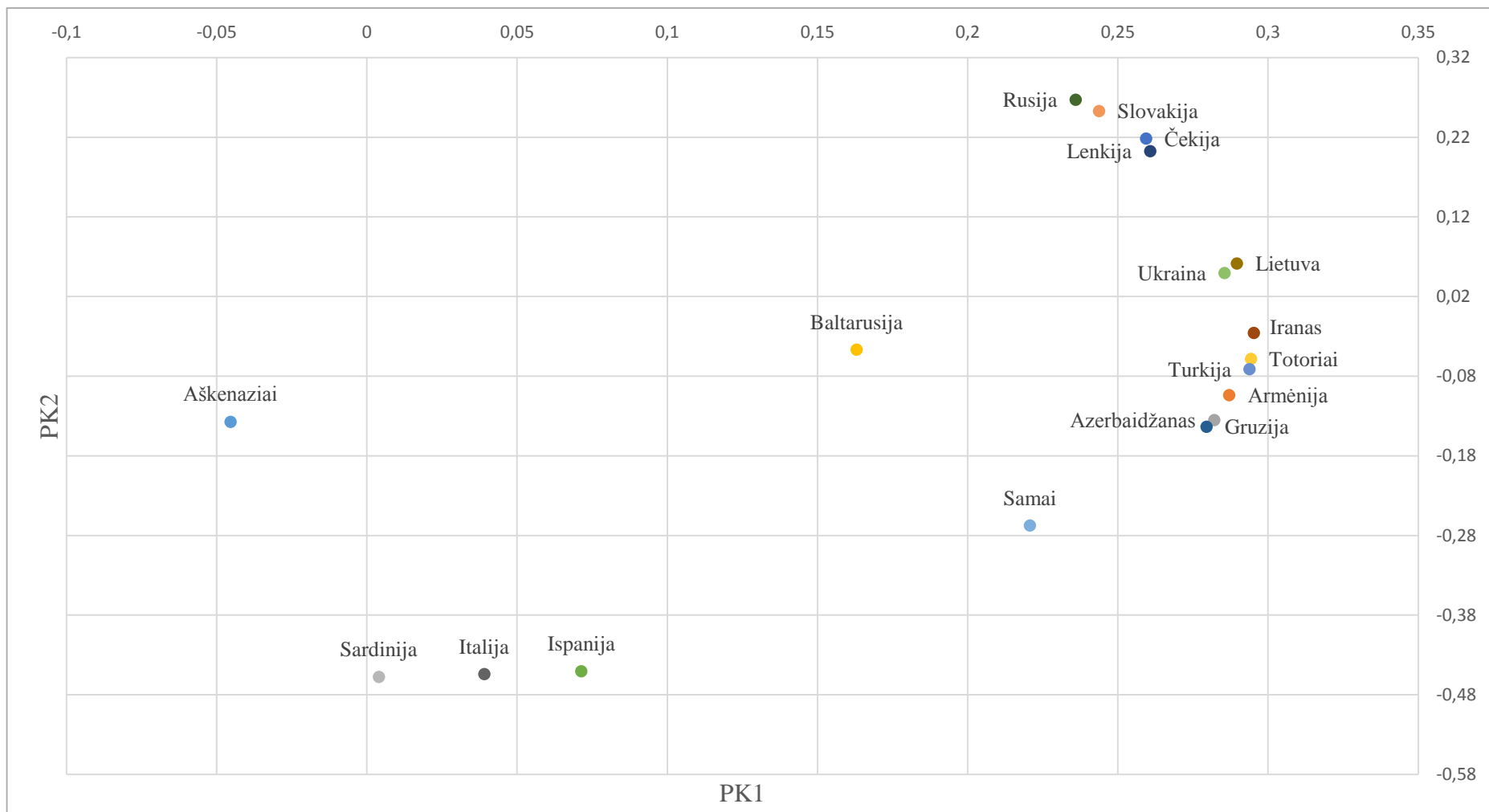
Vizualizavus PKA, remiantis poriniais skirtumais tarp mtDNR haplotipų (577–16 023 np) Europos populiacijose, rezultatus, matyti, kad Europos populiacijų grupavimasis artimas vizualizuotiems MDS rezultatams, tačiau lietuvių populiacija išsidėsto arčiau slavų populiacijų grupės, arčiausiai analizuotos Ukrainos populiacijos. PK1 aprašo 62,73 %, PK2 – 24,65 %, matomos mtDNR haplotipų (577–16 023 np) įvairovės tarp Europos populiacijų (4.26 pav.). Vizualizuoti PKA, remiantis  $\Phi_{ST}$  pagal mtDNR haplotipų (577–16 023 np) pasiskirstymą Europos populiacijose, rezultatai (4.27 pav.). PK1 aprašo 88,02 %, PK2 – 10,53 % matomos mtDNR haplotipų (577–16 023 np) įvairovės tarp Europos populiacijų.



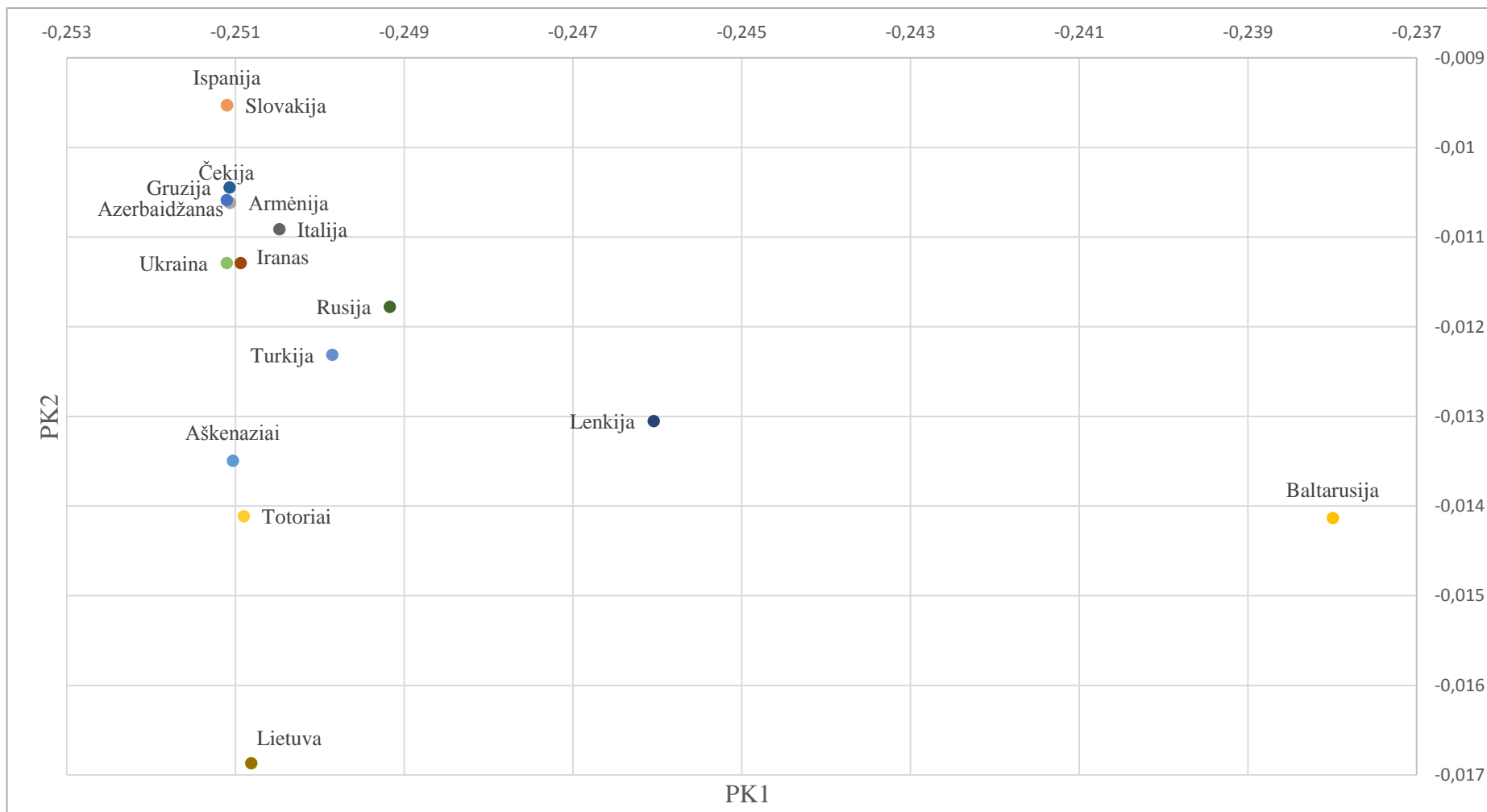
4.24 pav. MDS, remiantis poriniais skirtumais tarp mtDNR haplotipų (577–16 023 np) Europos populiacijose, vizualizuoti rezultatai. MDS paklaidos funkcija (angl. *Stress*) ir koreliacijos koeficientas ( $r^2$ ):  $stress = 0,12201$ ,  $r^2 = 0,94647$



4.25 pav. MDS, remiantis  $\Phi_{ST}$  atstumais pagal mtDNR haplotipų (577–16 023 np) pasiskirstymą Europos populacijose (pašalinus samų, Sardinijos populiacijas), vizualizuoti rezultatai. MDS paklaidos funkcija (angl. *Stress*) ir koreliacijos koeficientas ( $r^2$ ):  $stress = 0,07220$ ,  $r^2 = 0,99308$



4.26 pav. PKA, remiantis poriniais skirtumais tarp mtDNR haplotipų (577–16 023 np) Europos populiacijose, vizualizuoti rezultatai. PK1 aprašo 62,73 %, PK2 – 24,65 % mtDNR haplotipų (577–16 023 np) įvairovės tarp Europos populiacijų



4.27 pav. PKA, remiantis  $\Phi_{ST}$  atstumais pagal mtDNR haplotipų (577–16 023 np) pasiskirstymą Europos populacijose (pašalinus samų, Sardinijos populiacijas), vizualizuoti rezultatai. PK1 aprašo 88,02 %, PK2 – 10,53 % mtDNR haplotipų (577–16 023 np) įvairovės tarp Europos populiacijų

## APTARIMAS

D. Kasperavičiūtė kartu su kitais mokslininkais aprašė Lietuvos etnolingvistinių grupių struktūrą, remdamiesi nustatyta mtDNR haplogrupių ir HVRI haplotipų įvairove. Atlikta nustatytos Lietuvos ir Pabaltijo populiacijų (Latvija, Estija) mtDNR įvairovės analizė (Kasperavičiūtė *et al.* 2004). Kitų mokslininkų atlikti tyrimai, kuriuose įtraukta ir Lietuvos populiacija, pateikia analizės rezultatus, naudojant tik pagrindinių filogenetinių linijų ar HVRI/HVRII haplotipų duomenis (Kushniarevich *et al.* 2013; Lappalainen *et al.* 2008).

Šiuo darbu pirmą kartą nustatyta visa mtDNR seka didžiausioje, lyginant su ankstesniais tyrimais, šešių Lietuvos etnolingvistinių grupių imtyje. Tikslus mtDNR haplogrupių nustatymas, naudojant visos mtDNR sekos informaciją, leido detalai įvertinti haplogrupių pasiskirstymą lietuvių populiacijoje. Nustatytos dvi dažniausios filogenetinės linijos H ir U, kurių kiekviena sudaro atitinkamai 45,3 % ir 19,2 % tirtos populiacijos. Gauti duomenys artimi ankstesnių tyrimų rezultatams – atitinkamai 46,1 % ir 19,5 % (Kasperavičiūtė *et al.* 2004). Filogenetinės linijos T, J1, V sudarė atitinkamai 7,6 % ir po 5,8 %. Penkios dažniausios filogenetinės linijos (H, J1, T, U, V) sudaro ~84 % tirtos lietuvių populiacijos. Retaus filogenetinės linijos A, HV, I, K, M, N, W nustatomos <5 % dažniu ir kartu sudaro ~16 % tirtos lietuvių populiacijos. Filogenetinės linijos H dažnis (45,3 %) tirtoje lietuvių populiacijos grupėje artimas Šiaurės, Centrinėje ir Rytų Europoje nustatytam dažniui (40–50 %) (Loogvali *et al.* 2004; Richards *et al.* 2000). Haplogrupės H1 dažnis (19,6 %) taip pat artimas Šiaurės, Centrinėje ir Rytų Europoje matomam dažniui (10–18 %). Pastarosios, kaip ir filogenetinės linijos H, didelis dažnis būdingas ir Pietvakarių Europai, kur, manoma, buvo šios filogenetinės linijos prieglobstis PL metu. Dabartinėse populiacijose filogenetinės linijos H paplitimas, manoma, pasiektas viduriniame neolite, o galimas židinys – Centrinėje Europoje. Tikslus filogenetinės linijos H plitimo procesas vis dar lieka diskusijų klausimas. Filogenetinės linijos H dažnis tirtoje lietuvių populiacijos grupėje varijuoja 18 % ribose. Mažiausias dažnis Vakarų Žemaitijoje (~35 %), didžiausias Pietų

Žemaitijoje (~53 %) ir Vakarų Aukštaitijoje (~51 %). Ankstesniais tyrimais mažiausias dažnis nustatytas Vakarų Žemaitijoje ir Pietų Aukštaitijoje (~40 %), didžiausias – Šiaurės ir Pietų Žemaitijoje (~53 %) (Kasperavičiūtė *et al.* 2004). Filogenetinės linijos U dažnis tirtoje lietuvių populiacijoje siekia 19,2 %, kur U5 sudaro 12,7 %. Nustatytas dažnis artimas Centrinei ir Rytų Europai būdingam dažniui (21 %). Filogenetinė linija U, dominavusi tarp pasisavinamojo ūkio atstovų (82 %), dabartinėse populiacijose aptinkama 12–20 % dažniu. Filogenetinės linijos U dažnis tirtoje lietuvių populiacijos grupėje taip pat varijuoja 18 % ribose. Didžiausias dažnis matomas Šiaurės (~29 %) ir Vakarų (~23 %) Žemaitijoje, kitose trijose Lietuvos etnolingvistinėse grupėse dažnis yra tolygiai mažesnis (~16 %) ir mažiausias Pietų Žemaitijoje (~11 %). Ankstesniais tyrimais nustatyta mažesnė variacija tarp grupių (~13,3 %), kur didžiausias dažnis buvo Pietų Žemaitijoje (26,6 %), mažiausias Šiaurės Žemaitijoje ir Vakarų Aukštaitijoje (13,3 %) (Kasperavičiūtė *et al.* 2004).

Nustatyta haplotipų įvairovė tolygi tarp tirtų šešių Lietuvos etnolingvistinių grupių. Nukleotidų įvairovė ir vidutinis porinių skirtumų skaičius artimas tarp tirtų etnolingvistinių grupių, išskyrus dvi: (1) didesni įverčiai nustatyti Vakarų Žemaitijoje ( $ND = 31,69 (\pm 14,2)$  ir  $\pi = 1,915 \cdot 10^{-3} (\pm 0,95 \cdot 10^{-3})$ ); (2) mažesni įverčiai nustatyti Pietų Žemaitijoje ( $ND = 22,55 (\pm 10,1)$  ir  $\pi = 1,363 \cdot 10^{-3} (\pm 0,68 \cdot 10^{-3})$ ). Pastarosios populiacijos pasižymi šių įverčių skirtumais ir HVRI regione (4.5 lentelė). Esant tolygiai haplotipų įvairovei tarp etnolingvistinių grupių, Vakarų Žemaitijoje matoma daugiau, Pietų Žemaitijoje – mažiau skirtumų tarp tiriamųjų. Nustatytas statistiškai reikšmingas Tajima D įvertis ( $p < 0,05$ ) visose tirtose šešiose etnolingvistinėse grupėse dėl galimo populiacijos augimo. Nustatyta ~8 kartus mažesnė nukleotidų įvairovė koduojančioje srityje lyginant su reguliacine sritimi (HVRI regionas). Šis rezultatas rodo kryptingą mutacijų procesą, leidžiamas didesnis mutacijų dažnis reguliacinėje srityje. Mutacijų dažnio skirtumai matomi vertinant segreguojančias pozicijas, tranzicijas, transversijas, sinonimines ir nesinonimines pakaitas mtDNR sekoje. Reguliacinėje srityje ir *MT-ATP6* gene nustatyta daugiau, *MT-RNR1* (12S), *MT-RNR2* (16S), tRNR, *MT-CO1*,



*MT-ND4L* genuose – mažiau lyginant su tikėtiniu segreguojančių pozicijų skaičiumi pagal šių sričių ilgį. Reguliacinė sritis ir *MT-ATP6* genas pasižymi didesniu tranzicijų skaičiumi, nei tikėtina pagal šių sričių ilgį ( $p < 0,05$ ). *MT-RNR1* (12S), *MT-RNR2* (16S), tRNR, *MT-CO1*, *MT-ND4L* genams nustatytas statistiškai reikšmingai mažesnis tranzicijų skaičius ( $p < 0,05$ ). Transversijų atveju skirtumai mažesni. Transversijų reguliacinėje srityje statistiškai reikšmingai daugiau, *MT-RNR2* (16S) gene – mažiau ( $p < 0,05$ ) nei pagal šių sričių ilgį tikėtinas transversijų skaičius. Mažesnis transversijų skaičius susijęs su šių pakaitų galimu didesniu poveikiu fenotipui. Kryptingumas matomas ir nesinoniminių, sinoniminių pakaitų atveju (santykis ~1:2). Netolygus nesinoniminių ir sinoniminių pakaitų pasiskirstymas *MT-ATP6* ir *MT-ND6* genuose lemia statistiškai reikšmingą ( $p < 0,05$ ) variaciją. *MT-ATP6* gene nustatyta daugiau, *MT-ND6* gene – mažiau nesinoniminių pakaitų, nei tikėtina pagal šių genų ilgį. Didesnis nesinoniminių pakaitų skaičius *MT-ATP6* gene yra daugelyje populiacijų, mažesnis nesinoniminių pakaitų skaičius nustatomas skirtinguose genuose priklausomai nuo tirtos imties (Gunnarsdottir *et al.* 2011; Schonberg *et al.* 2011). Manoma, kad nesinoniminių pakaitų dominavimas yra jaunų filogenetinių linijų (Kivisild *et al.* 2006), bet ne klimato zonos, iš kurios yra tirta populiacija, požymis (Mishmar *et al.* 2003).

Nustatyti statistiškai reikšmingi  $F_{ST}$  atstumai ( $p < 0,05$ ) tarp Rytų Aukštaitijos ir Pietų Žemaitijos bei tarp Šiaurės Žemaitijos ir Pietų Žemaitijos pagal haplogrupių pasiskirstymą (69 grupės). Statistiškai reikšmingų skirtumų nerasta tarp Lietuvos etnolingvistinių grupių, remiantis poriniais skirtumais tarp haplotipų (16 568 np).

AMOVA rezultatai parodė statistiškai reikšmingus haplogrupių (23 grupės) įvairovės skirtumus tarp keturių grupių, sudarytų pagal geografinę padėtį (vakarų–rytų ašis): (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų Aukštaitija; (4) Rytų Aukštaitija. AMOVA, remiantis haplotipų (16 569 np) poriniais skirtumais, rezultatai išlieka statistiškai reikšmingi tarp trijų grupių (vakarų–rytų ašis): (1) Šiaurės, Vakarų Žemaitija; (2) Pietų Žemaitija; (3) Vakarų, Pietų, Rytų Aukštaitija. Vakarų–rytų ašis taip

pat matoma analizuojant haplotipų (577–16 023 np) įvairovės barjerus. Pirmas barjeras atskiria Šiaurės ir Vakarų Žemaitiją nuo likusių grupių, antras barjeras – Rytų Aukštaitiją nuo kitų grupių ir trečias barjeras Pietų Žemaitiją nuo Vakarų ir Pietų Aukštaitijos. Vizualizavus gautus rezultatus MDS metodu, pagal haplogrupių (23 grupės) pasiskirstymą, matyti, kad Lietuvos etnolingvistinės grupės išsidėsto atokiai viena nuo kitos ir nesudaro klasterių. Vizualizuoti MDS metodu, pagal haplogrupių (69 grupės) pasiskirstymą, gauti rezultatai rodo, kad atokiausiai išdėsto Rytų Aukštaitija, Šiaurės Žemaitija ir Pietų Žemaitija, tarp kurių nustatyti statistiškai reikšmingi  $F_{ST}$  atstumai. Kitos etnolingvistinės grupės (Vakarų Žemaitija, Vakarų, Pietų Aukštaitija) sudaro vieną klasterį.

PKA nevertinta vidinės įvairovės, kuri sudaro didžiąją dalį (>99 %), variacijos. PK1 aprašo 63,63 %, PK2 – 21,49 % mtDNR haplogrupių (69 grupės) įvairovės tarp tirtų šešių Lietuvos etnolingvistinių grupių. PK analizės, remiantis poriniais skirtumais tarp haplotipų (577–16 023 np), rezultatai rodo, kad PK1 aprašo 50,0 %, PK2 – 27,96 % mtDNR haplotipų įvairovės tarp tirtų šešių Lietuvos etnolingvistinių grupių. PK1 aprašo Pietų Žemaitijos mtDNR haplotipų įvairovės skirtumus nuo kitų penkių etnolingvistinių grupių.

Filogenetiniame medyje, pagal filogenetinės linijos H haplotipų (577–16 023 np) įvairovę, haplogrupė H1 formuoja žvaigždės struktūrą, kuri rodo šios haplogrupės populiacijos augimą. Manoma, kad šios filogenetinės linijos aktyvaus augimo, išplitusio po Europą, židinys galėjo būti Centrinėje Europoje vidurinio neolito laikotarpiu. Filogenetinio medžio struktūra išnyksta sudarius filogenetinį medį pagal mtDNR haplotipų (16 564 np) įvairovę. Filogenetiniam medžiui, pagal filogenetinės linijos U haplotipų (577–16 023 np) įvairovę, išskyrus haplogrupę U5a2, būdinga laisvesnė struktūra, ilgos haplotipų šakos, t. y. haplotipus skiria daug mutacijos žingsnių. Matoma filogenetinio medžio struktūra nerodo galimo neseno šios filogenetinės linijos augimo. Filogenetinė linija U galėjo dominuoti tarp pirmųjų gyventojų dabartinėje Lietuvos teritorijoje (Bramanti *et al.* 2009). Pastarieji gyventojai galėjo būti pakeisti ir (ar) priversti trauktis šiaurės vakarų kryptimi (didžiausias haplogrupės U dažnis Šiaurės ir Vakarų Žemaitijoje) plintant kartu su gamybiniu ūkiu

augančios filogenetinės linijos H populiacijai, kuri galėjo pasiekti dabartinę Lietuvos teritoriją iš Centrinės ar Rytų Europos. Dalis mtDNR analizės rezultatų rodo vakarų–rytų ašį galbūt dėl dviejų tirtoje lietuvių populiacijos grupėje pagrindinių filogenetinių linijų (H ir U) genetinio fondo judėjimo krypčių. Ši hipotezė neatskleidžia motininės linijos populiacijos formavimosi proceso sudėtingumo.

Filogenetinės linijos U galimą dominavimą tarp pirmųjų gyventojų patvirtina nustatytas šios filogenetinės linijos haplogrupių amžius (~15 000–26 000 metų). Filogenetinės linijos H neseną augimą patvirtina apskaičiuotas haplogrupių amžius. Daugelio haplogrupių (H, H1a, H1b, H1c, H1e, H2a, H6a, H7, H10, H11, H16+152, H17a) apskaičiuoto amžiaus variacija telpa į laikotarpį po PL. Dalies haplogrupių (H, H1a, H1c, H2a, H16+152, H17a) amžius siekia iki ~5 000 metų, t. y. laikotarpis, kai, manoma, pradėjo formotis baltų gentys.

Ribota kaimyninių populiacijų visos mtDNR sekos informacija neleido nustatyti tikslios lietuvių populiacijos padėties kitų populiacijų atžvilgiu. Pateiktas apibendrinimas apima tik į analizę įtrauktas populiacijas ir neatskleidžia esamos tikros lietuvių populiacijos padėties kaimyninių ar kitų Europos populiacijų atžvilgiu. Taip pat tikslingi atskirų haplogrupių tyrimai, kurių informacija yra prieinamas kitoms tyrėjų grupėms, galėjo turėti įtakos gautiems rezultatams taikant MDS ir PKA rezultatams. Vizualizavus gautus rezultatus MDS metodu pagal porinius skirtumus tarp mtDNR haplotipų (577–16 023 np) analizuotose Europos ir kitose populiacijose, matomas populiacijų grupavimasis pagal geografinę padėtį, išskyrus Baltarusiją, kuri yra kiek atokiau nuo kitų slavų populiacijų. Tirtos lietuvių populiacijos padėtis yra tarp slavų ir Vidurio Azijos populiacijų.

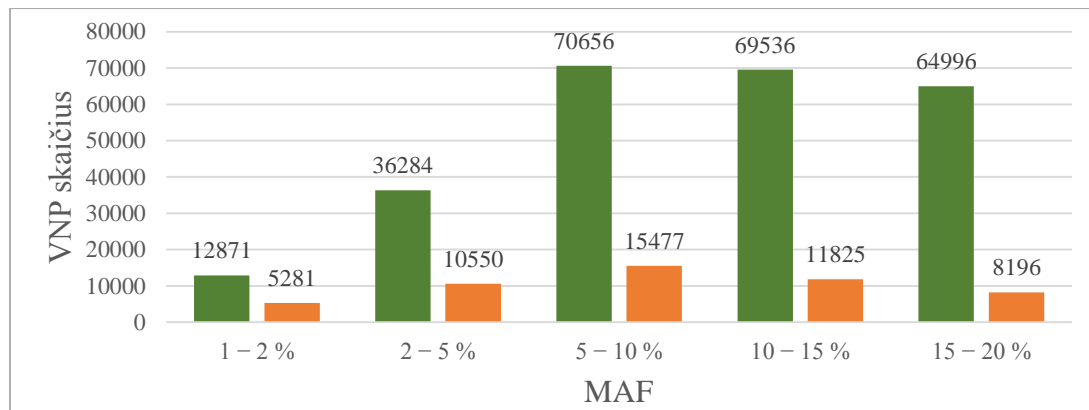
PKA, remiantis poriniais skirtumais tarp mtDNR haplotipų (577–16 023 np) analizuotose Europos ir kitose populiacijose, vizualizuoti rezultatai rodo panašų populiacijų išsidėstymą, tačiau lietuvių populiacijos padėtis artimesnė slavų populiacijoms. PK1 aprašo 62,73 % mtDNR haplotipų variacijos, kurią lemia įvairovės skirtumai tarp Artimųjų Rytų aškenazių, Pietų Europos ir kitų

analizuotų populiacijų. PK2 aprašo 24,65 % esamos mtDNR haplotipų variacijos, kurią lemia įvairovės skirtumai tarp slavų populiacijų (Rusija, Lenkija, Slovakija, Čekija), Pietų Europos (Italija, Sardinija, Ispanija) ir Artimųjų, Vidurio Rytų populiacijų. Vizualizavus gautus PKA rezultatus matyti, kad mtDNR haplotipų įvairovė lietuvių ir Ukrainos populiacijose lemia jų padėtį tarp slavų ir Artimųjų Rytų populiacijų.

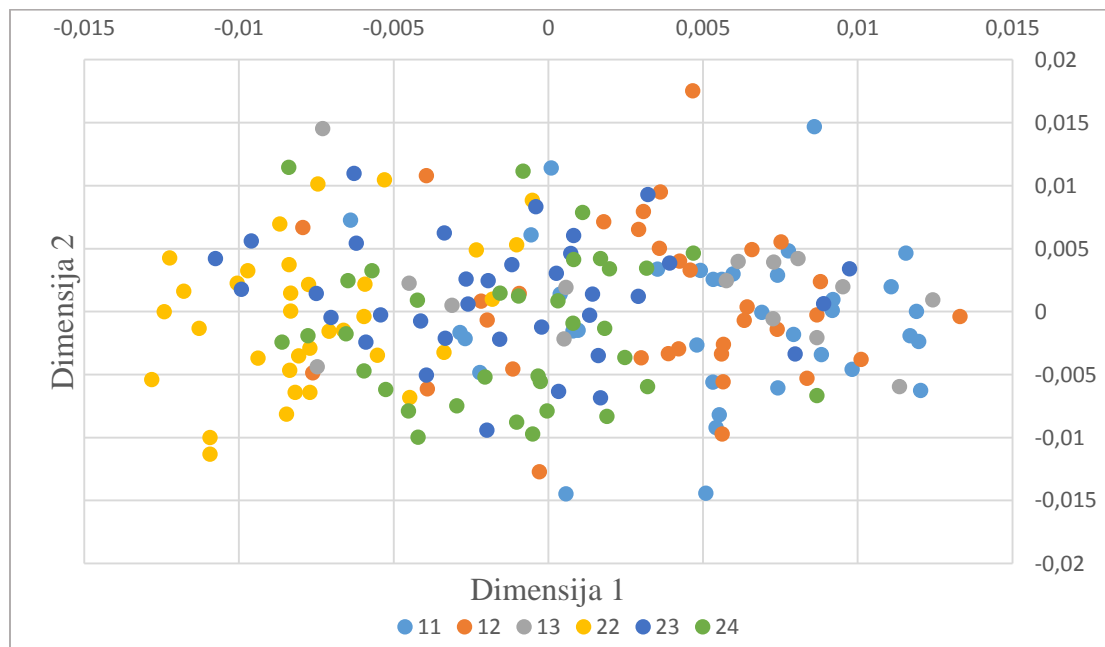
Šio darbo metu nustatyti mažesni mtDNR haplogrupių ir (ar) haplotipų įvairovės skirtumai palyginti su Y chromosomos haplogrupių ir haplotipų įvairove. Mažesnė mtDNR variacija nustatyta ir kituose populiacijų tyrimuose (Nasidze *et al.* 2004). Iškelta matriarchalinės migracijos hipotezė, teigianti, kad dėl didesnės moterų migracijos mtDNR įvairovės skirtumai yra mažesni. Pastaroji hipotezė paneigta ištyrus pasaulio populiacijas, kuriose nepastebėta mažesnės mtDNR įvairovės lyginant su Y chromosoma (Wilder *et al.* 2004). Naujausias tyrimas parodė, kad pasauliniu mastu didesni skirtumai tarp populiacijų nustatomi tiriant Y chromosomą nei mtDNR, tačiau šie skirtumai nėra tokie dideli, kaip teigta ankstesnių tyrimų išvadose. Taip pat pastarasis tyrimas parodė, kad yra regioniniai Y chromosomos ir mtDNR variacijos skirtumai (Lippold *et al.* 2014). Genetinės įvairovės skirtumai yra daugelio procesų – migracijos, mutacijų, gamtinės atrankos, genų dreifo – sąveikos tiriamoje populiacijoje rezultatas.

#### 4.3. Lietuvių populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis autosomų VNP pasiskirstymu

Tirtą lietuvių populiacijos grupę (3.5 pav.) sudarė 253 asmenys, kuriems genotipuoti tolygiai po visą genomą pasiskirstę 719 666 VNP. Atlikus pirminį ir antrinį gautų duomenų kokybės įvertinimą, tolesnei analizei naudoti 188 asmenų (Šiaurės Žemaitija – 36, Pietų Žemaitija – 34, Vakarų Žemaitija – 15, Pietų Aukštaitija – 34, Vakarų Aukštaitija – 35, Rytų Aukštaitija – 34; 123 vyriškosios lyties ir 65 moteriškosios lyties asmenys) 590 665 autosomų VNP (4.28 pav.).

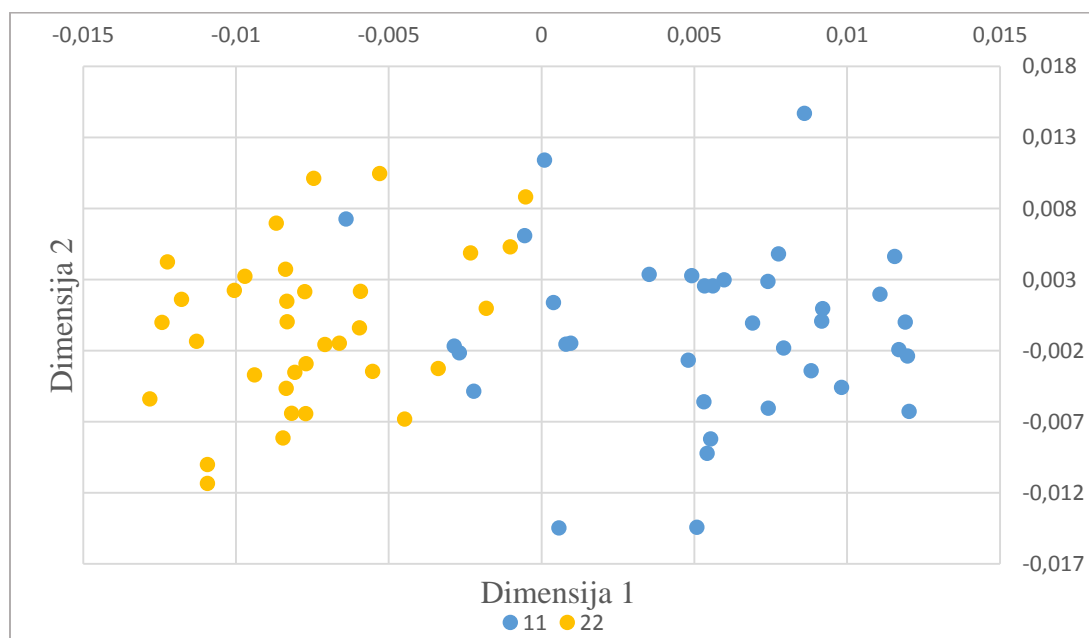


4.28 pav. Atlikus pirminį ir antrinį kokybės kontrolės vertinimą, tolesnei analizei naudotų 590 665 (žalia spalva) ir 105 387 (oranžinė spalva) autosomų VNP pasiskirstymas pagal MAF. Kokybės kontrolės metu pašalinti MAF < 0,01

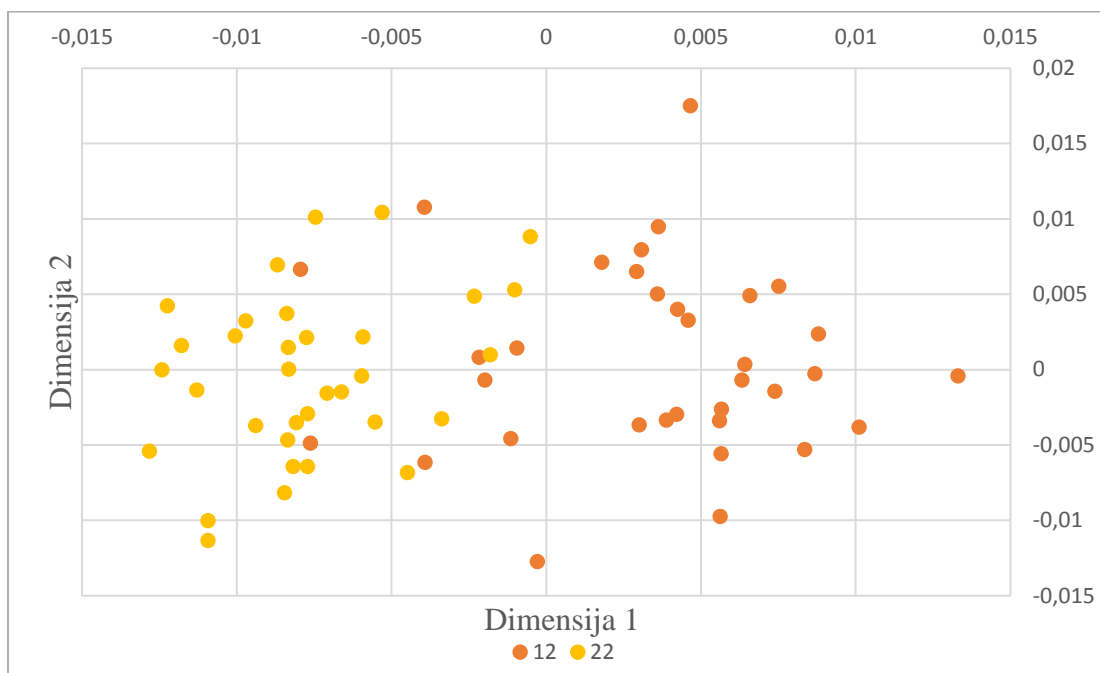


4.29 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 11 – Šiaurės Žemaitija, 12 – Pietų Žemaitija, 13 – Vakarų Žemaitija, 22 – Pietų Aukštaitija, 23 – Vakarų Aukštaitija, 24 – Rytų Aukštaitija

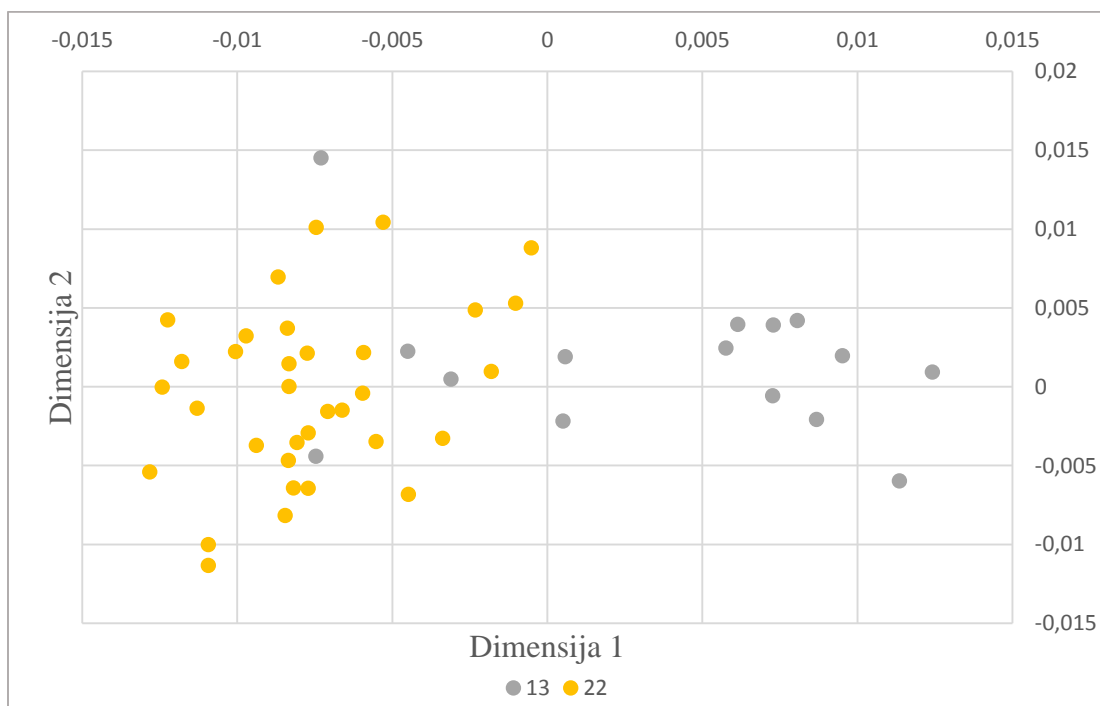
MDS metodu, naudojant 105 387 autosomų VNP, kuriems nustatyta pusiausvira sankiba ( $r^2 < 0,2$ ), t. y. nėra alelių asociacijos, gauti rezultatai vizualizuoti. MDS metodu analizuojant vienu metu visas Lietuvos etnolingvistines grupes ir vizualizavus gautus rezultatus, pirmoje ir antroje dimensijoje nėra aiškių klasterių, t. y. asmenys, iš tirtų šešių Lietuvos etnolingvistinių grupių išsidėsto arti vienas kito (4.29 pav.). Pašalinus iš vizualizuotų MDS rezultatų visus, išskyrus asmenis iš Šiaurės Žemaitijos ir Pietų Aukštaitijos, pirmoje dimensijoje gaunami du atsiskiriantys klasteriai. Matomi tik keli asmenys, kurie išsidėsto tarp dviejų susiformavusių klasterių (4.30 pav.). Taip pat analizuojant atskiras Lietuvos etnolingvistinių grupių poras taikant MDS ir vizualizavus rezultatus, pirmoje dimensijoje atsiskiriantys klasteriai susidaro tarp: (1) Pietų Aukštaitijos ir Vakarų Žemaitijos; (2) Pietų Aukštaitijos ir Pietų Žemaitijos. Vizualizavus pirmoje dimensijoje didesniu plotu sutampantys klasteriai matomi tarp: (1) Pietų ir Vakarų Aukštaitijos; (2) Pietų ir Rytų Aukštaitijos (4.31 pav., 4.32 pav., 4.33 pav., 4.34 pav.). Vizualizavus MDS metodu gautus rezultatus, visu plotu sutampančios tiriamųjų imtys gautos tarp Žemaitijos etnolingvistinių grupių.



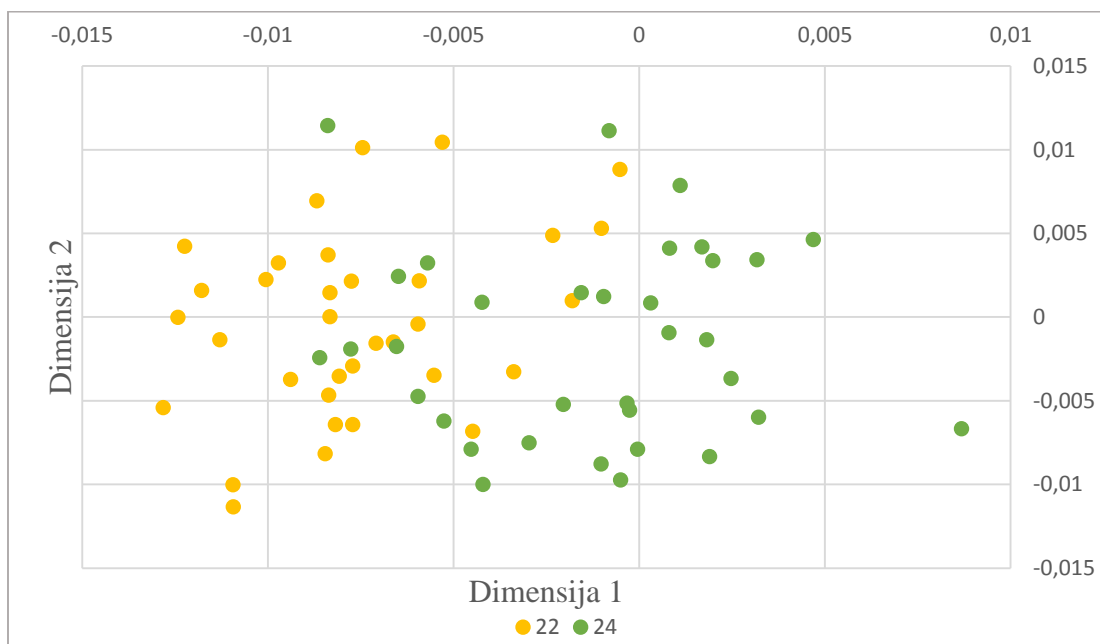
4.30 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 11 – Šiaurės Žemaitija, 22 – Pietų Aukštaitija



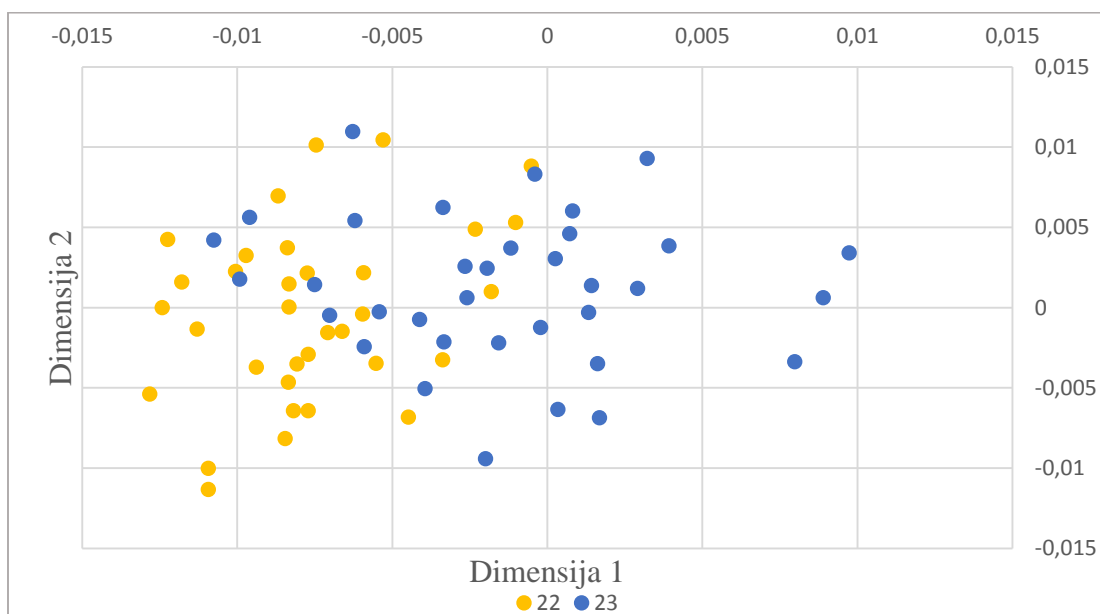
4.31 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 12 – Pietų Žemaitija, 22 – Pietų Aukštaitija



4.32 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 13 – Vakarų Žemaitija, 22 – Pietų Aukštaitija



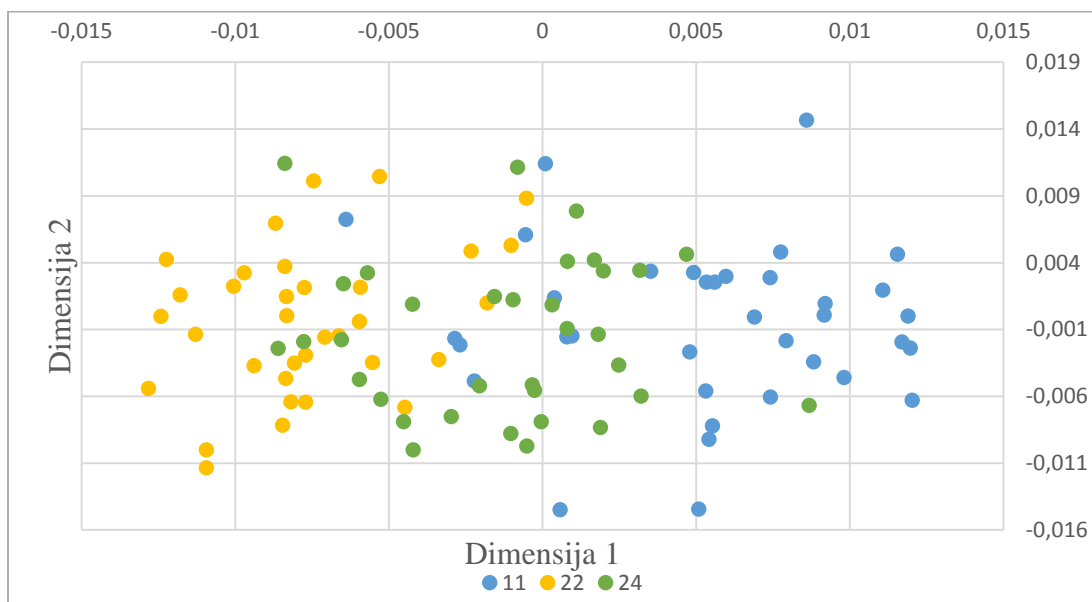
4.33 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 22 – Pietų Aukštaitija, 24 – Rytų Aukštaitija



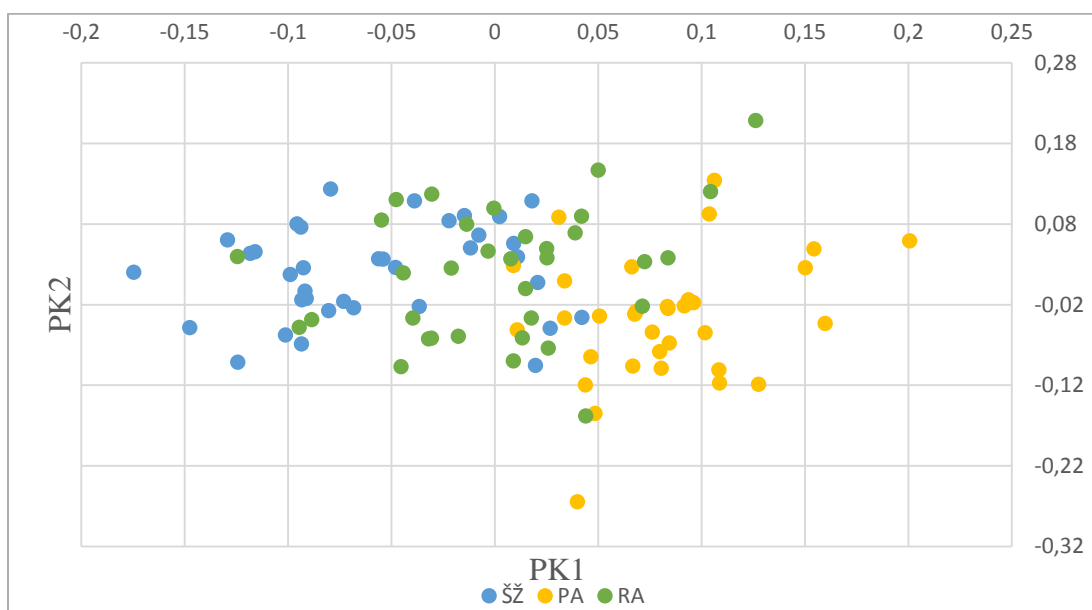
4.34 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 22 – Pietų Aukštaitija, 23 – Vakarų Aukštaitija

Analizuojant Pietų Aukštaitijos, Rytų arba Vakarų Aukštaitijos ir Šiaurės Žemaitijos populiacijas MDS metodu ir vizualizavus gautus rezultatus, pirmoje dimensijoje tirtos Lietuvos etnolingvistinės grupės išsidėsto pietryčių–šiaurės vakarų ašies atžvilgiu (4.35 pav.).





4.35 pav. MDS, remiantis atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, vizualizuoti rezultatai. Trumpiniai: 11 – Šiaurės Žemaitija, 22 – Pietų Aukštaitija, 24 – Rytų Aukštaitija



4.36 pav. PKA, remiantis 105 387 autosomų VNP alelių įvairovės pasiskirstymu, vizualizuoti rezultatai. PK1 aprašė 0,61 %, PK2 – 0,59 % stebimos autosomų VNP alelių įvairovės tarp tirtų asmenų (188) iš šešių Lietuvos etnolingvistinių grupių

PKA, remiantis 105 387 autosomų VNP alelių įvairovės pasiskirstymu, rezultatai rodo, kad PK1 aprašė 0,61 %, PK2 – 0,59 % autosomų VNP alelių įvairovės tarp tirtų asmenų (188) iš šešių Lietuvos etnolingvistinių grupių. Vizualizavus gautus PKA rezultatus, matomas PK1 aprašomos variacijos

pietryčių–šiaurės vakarų ašies gradientas. Nustatytas pirmos komponentės statistinis reikšmingumas (Tracy-Widom kriterijus,  $p < 0,05$ ) (4.36 pav.).

4.10 lentelė. Statistiškai reikšmingos PK1 pirmi 20 VNP, kurių esama genetinė variacija tarp tirtų šešių Lietuvos etnolingvistinių grupių yra didžiausia

VNP ID <sup>1</sup>	chr. <sup>3</sup>	Pozicija	PA	VA	RA	PŽ	VŽ	ŠŽ
rs806269	9		0,3088 (A) <sup>4</sup>	0,1429 (A)	0,0735 3 (A)	0,1176 (A)	0,0666 7 (A)	0,0277 8 (A)
rs2247168	11		0,3824 (A)	0,4143 (G)	0,3971 (G)	0,3235 (G)	0,4 (G)	0,3889 (G)
rs927694	13		0,3676 (A)	0,5 (C)	0,4853 (A)	0,4091 (C)	0,4 (C)	0,4444 (C)
rs2535584	17		0,3971 (A)	0,4714 (G)	0,4265 (A)	0,4848 (A)	0,4333 (A)	0,4167 (G)
rs1162608	9		0,3529 (G)	0,2857 (G)	0,2941 (G)	0,2206 (G)	0,0666 7 (G)	0,1806 (G)
rs645121	9		0,3529 (A)	0,4714 (A)	0,4265 (C)	0,4559 (C)	0,5 (A)	0,4861 (A)
rs2193769	5		0,3088 (G)	0,2143 (G)	0,2794 (G)	0,1618 (G)	0,1667 (G)	0,0694 4 (G)
rs17122441	14		0,3088 (A)	0,1286 (A)	0,0757 6 (A)	0,1029 (A)	0,1333 (A)	0,0694 4 (A)
rs10742969	11		0,1176 (A)	0,0571 4 (A)	0,0147 1 (A)	0 (A)	0 (A)	0,0277 8 (A)
rs17746916	10		0,2059 (A)	0,2857 (A)	0,2941 (A)	0,3088 (A)	0,5 (A)	0,4444 (A)
rs10769684	11		0,3382 (G)	0,3571 (G)	0,3971 (G)	0,4412 (A)	0,4333 (A)	0,3889 (A)
rs631844	5		0,25 (G)	0,3571 (G)	0,3971 (G)	0,4853 (G)	0,5 (G)	0,4861 (A)
rs221748 <sup>2</sup>	6	<i>PDE10A</i> (intronas)	0,2647 (A)	0,1714 (A)	0,1029 (A)	0,0735 3 (A)	0,1333 (A)	0,1111 (A)
rs6131629	20	<i>MACROD2</i> (intronas)	0,1029 (A)	0,2143 (A)	0,1667 (A)	0,2353 (A)	0,3 (A)	0,2639 (A)
rs6997954	8		0,2941 (G)	0,2286 (G)	0,1324 (G)	0,1471 (G)	0,1667 (G)	0,0555 6 (G)
rs4779631	15	<i>RYR3</i> (intronas)	0,3382 (A)	0,4714 (A)	0,4706 (G)	0,4412 (G)	0,4333 (G)	0,4444 (G)
rs9609889	22		0,1765 (G)	0,0714 3 (G)	0,1176 (G)	0,0882 4 (G)	0 (G)	0,0277 8 (G)
rs9511987	13	<i>ATP8A2</i> (intronas)	0,2647 (A)	0,4143 (A)	0,4242 (A)	0,4265 (G)	0,4333 (G)	0,3889 (A)
rs6007216	22	<i>LINC00229</i> (intronas)	0,3824 (A)	0,2143 (A)	0,1176 (A)	0,2353 (A)	0,2333 (A)	0,1667 (A)
rs6733159	2	<i>PARD3B</i> (intronas)	0,1324 (C)	0,2429 (C)	0,2059 (C)	0,2794 (C)	0,2333 (C)	0,2361 (C)

<sup>1</sup>DNR sekos varianto identifikacinis numeris db *NCBI dbSNP*

<sup>2</sup>Hardžio ir Vainbergo nepusiausvira ( $p = 0,03084$ )

<sup>3</sup>Chromosomos numeris

<sup>4</sup>MAF ir alelis

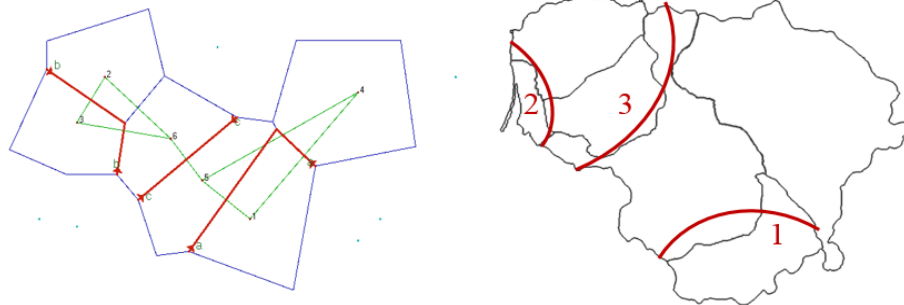
Nustatyti statistiškai reikšmingos PK1 VNP, kurių genetinė variacija tarp tirtų šešių Lietuvos etnolingvistinių grupių yra didžiausia. Apskaičiavus pirmų 20 variabiliausių VNP MAF ir įvertinus genomine padėtį, matyti, kad pirmi 12 VNP yra nekoduojančioje dalyje. Iš vertintų 20 VNP šeši VNP yra skirtingų genų intronuose (4.10 lentelė).

Apskaičiuoti  $F_{ST}$  atstumai ir atlikta statistiškai reikšmingos PK1 vektorių dispersijos analizė tarp tirtų visų šešių Lietuvos etnolingvistinių grupių (4.11 lentelė). Nustatyta statistiškai reikšminga  $F_{ST}$  ir geografinių atstumų teigiama koreliacija (Mantelio kriterijus,  $r = 0,60$ ;  $p = 0,0127$ ; 10 000 kėlinių).

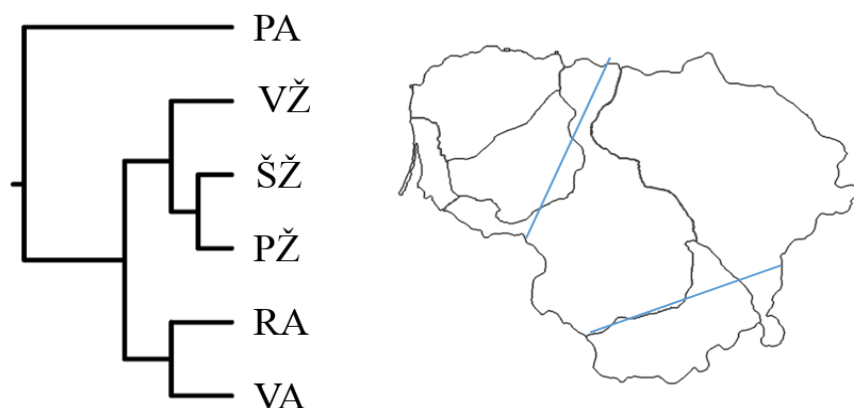
4.11 lentelė.  $F_{ST}$  atstumų matrica pagal 105 387 autosomų VNP alelių pasiskirstymą (žemiau įstrižainės). PK1 vektorių dispersijos analizės p reikšmės (aukščiau įstrižainės) tarp tirtų šešių Lietuvos etnolingvistinių grupių

	PA	ŠŽ	VŽ	RA	VA	PŽ
PA	0	$3,33 \cdot 10^{-11}$	$1,66 \cdot 10^{-6}$	$2,84 \cdot 10^{-7}$	$3,36 \cdot 10^{-3}$	$1,11 \cdot 10^{-11}$
ŠŽ	0,000909	0	0,2082	$4,09 \cdot 10^{-3}$	$2,06 \cdot 10^{-2}$	<b>0,0343</b>
VŽ	0,000795	0,00024	0	<b>0,0069</b>	<b>0,0092</b>	0,8915
RA	0,000499	0,000347	0,000488	0	0,5425	<b>0,0001</b>
VA	0,00017	0,000336	0,000282	0,000116	0	<b>0,0002</b>
PŽ	0,000717	0,000081	0,00025	0,000403	0,000143	0

Nustatyti 105 387 autosomų VNP alelių įvairovės barjerai tarp tirtų šešių Lietuvos etnolingvistinių grupių. Pirmas barjeras atskiria Pietų Aukštaitiją nuo kitų etnolingvistinių grupių. Antras barjeras atskiria Vakarų Žemaitiją nuo kitų etnolingvistinių grupių. Trečias barjeras – tarp Žemaitijos ir Aukštaitijos etnolingvistinių grupių (4.37 pav.).



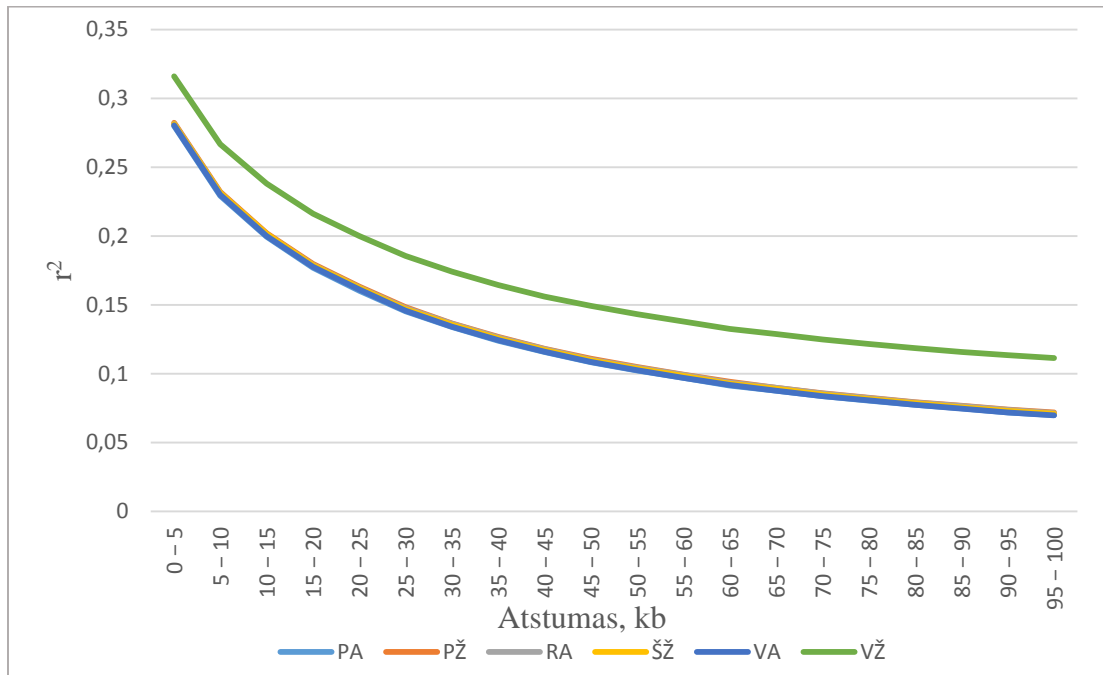
4.37 pav. Autosomų VNP (105 387) alelių įvairovės barjerai tarp tirtų Lietuvos etnolingvistinių grupių. Kairėje šešių Lietuvos etnolingvistinių grupių schema pagal pateiktas koordinatas (2 priedo 1 lentelė) ir matomi barjerai (raudona linija, eiliškumo tvarka nurodyta skaičiais). Analizei naudoti  $F_{ST}$  atstumai



4.38 pav. Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą, sudaryta šešių Lietuvos etnolingvistinių grupių dendrograma. Paveikslo dešinėje linijomis atskirti dendrogramoje matomi etnolingvistinių grupių klasteriai

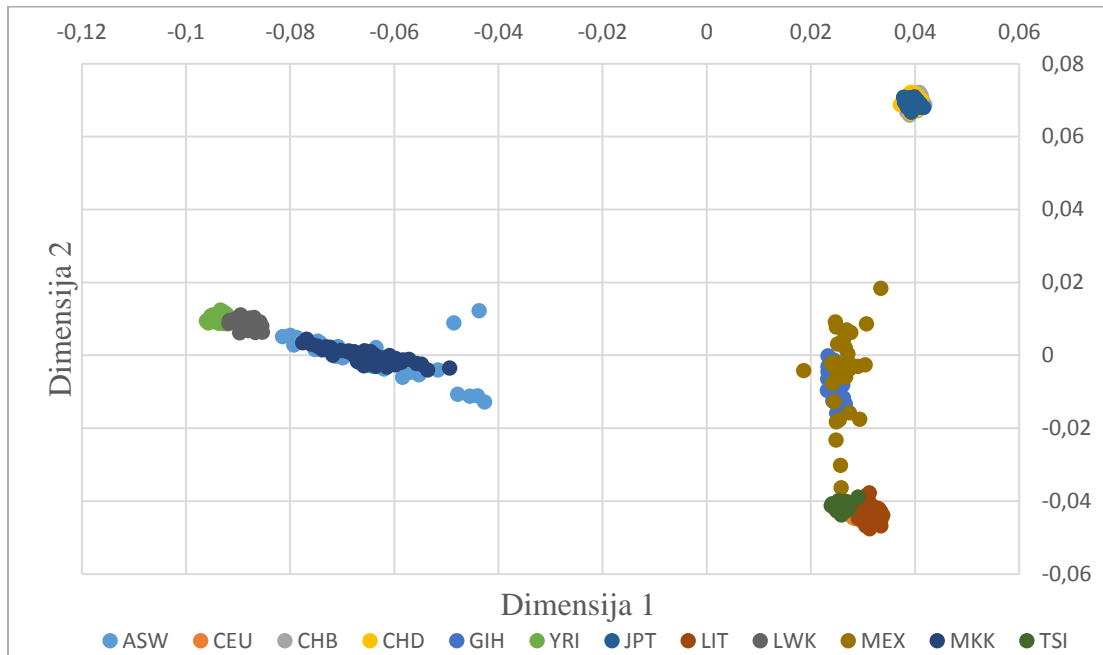
Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą, sudaryta tirtų šešių Lietuvos etnolingvistinių grupių dendrograma. Gautos trys grupės: (1) Šiaurės, Vakarų, Pietų Žemaitija; (2) Rytų, Vakarų Aukštaitija; (3) Pietų Aukštaitija. Pastaroji grupė (Pietų Aukštaitija) nutolusi didžiausiu atstumu nuo kitų dviejų grupių (4.38 pav.).

Nustatyta LD struktūra, vertinant  $r^2$  įverčio vidurkį 5 kb intervalais, tarp tirtų šešių Lietuvos etnolingvistinių grupių. LD  $r^2$  įverčiai artimi visoms tirtoms Lietuvos etnolingvistinėms grupėms, išskyrus Vakarų Žemaitiją. Pastarosios etnolingvistinės grupės  $r^2$  įvertis yra visuose intervaluose, t. y. nuo 0 kb iki 100 kb, didesnis lyginant su kitomis penkiomis etnolingvistinėmis grupėmis. Vakarų Žemaitijos imtyje 0–5 kb intervalo  $r^2 = 0,316$ , kitų grupių to paties intervalo vidutinis  $r^2 = 0,281$ . Vakarų Žemaitijos imtyje 95–100 kb intervalo  $r^2 = 0,111$ , kitų grupių 95–100 kb intervalo vidutinis  $r^2 = 0,071$ . Vakarų Žemaitijos visų intervalų  $r^2$  įverčiai statistiškai reikšmingai skyrėsi nuo kitų penkių etnolingvistinių grupių (nepriklausomų imčių t kriterijus,  $p < 0,05$ ) (4.39 pav.).

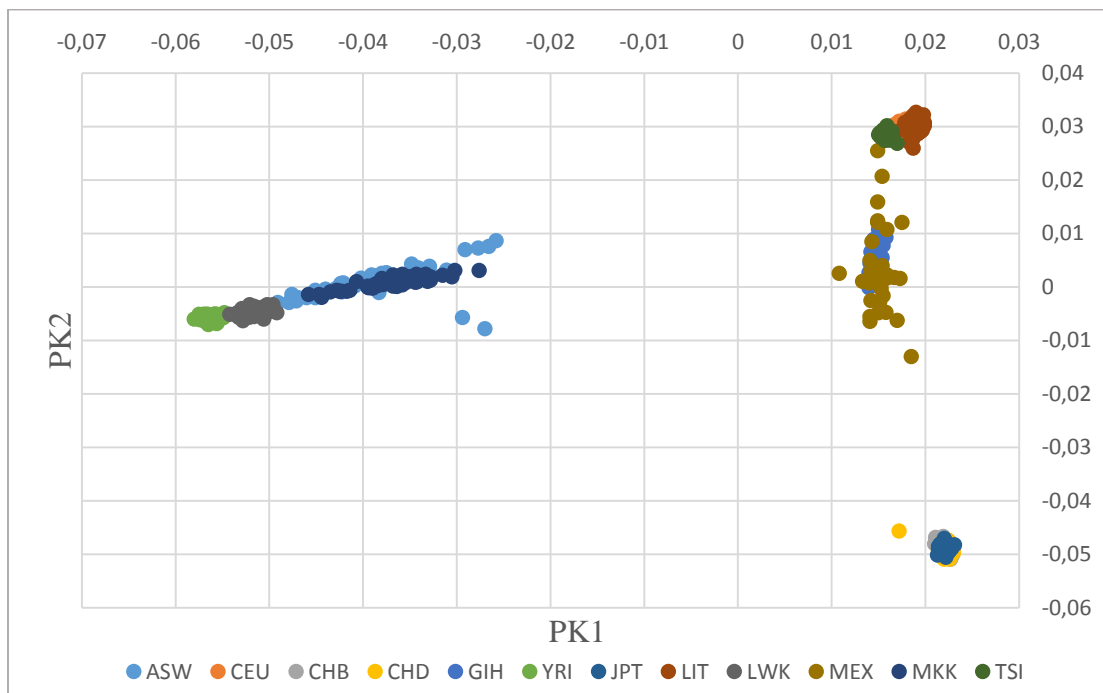


4.39 pav. LD struktūra (590 665 autosomų VNP) tirtoje lietuvių populiacijos grupėje

MDS metodu, remiantis 106 545 autosomų VNP, kuriems nustatyta pusiausvira sankiba ( $r^2 < 0,2$ ), t. y. nėra alelių asociacijos, gauti rezultatai vizualizuoti. Susidarė šios populiacijų grupės: (1) Azijos populiacijos (CHB – kinų populiacija (84 tiriamųjų) iš Pekino miesto, Kinija; JPT – japonų populiacija (86) iš Japonijos; CHD – kinų populiacija (85) iš Denverio miesto); (2) Afrikos populiacijos (YRI – jorubų populiacija (78) iš Nigerijos, LWK – luchjų populiacija (88) iš Kenijos, ASW – Afrikos kilmės Šiaurės Rytų JAV gyventojai (46), MKK – Kinyawa masajų populiacija (118) iš Kenijos); (3) Europos populiacijos (CEU – Šiaurės ar Vakarų Europos kilmės Jutos gyventojai (84), TSI – Toskanos regiono populiacija (88) iš Italijos, LIT – lietuvių populiacija (188)); (4) Šiaurės Amerikos ir Pietų Azijos populiacijos (MXL/MEX – meksikiečių kilmės Los Andželo gyventojai (31), GIH – indų kilmės Hiustono gyventojai (85)). Šiaurės Amerikos (MXL/MEX) ir Pietų Azijos (GIH) populiacijų klasteriai susikloja ir išsidėsto tarp susidariusių Europos ir Azijos populiacijų grupių. Lietuvos populiacija išsidėsto arčiausiai kitos Europos populiacijos (CEU). Didžiausi populiacijų genetiniai skirtumai yra Afrikos grupėje (4.40 pav.).



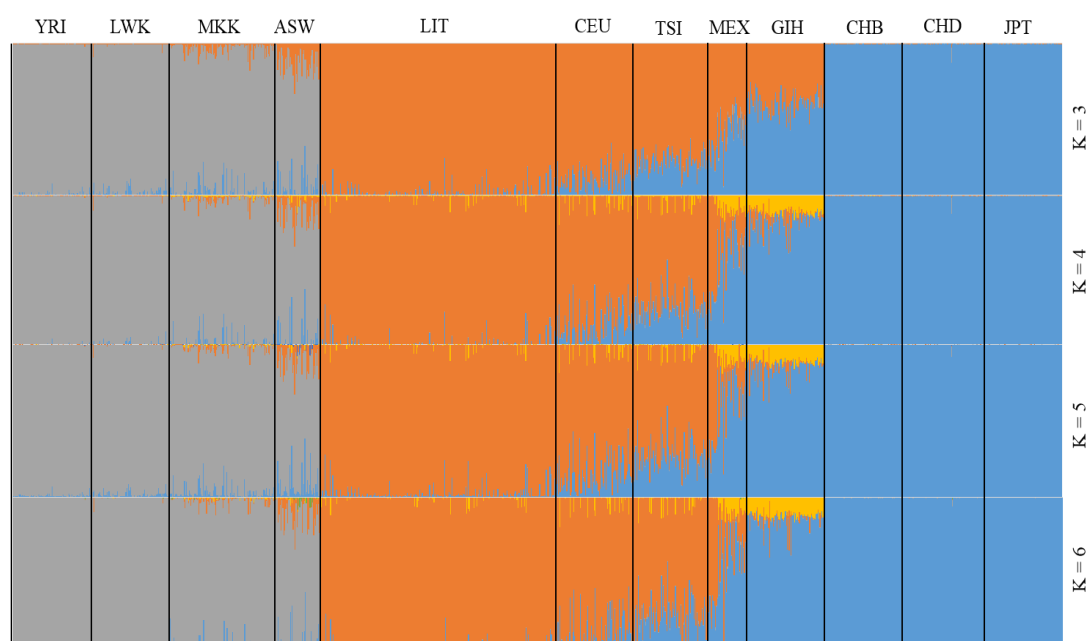
4.40 pav. MDS, remiantis atstumais pagal 106 545 autosomų VNP alelių pasiskirstymą lietuvių ir HapMap3 populiacijose (1 126 asmenys). Trumpiniai: Europos populiacijos (CEU, TSI, LIT), Azijos populiacijos (CHB, JPT, CHD), Afrikos populiacijos (YRI, LWK, ASW, MKK), Šiaurės Amerikos populiacija (MXL/MEX), Pietų Azijos populiacija (GIH). Detalesnį populiacijų aprašą žr. tekste, santrumpose



4.41 pav. PKA pagal 106 545 autosomų VNP alelių pasiskirstymą, PK1 aprašė 5,71 %, PK2 – 3,62 % autosomų VNP alelių įvairovės tarp tirtų lietuvių ir HapMap3 populiacijų (1 126 asmenys)

Atlikus PKA, remiantis 106 545 autosomų VNP alelių, kuriems nustatyta pusiausvira sankiba ( $r^2 < 0,2$ ), t. y. nėra alelių asociacijos, pasiskirstymu, gauti

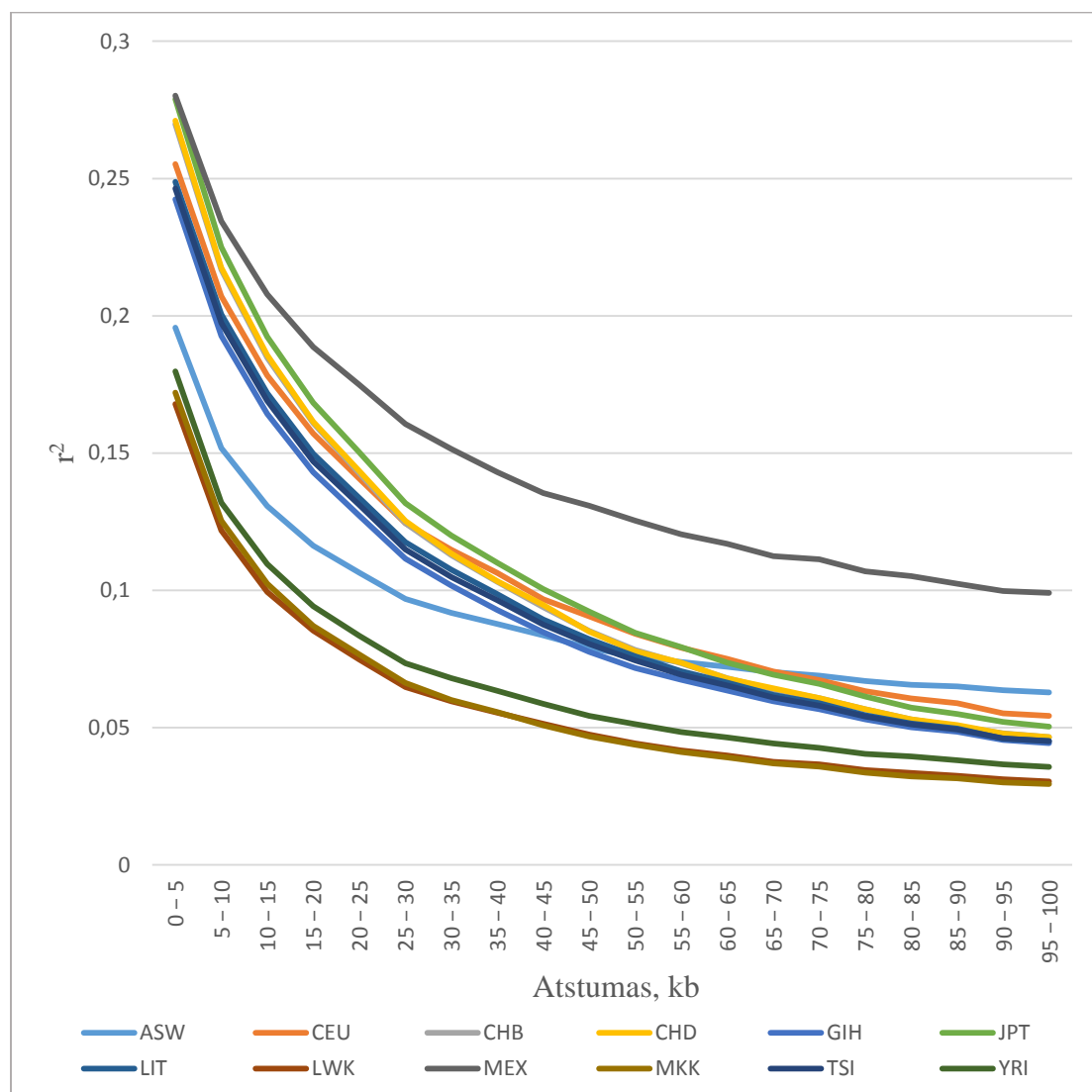
rezultatai vizualizuoti. PK1 aprašė 5,71 %, PK2 – 3,62 % esamos autosomų VNP alelių įvairovės tarp tirtų asmenų iš Lietuvos ir HapMap3 populiacijų. Populiacijų grupavimasis artimas vizualizuotiems rezultatams, gautiems taikant MDS. PK1 aprašoma variacijos įvairovė tarp Afrikos, Azijos populiacijų ir kitų analizuotų HapMap3 populiacijų. PK2 aprašoma variacijos įvairovė tarp Afrikos, Azijos, Europos, Šiaurės Amerikos ir Pietų Azijos populiacijų (4.41 pav.). Pirmos 38 komponentės statistiškai reikšmingos (Tracy-Widom kriterijus,  $p < 0,05$ ). Statistiškai reikšmingi skirtumai nustatyti tarp visų populiacijų ( $\chi^2$  kriterijus,  $p < 0,05$ ).



4.42 pav. Lietuvių ir HapMap3 populiacijų struktūros modeliai, aprašantys esamą genetinę variaciją tarp populiacijų galimu grupių skaičiumi ( $K = 2-6$ )

Atrinkti 1 569 VNP ( $r^2 < 0,2$ ), kuriems nustatyta didžiausia variacija (PK1–200) tarp tirtų lietuvių ir HapMap3 populiacijų (Huckins *et al.* 2014; Tian *et al.* 2008). Nustatytas grupių skaičius ( $K > 2$ ), kuris geriausiai aprašo esamą variaciją tarp tirtų populiacijų (2 priedo 1 pav.). Išanalizavus modelius su skirtingu grupių skaičiumi ( $K = 2-6$ ), matyti, kad tirtos populiacijos geriausiai aprašomos keturių grupių ( $K = 4$ ) modeliu (4.42 pav.). Šiuo modeliu galima atskirti pagrindines grupes (Afrikos, Azijos, Pietų Azijos, Šiaurės Amerikos, Europos populiacijas). Keturių grupių ( $K = 4$ ) modelis geriausiai aprašo esamą variaciją tarp populiacijų Europos grupėje. Matyti, kad didžiausia Azijos

genetinės įvairovės dalis matoma TSI populiacijoje, mažesnė – CEU ir mažiausia – tirtoje lietuvių populiacijoje. Pasirinktu modeliu esama variacija tarp Azijos populiacijų aprašoma prasčiausiai, t. y. neįmanoma atskirti populiacijų Azijos grupės viduje. Keturių grupių ( $K = 4$ ) modeliu aprašoma tirtų grupių populiacijų struktūra atitinka PKA gautus ir vizualizuotus rezultatus (4.41 pav.).

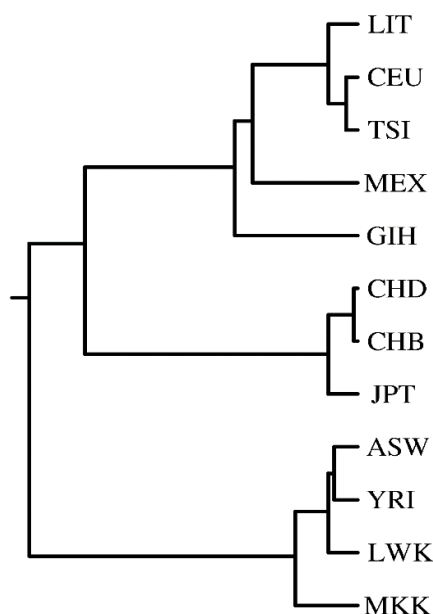


4.43 pav. LD struktūra (202 997 autosomų VNP) tirtoje lietuvių ir HapMap3 populiacijose (1 126 asmenys)

Nustatyta LD struktūra, vertinant  $r^2$  įverčio vidurkį 5 kb intervalais, tarp tirtų lietuvių ir HapMap3 populiacijų. Matomas artimas LD  $r^2$  įverčių grupavimasis. Mažiausia alelių asociacija ir trumpiausi LD blokai – tarp trijų Afrikos populiacijų (YRI, LWK, MKK). Afrikos populiacijai ASW būdinga didesnė



alelių asociacija ir ilgesni LD blokai, palyginti su kitomis Afrikos populiacijomis. Didžiausia alelių asociacija ir ilgiausi LD blokai matomi Šiaurės Amerikos (MXL/MEX) populiacijoje. Azijos (CHB, JPT, CHD), Europos (CEU, TSI, LIT), Pietų Azijos (GIH) populiacijoms būdingi tarpiniai  $r^2$  įverčiai. Matomas, nors ir nedidelis, grupavimasis tarp pastarųjų populiacijų trumpesniuose intervaluose ir jis išnyksta didėjant atstumui: (1) Azijos (CHB, JPT, CHD) populiacijoms būdinga didžiausia alelių asociacija; (2) Pietų Azijos (GIH) mažiausia. Lietuvių populiacijos LD struktūra artimiausia kitai Europos populiacijai (TSI) ir kiek daugiau skiriasi nuo CEU populiacijos. Lietuvių populiacijos visų intervalų  $r^2$  įverčiai statistiškai reikšmingai skyrėsi nuo Šiaurės Amerikos (MXL/MEX), Afrikos (YRI, MKK, LWK) populiacijų (nepriklausomų imčių  $t$  kriterijus,  $p < 0,05$ ) (4.43 pav.).



4.44 pav. Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumais pagal 106 545 autosomų VNP alelių pasiskirstymą, sudaryta Lietuvos ir HapMap3 populiacijų dendrograma

Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumais (2 priedo 15 lentelė), sudaryta tirtų lietuvių ir HapMap3 populiacijų dendrograma. Gautos penkios grupės: (1) Europos populiacijos (CEU, TSI, LIT); (2) Šiaurės Amerikos populiacija (MXL/MEX); (3) Pietų Azijos

populiacija (GIH); (4) Azijos populiacijos (CHB, JPT, CHD); (5) Afrikos populiacijos (YRI, LWK, ASW, MKK) (4.44 pav.).

## APTARIMAS

M. Nelis kartu su kitais mokslininkais, remdamiesi 273 454 autosomų VNP genotipais, išsamiai išanalizavo Lietuvos populiacijos padėtį artimiausių ir kitų Europos populiacijų atžvilgiu. Šio darbo metu lietuvių populiacijos imtį sudarė asmenys iš visų šešių Lietuvos etnolingvistinių grupių. Nustatyta lietuvių populiacijos padėtis atitiko esamą geografinę padėtį kitų populiacijų atžvilgiu. Arčiausiai išsidėstė kitos Pabaltijo populiacijos (Latvija, Estija) ir slavų populiacijos (Lenkija, Rusija) (Nelis *et al.* 2009). Kituose tyrimuose įtrauktų Lietuvos populiacijos mėginių pasiskirstymas tarp šešių Lietuvos etnolingvistinių grupių nėra aiškiai nurodytas arba įtraukti į tyrimą mėginiai atstovauja tik tam tikrai etninei grupei iš Lietuvos (Behar *et al.* 2010; Mendizabal *et al.* 2012). Šis darbas pirmą kartą analizuoja Lietuvos populiacijos etnolingvistinių grupių genetinę struktūrą ir padėtį HapMap3 populiacijų atžvilgiu, naudojant tolygiai po genomą pasiskirsčiusius autosomų VNP.

Plataus masto tolygiai po genomą pasiskirsčiusių VNP tyrimas leidžia nustatyti daugelio polimorfinių variantų alelių įvairovės pasiskirstymą tiriamoje populiacijoje. Gauti pirminiai duomenys reikalauja griežto kokybės įvertinimo, po kurio dalis tiriamųjų ir daugelis VNP turi būti pašalinti iš tolesnės analizės. Atlikta gautų pirminių duomenų kokybės kontrolė sumažina analizės metu gautų rezultatų paklaidą. Atlikus pirminį ir antrinį tirtos lietuvių populiacijos gautų duomenų kokybės įvertinimą, į tolesnę analizę įtraukti 188 asmenys ir 590 655 arba 105 387 ( $r^2 < 0,2$ ) autosomų VNP. Iš tolesnei analizei įtrauktų VNP pasiskirstymo pagal MAF (4.28 pav.) matyti, kad maža dalis VNP yra reti (MAF < 5 %) variantai – atitinkamai ~8 % ir 15 % visų VNP.

Atlikus visų šešių Lietuvos etnolingvistinių grupių analizę MDS metodu ir vizualizavus gautus rezultatus, matoma tiriamųjų iš skirtingų grupių sanklota. Analizuojant atskiras etnolingvistinių grupių poras ((1) Pietų Aukštaitija ir

Vakarų Žemaitija; (2) Pietų Aukštaitija ir Pietų Žemaitija; (3) Pietų ir Vakarų Aukštaitija; (4) Pietų ir Rytų Aukštaitija) matyti atsiskiriantys klasteriai. Didžiausias skirtumas – tarp Šiaurės Žemaitijos ir Pietų Aukštaitijos. Skirtumai tarp šių grupių gauti ir analizuojant Y chromosomos haplogrupių ir haplotipų įvairovę. Atskirus klasterius taip pat galima identifikuoti tarp Vakarų Žemaitijos ir Pietų Aukštaitijos, Pietų Žemaitijos ir Pietų Aukštaitijos. Didesniu plotu sutampantys klasteriai susidaro tarp Pietų ir Vakarų Aukštaitijos, Pietų ir Rytų Aukštaitijos. Visu plotu sutampančios grupės matomos Žemaitijos viduje. Šie rezultatai leidžia manyti, kad Aukštaitijos etnolingvistinių grupių diferenciacija yra didesnė, palyginti su homogenišku pasižyminčios Žemaitijos etnolingvistinėmis grupėmis. Analizuojant tris ar keturias etnolingvistines grupes (Šiaurės Žemaitija, Rytų Aukštaitija/Vakarų Aukštaitija, Pietų Aukštaitija) taikant MDS ir vizualizavus rezultatus, pirmoje dimensijoje matomas išsidėstymas pagal pietryčių–šiaurės vakarų ašį.

Vizualizuoti PKA rezultatai artimi rezultatams, gautiems MDS metodu. PKA vertinta vidinė, t. y. tarp tirtų asmenų etnolingvistinių grupių viduje, įvairovės variacija. Vizualizavus PKA rezultatus, keturių etnolingvistinių grupių (Šiaurės Žemaitija, Rytų Aukštaitija/ Vakarų Aukštaitija, Pietų Aukštaitija) klasteriai PK1 išsidėsto pagal pietryčių–šiaurės vakarų ašį. Statistiškai reikšmingos PK1 aprašoma variacija sudaro 0,61 % visos variacijos (188 asmenys, 105 387 VNP). Atrinkus pirmus 20 VNP, kurie pasižymi didžiausia variacija statistiškai reikšmingoje PK1, tik šeši yra skirtingų genų intronuose. Šiuos rezultatus lemia duomenims gauti naudotas bendrinis DNR genotipavimo lustas, kuriuo nustatomi tolygiai po visą genomą pasiskirsčiusių VNP genotipai. Neatmestinas nustatytų VNP dalyvavimas genomo reguliacijoje.

Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumais pagal 105 387 autosomų VNP alelių pasiskirstymą, sudarytoje dendrogramoje tirtos šešios etnolingvistinės grupės išsidėsto pagal savo geografinę padėtį – Žemaitija formuoja vieną klasterį, Rytų ir Vakarų Aukštaitija – antrą, labiausiai nutolusi Pietų Aukštaitija – trečią klasterį. Pirmas genetinės įvairovės barjeras atskiria Pietų Aukštaitiją nuo kitų etnolingvistinių grupių. Antras barjeras, nors ir matant

homogeniškumą Žemaitijoje, atskiria Vakarų Žemaitiją nuo kitų etnolingvistinių grupių, trečias barjeras atskiria dvi pagrindines grupes Aukštaitija ir Žemaitija.

Analizuojant LD struktūrą tirtos lietuvių populiacijos grupėje, matomas didelis homogeniškumas tarp etnolingvistinių grupių, išskyrus Vakarų Žemaitiją, kuri pasižymi statistiškai reikšmingai ( $p < 0,05$ ) didesne alelių asociacija ir ilgesniais LD blokais. Rezultatams galėjo turėti įtakos mažesnė tiriamųjų imtis (15 arba 20) lyginant su kitų etnolingvistinių grupių tolygiomis tiriamųjų imtimis (po ~34).

Gautus rezultatus: (1) tėvinės linijos genų fondo galimą judėjimą šiaurės–pietų ašimi; (2) motininės linijos genų fondo galimą judėjimą rytų–vakarų ašimi, papildoma tolygiai po genomą pasiskirsčiusių autosomų žymenų alelių įvairovės analizės rezultatai. Nustatytas tirtų šešių Lietuvos etnolingvistinių grupių išsidėstymas pagal pietryčių–šiaurės vakarų ašį, rodo, ne atskirų populiacijų (tėvinės, motininės linijų), bet galimą bendrą ir pagrindinę migracijos, t. y. genų fondo judėjimo, kryptį.

Prieinamos informacijos trūkumas apribojo tirtos lietuvių populiacijos grupės palyginimą su artimiausiomis populiacijomis. Lietuvių populiacijos grupės padėtis nustatyta HapMap3 projekto populiacijų, kurioms būdingi dideli genetiniai tarpusavio ir su lietuvių populiacija atstumai, atžvilgiu.

Vizualizavus MDS metodu gautus rezultatus, matomos trikampi formuojančios keturios grupės. Afrikos grupės, ypač MKK ir ASW, populiacijos pasižymi didesniais genetiniais skirtumais, palyginti su kitomis grupėmis. Pastarosios populiacijos išsidėsto arčiau Europos klasterio. Didesni genetiniai skirtumai taip pat matomi Šiaurės Amerikos (MXL/MEX) ir Pietų Azijos (GIH) populiacijose. Vizualizavus MDS metodu gautus rezultatus, jos susikloja ir išsidėsto tarp Europos ir Azijos grupių. Azijos grupė pasižymi didžiausiu homogeniškumu – sunku diferencijuoti populiacijas grupės viduje. Skirtumai tarp Europos grupės populiacijų yra tarpiniai tarp Afrikos ir Azijos grupių. Tirta lietuvių populiacija artimesnė CEU populiacijai ir vizualizavus MDS rezultatus, matyti, kad išsidėsto atokiausiai nuo kitų HapMap3 populiacijų. Remiantis genetinė įvairove, nustačius CEU populiacijos padėtį kitų populiacijų atžvilgiu,

matyti, kad CEU gerai reprezentuoja Šiaurės ir Vakarų Europos populiacijas (Lao *et al.* 2008; McEvoy *et al.* 2009). Taip pat nustatyta, kad genetinės įvairovės skirtumai atitinka geografinius atstumus tarp Europos populiacijų (Nelis *et al.* 2009). Trumpiausias geografinis atstumas yra tarp lietuvių ir CEU populiacijos, lyginat su kitomis tirtomis HapMap3 populiacijomis.

Vizualizuotus MDS metodu gautus rezultatus, patvirtina PKA rezultatai. Matomi dideli genetinės įvairovės skirtumai tarp tirtų lietuvių ir HapMap3 populiacijų. PK1 aprašė 5,71 %, PK2 – 3,62 % autosomų VNP alelių įvairovės tarp tirtų asmenų (1 126) iš Lietuvos ir HapMap3 populiacijų. Didžiausia PK1 aprašoma genetinė įvairovė tarp Afrikos ir kitų populiacijų, mažesnė – tarp Azijos ir Europos. PK2 aprašomi genetinės įvairovės skirtumai, iš kurių didžiausias tarp Azijos ir likusių populiacijų, mažesnis tarp Pietų Azijos, Pietų Amerikos ir Europos populiacijų, mažiausias tarp Afrikos ir likusių populiacijų.

Analizuota LD struktūra parodė, kad Šiaurės Amerikos (MXL/MEX) populiacija pasižymi didžiausia, Afrikos grupė mažiausia alelių asociacija ir LD blokais. Afrikos grupės ASW populiacijos LD struktūra išsiskiria iš kitų. Pastarajai populiacijai būdinga mažesnė alelių asociacija, bet didesni LD blokai, palyginti su Europos, Pietų Azijos grupių populiacijomis. Lietuvių populiacijos LD struktūra yra artima kitų Europos ir Pietų Azijos (GIH) populiacijų struktūrai. Rezultatams gali turėti įtakos vidinės lietuvių populiacijos LD struktūros ypatybės. Nustatyta statistiškai reikšmingai didesnė alelių asociacija ir ilgesni LD blokai Vakarų Žemaitijoje lemia tirtos visos lietuvių populiacijos didesnę vidutinę LD  $r^2$  reikšmę.

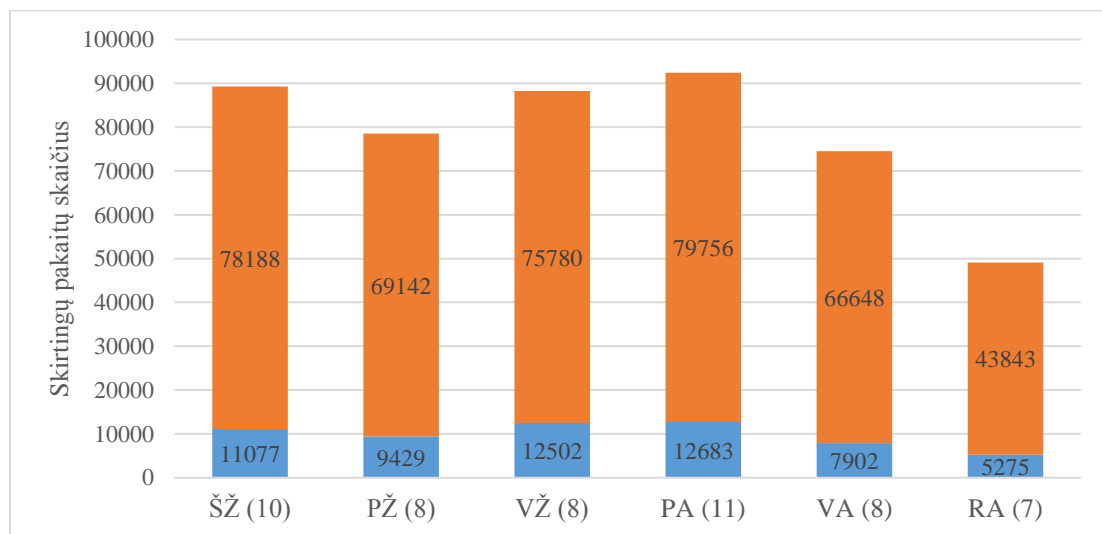
Atrinkus VNP, pasižyminčius didžiausia genetinė įvairove analizuotose PK1–100, sudarytas keturių grupių ( $K = 4$ ) modelis, kuris geriausiai aprašo genetinę variaciją tarp tirtų populiacijų. Šis modelis leidžia atskirti pagrindines tirtas grupes (Afrikos, Azijos, Šiaurės Amerikos, Pietų Azijos, Europos). Keturių grupių modelis leidžia atskirti ir populiacijų kilmę: geriausiai Europos, sunkiau Afrikos ir prastai Azijos grupėse. Matyti, kad tirta lietuvių populiacija lyginant su kitomis Europos populiacijomis (CEU ir TSI), yra toliausiai nuo Azijos populiacijų, t. y. Azijos populiacijų genetinės įvairovės dalis yra

mažiausia. Sudaryto modelio rezultatai artimi vizualizuotiems MDS ir PKA rezultatams, t. y. nustatytiems skirtumams tarp grupių ar jų viduje.

Hierarchinio klasterizavimo UPGMA metodu sudarytoje tirtos lietuvių ir HapMap3 populiacijų dendrogramoje matomos ir kitų analizių metu nustatytos grupės (Afrikos, Azijos, Šiaurės Amerikos, Pietų Azijos, Europos). Europos grupėje pirmas klasteris susidaro tarp CEU ir TSI populiacijų, prie kurių prisijungia ir tirta lietuvių populiacija. Prie Europos grupės prisijungia Šiaurės Amerikos, Pietų Azijos populiacijos, kurių padėtis panaši vizualizuotuose rezultatuose, gautus MDS metodu ir atlikus PKA. Susidaręs klasteris prisijungia prie Azijos grupės ir galiausiai prie labiausiai atsiskyrusios Afrikos grupės.

#### 4.4. Lietuvių populiacijos genetinė struktūra, remiantis egzomo DNR sekos variantų pasiskirstymu

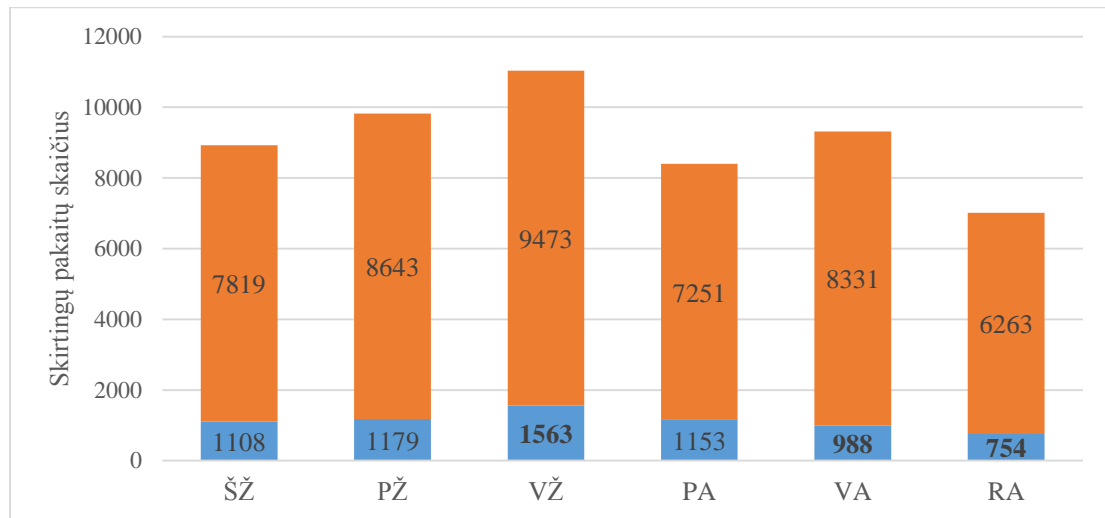
Tirtą lietuvių populiacijos grupę (3.6 pav.) sudarė 52 asmenys, kuriems nustatyta genomo koduojanti dalis – egzomas. Atlikus pirminę, antrinę ir tretinę gautų duomenų analizę, tolesnei analizei atrinkti autosomų VNP, kuriems nustatytas sekos padengimas  $\geq 10$  kartų (referentinio alelio vidutinis kokybės įvertis 28 ( $\pm 2,3$ ), naujo alelio vidutinis kokybės įvertis 28,4 ( $\pm 1,8$ )) (Casals *et al.* 2013). Nustatytų vieno nukleotido pakaitų žmogui vidutinis sekos padengimas 40,5 ( $\pm 49,2$ ), vidutinis pakaitų skaičius tiriamajam  $\sim 28\,890$  ( $\pm 6\,528$ ). Tiroje Lietuvos populiacijos grupėje nustatytos 163 203 skirtingos vieno nukleotido pakaitos, iš kurių 58 865 ( $\sim 36\%$ ) vienetiniai atvejai (angl. *Singleton*), t.y. nustatyti tik vienam tiriamajam. Rastas vienetinių DNR sekos variantų skaičius Aukštaitijoje (26 tiriamieji) – 25 860, Žemaitijoje (26 tiriamieji) – 33 008 pakaitos.



4.45 pav. Skirtingų vieno nukleotido pakaitų (oranžinė spalva) ir vienetinių pakaitų (mėlyna spalva) pasiskirstymas tarp tirtų šešių etnolingvistinių grupių

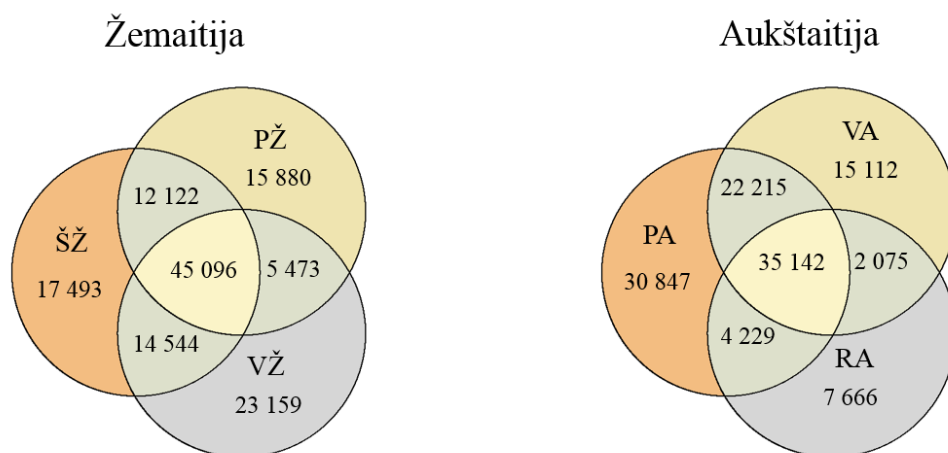
Nustatytas skirtingų vieno nukleotido pakaitų ir vienetinių pakaitų pasiskirstymas tarp tirtų šešių Lietuvos etnolingvistinių grupių. Gauti rezultatai rodo, kad daugiausiai skirtingų variantų (92 439) yra Pietų Aukštaitijoje (11 asmenų), mažiausiai (49 118) – Rytų Aukštaitijoje (7 asmenys) (4.45 pav.). Apskaičiuoti statistiškai reikšmingi skirtumai tarp asmeniui nustatytų skirtingų vienetinių pakaitų skaičiaus ir tikėtino skaičiaus ( $\chi^2$  kriterijus,  $p < 0,05$ ):

(1) didesnis Vakarų Žemaitijoje; (2) mažesnis Vakarų ir Rytų Aukštaitijoje (4.46 pav.).



4.46 pav. Skirtingų vieno nukleotido pakaitų (oranžinė spalva) ir skirtingų vienetinių pakaitų (mėlyna spalva) asmeniui pasiskirstymas tarp tirtų šešių etnolingvistinių grupių. Paryškinti statistiškai reikšmingi skirtumai nuo tikėtino pakaitų skaičiaus

Įvertintas skirtingų vieno nukleotido pakaitų, nustatytų dvejose ir daugiau Lietuvos etnolingvistinių grupių, skaičius. Daugiausia tų pačių vieno nukleotido pakaitų nustatyta tarp Pietų Aukštaitijos, mažiausiai – tarp Rytų Aukštaitijos ir kitų etnolingvistinių grupių. Lyginant skirtingų vieno nukleotido pakaitų pasiskirstymą tarp Lietuvos etnolingvistinių grupių Aukštaitijos ir Žemaitijos (po 26 asmenis) viduje, matyti, kad Žemaitijai būdingas didesnis skaičius pakaitų, nustatomų tarp visų trijų Lietuvos etnolingvistinių grupių, nei Aukštaitijai (4.47 pav.).

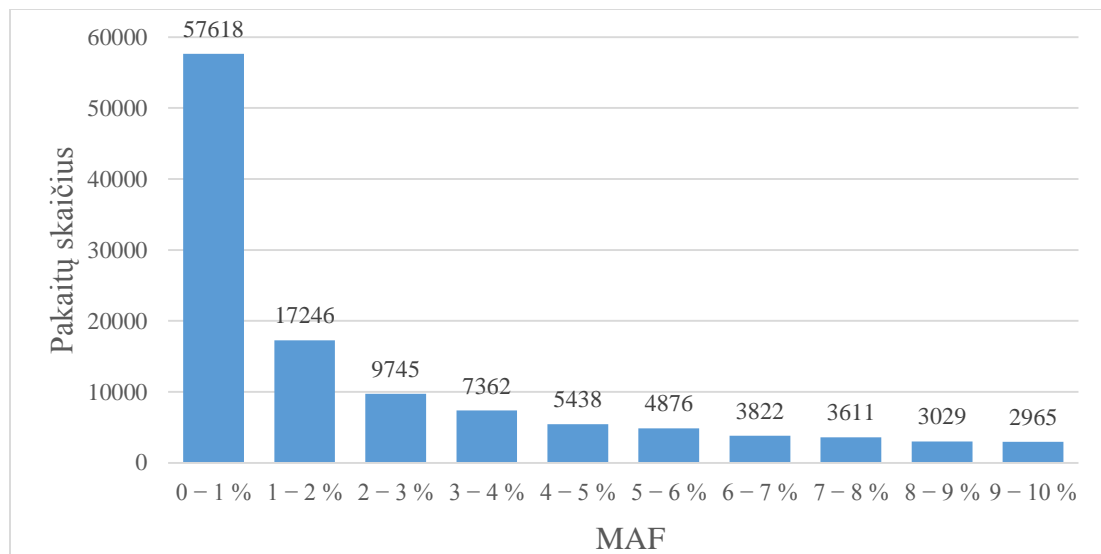


4.47 pav. Skirtingų vieno nukleotido pakaitų, nustatytų dvejose ir daugiau etnolingvistinių grupių, pasiskirstymas



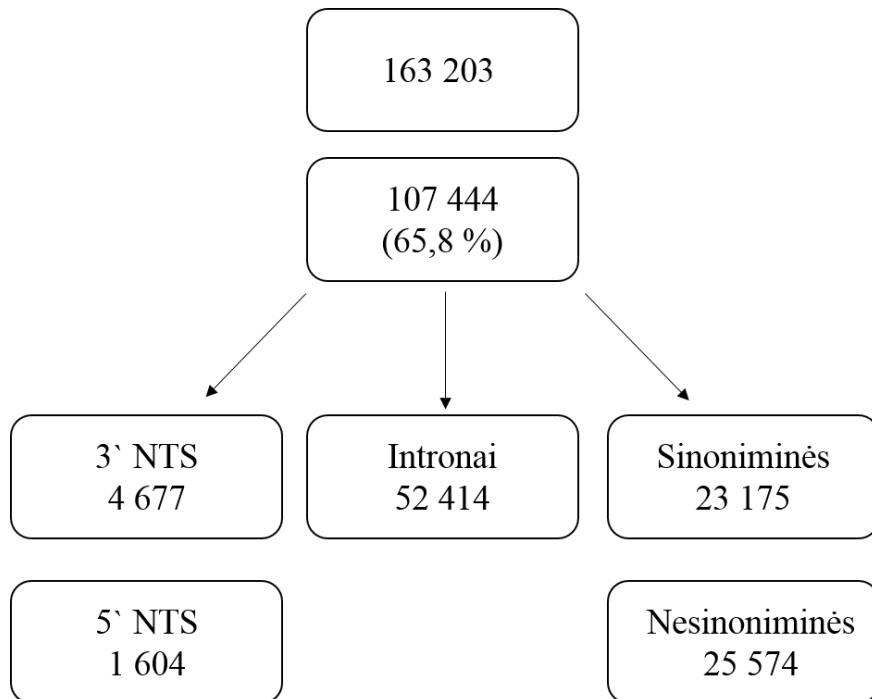
Vidutinis heterozigotinių pakaitų skaičius tiriamajam yra 18 574 atvejai (~64 %), homozigotinių – 10 325 atvejai (~36 %). Vidutinis tranzicijų skaičius tiriamajam yra 21 039 (vidutinis heterozigotinių ir homozigotinių pakaitų santykis – 1,78), transversijų – 7 851 (vidutinis heterozigotinių ir homozigotinių pakaitų santykis – 1,84).

Nustatytas skirtingų vieno nukleotido pakaitų pasiskirstymas pagal MAF. Vieno nukleotido pakaitos, kurių MAF <1 %, sudaro ~35 %, vieno nukleotido pakaitos, kurių MAF <5 %, sudaro ~60 % visų nustatytų skirtingų pakaitų (4.48 pav.).



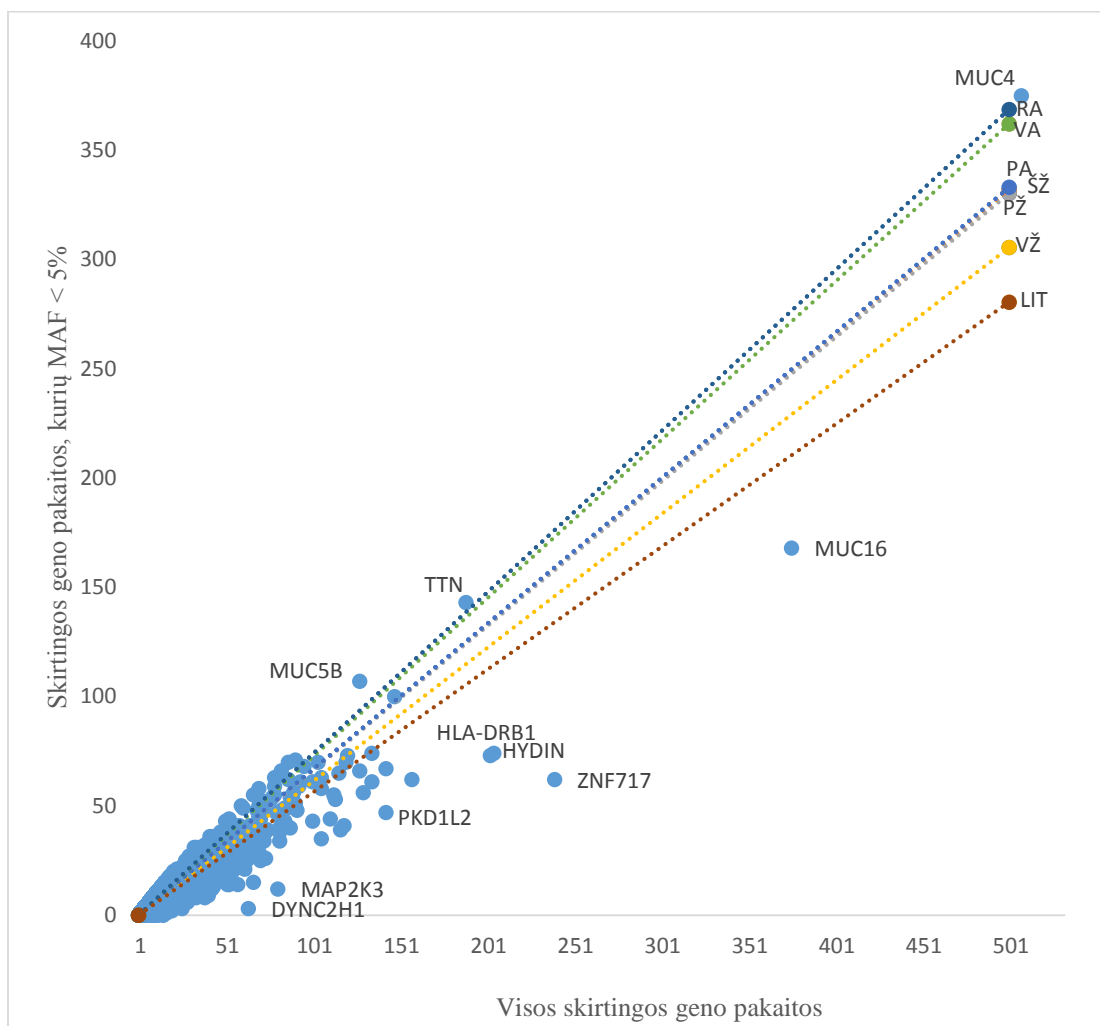
4.48 pav. Egzomo skirtingų pakaitų skaičiaus pasiskirstymas pagal MAF

Grupei priskirti 65,8 % visų skirtingų nustatytų pakaitų (4.49 pav.). Vidutiniškai tiriamajam vienai iš penkių grupių priskirtas pakaitų skaičius: (1) 3' netransliuojama sritis (3' NTS) – 1 008 pakaitos; (2) 5' NTS – 337 pakaitos; (3) intronuose – 10 455 pakaitos; (4) koduojančioje dalyje – 6 864 sinoniminės pakaitos; (5) koduojančioje dalyje – 7 057 nesinoniminės pakaitos. Vienai iš penkių grupių priskiriama ~89 % visų tiriamojo nustatytų pakaitų.



4.49 pav. Skirtingų vieno nukleotido pakaitų pasiskirstymas tarp grupių. Grupėms priskirta 65,8 % visų skirtingų vieno nukleotido pakaitų

Sudarytas tiesinės regresijos modelis tarp gene nustatytų visų ir retų pakaitų, kurių  $MAF < 5\%$ , tirtoje lietuvių populiacijos grupėje. Gauti rezultatai rodo genus, kuriems nustatytas didesnis retų variantų ( $MAF < 5\%$ ) (*MUC4*, *MUC5B*, *TTN*) arba didesnis dažnų (*MUC16*, *ZNF717*, *HLA-DRB1*, *PKD1L2*, *MAP2K3*, *DYNC2H1*) variantų skaičius iš visų to geno variantų skaičiaus. Taip pat tiesinės regresijos modelis tarp geno visų ir retų pakaitų, kurių  $MAF < 5\%$ , sudarytas kiekvienai Lietuvos etnolingvistinei grupei atskirai. Gauti rezultatai rodo, kad trims Lietuvos etnolingvistinėms grupėms (Šiaurės, Pietų Žemaitija, Pietų Aukštaitija) būdingas artimas retų geno variantų skaičius. Rytų ir Vakarų Aukštaitijai būdinga daugiau retesnių, Vakarų Žemaitijai – daugiau dažnesnių geno variantų (4.50 pav.).



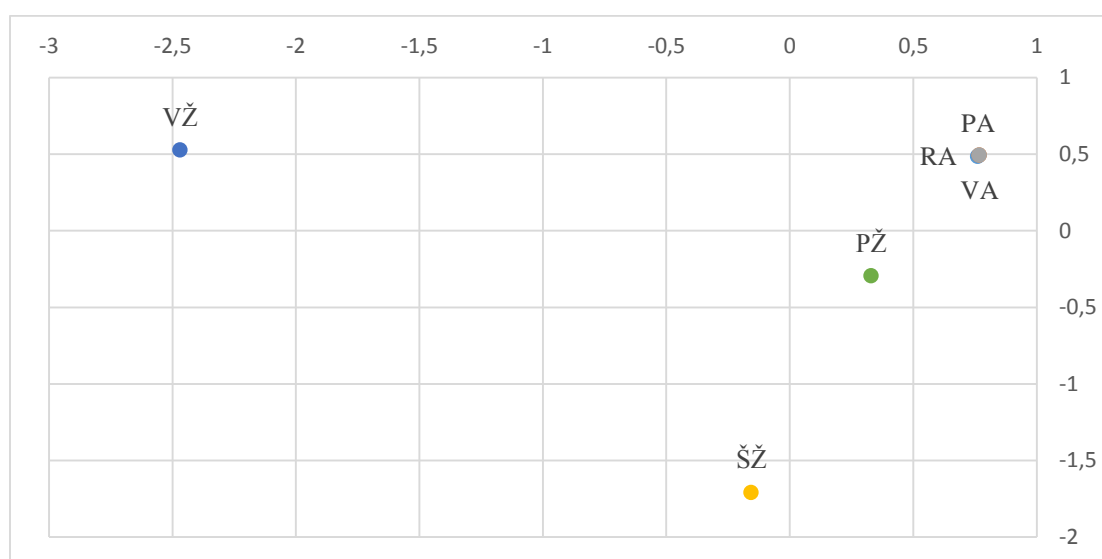
4.50 pav. Visos skirtingos geno pakaitos ir skirtingos geno pakaitos, kurių MAF < 5 %. Linija – atskiros etnolingvistinės grupės ar tirtos Lietuvos populiacijos tiesinė regresinė kreivė. Trumpinys: LIT – tirta lietuvių populiacijos grupė

MDS ir PKA atrinktos variabilios DNR sekos pozicijos (6 939), kurios nustatytos bent 80 % tirtos imties, norint išvengti skirtumų dėl trūkstamos informacijos. Nustatyti poriniai skirtumai tarp šešių Lietuvos etnolingvistinių grupių.

Taikant MDS, remiantis poriniais skirtumais pagal 6 939 autosomų vieno nukleotido pakaitų alelių pasiskirstymą (4.12 lentelė), gauti rezultatai, kuriuos vizualizavus iš tirtų šešių Lietuvos etnolingvistinių grupių Aukštaitijos grupė (Pietų, Vakarų, Rytų Aukštaitija) sudaro vieną klasterį. Žemaitijos etnolingvistinės grupės (Šiaurės ir Vakarų Žemaitija) išsidėsto atokiau – toliausiai nuo Aukštaitijos klasterio ir viena kitos, arčiausiai Aukštaitijos klasterio – Pietų Žemaitija (4.51 pav.).

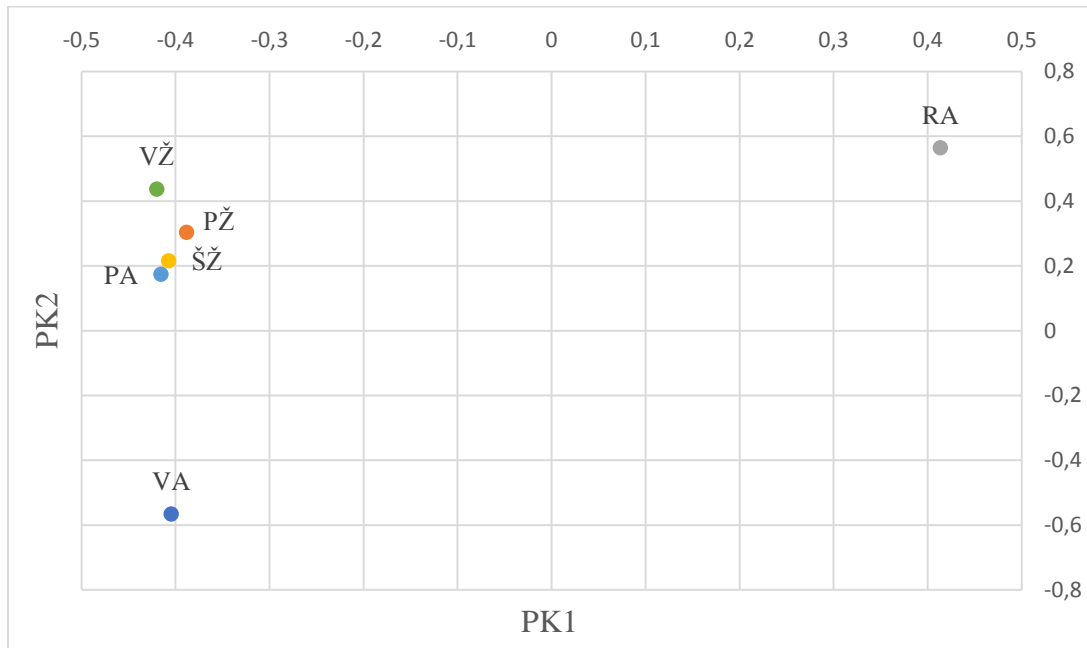
4.12 Porinių skirtumų matrica pagal 6 939 autosomų vieno nukleotido pakaitų alelių pasiskirstymą (žemiau įstrižainės), p reikšmės (aukščiau įstrižainės) tarp tirtų šešių Lietuvos etnolingvistinių grupių

	PA	RA	VA	ŠŽ	VŽ	PŽ
PA	0	<b>0,0009</b>	0,7916	0,7415	<b>0,0065</b>	0,6282
RA	0,0152	0	<b>0,0002</b>	<b>0,0002</b>	<b>0,0002</b>	<b>0,0003</b>
VA	-0,0208	0,0132	0	0,2349	0,0745	0,1782
ŠŽ	-0,0195	0,0258	-0,0209	0	<b>0,0289</b>	0,8319
VŽ	-0,0160	0,0302	-0,0210	-0,0209	0	<b>0,0005</b>
PŽ	-0,0225	0,0232	-0,0240	-0,0258	-0,0218	0

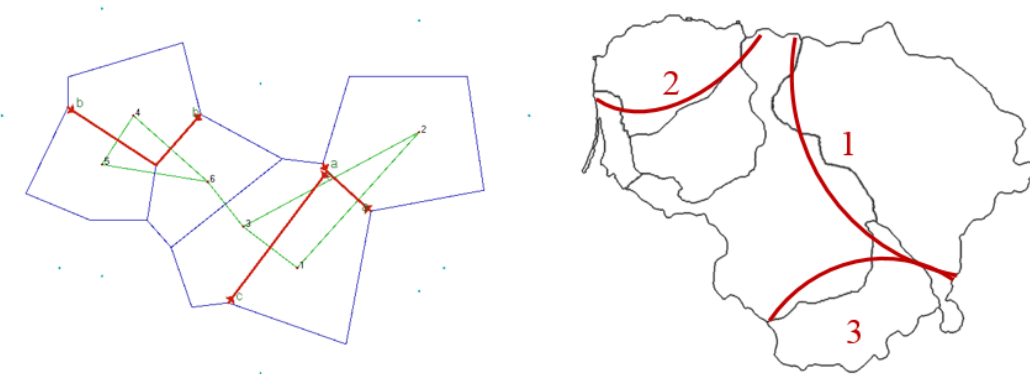


4.51 pav. MDS, remiantis poriniais skirtumais pagal 6 939 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, rezultatai

Atlikus PKA, remiantis poriniais skirtumais pagal 6 939 autosomų vieno nukleotido pakaitų alelių pasiskirstymą, nustatyta, kad PK1 aprašo 75,53 %, PK2 – 9,30 % variacijos tarp tirtų šešių Lietuvos etnolingvistinių grupių. Vizualizavus PKA rezultatus, matyti, kad PK1 aprašoma variacija dėl Rytų Aukštaitijos, PK2 – Vakarų Aukštaitijos alelių įvairovės skirtumų nuo kitų etnolingvistinių grupių (4.52 pav.).

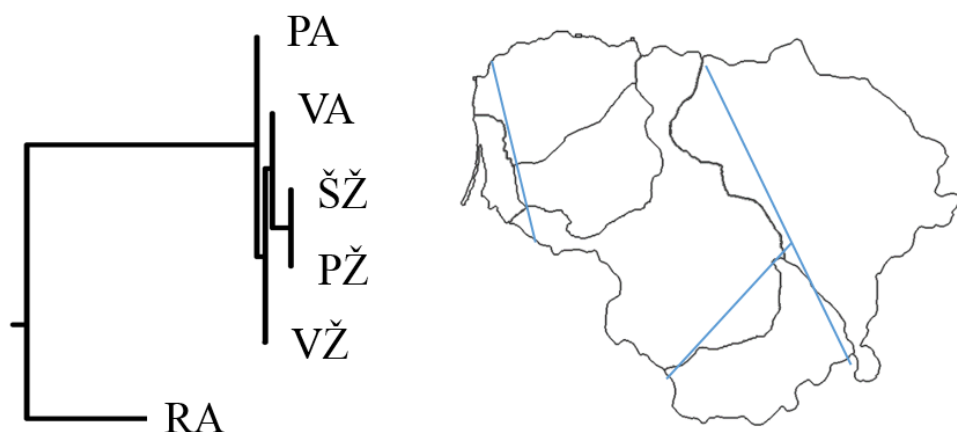


4.52 pav. PK analizės, remiantis poriniais skirtumais pagal 6 939 autosomų VNP alelių pasiskirstymą tirtose šešiose Lietuvos etnolingvistinėse grupėse, rezultatai



4.53 pav. Egzomo autosomų sekos variantų (6 939) alelių įvairovės barjerai tarp tirtų Lietuvos etnolingvistinių grupių. Kairėje šešių Lietuvos etnolingvistinių grupių schema pagal pateiktas koordinatas (2 priedo 1 lentelė) ir barjerai (raudona linija, eiliškumo tvarka nurodyta skaičiais). Analizei naudota  $F_{ST}$  atstumų matrica

Nustatyti egzomo autosomų sekos variantų alelių įvairovės barjerai tarp tirtų šešių Lietuvos etnolingvistinių grupių. Pirmas barjeras atskiria Rytų Aukštaitiją, antras barjeras – Šiaurės Žemaitiją, trečias barjeras – Pietų Aukštaitiją nuo kitų etnolingvistinių grupių (4.53 pav.).



4.54 pav. Hierarchinio klasterizavimo UPGMA metodu, remiantis  $F_{ST}$  atstumų matrica pagal 6 939 egzomo autosomų sekos variantų alelių pasiskirstymą, sudaryta šešių Lietuvos etnolingvistinių grupių dendrograma. Paveikslo dešinėje linijomis atskirti dendrogramoje matomi etnolingvistinių grupių klasteriai

Hierarchinio klasterizavimo UPGMA metodu, remiantis poriniais skirtumais pagal 6 939 egzomo autosomų sekos variantų alelių pasiskirstymą, sudaryta tirtų šešių Lietuvos etnolingvistinių grupių dendrograma. Gautas dvi grupės: (1) Pietų, Vakarų Aukštaitija, Vakarų, Pietų, Šiaurės Žemaitija; (2) Rytų Aukštaitija. Pastaroji grupė (Rytų Aukštaitija) nutolusi didžiausiu atstumu nuo kitų etnolingvistinių grupių (4.54 pav.).

## APTARIMAS

S. Giannatassio kartu su kitais mokslininkais aprašė dažniausių *CFTR* geno mutacijų paplitimą Lietuvos populiacijos cistine fibroze sergančių pacientų grupėje (Giannatassio *et al.* 2006). O. Tighe kartu su kitais mokslininkais išanalizavo dažniausios *PAH* geno mutacijos paplitimą Lietuvos populiacijos fenilketonurija sergančių pacientų grupėje (Tighe *et al.* 2003). Šis mūsų darbas pirmą kartą analizuoja Lietuvos populiacijos etnolingvistinių grupių genetinę struktūrą pagal egzomo autosomų genetinių žymenų alelių pasiskirstymą.

Tirtoje Lietuvos populiacijos grupėje vienam asmeniui nustatyta vidutiniškai ~28 890 ( $\pm 6 528$ ) DNR sekos pakaitų, iš kurių ~64 % heterozigotinio genotipo. Vidutinis tranzicijų ir transversijų santykis žmogui yra

2,7. Nustatyti sekos variantai priskirti penkioms grupėms: (1) 3` NTS – 3,5 %; (2) 5` NTS – 1,2 %; (3) intronuose – 36,2 %; (4) koduojančioje dalyje sinoniminiai – 23,8 %; (5) koduojančioje dalyje nesinoniminiai – 24,4 %, kurie sudaro ~89 % visų vienam asmeniui nustatytų pakaitų.

Tirtoje Lietuvos populiacijos grupėje nustatytos 163 203 skirtingos vieno nukleotido pakaitos, iš kurių 58 865 (~36 %) vienetiniai atvejai, t. y. tik vienam asmeniui iš visos tirtos imties. Nustatytas pastarųjų pakaitų netolygus pasiskirstymas tarp Aukštaitijos ir Žemaitijos (po 26 asmenis) santykiu 5:3. Rasti statistiškai reikšmingi skirtumai tarp nustatomo ir tikėtino DNR sekos vienetinių pakaitų (angl. *Singleton*) skaičiaus žmogui tarp etnolingvistinių grupių: (1) daugiau Vakarų Žemaitijoje; (2) mažiau Vakarų ir Rytų Aukštaitijoje. Vienetinės DNR sekos pakaitos sudaro didžiąją dalį (~60 %) retų variantų (MAF < 5 %). Reti sekos variantai (MAF < 5 %) sudaro didžiąją dalį (~60 %) nustatytų visų skirtingų DNR sekos variantų. Matomas nevienodas skirtingų vieno nukleotido pakaitų, nustatytų dvejose ir daugiau etnolingvistinių grupių, pasiskirstymas tarp Aukštaitijos ir Žemaitijos. Žemaitijai, kuriai būdingas ir didesnis vienetinių DNR sekos pakaitų skaičius (33 008 pakaitos), būdingas didesnis pakaitų, nustatomų tarp visų trijų etnolingvistinių grupių, skaičius. Mažesnis – Aukštaitijai, kuriai būdingas ir mažesnis vienetinių DNR sekos pakaitų skaičius (25 860 pakaitos). Žemaitijai būdingas didesnis skirtingų pakaitų skaičius lyginant su Aukštaitija – atitinkamai 133 777 ir 117 292 pakaitos. Rezultatams reikšmingos įtakos galėjo turėti netolygus gautų duomenų pasiskirstymas.

Retų variantų skaičius tarp skirtingų populiacijų yra nevienodas. Skirtumų priežastis gali būti gamtinė atranka arba populiacijos demografiniai įvykiai. Gamtinė atranka veikia tam tikrus funkcinius regionus, o demografiniai įvykiai turi įtakos visam genomui. Sudarius tiesinės regresijos modelį tarp gene nustatytų visų ir retų pakaitų, kurių MAF < 5 %, tiesės nuolydis atspindi populiacijos demografinius įvykius (Raska, Zhu 2011). Matyti, kad Rytų ir Vakarų Aukštaitijai būdinga daugiau retų DNR sekos variantų gene, palyginti su kitomis Lietuvos etnolingvistinėmis grupėmis. Vakarų Žemaitijai būdingi

dažnesni sekos variantai gene, kitomis grupėms būdingas panašus santykis. Gauti rezultatai rodo galimas skirtingas Lietuvos etnolingvistinių grupių istorijas ir demografinius įvykius. Genai, nutolę nuo regresijos tiesės, rodo galimą gamtinės atrankos poveikį. Nustatyti genai, kuriems būdingas didesnis retų (*MUC4*, *MUC5B*, *TTN*) arba dažnų (*MUC16*, *ZNF717*, *HLA-DRB1*, *PKD1L2*, *MAP2K3*, *DYNC2H1*) DNR sekos variantų skaičius. Mucinų (angl. *MUC gene family*) genų šeimos nariai koduoja epitelio ląstelių paviršiaus glikoproteinus, reguliuojančius patogenų patekimą į ląstelę. Titino genas (*TTN*) šiuo metu žinomas kaip ilgiausią koduojančią dalį turintis genas. Didesnis sekos variantų, turinčių retus alelius (MAF < 5 %), skaičius gene galimas dėl gamtinės atrankos poveikio geno produktams, dalyvaujantiems specifinėse organizmo apsaugos funkcijose, t. y. susiduriant su skirtingais išorės patogenais, ar dėl fizinio geno ilgio. *HLA-DRB1* (angl. *Major Histocompatibility Complex, Class II, DR Beta 1*) geno produktas dalyvauja imuninėje sistemoje. *ZNF717* (angl. *Zinc Finger Protein 717*) geno produktas – transkripcijos veiksnys. *PKD1L2* (angl. *Polycystic Kidney Disease 1-Like 2*) geno produktas – katijonų kanalo sudedamoji dalis. *MAP2K3* (angl. *Mitogen-activated Protein Kinase Kinase 3*) geno produktas – MAP kinazių reguliuojamo signalinio kelio narys. *DYNC2H1* (angl. *Dynein, Cytoplasmic 2, Heavy Chain 1*) geno produktas – dineinas, dalyvaujanti pernašoje ląstelės viduje. Didesnis sekos variantų, turinčių dažnus alelius (MAF > 5 %), skaičius gene galimas dėl gamtinės atrankos poveikio geno produktams, kurie sąveikauja su daugeliu kitų genų produktų ir dalyvauja bendrinėse organizmo funkcijose.

Siekiant išvengti trūkstamos informacijos įtakos analizės rezultatams, buvo atrinkti egzomo DNR sekos variantai (6 939), kurie nustatyti  $\geq 80$  % asmenų. Šių egzomo autosomų žymenų pasiskirstymas pagal MAF yra tolygus. Žymenys, kurių MAF < 5 %, sudaro 12,8 %. Remiantis apskaičiuotais poriniais skirtumais pagal 6 939 autosomų VNP alelių įvairovę, sudaryta MDS ir dendrograma, atlikta PKA analizė, nustatyti genetinės įvairovės barjerai.

Vizualizavus MDS metodu gautus rezultatus, Aukštaitijos etnolingvistinės grupės formuoja glaustą klasterį, kitos Lietuvos etnolingvistinės grupės



(Žemaitijos) išsidėsto atokiau. Toliausiai nuo Aukštaitijos klasterio ir viena kitos išsidėsto Šiaurės ir Vakarų Žemaitija, arčiausiai Aukštaitijos klasterio – Pietų Žemaitija.

PKA vizualizuoti rezultatai skiriasi nuo rezultatų, gautų MDS metodu. PK analize nevertinta vidinės įvairovės variacija. PK1 (75,53 %) aprašo Rytų Aukštaitijos įvairovės skirtumus nuo kitų etnolingvistinių grupių. PK2 (9,30 %) aprašo Vakarų Aukštaitijos įvairovės skirtumus nuo kitų etnolingvistinių grupių.

Pirmas genetinės įvairovės barjeras atskiria Rytų Aukštaitiją nuo kitų etnolingvistinių grupių. Šis rezultatas patvirtina PKA rezultatus. Antras genetinės įvairovės barjeras atskiria Šiaurės Žemaitiją, trečias – Pietų Aukštaitiją nuo kitų etnolingvistinių grupių. Dendrogramoje Rytų Aukštaitija išsidėsto atokiausiai nuo kitų grupių, kurios sudaro bendrą glaustą šaką. Pirmiausia susijungia Šiaurės ir Pietų Žemaitija, prie jų prisijungia Vakarų Aukštaitija – visos trys grupės iš Šiaurės Vakarų ir Vidurio Lietuvos. Vakarų Žemaitija ir Pietų Aukštaitija prisijungia prie Vidurio Lietuvos grupės pamato.

Skirtingais analizės metodais gauti rezultatai pagal egzomo autosomų žymenis yra netolygūs. Vizualizavus rezultatus, gautus taikant MDS, matomas Aukštaitijos etnolingvistinių grupių išskirtinis homogeniškumas, tačiau PK1 aprašoma variacija yra dėl Rytų Aukštaitijos įvairovės skirtumų nuo kitų etnolingvistinių grupių. Rytų Aukštaitijos genetinės įvairovės skirtumai suformuoja pirmiausią barjerą, kuris atskiria ją nuo kitų etnolingvistinių grupių. Dendrogramoje Rytų Aukštaitija taip pat formuoja atskirą šaką nuo kitų etnolingvistinių grupių suformuoto glaudaus klasterio. Daugelis rezultatų patvirtina Rytų Aukštaitijos įvairovės skirtumus nuo kitų etnolingvistinių grupių. Naudoti analizės metodai neleidžia nustatyti genetinės įvairovės skirtumų tarp kitų etnolingvistinių grupių.

#### 4.5. Apibendrinimas

Pirmieji skirtumai tarp Lietuvos populiacijos atskirų geografinių regionų nustatyti remiantis fenetiniais atstumais – pirštų ir delnų dermatoglifiniai požymiai (Klevcova, Kučinskas 1987). Skirtumai tarp etnolingvistinių grupių

taip pat nustatyti tiriant AB0 ir Rh(D) kraujo grupių pasiskirstymą Lietuvos populiacijoje. Remiantis gautais rezultatais, matyti, kad Žemaitijai būdingas didesnis homogeniškumas negu Aukštaitijai (Kučinskas *et al.* 1994). Pastarieji rezultatai aiškinami Žemaitijos autochtoniškumu, o Aukštaitijos populiacijos vėlesniu atsikraustymu (Rytų baltai) (Kučinskas 2004). Kraujo grupių sistemų (AB0, Rh, MNS, P, LU, KELL, Co, LE, FUT2, FY, JK, LW) alelių dažnių pasiskirstymo analizės rezultatai parodė Pietų Aukštaitijos statistiškai reikšmingus skirtumus nuo visų kitų etnolingvistinių grupių, remiantis P1 ir LW<sup>b</sup> alelių dažniais. Tiriant serumo baltymų genetinius polimorfizmus (TF, GC, PI) statistiškai reikšmingų skirtumų nerasta. Analizuojant *Alu* sekas, nustatyti statistiškai reikšmingi skirtumai tarp Šiaurės Žemaitijos ir Pietų Aukštaitijos, remiantis *Alu* TPA25 pasiskirstymu tirtoje Lietuvos populiacijoje (Kučinskas 2001). Pirmų mtDNR restrikcijos fragmentų ilgio polimorfizmo (RFIP) tyrimų rezultatai parodė minimalius skirtumus tarp Aukštaitijos ir Žemaitijos (Kučinskas 1994). Tolesnius tyrimus lėmė metodų ir technologijų vystymasis. Tai leido detaliau įvertinti Lietuvos populiacijos genetinę įvairovę naudojant skirtingus žymenis (Kasperavičiūtė *et al.* 2004). Ankstesniais tyrimais analizuota ne tik Lietuvos populiacijos vidinė genetinė struktūra, bet ir padėtis kitų populiacijų atžvilgiu, remiantis skirtingų žymenų pasiskirstymu. Dalies žymenų nustatytas dažnis Lietuvos populiacijoje atitinka dažnio gradientą Europoje, t. y. geografinę padėtį, tačiau esama ir išimčių (Kučinskas 2004). Nustatyta Lietuvos populiacijos padėtis, remiantis mtDNR ir Y chromosomos įvairovės pasiskirstymu, atitinka geografinę padėtį, t. y. ji artimiausia kaimyninėms populiacijoms (Kasperavičiūtė *et al.* 2004).

Detalesnė žymenų analizė suteikia daugiau informacijos apie tiriamos populiacijos genetinę struktūrą, haplotipų, haplogrupių pasiskirstymą geografinėje teritorijoje, leidžia tiksliau įvertinti populiacijos ar filogenetinių linijų amžių (Gunnarsdottir *et al.* 2011; Underhill *et al.* 2010; Underhill *et al.* 2014; Wei *et al.* 2013). Nustatytas ne tik Y chromosomos ar mtDNR, taip pat ir autosomų, žymenų alelių netolygus pasiskirstymas tarp populiacijų (Menozzi *et al.* 1978). Naujų metodų ir technologijų kūrimas leido populiacijų tyrimus

išplėsti nuo kelių iki kelių šimtų tūkstančių žymenų genotipavimo vienu metu ar stambių genomo sričių sekoskaitos (IHC 2003; Lander *et al.* 2001). Skirtingi žymenys leidžia analizuoti atskirus populiacijų evoliucijos aspektus.

Šiame darbe Lietuvos populiacijos etnolingvistinių grupių struktūra ir padėtis kitų populiacijų atžvilgiu analizuota remiantis keturiomis genomo žymenų grupėmis. Tai: (1) Y chromosomos haplogrupės ir haplotipai; (2) mtDNR haplogrupės ir haplotipai; (3) tolygiai po visą genomą pasiskirstę VNP; (4) egzomo DNR sekos variantai.

Aiškiausi skirtumai matomi analizuojant šešias Lietuvos etnolingvistines grupes pagal Y chromosomos haplogrupių ir haplotipų įvairovės pasiskirstymą. Gauti panašūs rezultatai analizuojant skirtingo tipo žymenis: (1) haplogrupės, kurių nustatoma įvairovė susidarė dėl sukauptos genetinės variacijos sąlyginai ilgesniu laikotarpiu; (2) haplotipai, kurių įvairovė susidarė sąlyginai trumpesniu laikotarpiu. Matomas pagrindinių haplogrupių N1c1, R1a1a ir R1a1a1g šiaurės–pietų ašies gradientas tirtoje Lietuvos populiacijoje. Pastarųjų haplogrupių paskirstymas ir lemia gautus šio darbo rezultatus. Analizuojant Y chromosomos haplogrupių ir haplotipų pasiskirstymą, galima manyti, kad genetinės įvairovės skirtumus tarp tirtų šešių etnolingvistinių grupių lėmė ne kalbinės ar kultūrinės priežastys, bet geografinė padėtis Lietuvos teritorijoje. Gauti rezultatai rodo, kad Lietuvos etnolingvistinių grupių Y chromosomos genetinės įvairovės pasiskirstymas pagal geografinę padėtį galėjo susidaryti dėl genetinio fondo judėjimo šiaurės–pietų ašimi Lietuvos teritorijoje. Naudojant Y chromosomos haplotipų duomenis ir pasirinktą modelį apskaičiuotas dažniausių haplogrupių amžius yra artimas, tik R1a1a haplogrupės amžius šiek tiek didesnis. Šie duomenys neleidžia daryti išvadų apie skirtingą pastarųjų haplogrupių formavimosi pradžią, bet leidžia manyti, kad formavimosi istorijos buvo skirtingos.

Mažesni, tačiau tolygūs skirtumai matomi analizuojant šešias Lietuvos etnolingvistinės grupes pagal tolygiai po visą genomą pasiskirsčiusių VNP alelių įvairovę. Naudojami tolygiai po genomą pasiskirstę žymenys, kurių MAF > 5 %, leidžia analizuoti ilgesniu laikotarpiu susidariusią populiacijos struktūrą.

Nustačius griežtus kokybės įverčius turimiems duomenims, tolesnei analizei įtraukti negiminingų asmenų VNP, kuriems nustatyta pusiausvira sankiba su kitais VNP. Skirtingais analizės metodais gauti rezultatai rodo genetinės įvairovės skirtumus šiaurės vakarų–pietryčių ašimi. Galima skirtumų priežastis yra ne kalbinė ar kultūrinė, bet geografinė padėtis. Nustatyta etnolingvistinių grupių išsidėstymo ašis artima, bet netapati etnolingvistinių grupių padėčiai pagal Y chromosomos įvairovės skirtumus. Nustatytas gradientas rodo galimą pirmųjų gyventojų genetinio fondo pagrindinę judėjimo dabartinėje Lietuvos teritorijoje ašį. Minimalius skirtumus tarp gautų rezultatų, naudojant Y chromosomos ir tolygiai po genomą išsidėsčiusius žymenis, galėjo lemti imties sudarymo paklaidos. Paskutinis iš tirtų trijų genetinės įvairovės barjerų matomas tarp Aukštaitijos ir Žemaitijos, tačiau jis yra labiau dėl geografinės grupių padėties nei kalbinių ar kultūrinių skirtumų. Taip pat matyti, kad Žemaitija yra homogeniška, o asmenis iš Aukštaitijos, nors ir esant didelei paklaidos tikimybei, galima priskirti etnolingvistinei grupei. Šie rezultatai patvirtina ankstesnius kraujo grupių tyrimus (Kučinskas *et al.* 1994).

Dalis rezultatų rodo skirtumus tarp tirtų šešių Lietuvos etnolingvistinių grupių pagal mtDNR haplogrupių ir haplotipų pasiskirstymą. Tačiau rezultatai, gauti analizuojant skirtingais metodais, nėra tolygūs. Nenustatytas mtDNR haplogrupių gradientas tarp tirtų Lietuvos etnolingvistinių grupių. Šiame darbe nustatytas filogenetinės linijos U netolygus pasiskirstymas nebuvo pastebėtas ankstesniuose tyrimuose (Kasperavičiūtė *et al.* 2004) ir gali būti dėl imties sudarymo paklaidos. Dažniausių filogenetinių linijų H (45,3 %) ir U (19,2 %) pasiskirstymas neturi aiškios gradiento ašies, ir skirtingų analizės metodų rezultatai priklauso nuo daugelio haplogrupių pasiskirstymo tarp tirtų šešių etnolingvistinių grupių. AMOVA rezultatai rodo mtDNR haplogrupių ir haplotipų pasiskirstymo vakarų–rytų ašį. Vizualizavus rezultatus, gautus MDS metodu ar atlikus PKA, etnolingvistinės grupės išsidėsto atokiai viena nuo kitos arba sudaro klasterius, kurie nepatvirtina AMOVA rezultatų. Pirmas mtDNR įvairovės barjeras atskiria šiaurės vakarų etnolingvistines grupes (Vakarų ir Šiaurės Žemaitija), antras – Rytų Aukštaitiją, trečias Aukštaitiją ir Žemaitiją.

Nustatyti mtDNR įvairovės barjerai artimi AMOVA nustatyti vakarų–rytų ašiai. Filogenetinių linijų H ir U filogenetiniai medžiai, sudaryti pagal koduojančios dalies, kuriai būdingas mažesnis mutacijų dažnis, haplotipus (577–16 023 np), rodo haplogrupės H1 augimą. Gauti rezultatai išnyksta sudarius filogenetinį medį pagal visos mtDNR haplotipų (16 564 np) įvairovę. mtDNR haplotipų (16 564 np) įvairovė leido įvertinti atskirų haplogrupių amžių. Gauti rezultatai neprieštaruoja galimam filogenetinės linijos U dominavimui tarp pirmųjų gyventojų dabartinėje Lietuvos teritorijoje ir vėlesniam filogenetinės linijos H populiacijos atsikraustymui.

Egzomo autosomų DNR sekos variantų apibendrinimui turi įtakos duomenų trūkumas (angl. *Missing data*) dėl netolygaus gautų duomenų pasiskirstymo genome. Nustatytų skirtumų patvirtinimui reikalinga išplėsta tiriamųjų imtis ir tolygus duomenų pasiskirstymas. Norint išvengti pastarosios problemos, atrinkti egzomo autosomų DNR sekos variantai, kurie nustatyti mažiausiai 80 % tirtų asmenų. Šių DNR sekos variantų pasiskirstymas pagal MAF yra tolygus ir nelieta retų sekos variantų ( $MAF < 5\%$ ) dominavimo (12,8 %). Rezultatai, gauti naudojant skirtingus analizės metodus, nėra tolygūs ir matomi tik Rytų Aukštaitijos genetinės įvairovės skirtumai nuo kitų etnolingvistinių grupių, kurie pasikartoja atlikus analizę keliais metodais. Retų variantų ( $MAF < 5\%$ ), kurie genome susidarė palyginti neseniai ir leidžia įvertinti vėlesnius populiacijos evoliucijos etapus, pasiskirstymo lietuvių populiacijoje įvertinimas reikalauja didinti tolygios apimties duomenų imtį.

## 5. IŠVADOS

1. Remiantis Y chromosomos haplogrupių ir haplotipų įvairovės pasiskirstymu:
  - 1.1. Tirta lietuvių populiacija pasižymi homogeniškumu, skirtumai tarp grupių, sudarytų remiantis geografine Lietuvos etnolingvistinių grupių padėtimi, sudaro <2 %.
  - 1.2. Nustatyti genetinės įvairovės skirtumai, nulemti pagrindinių haplogrupių (N1c1, R1a1a, R1a1a1g) gradiento šiaurės–pietų ašimi.
  - 1.3. Galimos skirtingos pagrindinių haplogrupių N1c1, R1a1a populiacijų formavimosi Lietuvos teritorijoje istorijos.
  - 1.4. Tirta lietuvių populiacija yra artimiausia Pabaltijo populiacijoms (Latvija ir Estija) ir padėtis kitų analizuotų populiacijų atžvilgiu priklauso nuo geografinių atstumų.
2. Remiantis mtDNR haplogrupių ir haplotipų įvairovės pasiskirstymu:
  - 2.1. Tiroje lietuvių populiacijoje didžiausia dalį sudaro vidinė genetinė įvairovė, t. y. tarp tiriamų asmenų etnolingvistinių grupių viduje, kuri siekia ~99 %.
  - 2.2. Nustatyti genetinės įvairovės vakarų–rytų ašimi skirtumai, kuriuos lemia pagrindinių (H ir U) ir kitų filogenetinių linijų (A, HV, I, J, K, M, N, T, V, W) haplogrupių pasiskirstymas tiroje lietuvių populiacijoje.
  - 2.3. Filogenetinė linija U galėjo dominuoti tarp pirmųjų dabartinėje Lietuvos teritorijoje gyventojų, kuriuos galėjo pakeisti filogenetinės linijos H paplitimas vėlesniu laikotarpiu.
  - 2.4. Tirta lietuvių populiacija išsidėsto tarp analizuotų slavų ir Vidurio Azijos populiacijų.
3. Remiantis tolygiai po genomą pasiskirsčiusių VNP alelių įvairove:
  - 3.1. Žemaitijos etnolingvistinės grupės yra homogeniškesnės, t. y. nustatyti mažesni skirtumai, palyginti su Aukštaitijos etnolingvistinėmis grupėmis.
  - 3.2. Nustatyti šiaurės vakarų–pietryčių ašimi genetinės įvairovės skirtumai tiroje lietuvių populiacijoje.

- 3.3. Tirta lietuvių populiacija, kuriai nustatyta mažiausia Azijos genetinio fondo dalis, yra artima HapMap3 Europos populiacijų grupei.
4. Remiantis egzomo autosomų DNR sekos variantų alelių įvairove tarp tirtų šešių Lietuvos etnolingvistinių grupių, nustatytas Rytų Aukštaitijos genetinės įvairovės skirtumas nuo kitų penkių Lietuvos etnolingvistinių grupių.
  5. Genetinės įvairovės skirtumai tarp tirtų Lietuvos etnolingvistinių grupių matomi pagal tolygiai po genomą pasiskirsčiusių VNP alelių, mtDNR, Y chromosomos haplogrupių ir haplotipų pasiskirstymą, o didžiausi skirtumai būdingi Y chromosomos genetinei įvairovei.

## 6. LITERATŪROS SĄRAŠAS

1. Abdellaoui A., Hottenga J. J., de Knijff P., Nivard M. G., Xiao X., Scheet P., Brooks A., Ehli E. A., Hu Y., Davies G. E., Hudziak J. J., Sullivan P. F., van Beijsterveldt T., Willemsen G., de Geus E. J., Penninx B. W. and Boomsma D. I. (2013). Population Structure, Migration, and Diversifying Selection in the Netherlands. *Eur J Hum Genet* 21(11): 1277-85.
2. Abecasis G. R., Altshuler D., Auton A., Brooks L. D., Durbin R. M., Gibbs R. A., Hurles M. E. and McVean G. A. (2010). A Map of Human Genome Variation from Population-Scale Sequencing. *Nature* 467(7319): 1061-73.
3. Abecasis G. R., Auton A., Brooks L. D., DePristo M. A., Durbin R. M., Handsaker R. E., Kang H. M., Marth G. T. and McVean G. A. (2012). An Integrated Map of Genetic Variation from 1,092 Human Genomes. *Nature* 491(7422): 56-65.
4. Achilli A., Rengo C., Battaglia V., Pala M., Olivieri A., Fornarino S., Magri C., Scozzari R., Babudri N., Santachiara-Benerecetti A. S., Bandelt H. J., Semino O. and Torroni A. (2005). Saami and Berbers--an Unexpected Mitochondrial DNA Link. *Am J Hum Genet* 76(5): 883-6.
5. Altshuler D. M., Gibbs R. A., Peltonen L., Dermitzakis E., Schaffner S. F., Yu F., Bonnen P. E., de Bakker P. I., Deloukas P., Gabriel S. B., Gwilliam R., Hunt S., Inouye M., Jia X., Palotie A., Parkin M., Whittaker P., Chang K., Hawes A., Lewis L. R., Ren Y., Wheeler D., Muzny D. M., Barnes C., Darvishi K., Hurles M., Korn J. M., Kristiansson K., Lee C., McCarroll S. A., Nemesh J., Keinan A., Montgomery S. B., Pollack S., Price A. L., Soranzo N., Gonzaga-Jauregui C., Anttila V., Brodeur W., Daly M. J., Leslie S., McVean G., Moutsianas L., Nguyen H., Zhang Q., Ghorri M. J., McGinnis R., McLaren W., Takeuchi F., Grossman S. R., Shlyakhter I., Hostetter E. B., Sabeti P. C., Adebamowo C. A., Foster M. W., Gordon D. R., Licinio J., Manca M. C., Marshall P. A., Matsuda I., Ngare D., Wang V. O., Reddy D., Rotimi C. N., Royal C. D., Sharp R. R., Zeng C., Brooks L. D. and McEwen J. E. (2010). Integrating Common and Rare Genetic Variation in Diverse Human Populations. *Nature* 467(7311): 52-8.
6. Athey T. W. (2005). Haplogroup Prediction from Y-STR Values Using an Allele-Frequency Approach. *Journal of Genetic Genealogy*(1): 1-7.
7. Atkinson Q. D., Gray R. D. and Drummond A. J. (2009). Bayesian Coalescent Inference of Major Human Mitochondrial DNA Haplogroup Expansions in Africa. *Proc Biol Sci* 276(1655): 367-73.
8. Baer C. F., Miyamoto M. M. and Denver D. R. (2007). Mutation Rate Variation in Multicellular Eukaryotes: Causes and Consequences. *Nat Rev Genet* 8(8): 619-31.
9. Baye T. M., He H., Ding L., Kurowski B. G., Zhang X. and Martin L. J. (2011). Population Structure Analysis Using Rare and Common Functional Variants. *BMC Proc* 5 Suppl 9: S8.
10. Balanovsky O., Rootsi S., Pshenichnov A., Kivisild T., Churnosov M., Evseeva I., Pocheshkhova E., Boldyreva M., Yankovsky N., Balanovska



- E. and Villems R. (2008). Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context. *Am J Hum Genet* 82(1): 236-50.
11. Balanovsky O., Utevska O. and Balanaovska E. (2013). Genetics of Indo-European Populations: The Past, the Future. *Journal of Language Relationship*(9): 23-35.
  12. Ballantyne K. N., Goedbloed M., Fang R., Schaap O., Lao O., Wollstein A., Choi Y., van Duijn K., Vermeulen M., Brauer S., Decorte R., Poetsch M., von Wurmb-Schwark N., de Knijff P., Labuda D., Vezina H., Knoblauch H., Lessig R., Roewer L., Ploski R., Dobosz T., Henke L., Henke J., Furtado M. R. and Kayser M. (2010). Mutability of Y-chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications. *Am J Hum Genet* 87(3): 341-53.
  13. Barac L., Pericic M., Klaric I. M., Rootsi S., Janicijevic B., Kivisild T., Parik J., Rudan I., Villems R. and Rudan P. (2003). Y chromosomal Heritage of Croatian Population and Its Island Isolates. *Eur J Hum Genet* 11(7): 535-42.
  14. Barbieri C., Whitten M., Beyer K., Schreiber H., Li M. and Pakendorf B. (2012). Contrasting Maternal and Paternal Histories in the Linguistic Context of Burkina Faso. *Mol Biol Evol* 29(4): 1213-23.
  15. Battaglia V., Grugni V., Perego U. A., Angerhofer N., Gomez-Palmieri J. E., Woodward S. R., Achilli A., Myres N., Torroni A. and Semino O. (2013). The First Peopling of South America: New Evidence from Y-chromosome Haplogroup Q. *PLoS One* 8(8): e71390.
  16. Behar D. M., Yunusbayev B., Metspalu M., Metspalu E., Rosset S., Parik J., Rootsi S., Chaubey G., Kutuev I., Yudkovsky G., Khusnutdinova E. K., Balanovsky O., Semino O., Pereira L., Comas D., Gurwitz D., Bonne-Tamir B., Parfitt T., Hammer M. F., Skorecki K. and Villems R. (2010). The Genome-Wide Structure of the Jewish People. *Nature* 466(7303): 238-42.
  17. Behar D. M., van Oven M., Rosset S., Metspalu M., Loogvali E. L., Silva N. M., Kivisild T., Torroni A. and Villems R. (2012). A "Copernican" Reassessment of the Human Mitochondrial DNA Tree from Its Root. *Am J Hum Genet* 90(4): 675-84.
  18. Beyin A. (2011). Upper Pleistocene Human Dispersals out of Africa: A Review of the Current State of the Debate. *Int J Evol Biol* 2011: 615094.
  19. Bellance N., Lestienne P. and Rossignol R. (2009). Mitochondria: From Bioenergetics to the Metabolic Regulation of Carcinogenesis. *Front Biosci (Landmark Ed)* 14: 4015-34.
  20. Belle E. M. S., Shah S., Parfitt T. and Thomas M. G. (2010). Y chromosomes of Self-Identified Syeds from the Indian Subcontinent Show Evidence of Elevated Arab Ancestry but not of a Recent Common Patrilineal Origin. *Archaeol Anthropol Sci*(2): 217-24.
  21. Bosch E., Calafell F., Gonzalez-Neira A., Flaiz C., Mateu E., Scheil H. G., Huckenbeck W., Efremovska L., Mikerezi I., Xirotiris N., Grasa C., Schmidt H. and Comas D. (2006). Paternal and Maternal Lineages in the

- Balkans Show a Homogeneous Landscape over Linguistic Barriers, except for the Isolated Aromuns. *Ann Hum Genet* 70(Pt 4): 459-87.
22. Bramanti B., Thomas M. G., Haak W., Unterlaender M., Jores P., Tambets K., Antanaitis-Jacobs I., Haidle M. N., Jankauskas R., Kind C. J., Lueth F., Terberger T., Hiller J., Matsumura S., Forster P. and Burger J. (2009). Genetic Discontinuity between Local Hunter-Gatherers and Central Europe's First Farmers. *Science* 326(5949): 137-40.
  23. Brisighelli F., Alvarez-Iglesias V., Fondevila M., Blanco-Verea A., Carracedo A., Pascali V. L., Capelli C. and Salas A. (2012). Uniparental Markers of Contemporary Italian Population Reveals Details on Its Pre-Roman Heritage. *PLoS One* 7(12): e50794.
  24. Brotherton P., Haak W., Templeton J., Brandt G., Soubrier J., Jane Adler C., Richards S. M., Sarkissian C. D., Ganslmeier R., Friederich S., Dresely V., van Oven M., Kenyon R., Van der Hoek M. B., Korfach J., Luong K., Ho S. Y., Quintana-Murci L., Behar D. M., Meller H., Alt K. W., Cooper A., Adhikarla S., Ganesh Prasad A. K., Pitchappan R., Varatharajan Santhakumari A., Balanovska E., Balanovsky O., Bertranpetit J., Comas D., Martinez-Cruz B., Mele M., Clarke A. C., Matisoo-Smith E. A., Dulik M. C., Gaieski J. B., Owings A. C., Schurr T. G., Vilar M. G., Hobbs A., Soodiyall H., Javed A., Parida L., Platt D. E., Royyuru A. K., Jin L., Li S., Kaplan M. E., Merchant N. C., John Mitchell R., Renfrew C., Lacerda D. R., Santos F. R., Soria Hernanz D. F., Spencer Wells R., Swamikrishnan P., Tyler-Smith C., Paulo Vieira P. and Ziegler J. S. (2013). Neolithic Mitochondrial Haplogroup H Genomes and the Genetic Origins of Europeans. *Nat Commun* 4: 1764.
  25. Brown M. D., Hosseini S. H., Torroni A., Bandelt H. J., Allen J. C., Schurr T. G., Scozzari R., Cruciani F. and Wallace D. C. (1998). mtDNA Haplogroup X: An Ancient Link between Europe/Western Asia and North America? *Am J Hum Genet* 63(6): 1852-61.
  26. Butrimas A., Kazakevičius V., Česnys G., Balčiūnienė I. and Jankauskas R. (1985). Ankstyvieji Virvelinės Keramikos Kultūros Kapai Lietuvoje. *Lietuvos archeologija* 4: 14-19.
  27. Cardoso S., Valverde L., Alfonso-Sanchez M. A., Palencia-Madrid L., Elcoroaristizabal X., Algorta J., Catarino S., Arteta D., Herrera R. J., Zarrabeitia M. T., Pena J. A. and de Pancorbo M. M. (2013). The Expanded mtDNA Phylogeny of the Franco-Cantabrian Region Upholds the Pre-Neolithic Genetic Substrate of Basques. *PLoS One* 8(7): e67835.
  28. Casals F., Hodgkinson A., Hussin J., Idaghdour Y., Bruat V., de Maillard T., Grenier J. C., Gbeha E., Hamdan F. F., Girard S., Spinella J. F., Lariviere M., Saillour V., Healy J., Fernandez I., Sinnott D., Michaud J. L., Rouleau G. A., Haddad E., Le Deist F. and Awadalla P. (2013). Whole-Exome Sequencing Reveals a Rapid Change in the Frequency of Rare Functional Variants in a Founding Population of Humans. *PLoS Genet* 9(9): e1003815.
  29. Chandrasekar A., Kumar S., Sreenath J., Sarkar B. N., Urade B. P., Mallick S., Bandopadhyay S. S., Barua P., Barik S. S., Basu D., Kiran U.,

- Gangopadhyay P., Sahani R., Prasad B. V., Gangopadhyay S., Lakshmi G. R., Ravuri R. R., Padmaja K., Venugopal P. N., Sharma M. B. and Rao V. R. (2009). Updating Phylogeny of Mitochondrial DNA Macrohaplogroup M in India: Dispersal of Modern Human in South Asian Corridor. *PLoS One* 4(10): e7447.
30. Chiaroni J., Underhill P. A. and Cavalli-Sforza L. L. (2009). Y Chromosome Diversity, Human Expansion, Drift, and Cultural Evolution. *Proc Natl Acad Sci U S A* 106(48): 20174-9.
  31. Costa M. D., Pereira J. B., Pala M., Fernandes V., Olivieri A., Achilli A., Perego U. A., Rychkov S., Naumova O., Hatina J., Woodward S. R., Eng K. K., Macaulay V., Carr M., Soares P., Pereira L. and Richards M. B. (2013). A Substantial Prehistoric European Ancestry Amongst Ashkenazi Maternal Lineages. *Nat Commun* 4: 2543.
  32. Cvjetan S., Tolk H. V., Lauc L. B., Colak I., Dordevic D., Efremovska L., Janicijevic B., Kvesic A., Klaric I. M., Metspalu E., Pericic M., Parik J., Popovic D., Sijacki A., Terzic R., Villems R. and Rudan P. (2004). Frequencies of mtDNA Haplogroups in Southeastern Europe--Croatians, Bosnians and Herzegovinians, Serbians, Macedonians and Macedonian Romani. *Coll Antropol* 28(1): 193-8.
  33. Dahmany Y., Marcuello A., Montiel-Sosa F. J., Montoya J., Díez-Sánchez C., López-Pérez M. J. and Ruiz-Pesini E. (2006). Mitochondrial Lineages Distribution in the Spanish Population: Anticipating Association Studies. *An R Acad Nac Farm*(72): 37-47.
  34. de Knijff P. (2000). Messages through Bottlenecks: On the Combined Use of Slow and Fast Evolving Polymorphic Markers on the Human Y chromosome. *Am J Hum Genet* 67(5): 1055-61.
  35. Delfin F., Min-Shan Ko A., Li M., Gunnarsdottir E. D., Tabada K. A., Salvador J. M., Calacal G. C., Sagum M. S., Datar F. A., Padilla S. G., De Ungria M. C. and Stoneking M. (2014). Complete Mtdna Genomes of Filipino Ethnolinguistic Groups: A Melting Pot of Recent and Ancient Lineages in the Asia-Pacific Region. *Eur J Hum Genet* 22(2): 228-37.
  36. Derenko M., Malyarchuk B., Denisova G., Perkova M., Rogalla U., Grzybowski T., Khusnutdinova E., Dambueva I. and Zakharov I. (2012). Complete Mitochondrial DNA Analysis of Eastern Eurasian Haplogroups Rarely Found in Populations of Northern Asia and Eastern Europe. *PLoS One* 7(2): e32179.
  37. Derenko M., Malyarchuk B., Grzybowski T., Denisova G., Rogalla U., Perkova M., Dambueva I. and Zakharov I. (2010). Origin and Post-Glacial Dispersal of Mitochondrial DNA Haplogroups C and D in Northern Asia. *PLoS One* 5(12): e15214.
  38. Derenko M. V., Grzybowski T., Malyarchuk B. A., Dambueva I. K., Denisova G. A., Czarny J., Dorzhu C. M., Kakpakov V. T., Miscicka-Sliwka D., Wozniak M. and Zakharov I. A. (2003). Diversity of Mitochondrial DNA Lineages in South Siberia. *Ann Hum Genet* 67(Pt 5): 391-411.

39. Di Giacomo F., Luca F., Popa L. O., Akar N., Anagnou N., Banyko J., Brdicka R., Barbujani G., Papola F., Ciavarella G., Cucci F., Di Stasi L., Gavrila L., Kerimova M. G., Kovatchev D., Kozlov A. I., Loutradis A., Mandarino V., Mammi C., Michalodimitrakis E. N., Paoli G., Pappa K. I., Pedicini G., Terrenato L., Tofanelli S., Malaspina P. and Novelletto A. (2004). Y chromosomal Haplogroup J as a Signature of the Post-Neolithic Colonization of Europe. *Hum Genet* 115(5): 357-71.
40. Dupuy B. M., Stenersen M., Lu T. T. and Olaisen B. (2006). Geographical Heterogeneity of Y-chromosomal Lineages in Norway. *Forensic Sci Int* 164(1): 10-9.
41. Edgar R. C. (2004). Muscle: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res* 32(5): 1792-7.
42. Excoffier L. and Lischer H. E. (2010). Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Mol Ecol Resour* 10(3): 564-7.
43. Fan H. and Chu J. Y. (2007). A Brief Review of Short Tandem Repeat Mutation. *Genomics Proteomics Bioinformatics* 5(1): 7-14.
44. Felsenstein J. (2004). Phylip (Phylogeny Inference Package) Version 3.6. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle.
45. Flicek P., Amode M. R., Barrell D., Beal K., Billis K., Brent S., Carvalho-Silva D., Clapham P., Coates G., Fitzgerald S., Gil L., Giron C. G., Gordon L., Hourlier T., Hunt S., Johnson N., Juettemann T., Kahari A. K., Keenan S., Kulesha E., Martin F. J., Maurel T., McLaren W. M., Murphy D. N., Nag R., Overduin B., Pignatelli M., Pritchard B., Pritchard E., Riat H. S., Ruffier M., Sheppard D., Taylor K., Thormann A., Trevanion S. J., Vullo A., Wilder S. P., Wilson M., Zadissa A., Aken B. L., Birney E., Cunningham F., Harrow J., Herrero J., Hubbard T. J., Kinsella R., Muffato M., Parker A., Spudich G., Yates A., Zerbino D. R. and Searle S. M. (2014). Ensembl 2014. *Nucleic Acids Res* 42(Database issue): D749-55.
46. Freije D., Helms C., Watson M. S. and Donis-Keller H. (1992). Identification of a Second Pseudoautosomal Region near the Xq and Yq Telomeres. *Science* 258(5089): 1784-7.
47. Fu Q., Rudan P., Paabo S. and Krause J. (2012). Complete Mitochondrial Genomes Reveal Neolithic Expansion into Europe. *PLoS One* 7(3): e32473.
48. Galtier N., Enard D., Radondy Y., Bazin E. and Belkhir K. (2006). Mutation Hot Spots in Mammalian Mitochondrial DNA. *Genome Res* 16(2): 215-22.
49. Garcia O., Fregel R., Larruga J. M., Alvarez V., Yurrebaso I., Cabrera V. M. and Gonzalez A. M. (2011). Using Mitochondrial DNA to Test the Hypothesis of a European Post-Glacial Human Recolonization from the Franco-Cantabrian Refuge. *Heredity (Edinb)* 106(1): 37-45.
50. Giannattasio S., Bobba A., Jurgelevičius V., Vacca R. A., Lattanzio P., Merafina R. S., Utkus A., Kučinskis V. and Marra E. (2006). Molecular Basis of Cystic Fibrosis in Lithuania: Incomplete CFTR Mutation

- Detection by PCR-Based Screening Protocols. *Genetic Testing* 10(3): 169-73.
51. Gilissen C., Hoischen A., Brunner H. G. and Veltman J. A. (2011). Unlocking Mendelian Disease Using Exome Sequencing. *Genome Biol* 12(9): 228.
  52. Girdenis A. and Zinkevičius Z. (1966). Dėl Lietuvių Tarmių Klasifikacijos. *Kalbotyra*(14): 139-47.
  53. Grady J. P., Campbell G., Ratnaike T., Blakely E. L., Falkous G., Nesbitt V., Schaefer A. M., McNally R. J., Gorman G. S., Taylor R. W., Turnbull D. M. and McFarland R. (2014). Disease Progression in Patients with Single, Large-Scale Mitochondrial DNA Deletions. *Brain* 137(Pt 2): 323-34.
  54. Gravel S., Zakharia F., Moreno-Estrada A., Byrnes J. K., Muzzio M., Rodriguez-Flores J. L., Kenny E. E., Gignoux C. R., Maples B. K., Guiblet W., Dutil J., Via M., Sandoval K., Bedoya G., Oleksyk T. K., Ruiz-Linares A., Burchard E. G., Martinez-Cruzado J. C. and Bustamante C. D. (2013). Reconstructing Native American Migrations from Whole-Genome and Whole-Exome Data. *PLoS Genet* 9(12): e1004023.
  55. Grzybowski T., Malyarchuk B. A., Derenko M. V., Perkova M. A., Bednarek J. and Wozniak M. (2007). Complex Interactions of the Eastern and Western Slavic Populations with Other European Groups as Revealed by Mitochondrial DNA Analysis. *Forensic Sci Int Genet* 1(2): 141-7.
  56. Gunnarsdottir E. D., Li M., Bauchet M., Finstermeier K. and Stoneking M. (2011). High-Throughput Sequencing of Complete Human mtDNA Genomes from the Philippines. *Genome Res* 21(1): 1-11.
  57. Haak W., Balanovsky O., Sanchez J. J., Koshel S., Zaporozhchenko V., Adler C. J., Der Sarkissian C. S., Brandt G., Schwarz C., Nicklisch N., Dresely V., Fritsch B., Balanovska E., Villems R., Meller H., Alt K. W. and Cooper A. (2010). Ancient DNA from European Early Neolithic Farmers Reveals Their near Eastern Affinities. *PLoS Biol* 8(11): e1000536.
  58. Hall T. A. (1999). Bioedit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/Nt. *Nucl. Acids. Symp. Ser*(41): 95-98.
  59. Heyer E., Zietkiewicz E., Rochowski A., Yotova V., Puymirat J. and Labuda D. (2001). Phylogenetic and Familial Estimates of Mitochondrial Substitution Rates: Study of Control Region Mutations in Deep-Rooting Pedigrees. *Am J Hum Genet* 69(5): 1113-26.
  60. Henn B. M., Gravel S., Moreno-Estrada A., Acevedo-Acevedo S. and Bustamante C. D. (2010). Fine-Scale Population Structure and the Era of Next-Generation Sequencing. *Hum Mol Genet* 19(R2): R221-6.
  61. Hernandez C. L., Reales G., Dugoujon J. M., Novelletto A., Rodriguez J. N., Cuesta P. and Calderon R. (2014). Human Maternal Heritage in Andalusia (Spain): Its Composition Reveals High Internal Complexity and Distinctive Influences of mtDNA Haplogroups U6 and L in the Western and Eastern Side of Region. *BMC Genet* 15: 11.

62. Hervella M., Izagirre N., Alonso S., Fregel R., Alonso A., Cabrera V. M. and de la Rúa C. (2012). Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. *PLoS One* 7(4): e34417.
63. Huckins L. M., Boraska V., Franklin C. S., Floyd J. A. B., Southam L., GCAN, WTCCC, Sullivan P. F., Bulik C. M., Collier D. A., Tyler-Smith C., Zeggini E. and Tachmazidou I. (2014). Using Ancestry-Informative Markers to Identify Fine Structure across 15 Populations of European Origin. *Eur J Hum Genet*.
64. Huyghe J. R., Fransen E., Hannula S., Van Laer L., Van Eyken E., Maki-Torkko E., Aikio P., Sorri M., Huentelman M. J. and Van Camp G. (2011). A Genome-Wide Analysis of Population Structure in the Finnish Saami with Implications for Genetic Association Studies. *Eur J Hum Genet* 19(3): 347-52.
65. IHC (2003). The International Hapmap Project. *Nature* 426(6968): 789-96.
66. Ingman M. and Gyllensten U. (2007). A Recent Genetic Link between Sami and the Volga-Ural Region of Russia. *Eur J Hum Genet* 15(1): 115-20.
67. Ingman M., Kaessmann H., Paabo S. and Gyllensten U. (2000). Mitochondrial Genome Variation and the Origin of Modern Humans. *Nature* 408(6813): 708-13.
68. International Society of Genetic Genealogy (2014). Y-DNA Haplogroup Tree 2014, Version: 9.44, Date: 20/03/2014, <http://www.isogg.org/tree/23/03/2014>.
69. YCC (2002). A Nomenclature System for the Tree of Human Y-chromosomal Binary Haplogroups. *Genome Res* 12(2): 339-48.
70. Jakkula E., Rehnstrom K., Varilo T., Pietilainen O. P., Paunio T., Pedersen N. L., deFaire U., Jarvelin M. R., Saharinen J., Freimer N., Ripatti S., Purcell S., Collins A., Daly M. J., Palotie A. and Peltonen L. (2008). The Genome-Wide Patterns of Variation Expose Significant Substructure in a Founder Population. *Am J Hum Genet* 83(6): 787-94.
71. Jansen R. P. (2000). Origin and Persistence of the Mitochondrial Genome. *Hum Reprod* 15 Suppl 2: 1-10.
72. Jobling M. A. and Tyler-Smith C. (2003). The Human Y chromosome: An Evolutionary Marker Comes of Age. *Nat Rev Genet* 4(8): 598-612.
73. Kayser M. (2010). The Human Genetic History of Oceania: Near and Remote Views of Dispersal. *Curr Biol* 20(4): R194-201.
74. Kayser M., Lao O., Anslinger K., Augustin C., Bargel G., Edelmann J., Elias S., Heinrich M., Henke J., Henke L., Hohoff C., Illing A., Jonkisz A., Kuzniar P., Lebioda A., Lessig R., Lewicki S., Maciejewska A., Monies D. M., Pawlowski R., Poetsch M., Schmid D., Schmidt U., Schneider P. M., Stradmann-Bellinghausen B., Szibor R., Wegener R., Wozniak M., Zoledziewska M., Roewer L., Dobosz T. and Ploski R. (2005). Significant Genetic Differentiation between Poland and Germany Follows Present-

- Day Political Borders, as Revealed by Y-chromosome Analysis. *Hum Genet* 117(5): 428-43.
75. Karafet T. M., Hallmark B., Cox M. P., Sudoyo H., Downey S., Lansing J. S. and Hammer M. F. (2010). Major East-West Division Underlies Y chromosome Stratification across Indonesia. *Mol Biol Evol* 27(8): 1833-44.
  76. Karafet T. M., Mendez F. L., Meilerman M. B., Underhill P. A., Zegura S. L. and Hammer M. F. (2008). New Binary Polymorphisms Reshape and Increase Resolution of the Human Y chromosomal Haplogroup Tree. *Genome Res* 18(5): 830-8.
  77. Kasperavičiūtė D., Kučinskas V. and Stoneking M. (2004). Y chromosome and Mitochondrial DNA Variation in Lithuanians. *Ann Hum Genet* 68(Pt 5): 438-52.
  78. Kharkov V. N., Stepanov V. A., Feshchenko S. P., Borinskaia S. A., Iankovskii N. K. and Puzyrev V. P. (2005). Frequencies of Y chromosome Binary Haplogroups in Belarussians. *Genetika* 41(8): 1132-6.
  79. Kimura M. and Crow J. F. (1964). The Number of Alleles That Can Be Maintained in a Finite Population. *Genetics* 49: 725-38.
  80. King R. J., Di Cristofaro J., Kouvatsi A., Triantaphyllidis C., Scheidel W., Myres N. M., Lin A. A., Eissautier A., Mitchell M., Binder D., Semino O., Novelletto A., Underhill P. A. and Chiaroni J. (2011). The Coming of the Greeks to Provence and Corsica: Y-Chromosome Models of Archaic Greek Colonization of the Western Mediterranean. *BMC Evol Biol* 11: 69.
  81. King T. E. and Jobling M. A. (2009). Founders, Drift, and Infidelity: The Relationship between Y chromosome Diversity and Patrilineal Surnames. *Mol Biol Evol* 26(5): 1093-102.
  82. Kivisild T., Shen P., Wall D. P., Do B., Sung R., Davis K., Passarino G., Underhill P. A., Scharfe C., Torroni A., Scozzari R., Modiano D., Coppa A., de Knijff P., Feldman M., Cavalli-Sforza L. L. and Oefner P. J. (2006). The Role of Selection in the Evolution of Human Mitochondrial Genomes. *Genetics* 172(1): 373-87.
  83. Klevcova N. I. and Kučinskas V. (1987). Dermatoglifičeskije Ocobenosti Litovcev. *Voprosy Antropologii*(81): 74-88.
  84. Kloss-Brandstatter A., Pacher D., Schonherr S., Weissensteiner H., Binna R., Specht G. and Kronenberg F. (2011). Haplogrep: A Fast and Reliable Algorithm for Automatic Classification of Mitochondrial DNA Haplogroups. *Hum Mutat* 32(1): 25-32.
  85. Kraytsberg Y., Schwartz M., Brown T. A., Ebralidse K., Kunz W. S., Clayton D. A., Vissing J. and Khrapko K. (2004). Recombination of Human Mitochondrial DNA. *Science* 304(5673): 981.
  86. Kučinskas V. (1994). Human Mitochondrial DNA Variation in Lithuania. *Anthropol Anz* 52(4): 289-95.
  87. Kučinskas V. (2001). Population Genetics of Lithuanians. *Ann Hum Biol* 28(1): 1-14.
  88. Kučinskas V. (2004). Genomo Įvairovė: Lietuviai Europoje. Vilnius, Splanvų šalis.

89. Kučinskas V., Radikas J. and Rasmuson M. (1994). Genetic Diversity in the Lithuanian Rural Population as Illustrated by Variation in the ABO and Rh(D) Blood Groups. *Hum Hered* 44(6): 344-9.
90. Kuroki Y., Toyoda A., Noguchi H., Taylor T. D., Itoh T., Kim D. S., Kim D. W., Choi S. H., Kim I. C., Choi H. H., Kim Y. S., Satta Y., Saitou N., Yamada T., Morishita S., Hattori M., Sakaki Y., Park H. S. and Fujiyama A. (2006). Comparative Analysis of Chimpanzee and Human Y chromosomes Unveils Complex Evolutionary Pathway. *Nat Genet* 38(2): 158-67.
91. Kushniarevich A., Sivitskaya L., Danilenko N., Novogrodskii T., Tsybovsky I., Kiseleva A., Kotova S., Chaubey G., Metspalu E., Sahakyan H., Bahmanimehr A., Reidla M., Rootsi S., Parik J., Reisberg T., Achilli A., Hooshiar Kashani B., Gandini F., Olivieri A., Behar D. M., Torroni A., Davydenko O. and Villems R. (2013). Uniparental Genetic Heritage of Belarusians: Encounter of Rare Middle Eastern Matrilineages with a Central European Mitochondrial DNA Pool. *PLoS One* 8(6): e66499.
92. Laitinen V., Lahermo P., Sistonen P. and Savontaus M. L. (2002). Y-chromosomal Diversity Suggests That Baltic Males Share Common Finno-Ugric-Speaking Forefathers. *Hum Hered* 53(2): 68-78.
93. Lander E. S., Linton L. M., Birren B., Nussbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissole S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R. A., Muzny D. M., Scherer S. E., Bouck J. B., Sodergren E. J., Worley K. C., Rives C. M., Gorrell J. H., Metzker M. L., Naylor S. L., Kucherlapati R. S., Nelson D. L., Weinstock G. M., Sakaki Y., Fujiyama A., Hattori M., Yada T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Smith D. R., Doucette-Stamm L., Rubenfield M., Weinstock K., Lee H. M., Dubois J., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H., Yu J., Wang J., Huang G., Gu J., Hood L., Rowen L., Madan A., Qin S., Davis R. W., Federspiel N. A., Abola A. P., Proctor



- M. J., Myers R. M., Schmutz J., Dickson M., Grimwood J., Cox D. R., Olson M. V., Kaul R., Shimizu N., Kawasaki K., Minoshima S., Evans G. A., Athanasiou M., Schultz R., Roe B. A., Chen F., Pan H., Ramser J., Lehrach H., Reinhardt R., McCombie W. R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J. A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D. G., Burge C. B., Cerutti L., Chen H. C., Church D., Clamp M., Copley R. R., Doerks T., Eddy S. R., Eichler E. E., Furey T. S., Galagan J., Gilbert J. G., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W., Johnson L. S., Jones T. A., Kasif S., Kasprzyk A., Kennedy S., Kent W. J., Kitts P., Koonin E. V., Korf I., Kulp D., Lancet D., Lowe T. M., McLysaght A., Mikkelsen T., Moran J. V., Mulder N., Pollara V. J., Ponting C. P., Schuler G., Schultz J., Slater G., Smit A. F., Stupka E., Szustakowski J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y. I., Wolfe K. H., Yang S. P., Yeh R. F., Collins F., Guyer M. S., Peterson J., Felsenfeld A., Wetterstrand K. A., Patrinos A., Morgan M. J., de Jong P., Catanese J. J., Osoegawa K., Shizuya H., Choi S. and Chen Y. J. (2001). Initial Sequencing and Analysis of the Human Genome. *Nature* 409(6822): 860-921.
94. Lao O., Lu T. T., Nothnagel M., Junge O., Freitag-Wolf S., Caliebe A., Balascakova M., Bertranpetit J., Bindoff L. A., Comas D., Holmlund G., Kouvatsi A., Macek M., Mollet I., Parson W., Palo J., Ploski R., Sajantila A., Tagliabracci A., Gether U., Werge T., Rivadeneira F., Hofman A., Uitterlinden A. G., Gieger C., Wichmann H. E., Ruther A., Schreiber S., Becker C., Nurnberg P., Nelson M. R., Krawczak M. and Kayser M. (2008). Correlation between Genetic and Geographic Structure in Europe. *Curr Biol* 18(16): 1241-8.
95. Lappalainen T., Hannelius U., Salmela E., von Döbeln U., Lindgren C. M., Huoponen K., Savontaus M. L., Kere J. and Lahermo P. (2009). Population Structure in Contemporary Sweden - a Y-chromosomal and Mitochondrial DNA Analysis. *Ann Hum Genet* 73(1): 61-73.
96. Lappalainen T., Laitinen V., Salmela E., Andersen P., Huoponen K., Savontaus M. L. and Lahermo P. (2008). Migration Waves to the Baltic Sea Region. *Ann Hum Genet* 72(Pt 3): 337-48.
97. Larmuseau M. H., Ottoni C., Raeymaekers J. A., Vanderheyden N., Larmuseau H. F. and Decorte R. (2012). Temporal Differentiation across a West-European Y-chromosomal Cline: Genealogy as a Tool in Human Population Genetics. *Eur J Hum Genet* 20(4): 434-40.
98. Levin L., Zhidkov I., Gurman Y., Hawlena H. and Mishmar D. (2013). Functional Recurrent Mutations in the Human Mitochondrial Phylogeny: Dual Roles in Evolution and Disease. *Genome Biol Evol* 5(5): 876-90.
99. Librado P. and Rozas J. (2009). Dnasp V5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* 25(11): 1451-2.
100. Lippold S., Xu H., Ko A., Li M., Renaud G., Butthof A., Schroder R. and Stoneking M. (2014). Human Paternal and Maternal Demographic

- Histories: Insights from High-Resolution Y chromosome and Mtdna Sequences. *Investig Genet* 5: 13.
101. Loogvali E. L., Roostalu U., Malyarchuk B. A., Derenko M. V., Kivisild T., Metspalu E., Tambets K., Reidla M., Tolk H. V., Parik J., Pennarun E., Laos S., Lunkina A., Golubenko M., Barac L., Pericic M., Balanovsky O. P., Gusar V., Khusnutdinova E. K., Stepanov V., Puzyrev V., Rudan P., Balanovska E. V., Grechanina E., Richard C., Moisan J. P., Chaventre A., Anagnou N. P., Pappa K. I., Michalodimitrakis E. N., Claustres M., Golge M., Mikerezi I., Usanga E. and Villems R. (2004). Disuniting Uniformity: A Pied Cladistic Canvas of mtDNA Haplogroup H in Eurasia. *Mol Biol Evol* 21(11): 2012-21.
  102. Luca F., Di Giacomo F., Benincasa T., Popa L. O., Banyko J., Kracmarova A., Malaspina P., Novelletto A. and Brdicka R. (2007). Y-chromosomal Variation in the Czech Republic. *Am J Phys Anthropol* 132(1): 132-9.
  103. Lukes A., Zvelebil M. and European Association of Archaeologists. Annual Meeting (2004). Lbk Dialogues: Studies in the Formation of the Linear Pottery Culture. Oxford, England, Archaeopress.
  104. MacArthur D. G., Balasubramanian S., Frankish A., Huang N., Morris J., Walter K., Jostins L., Habegger L., Pickrell J. K., Montgomery S. B., Albers C. A., Zhang Z. D., Conrad D. F., Lunter G., Zheng H., Ayub Q., DePristo M. A., Banks E., Hu M., Handsaker R. E., Rosenfeld J. A., Fromer M., Jin M., Mu X. J., Khurana E., Ye K., Kay M., Saunders G. I., Suner M. M., Hunt T., Barnes I. H., Amid C., Carvalho-Silva D. R., Bignell A. H., Snow C., Yngvadottir B., Bumpstead S., Cooper D. N., Xue Y., Romero I. G., Wang J., Li Y., Gibbs R. A., McCarroll S. A., Dermitzakis E. T., Pritchard J. K., Barrett J. C., Harrow J., Hurles M. E., Gerstein M. B. and Tyler-Smith C. (2012). A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes. *Science* 335(6070): 823-8.
  105. Madrigal L., Posthumously L. C., Melendez-Obando M., Villegas-Palma R., Barrantes R., Raventos H., Pereira R., Luiselli D., Pettener D. and Barbujani G. (2012). High Mitochondrial Mutation Rates Estimated from Deep-Rooting Costa Rican Pedigrees. *Am J Phys Anthropol* 148(3): 327-33.
  106. Malyarchuk B., Derenko M., Denisova G. and Kravtsova O. (2010). Mitogenomic Diversity in Tatars from the Volga-Ural Region of Russia. *Mol Biol Evol* 27(10): 2220-6.
  107. Malyarchuk B., Derenko M., Grzybowski T., Perkova M., Rogalla U., Vanecek T. and Tsybovsky I. (2010). The Peopling of Europe from the Mitochondrial Haplogroup U5 Perspective. *PLoS One* 5(4): e10285.
  108. Malyarchuk B. A. and Derenko M. V. (2008). [Gene Pool Structure of Russian Populations from the European Part of Russia Inferred from the Data on Y Chromosome Haplogroups Distribution]. *Genetika* 44(2): 226-31.
  109. Malmstrom H., Gilbert M. T., Thomas M. G., Brandstrom M., Stora J., Molnar P., Andersen P. K., Bendixen C., Holmlund G., Gotherstrom A. and Willerslev E. (2009). Ancient DNA Reveals Lack of Continuity

- between Neolithic Hunter-Gatherers and Contemporary Scandinavians. *Curr Biol* 19(20): 1758-62.
110. Manni F., Guerard E. and Heyer E. (2004). Geographic Patterns of (Genetic, Morphologic, Linguistic) Variation: How Barriers Can Be Detected by Using Monmonier's Algorithm. *Hum Biol* 76(2): 173-90.
  111. Marjanovic D., Fornarino S., Montagna S., Primorac D., Hadziselimovic R., Vidovic S., Pojskic N., Battaglia V., Achilli A., Drobnic K., Andjelinovic S., Torroni A., Santachiara-Benerecetti A. S. and Semino O. (2005). The Peopling of Modern Bosnia-Herzegovina: Y-chromosome Haplogroups in the Three Main Ethnic Groups. *Ann Hum Genet* 69(Pt 6): 757-63.
  112. McEvoy B. P., Montgomery G. W., McRae A. F., Ripatti S., Perola M., Spector T. D., Cherkas L., Ahmadi K. R., Boomsma D., Willemsen G., Hottenga J. J., Pedersen N. L., Magnusson P. K., Kyvik K. O., Christensen K., Kaprio J., Heikkila K., Palotie A., Widen E., Muilu J., Syvanen A. C., Liljedahl U., Hardiman O., Cronin S., Peltonen L., Martin N. G. and Visscher P. M. (2009). Geographical Structure and Differential Natural Selection among North European Populations. *Genome Res* 19(5): 804-14.
  113. Mendez F. L., Karafet T. M., Krahn T., Ostrer H., Soodyall H. and Hammer M. F. (2011). Increased Resolution of Y chromosome Haplogroup T Defines Relationships among Populations of the near East, Europe, and Africa. *Hum Biol* 83(1): 39-53.
  114. Mendizabal I., Lao O., Marigorta U. M., Wollstein A., Gusmao L., Ferak V., Ioana M., Jordanova A., Kaneva R., Kouvatsi A., Kucinkas V., Makukh H., Metspalu A., Netea M. G., de Pablo R., Pamjav H., Radojkovic D., Rolleston S. J., Sertic J., Macek M., Jr., Comas D. and Kayser M. (2012). Reconstructing the Population History of European Romani from Genome-Wide Data. *Curr Biol* 22(24): 2342-9.
  115. Menozzi P., Piazza A. and Cavalli-Sforza L. (1978). Synthetic Maps of Human Gene Frequencies in Europeans. *Science* 201(4358): 786-92.
  116. Mielnik-Sikorska M., Daca P., Malyarchuk B., Derenko M., Skonieczna K., Perkova M., Dobosz T. and Grzybowski T. (2013). The History of Slavs Inferred from Complete Mitochondrial Genome Sequences. *PLoS One* 8(1): e54360.
  117. Mishmar D., Ruiz-Pesini E., Golik P., Macaulay V., Clark A. G., Hosseini S., Brandon M., Easley K., Chen E., Brown M. D., Sukernik R. I., Olckers A. and Wallace D. C. (2003). Natural Selection Shaped Regional Mtdna Variation in Humans. *Proc Natl Acad Sci U S A* 100(1): 171-6.
  118. Myers S., Freeman C., Auton A., Donnelly P. and McVean G. (2008). A Common Sequence Motif Associated with Recombination Hot Spots and Genome Instability in Humans. *Nat Genet* 40(9): 1124-9.
  119. Myres N. M., Rootsi S., Lin A. A., Jarve M., King R. J., Kutuev I., Cabrera V. M., Khusnutdinova E. K., Pshenichnov A., Yunusbayev B., Balanovsky O., Balanovska E., Rudan P., Baldovic M., Herrera R. J., Chiaroni J., Di Cristofaro J., VILLEMS R., Kivisild T. and Underhill P. A. (2011). A Major

- Y-chromosome Haplogroup R1b Holocene Era Founder Effect in Central and Western Europe. *Eur J Hum Genet* 19(1): 95-101.
120. Moorjani P., Patterson N., Loh P. R., Lipson M., Kisfali P., Melegh B. I., Bonin M., Kadasi L., Riess O., Berger B., Reich D. and Melegh B. (2013). Reconstructing Roma History from Genome-Wide Data. *PLoS One* 8(3): e58633.
  121. Moskvina V., Smith M., Ivanov D., Blackwood D., Stclair D., Hultman C., Toncheva D., Gill M., Corvin A., O'Dushlaine C., Morris D. W., Wray N. R., Sullivan P., Pato C., Pato M. T., Sklar P., Purcell S., Holmans P., O'Donovan M. C., Owen M. J. and Kirov G. (2010). Genetic Differences between Five European Populations. *Hum Hered* 70(2): 141-49.
  122. Nasidze I., Ling E. Y., Quinque D., Dupanloup I., Cordaux R., Rychkov S., Naumova O., Zhukova O., Sarraf-Zadegan N., Naderi G. A., Asgary S., Sardas S., Farhud D. D., Sarkisian T., Asadov C., Kerimov A. and Stoneking M. (2004). Mitochondrial DNA and Y-chromosome Variation in the Caucasus. *Ann Hum Genet* 68(Pt 3): 205-21.
  123. Nei M. (1972). Genetic Distance between Populations. *American Naturalist*(106): 283-92.
  124. Nelis M., Esko T., Magi R., Zimprich F., Zimprich A., Toncheva D., Karachanak S., Piskackova T., Balascak I., Peltonen L., Jakkula E., Rehnstrom K., Lathrop M., Heath S., Galan P., Schreiber S., Meitinger T., Pfeufer A., Wichmann H. E., Melegh B., Polgar N., Toniolo D., Gasparini P., D'Adamo P., Klovins J., Nikitina-Zake L., Kucinskas V., Kasnauskiene J., Lubinski J., Debniak T., Limborska S., Khrunin A., Estivill X., Rabionet R., Marsal S., Julia A., Antonarakis S. E., Deutsch S., Borel C., Attar H., Gagnebin M., Macek M., Krawczak M., Remm M. and Metspalu A. (2009). Genetic Structure of Europeans: A View from the North-East. *PLoS One* 4(5): e5472.
  125. O'Dushlaine C. T., Morris D., Moskvina V., Kirov G., Consortium I. S., Gill M., Corvin A., Wilson J. F. and Cavalleri G. L. (2010). Population Structure and Genome-Wide Patterns of Variation in Ireland and Britain. *Eur J Hum Genet* 18(11): 1248-54.
  126. Ota T. and Kimura M. (1973). A Model of Mutation Appropriate to Estimate the Number of Electrophoretically Detectable Alleles in a Finite Population. *Genet Res* 22(2): 201-4.
  127. Pala M., Achilli A., Olivieri A., Hooshiar Kashani B., Perego U. A., Sanna D., Metspalu E., Tambets K., Tamm E., Accetturo M., Carossa V., Lancioni H., Panara F., Zimmermann B., Huber G., Al-Zahery N., Brisighelli F., Woodward S. R., Francalacci P., Parson W., Salas A., Behar D. M., Villems R., Semino O., Bandelt H. J. and Torroni A. (2009). Mitochondrial Haplogroup U5b3: A Distant Echo of the Epipaleolithic in Italy and the Legacy of the Early Sardinians. *Am J Hum Genet* 84(6): 814-21.
  128. Pala M., Olivieri A., Achilli A., Accetturo M., Metspalu E., Reidla M., Tamm E., Karmin M., Reisberg T., Hooshiar Kashani B., Perego U. A., Carossa V., Gandini F., Pereira J. B., Soares P., Angerhofer N., Rychkov

- S., Al-Zahery N., Carelli V., Sanati M. H., Houshmand M., Hatina J., Macaulay V., Pereira L., Woodward S. R., Davies W., Gamble C., Baird D., Semino O., VILLEMS R., TORRONI A. and RICHARDS M. B. (2012). Mitochondrial DNA Signals of Late Glacial Recolonization of Europe from near Eastern Refugia. *Am J Hum Genet* 90(5): 915-24.
129. Patterson N., Price A. L. and Reich D. (2006). Population Structure and Eigenanalysis. *PLoS Genet* 2(12): e190.
130. Pereira L., Richards M., Goios A., Alonso A., Albarran C., Garcia O., Behar D. M., Golge M., Hatina J., Al-Gazali L., Bradley D. G., Macaulay V. and Amorim A. (2005). High-Resolution mtDNA Evidence for the Late-Glacial Resettlement of Europe from an Iberian Refugium. *Genome Res* 15(1): 19-24.
131. Pereira L., Soares P., Radivojac P., Li B. and Samuels D. C. (2011). Comparing Phylogeny and the Predicted Pathogenicity of Protein Variations Reveals Equal Purifying Selection across the Global Human mtDNA Diversity. *Am J Hum Genet* 88(4): 433-9.
132. Poznik G. D., Henn B. M., Yee M. C., Sliwerska E., Euskirchen G. M., Lin A. A., Snyder M., Quintana-Murci L., Kidd J. M., Underhill P. A. and Bustamante C. D. (2013). Sequencing Y chromosomes Resolves Discrepancy in Time to Common Ancestor of Males Versus Females. *Science* 341(6145): 562-5.
133. Price A. L., Helgason A., Palsson S., Stefansson H., St Clair D., Andreassen O. A., Reich D., Kong A. and Stefansson K. (2009). The Impact of Divergence Time on the Nature of Population Structure: An Example from Iceland. *PLoS Genet* 5(6): e1000505.
134. Pritchard J. K., Stephens M. and Donnelly P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155(2): 945-59.
135. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A., Bender D., Maller J., Sklar P., de Bakker P. I., Daly M. J. and Sham P. C. (2007). Plink: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 81(3): 559-75.
136. Quintana-Murci L., Semino O., Bandelt H. J., Passarino G., McElreavey K. and Santachiara-Benerecetti A. S. (1999). Genetic Evidence of an Early Exit of Homo Sapiens Sapiens from Africa through Eastern Africa. *Nat Genet* 23(4): 437-41.
137. Raitio M., Lindroos K., Laukkanen M., Pastinen T., Sistonen P., Sajantila A. and Syvanen A. C. (2001). Y-chromosomal Snps in Finno-Ugric-Speaking Populations Analyzed by Minisequencing on Microarrays. *Genome Res* 11(3): 471-82.
138. Ralph P. and Coop G. (2013). The Geography of Recent Genetic Ancestry across Europe. *PLoS Biol* 11(5): e1001555.
139. Rappold G. A. (1993). The Pseudoautosomal Regions of the Human Sex Chromosomes. *Hum Genet* 92(4): 315-24.
140. Raska P. and Zhu X. (2011). Rare Variant Density across the Genome and across Populations. *BMC Proc* 5 Suppl 9: S39.

141. Rasteiro R. and Chikhi L. (2013). Female and Male Perspectives on the Neolithic Transition in Europe: Clues from Ancient and Modern Genetic Data. *PLoS One* 8(4): e60944.
142. Rebala K., Mikulich A. I., Tsybovsky I. S., Sivakova D., Dzapinkova Z., Szczerkowska-Dobosz A. and Szczerkowska Z. (2007). Y-STR Variation among Slavs: Evidence for the Slavic Homeland in the Middle Dnieper Basin. *J Hum Genet* 52(5): 406-14.
143. Richards M., Macaulay V., Hickey E., Vega E., Sykes B., Guida V., Rengo C., Sellitto D., Cruciani F., Kivisild T., Villems R., Thomas M., Rychkov S., Rychkov O., Rychkov Y., Golge M., Dimitrov D., Hill E., Bradley D., Romano V., Cali F., Vona G., Demaine A., Papiha S., Triantaphyllidis C., Stefanescu G., Hatina J., Belledi M., Di Rienzo A., Novelletto A., Oppenheim A., Norby S., Al-Zaheri N., Santachiara-Benerecetti S., Scozari R., Torroni A. and Bandelt H. J. (2000). Tracing European Founder Lineages in the near Eastern mtDNA Pool. *Am J Hum Genet* 67(5): 1251-76.
144. Richards M. B., Macaulay V. A., Bandelt H. J. and Sykes B. C. (1998). Phylogeography of Mitochondrial DNA in Western Europe. *Ann Hum Genet* 62(Pt 3): 241-60.
145. Rimantienė R. (1996). *Akmens Amžius Lietuvoje*. Vilnius, Žiburio leidykla.
146. Roewer L., Croucher P. J., Willuweit S., Lu T. T., Kayser M., Lessig R., de Knijff P., Jobling M. A., Tyler-Smith C. and Krawczak M. (2005). Signature of Recent Historical Events in the European Y-chromosomal Str Haplotype Distribution. *Hum Genet* 116(4): 279-91.
147. Rogozin I. B. and Pavlov Y. I. (2003). Theoretical Analysis of Mutation Hotspots and Their DNA Sequence Context Specificity. *Mutat Res* 544(1): 65-85.
148. Rootsi S., Behar D. M., Jarve M., Lin A. A., Myres N. M., Passarelli B., Poznik G. D., Tzur S., Sahakyan H., Pathak A. K., Rosset S., Metspalu M., Grugni V., Semino O., Metspalu E., Bustamante C. D., Skorecki K., Villems R., Kivisild T. and Underhill P. A. (2013). Phylogenetic Applications of Whole Y-chromosome Sequences and the near Eastern Origin of Ashkenazi Levites. *Nat Commun* 4: 2928.
149. Rootsi S., Magri C., Kivisild T., Benuzzi G., Help H., Bermisheva M., Kutuev I., Barac L., Pericic M., Balanovsky O., Pshenichnov A., Dion D., Grobei M., Zhivotovsky L. A., Battaglia V., Achilli A., Al-Zahery N., Parik J., King R., Cinnioglu C., Khusnutdinova E., Rudan P., Balanovska E., Scheffrahn W., Simonescu M., Brehm A., Goncalves R., Rosa A., Moisan J. P., Chaventre A., Ferak V., Furedi S., Oefner P. J., Shen P., Beckman L., Mikerezi I., Terzic R., Primorac D., Cambon-Thomsen A., Krumina A., Torroni A., Underhill P. A., Santachiara-Benerecetti A. S., Villems R. and Semino O. (2004). Phylogeography of Y-chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe. *Am J Hum Genet* 75(1): 128-37.

150. Rootsi S., Myres N. M., Lin A. A., Jarve M., King R. J., Kutuev I., Cabrera V. M., Khusnutdinova E. K., Varendi K., Sahakyan H., Behar D. M., Khusainova R., Balanovsky O., Balanovska E., Rudan P., Yepiskoposyan L., Bahmanimehr A., Farjadian S., Kushniarevich A., Herrera R. J., Grugni V., Battaglia V., Nici C., Crobu F., Karachanak S., Hooshiar Kashani B., Houshmand M., Sanati M. H., Toncheva D., Lisa A., Semino O., Chiaroni J., Di Cristofaro J., VILLEMS R., Kivisild T. and Underhill P. A. (2012). Distinguishing the Co-Ancestries of Haplogroup G Y-chromosomes in the Populations of Europe and the Caucasus. *Eur J Hum Genet* 20(12): 1275-82.
151. Rootsi S., Zhivotovsky L. A., Baldovic M., Kayser M., Kutuev I. A., Khusainova R., Bermisheva M. A., Gubina M., Fedorova S. A., Ilumae A. M., Khusnutdinova E. K., Voevoda M. I., Osipova L. P., Stoneking M., Lin A. A., Ferak V., Parik J., Kivisild T., Underhill P. A. and VILLEMS R. (2007). A Counter-Clockwise Northern Route of the Y-chromosome Haplogroup N from Southeast Asia Towards Europe. *Eur J Hum Genet* 15(2): 204-11.
152. Rosset S., Wells R. S., Soria-Hernanz D. F., Tyler-Smith C., Royyuru A. K. and Behar D. M. (2008). Maximum-Likelihood Estimation of Site-Specific Mutation Rates in Human Mitochondrial DNA from Partial Phylogenetic Classification. *Genetics* 180(3): 1511-24.
153. Saillard J., Forster P., Lynnerup N., Bandelt H. J. and Norby S. (2000). Mtdna Variation among Greenland Eskimos: The Edge of the Beringian Expansion. *Am J Hum Genet* 67(3): 718-26.
154. Salas A., Lareu V., Calafell F., Bertranpetit J. and Carracedo A. (2000). Mtdna Hypervariable Region Ii (Hvii) Sequences in Human Evolution Studies. *Eur J Hum Genet* 8(12): 964-74.
155. Salmela E., Lappalainen T., Fransson I., Andersen P. M., Dahlman-Wright K., Fiebig A., Sistonen P., Savontaus M. L., Schreiber S., Kere J. and Lahermo P. (2008). Genome-Wide Analysis of Single Nucleotide Polymorphisms Uncovers Population Structure in Northern Europe. *PLoS One* 3(10): e3519.
156. Salmela E., Lappalainen T., Liu J., Sistonen P., Andersen P. M., Schreiber S., Savontaus M. L., Czene K., Lahermo P., Hall P. and Kere J. (2011). Swedish Population Substructure Revealed by Genome-Wide Single Nucleotide Polymorphism Data. *PLoS One* 6(2): e16747.
157. Sanchez J. J., Børsting C., Hernandez A., Mengel-Jørgensen J. and Morling N. (2004). Y Chromosome Snp Haplogroups in Danes, Greenlanders and Somalis. *International Congress Series 1261*, 347-49 DOI: 10.1016/S0531-5131(03)01635-2.
158. Sarac J., Saric T., Augustin D. H., Jeran N., Kovacevic L., Cvjetan S., Lewis A. P., Metspalu E., Reidla M., Novokmet N., Vidovic M., Nevajda B., Glasnovic A., Marjanovic D., Missoni S., VILLEMS R. and Rudan P. (2014). Maternal Genetic Heritage of Southeastern Europe Reveals a New Croatian Isolate and a Novel, Local Sub-Branching in the X2 Haplogroup. *Ann Hum Genet.*

159. Schlebusch C. M., Lombard M. and Soodyall H. (2013). Mtdna Control Region Variation Affirms Diversity and Deep Sub-Structure in Populations from Southern Africa. *BMC Evol Biol* 13: 56.
160. Schonberg A., Theunert C., Li M., Stoneking M. and Nasidze I. (2011). High-Throughput Sequencing of Complete Human Mtdna Genomes from the Caucasus and West Asia: High Diversity and Demographic Inferences. *Eur J Hum Genet* 19(9): 988-94.
161. Schwartz M. and Vissing J. (2002). Paternal Inheritance of Mitochondrial DNA. *N Engl J Med* 347(8): 576-80.
162. Scozzari R., Massaia A., D'Atanasio E., Myres N. M., Perego U. A., Trombetta B. and Cruciani F. (2012). Molecular Dissection of the Basal Clades in the Human Y chromosome Phylogenetic Tree. *PLoS One* 7(11): e49170.
163. Semino O., Magri C., Benuzzi G., Lin A. A., Al-Zahery N., Battaglia V., Maccioni L., Triantaphyllidis C., Shen P., Oefner P. J., Zhivotovsky L. A., King R., Torroni A., Cavalli-Sforza L. L., Underhill P. A. and Santachiara-Benerecetti A. S. (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(5): 1023-34.
164. Shi H., Qi X., Zhong H., Peng Y., Zhang X., Ma R. Z. and Su B. (2013). Genetic Evidence of an East Asian Origin and Paleolithic Northward Migration of Y-Chromosome Haplogroup N. *PLoS One* 8(6): e66102.
165. Simoni L., Calafell F., Pettener D., Bertranpetit J. and Barbujani G. (2000). Geographic Patterns of mtDNA Diversity in Europe. *Am J Hum Genet* 66(1): 262-78.
166. Soares P., Ermini L., Thomson N., Mormina M., Rito T., Rohl A., Salas A., Oppenheimer S., Macaulay V. and Richards M. B. (2009). Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. *Am J Hum Genet* 84(6): 740-59.
167. Stathias V., Sotiris G. R., Karagiannidis I., Bourikas G., Martinis G., Papazoglou D., Tavridou A., Papanas N., Maltezos E., Theodoridis M., Vargemezis V., Manolopoulos V. G., Speed W. C., Kidd J. R., Kidd K. K., Drineas P. and Paschou P. (2012). Exploring Genomic Structure Differences and Similarities between the Greek and European Hapmap Populations: Implications for Association Studies. *Ann Hum Genet* 76(6): 472-83.
168. Summerer M., Horst J., Erhart G., Weissensteiner H., Schonherr S., Pacher D., Forer L., Horst D., Manhart A., Horst B., Sanguansermsri T. and Kloss-Brandstatter A. (2014). Large-Scale Mitochondrial DNA Analysis in Southeast Asia Reveals Evolutionary Effects of Cultural Isolation in the Multi-Ethnic Population of Myanmar. *BMC Evol Biol* 14: 17.
169. Tennessen J. A., O'Connor T. D., Bamshad M. J. and Akey J. M. (2011). The Promise and Limitations of Population Exomics for Human Evolution Studies. *Genome Biol* 12(9): 127.



170. Tian C., Plenge R. M., Ransom M., Lee A., Villoslada P., Selmi C., Klareskog L., Pulver A. E., Qi L., Gregersen P. K. and Seldin M. F. (2008). Analysis and Application of European Genetic Substructure Using 300 K Snp Information. *PLoS Genet* 4(1): e4.
171. Tighe O., Dunican D., O'Neill C., Bertorelle G., Beattie D., Graham C., Zschocke J., Cali F., Romano V., Hrabincova E., Kozak L., Nechyporenko M., Livshits L., Guldberg P., Jurkowska M., Zekanowski C., Perez B., Desviat L. R., Ugarte M., Kucinkas V., Knappskog P., Treacy E., Naughten E., Tyfield L., Byck S., Scriver C. R., Mayne P. D. and Croke D. T. (2003). Genetic Diversity within the R408w Phenylketonuria Mutation Lineages in Europe. *Hum Mutat* 21(4): 387-93.
172. Tilford C. A., Kuroda-Kawaguchi T., Skaletsky H., Rozen S., Brown L. G., Rosenberg M., McPherson J. D., Wylie K., Sekhon M., Kucaba T. A., Waterston R. H. and Page D. C. (2001). A Physical Map of the Human Y Chromosome. *Nature* 409(6822): 943-5.
173. Torroni A., Schurr T. G., Cabell M. F., Brown M. D., Neel J. V., Larsen M., Smith D. G., Vullo C. M. and Wallace D. C. (1993). Asian Affinities and Continental Radiation of the Four Founding Native American Mtdnas. *Am J Hum Genet* 53(3): 563-90.
174. Trombetta B., Cruciani F., Sellitto D. and Scozzari R. (2011). A New Topology of the Human Y Chromosome Haplogroup E1b1 (E-P2) Revealed through the Use of Newly Characterized Binary Polymorphisms. *PLoS One* 6(1): e16073.
175. Tully L. A., Parsons T. J., Steighner R. J., Holland M. M., Marino M. A. and Prenger V. L. (2000). A Sensitive Denaturing Gradient-Gel Electrophoresis Assay Reveals a High Frequency of Heteroplasmy in Hypervariable Region 1 of the Human Mtdna Control Region. *Am J Hum Genet* 67(2): 432-43.
176. Underhill P. A. and Kivisild T. (2007). Use of Y chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annu Rev Genet* 41: 539-64.
177. Underhill P. A., Myres N. M., Rootsi S., Metspalu M., Zhivotovsky L. A., King R. J., Lin A. A., Chow C. E., Semino O., Battaglia V., Kutuev I., Jarve M., Chaubey G., Ayub Q., Mohyuddin A., Mehdi S. Q., Sengupta S., Rogaev E. I., Khusnutdinova E. K., Pshenichnov A., Balanovsky O., Balanovska E., Jeran N., Augustin D. H., Baldovic M., Herrera R. J., Thangaraj K., Singh V., Singh L., Majumder P., Rudan P., Primorac D., VILLEMS R. and Kivisild T. (2010). Separating the Post-Glacial Coancestry of European and Asian Y chromosomes within Haplogroup R1a. *Eur J Hum Genet* 18(4): 479-84.
178. Underhill P. A., Poznik G. D., Rootsi S., Jarve M., Lin A. A., Wang J., Passarelli B., Kanbar J., Myres N. M., King R. J., Di Cristofaro J., Sahakyan H., Behar D. M., Kushniarevich A., Sarac J., Saric T., Rudan P., Pathak A. K., Chaubey G., Grugni V., Semino O., Yepiskoposyan L., Bahmanimehr A., Farjadian S., Balanovsky O., Khusnutdinova E. K., Herrera R. J., Chiaroni J., Bustamante C. D., Quake S. R., Kivisild T. and

- Villems R. (2014). The Phylogenetic and Geographic Structure of Y-Chromosome Haplogroup R1a. *Eur J Hum Genet*.
179. van Oven M. and Kayser M. (2009). Updated Comprehensive Phylogenetic Tree of Global Human Mitochondrial DNA Variation. *Hum Mutat* 30(2): E386-94.
  180. Varzari A., Kharkov V., Nikitin A. G., Raicu F., Simonova K., Stephan W., Weiss E. H. and Stepanov V. (2013). Paleo-Balkan and Slavic Contributions to the Genetic Pool of Moldavians: Insights from the Y Chromosome. *PLoS One* 8(1): e53731.
  181. Vidrova V., Tesarova M., Trefilova E., Honzik T., Magner M. and Zeman J. (2008). Mitochondrial DNA Haplogroups in the Czech Population Compared to Other European Countries. *Hum Biol* 80(6): 669-74.
  182. Wei W., Ayub Q., Chen Y., McCarthy S., Hou Y., Carbone I., Xue Y. and Tyler-Smith C. (2013). A Calibrated Human Y-chromosomal Phylogeny Based on Resequencing. *Genome Res* 23(2): 388-95.
  183. Wheeler D. A., Srinivasan M., Egholm M., Shen Y., Chen L., McGuire A., He W., Chen Y. J., Makhijani V., Roth G. T., Gomes X., Tartaro K., Niazi F., Turcotte C. L., Irzyk G. P., Lupski J. R., Chinault C., Song X. Z., Liu Y., Yuan Y., Nazareth L., Qin X., Muzny D. M., Margulies M., Weinstock G. M., Gibbs R. A. and Rothberg J. M. (2008). The Complete Genome of an Individual by Massively Parallel DNA Sequencing. *Nature* 452(7189): 872-6.
  184. Wiik K. (2008). Where Did European Men Come From? *Journal of Genetic Genealogy*(4): 35-85.
  185. Wilder J. A., Kingan S. B., Mobasher Z., Pilkington M. M. and Hammer M. F. (2004). Global Patterns of Human Mitochondrial DNA and Y-chromosome Structure Are Not Influenced by Higher Migration Rates of Females Versus Males. *Nat Genet* 36(10): 1122-5.
  186. Wilson I., Weale M. and Balding D. (2003). Inferences from DNA Data: Population Histories, Evolutionary Processes and Forensic Match Probabilities. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*(166): 155-88.
  187. Xue Y., Wang Q., Long Q., Ng B. L., Swerdlow H., Burton J., Skuce C., Taylor R., Abdellah Z., Zhao Y., MacArthur D. G., Quail M. A., Carter N. P., Yang H. and Tyler-Smith C. (2009). Human Y chromosome Base-Substitution Mutation Rate Measured by Direct Sequencing in a Deep-Rooting Pedigree. *Curr Biol* 19(17): 1453-7.
  188. Zawistowski M., Reppell M., Wegmann D., St Jean P. L., Ehm M. G., Nelson M. R., Novembre J. and Zollner S. (2014). Analysis of Rare Variant Population Structure in Europeans Explains Differential Stratification of Gene-Based Tests. *Eur J Hum Genet*.
  189. Zerjal T., Beckman L., Beckman G., Mikelsaar A. V., Krumina A., Kucinskas V., Hurles M. E. and Tyler-Smith C. (2001). Geographical, Linguistic, and Cultural Influences on Genetic Diversity: Y-chromosomal Distribution in Northern European Populations. *Mol Biol Evol* 18(6): 1077-87.

190. Zhao H., Pfeiffer R. and Gail M. H. (2003). Haplotype Analysis in Population Genetics and Association Studies. *Pharmacogenomics* 4(2): 171-8.
191. Zheng H. X., Yan S., Qin Z. D. and Jin L. (2012). Mtdna Analysis of Global Populations Support That Major Population Expansions Began before Neolithic Time. *Sci Rep* 2: 745.
192. Zhivotovsky L. A., Underhill P. A., Cinnioglu C., Kayser M., Morar B., Kivisild T., Scozzari R., Cruciani F., Destro-Bisol G., Spedini G., Chambers G. K., Herrera R. J., Yong K. K., Gresham D., Tournev I., Feldman M. W. and Kalaydjieva L. (2004). The Effective Mutation Rate at Y chromosome Short Tandem Repeats, with Application to Human Population-Divergence Time. *Am J Hum Genet* 74(1): 50-61.
193. Zinkevičius Z. (1996). The History of the Lithuanian Language. Vilnius, Mokslo ir enciklopedijų leidykla.
194. Zoosmann-Diskin A. (2010). The Origin of Eastern European Jews Revealed by Autosomal, Sex Chromosomal and Mtdna Polymorphisms. *Biol Direct* 5: 57.

PADĚKA

## 1 PRIEDAS. TYRIMO METODAI

### **Genominės DNR išskyrimas fenolio ir chloroformo mišiniu**

Protokolas skirtas genominės DNR (gDNR) išskyrimui iš veninio kraujo leukocitų ląstelių ėminių.

Genominės DNR išskyrimą atliko VU MF ŽMGK darbuotojai.

Protokolas parengtas remiantis:

Diagnostikos protokolu „DBR išskyrimas iš periferinio kraujo leukocitų fenolio-chloroformo-izoamilo alkoholio metodu“, patvirtintu VšĮ VUL SK Generalinio direktoriaus įsakymu Nr. 461, 7 priedas.

### **Reagentai ir cheminės medžiagos:**

- **10 % natrio dodecilsulfatas** [ $C_{12}H_{25}NaO_4S$ ] (angl. *SDS – solution 10%*, gamintojas *AppliChem GmbH*, Vokietija);
- **2 % proteinazė K** (angl. *Proteinase K*, gamintojas *Thermo Fisher Scientific, Inc.*, JAV);
- **3 M natrio acetatas**, pH = 5,2 [ $C_2H_3O_2Na$ ] (angl. *Sodium acetate – Solution*, gamintojas *AppliChem GmbH*, Vokietija);
- **70 % etanolis** (–20 °C); **96 % etanolis** [ $C_2H_6O$ ];
- **Branduolių pernešimo buferis** (ruošiamas laboratorijoje, sudėtis: 10 mM Tris – HCl, pH = 10,5; 1 mM EDTA; 0,15 mM NaCl);
- **Chloroformas** [ $CHCl_3$ ] (angl. *Trichlormethan/chloroform*, gamintojas *Carl Roth GmbH + Co.KG*, Vokietija);
- **Chloroformo – izoamilo alkoholio mišinys** (24:1) (angl. *Roti<sup>®</sup>-C/I*, gamintojas *Carl Roth GmbH + Co.KG*, Vokietija);
- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **Fenolis** [ $C_6H_6O$ ] (angl. *Roti<sup>®</sup>-Phenol*, gamintojas *Carl Roth GmbH + Co.KG*, Vokietija);
- **Lizuojantis buferis** (ruošiamas laboratorijoje, sudėtis: 0,32 M sacharozės; 5 mM MgCl<sub>2</sub>; 1 % Triton X – 100; 10 mM Tris–HCl, pH = 7,5);

- **TE buferis**, pH = 8 (angl. *TE buffer*, gamintojas *Invitrogen™ Life Technologies Corp.*, JAV);
- **Vakuuminis mėgintuvėlis su EDTA** [C<sub>10</sub>H<sub>16</sub>N<sub>2</sub>O<sub>8</sub>], kuriame 3–10 ml veninio kraujo.

**Įranga ir laboratoriniai reikmenys:**

- **Centrifuga** (50 ml);
- **Latekso pirštinės**;
- **Liofilizatorius** (angl. *Eppendorf Concentrator 5301*, gamintojas *Eppendorf AG*, Vokietija);
- **Mikrocentrifuga** (0,5 ml, 1,5 ml, 2 ml);
- **Mikropipetės**, vienkanalės (2–20 μl, 10–100 μl, 100–1000 μl);
- **Sterilios Pastero pipetės** (3 ml);
- **Sterilūs antgaliai** (20 μl, 100 μl, 1000 μl);
- **Sterilūs mėgintuvėliai** (0,5 ml, 1,5 ml, 2 ml, 50 ml);
- **Termostatas** (angl. *Amersham Hybridization Oven / Shaker*, gamintojas *GE Healthcare, General Electric Corp.*, Jungtinė Karalystė).

**Darbo eiga:**

- (1) Veninį kraują iš vakuuminio mėgintuvėlio (3–10 ml) su antikoaguliantu EDTA perpilti į sterilų 50 ml tūrio mėgintuvėlį.
- (2) Įpilti lizuojančio buferio iki galutinio 35 ml tūrio, centrifuguoti 10 °C temperatūroje 3000 aps./min 10 min.
- (3) Viršnuosėdinį skystį nupilti, nuosėdas sumaišyti Pastero pipete ir įpilti lizuojančio buferio iki galutinio 5 ml tūrio, centrifuguoti 10 °C temperatūroje 3000 aps./min 10 min.
- (4) Viršnuosėdinį skystį nupilti, nuosėdas sumaišyti Pastero pipete ir įpilti lizuojančio buferio iki galutinio 5 ml tūrio, centrifuguoti 10 °C temperatūroje 3000 aps./min 10 min.
- (5) Viršnuosėdinį skystį nupilti, į mėgintuvėlį įpilti 400 μl ląstelių branduolių pernešimo buferio.

- (6) Visą tūrį sterilia Pastero pipete perkelti į naują sterilų 2 ml tūrio mėgintuvėlį, įpilti 20 µl 10 % natrio dodecilsulfato tirpalo ir 10 µl 2 % proteinazės K tirpalo.
- (7) Mėgintuvėlį su mišiniu inkubuoti termostate 37 °C temperatūroje 16 val. arba 55 °C temperatūroje 3 val.
- (8) Po inkubavimo į mėgintuvėlį įpilti 400 µl fenolio, mėgintuvėlį su mišiniu vartyti 10 min, vėliau centrifuguoti 5 000 aps./min 2 min.
- (9) Viršutinį sluoksnį nusiurbti ir įpilti į naują sterilų 2 ml tūrio mėgintuvėlį.
- (10) Įpilti 200 µl chloroformo ir 200 µl fenolio, mėgintuvėlį su mišiniu vartyti 5–10 min, vėliau centrifuguoti 5 000 aps./min 2 min.
- (11) Viršutinį sluoksnį nusiurbti ir įpilti į naują sterilų 2 ml tūrio mėgintuvėlį.
- (12) Įpilti 500 µl chloroformo ir izoamilio alkoholio mišinio, mėgintuvėlį su mišiniu vartyti 5 min, vėliau centrifuguoti 5 000 aps./min 2 min.
- (13) Viršutinį sluoksnį nusiurbti ir įpilti į naują sterilų 2 ml tūrio mėgintuvėlį.
- (14) Įpilti 400 µl chloroformo, mėgintuvėlį su mišiniu vartyti 5 min, vėliau centrifuguoti 5 000 aps./min 2 min.
- (15) Viršutinį sluoksnį nusiurbti, visą mišinio tūrį padalinti į 2 lygias dalis, kurias įpilti į naujus sterilius atitinkamai 0,5 ml ir 1,5 ml tūrio mėgintuvėlius. Į kiekvieną iš mėgintuvėlių įpilti po 20 µl 3M natrio acetato tirpalo ir 500 µl 96 % etanolio.
- (16) Atsargiai vartant mėgintuvėlius išsodinti genominę DNR, mėgintuvėlius centrifuguoti 10 000–13 000 aps./min 5 min.
- (17) Viršnuosėdinį skystį nusiurbti, įpilti po 500 µl 70 % etanolio, centrifuguoti 10 000–13 000 aps./min 5 min.
- (18) Viršnuosėdinį skystį nusiurbti, įpilti po 500 µl 96 % etanolio, centrifuguoti 10 000–13 000 aps./min 5 min.
- (19) 0,5 ml tūrio mėgintuvėlį su etanolyje išsodinta gDNR laikyti –20 °C arba –80 °C temperatūroje.
- (20) Nusiurbti 1,5 ml tūrio mėgintuvėlio viršnuosėdinį sluoksnį, gDNR liofilizuoti 30 °C temperatūroje.

(21) Į mėgintuvėlį su liofilizuota gDNR įpilti 100 µl TE buferio ir laikyti 4 °C temperatūroje ne mažiau nei 24 val.

**Automatizuotas DNR išskyrimo, naudojant magnetines daleles, metodas (TECAN Freedom EVO® 200)**

Protokolas skirtas automatizuotam genominės DNR (gDNR) išskyrimui iš veninio kraujo leukocitų ląstelių ėminių gDNR prijungimo prie magnetinių dalelių metodu, naudojant *TECAN Freedom EVO® 200* sistemą. gDNR išskyrimą atliko VU MF ŽMGK darbuotojai.

Protokolas parengtas remiantis:

- DNR išskyrimo rinkinio techniniu aprašu (angl. *MagneSil® Genomic, Large Volume System, Instructions For Use of Products A4080, A4082, A4085 Technical Bulletin*).

**Reagentai ir cheminės medžiagos:**

- **96 % etanolis;**
- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- DNR išskyrimo (angl. *MagneSil® Genomic, Large Volume System*, gamintojas *Promega Corp.*, JAV) rinkinys:
  - **Magnetinės dalelės** (angl. *MagneSil® Paramagnetic Particles*);
  - **Plovimo buferis etanolio pagrindu** (angl. *Alcohol wash*);
  - **Lizuojantis buferis** (angl. *eLysis Buffer (Large Volume System)*);
  - **Išplovimo buferis** (angl. *Elution Buffer*);
- **Izopropanolis (2-propanolis)** [C<sub>3</sub>H<sub>8</sub>O] (angl. *2-propanol, iso-propyl alcohol*, gamintojas *Sigma-Aldrich Co.*, JAV);
- **Ląstelių lizės tirpalas** (angl. *Cell Lysis Solution (Genomic Purification)*, gamintojas *Promega Corp.*, JAV);
- **NaOH**, 0,25 M (ruošiamas laboratorijoje).



### **Įranga ir laboratoriniai reikmenys:**

- **DNR išskyrimo automatizuota sistema** (angl. *TECAN Freedom EVO<sup>®</sup> 200*, gamintojas *Tecan Schweiz AG*, Šveicarija);
- **Kaitinimo blokas su adapteriu** (angl. *VWR<sup>™</sup> Digital Heatblock*, gamintojas *VWR International LLC*, JAV; angl. *Heat Block Adapter*, gamintojas *Promega Corp.*, JAV);
- **Latekso pirštinės**;
- **Magnetinis stovas** (angl. *MagnaBot<sup>®</sup> Large Volume Magnetic Separation Device*, gamintojas *Promega Corp.*, JAV);
- **Mėgintuvėlių (50 ml) stovas** (angl. *Tube Holder, 50 ml Tubes*, gamintojas *Promega Corp.*, JAV);
- **NALGENE<sup>®</sup> indas**, 1 l indas (angl. *NALGENE<sup>®</sup> 1000 ml PP jar*, gamintojas *Nalge Nunc International Corp.*, JAV);
- **Purtyklė su adapteriu** (angl. *IKA<sup>®</sup> KS 130 control*, gamintojas *IKA-Werke GmbH & Co. KG*, Vokietija; angl. *Shaker Top Adapter*, gamintojas *Promega Corp.*, JAV);
- **Sterilūs mėgintuvėliai (1,5 ml)**;
- **Sterilūs mėgintuvėliai (50 ml)** (angl. *Corning 50 ml Centrifuge Tube*, gamintojas *Corning Inc.*, JAV).

### **Eiga:**

#### **Reagentų ir įrangos paruošimas**

- (1) Paruošti reagentus pagal veninio kraujo ėminio tūrį (ml) ir kiekį (8, 16 arba 24) (1 lentelė, 2 lentelė).
- (2) Paruošti plovimo buferį etanolio pagrindu, į buferio tirpalą (500 ml) įpilti 250 ml izopropanolio ir 250 ml 96 % etanolio.
- (3) Ląstelių lizės tirpalą skiesti ddH<sub>2</sub>O santykiu 1:4 (1 dalis ląstelių lizės tirpalo, 4 dalys ddH<sub>2</sub>O).
- (4) Magnetines daleles inkubuoti termostate 37 °C temperatūroje 35 min.
- (5) Ištuštinti, esant skysčio, atliekų indus (angl. *Waste bottle*).

- (6) Prietaiso skysčių sistemos (angl. *Liquid System*) indą pripildyti ddH<sub>2</sub>O.
- (7) Į prietaiso 10–14 pozicijas įstatyti tris NALGENE® 1 l talpos indus.
- (8) Patikrinti prietaiso švirkštus, esant vandens lašų, nuvalyti beplauše servetėle.
- (9) Įjungti 62–69 pozicijoje kaitinimo bloką su adapteriu, nustatyti 95 °C temperatūrą.
- (10) Įjungti 50–60 pozicijoje purtyklę su adapteriu.
- (11) Į prietaiso 17–25 pozicijas įstatyti du mėgintuvėlių (50 ml) stovus, kurių kiekviename įstatyta po aštuonis sterilius mėgintuvėlius (50 ml) be dangtelių.
- (12) Į prietaiso 17–20 poziciją įstatyti magnetinį stovą.
- (13) Į prietaiso 1 ir (ar) 2 pozicijas įstatyti sterilius 1,5 ml mėgintuvėlius, atsižvelgiant į kraujo ėminių skaičių.
- (14) Į prietaiso 34–39 pozicijas įstatyti stovą su aštuoniais mėgintuvėliais (50 ml), pripildytais magnetinių dalelių.
- (15) Į prietaiso 3 ir (ar) 4 pozicijas įstatyti vakuuminius mėgintuvėlius be dangtelių su veniniu krauju (10 mm/3 ml arba 16 mm/10 ml su EDTA).
- (16) Į prietaiso 42–43 poziciją įstatyti du indus: vieną su ląstelių lizės tirpalu, kitą su eliucijos buferiu.
- (17) Į prietaiso 28–29 poziciją įstatyti indą su lizuojančiu buferiu.
- (18) Į prietaiso 30–31 poziciją įstatyti indą su paruoštu plovimo buferiu etanolio pagrindu.

1 lentelė. Reagentų kiekiai, reikalingi išskirti gDNR iš 3 ml veninio kraujo ėminio

Ėminiai (3 ml)	Lizuojantis buferis (ml)	Plovimo buferis etanolio pagrindu (ml)	Išplovimo buferis (ml)	Magnetinės dalelės (ml)	Ląstelių lizės tirpalas (1:4)
8	140	350	25	9 (9/8 = 1,1)	200
16	280	700	50	18 (18/8 = 2,25)	–
24	420	1050	75	27 (27/8 = 3,375)	–

2 lentelė. Reagentų kiekiai, reikalingi išskirti gDNR iš 10 ml veninio kraujo ėminio

Ėminiai (10 ml)	Lizuojantis buferis (ml)	Plovimo buferis etanolio pagrindu (ml)	Išplovimo buferis (ml)	Magnetinės dalelės (ml)	Ląstelių lizės tirpalas (1:4)
8	270	350	25	24 (24/8 = 3)	200
16	540	700	50	48 (po 3 ml)	-
24	810	1050	75	72 (po 3 ml)	-

### Prietaiso ir programos aktyvavimas

- (19) Uždaryti apsauginį stiklą, įjungti prietaisą (mygtukas *ON*).
- (20) Įjungti programą *Freedom EVOware® Standard 2.3 (EVOware 2.3)*.
- (21) Pasirinkti *Run an existing script*.
- (22) Spausti piktogramą *Start your selection* (žalia rodyklė).
- (23) Kataloge *Favorite* pasirinkti dokumentą  
*NEW\_Tecan\_gDNA\_XL\_16\_mm\_tubes*.
- (24) Spausti piktogramą *Run*.
- (25) *Fill up system liquid!!!*: spausti *Ok*.
- (26) *Volume in waste*: įvesti 30.
- (27) *Volume in cleaner*: įvesti 30.
- (28) *Start recovery mode? (0 = no; 1 = yes)*: įvesti 0.
- (29) *Read barcode (1 = yes; 0 = no)*: įvesti 0.
- (30) *How many samples would you like to process?*: įvesti 8, 16 arba 24.
- (31) *Start with blood in 50 ml processing tube (0 = no; 1 = yes)*: įvesti 0.
- (32) *All required racks, plates and buffers correctly prepared on to worktable?*  
spausti *Ok*.
- (33) *Is the heat block turned on and set up 95 °C?* spausti *Ok*.
- (34) *Would you like 1 or 2 elution to be performed?* įvesti 1.

### DNR išskyrimas

- (35) Ėminio ląstelių lizavimas ląstelių lizės tirpalu.
- (36) gDNR prijungimas prie magnetinių dalelių.
- (37) gDNR plovimas tris kartus lizuojančio ir plovimo buferio etanolio pagrindu mišiniu.

(38) gDNR plovimas tris kartus plovimo buferiu etanolio pagrindu.

(39) gDNR pašalinimas nuo magnetinių dalelių išplovimo buferiu.

#### **Prietaiso priežiūra**

(40) Po kiekvieno gDNR išskyrimo atlikti prietaiso priežiūrą: prietaiso išorinius ir vidinius paviršius nuvalyti ir dezinfekuoti.

(41) Į prietaiso 42 poziciją įstatyti du indus, vieną su 96 % etanolium, kitą su 0,25 M NaOH.

(42) Įjungti *EVOware 2.3*.

(43) Pasirinkti *Run maintenance*, pasirinkti *Clean\_Fix\_Tips\_Inside*, spausti piktogramą *Start your selection* (žalia rodyklė), spausti piktogramą *Run*.

#### **Genominės DNR koncentracijos ir švarumo matavimas**

Protokolas skirtas genominės DNR koncentracijai ir švarumui nustatyti spektrofotometru. Protokolas parengtas remiantis:

- DNR koncentracijos ir švarumo nustatymo spektrofotometru vartotojo vadovu (angl. *NanoDrop® ND-1000 Spectrophotometer User Manual*).

#### **Reagentai ir cheminės medžiagos:**

- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **gDNR**, genominės DNR tirpalas;
- **TE buferis**, pH = 8 (angl. *TE buffer*, gamintojas *Invitrogen™ Life Technologies Corp.*, JAV).

#### **Įranga ir laboratoriniai reikmenys:**

- **Latekso pirštinės**;
- **Mikrocentrifuga** (0,5 ml, 1,5 ml);
- **Mikropipetės**, vienkanalės (2–20 µl, 10–100 µl);
- **Purtyklė** (angl. *Vortex-Genie 2, Model G560*, gamintojas *Scientific Industries Inc.*, JAV);

- **Spektrofotometras** (angl. *NanoDrop<sup>®</sup> ND-1000 Spectrophotometer*, gamintojas *NanoDrop Technologies Inc.*, JAV);
- **Sterilūs antgaliai** (20 µl, 100 µl);
- **Sterilūs mėgintuvėliai** (0,5 ml, 1,5 ml).

***Darbo eiga:***

- (22) Įjungti programą *ND-1000 v3.5.2*.
- (23) Pasirinkti *Nucleic Acid*.
- (24) Pakelti svirtelę (angl. *Sampling arm*), nuvalyti apatinį ir viršutinį matavimo paviršius (angl. *Lower and Upper Measurement Pedestal*).
- (25) Mikropipete užlašinti 2 µl ddH<sub>2</sub>O mėginio ir pateiktoje informacinėje lentelėje spausti *OK*, aktyvuojamas prietaisas.
- (26) Pasibaigus prietaiso aktyvacijai su ddH<sub>2</sub>O mėginiu, spausti *Blank*.
- (27) Pakelti svirtelę ir švaria beplauše servetėle nuvalyti apatinį ir viršutinį matavimo paviršius.
- (28) Mėgintuvėlį su išskirtos gDNR tirpalu supurtyti ir nucentrifuguoti.
- (29) Mikropipetei užlašinti 1 µl gDNR tirpalo mėginio ir spausti *Measure*.
- (30) Įvertinti išskirtos gDNR tirpalo koncentraciją (gDNR tirpalo sugertis esant 260 nm bangos ilgiui;  $A_{260} = \text{ng}/\mu\text{l}$ ) ir švarumą (gDNR tirpalo sugerties reikšmių santykis esant 260 ir 280 nm bangos ilgiams;  $A_{260/280} = 1,7-1,9$ ).
- (31) Pakartotinai spausti *Measure* ir nustatyti gautų reikšmių vidurkį.
- (32) Esant poreikiui, išskirtos gDNR tirpalą skiesti ddH<sub>2</sub>O arba TE buferiu iki reikiamos darbinės koncentracijos.

**Y chromosomos vieno nukleotido polimorfizmų genotipavimo TaqMan<sup>®</sup> zondais protokolas**

Protokolas skirtas Y chromosomos vieno nukleotido polimorfizmų (VNP) alelių TaqMan<sup>®</sup> zondais tyrimui naudojant *7900HT Fast Real-Time PCR System* prietaisą.

Darbą atliko dokt. Ingrida Uktverytė Populiacinės genetikos laboratorijoje, Rusijos medicinos mokslų akademijos Medicininės genetikos tyrimų centre

RAMN (Maskva, Rusijos Federacija) nuo 2010 m. spalio 18 d. iki 2010 m. lapkričio 19 d. Darbui vadovavo dr. Oleg Balanovsky, prof. Elena Balanovska.

Protokolas parengtas remiantis:

- Alelių *7900HT Fast Real-Time PCR System* prietaisu nustatymo vartotojo vadovu (angl. *Applied Biosystems 7300/7500/7500 Fast Real-Time PCR System Allelic Discrimination Getting Started Guide*).

#### **Reagentai ir cheminės medžiagos:**

- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **gDNR**, genominės DNR mėginys (konc. 2 ng/μl);
- **Nekomercinių pradmenų ir zondų mišiniai** (angl. *Custom primers and probes labeled FAM<sup>TM</sup>, VIC<sup>®</sup> dye*, gamintojas *Applied Biosystems<sup>®</sup> Life Technologies Corp., JAV*);
- **TaqMan<sup>®</sup> universalus PGR mišinys** (angl. *TaqMan<sup>®</sup> Universal PCR Master Mix*, gamintojas *Applied Biosystems<sup>®</sup> Life Technologies Corp., JAV*).

#### **Įranga ir laboratoriniai reikmenys:**

- **96 šulinėlių optinė plokštelė** (angl. *ABI PRISM<sup>®</sup> 96-Well Optical Reaction Plate with Barcode*, gamintojas *Applied Biosystems<sup>®</sup> Life Technologies Corp., JAV*);
- **96 šulinėlių optinės plokštelės dengimo plėvelė** (angl. *MicroAmp<sup>TM</sup> Optical Adhesive Film*, gamintojas *Applied Biosystems<sup>®</sup> Life Technologies Corp., JAV*);
- **Centrifuga** (96 šulinėlių plokštelėms);
- **Latekso pirštinės**;
- **Mikrocentrifuga** (0,2 ml, 0,5 ml, 1,5 ml);
- **Mikropipetės**, vienkanalės (0,2–2,5 μl, 2–20 μl, 10–100 μl, 20–200 μl, 100–1000 μl);

- **Purtyklė** (angl. *Vortex-Genie 2, Model G560*, gamintojas *Scientific Industries Inc., JAV*);
- **RL-PGR sistema**, realaus laiko PGR sistema (angl. *7900HT Fast Real-Time PCR System*, gamintojas *Applied Biosystems® Life Technologies Corp., JAV*);
- **Sterilūs antgaliai** (10 µl, 20 µl, 100 µl, 200 µl, 1000 µl);
- **Sterilūs mėgintuvėliai** (0,2 ml, 0,5 ml, 1,5 ml, 50 ml).

**Darbo eiga:**

**Optinės plokštelės su gDNR mėginiais paruošimas**

- (1) Įpilti po 5 µl gDNR (2 ng/µl) į kiekvieną iš 90 optinės plokštelės šulinėlių.
- (2) Palikti optinę plokštelę su mėginiais 4 °C temperatūroje 48 val. (fakultatyvus).
- (3) Paruošti RL-PGR reakcijos mišinį (3 lentelė).

3 lentelė. RL-PGR reakcijos mišinys

Reagentai	Tūris (µl)/ reakcijai (išdžiovintas DNR mėginys)	Tūris (µl)/ reakcijai
ddH <sub>2</sub> O	4,8	–
TaqMan® universalus PGR mišinys (angl. <i>TaqMan® Universal PCR Master Mix</i> )	5	5
Pradmens ir zondo mišinys	0,2	0,2
<b>Iš viso</b>	10	5,2

- (4) Įpilti po 5 µl gDNR k+ (teigiama kontrolė, 1 šulinėlis), k0 (nulinė kontrolė, 1 šulinėlis) ir 5 µl ddH<sub>2</sub>O – NTC (neigiama RL-PGR kontrolė, 4 šulinėliai);
- (5) Įpilti po 10 µl arba 5 µl (gDNR mėginiai liofilizuoti/neliofilizuoti) paruošto PGR reakcijos mišinio (3 lentelė) į kiekvieną optinės plokštelės šulinėlį su gDNR mėginiu;
- (6) Įpilti 5 µl paruošto PGR reakcijos mišinio (3 lentelė) į kiekvieną optinės plokštelės šulinėlį su k+, k0.
- (7) Įpilti 10 µl paruošto PGR reakcijos mišinio (3 lentelė) į kiekvieną optinės plokštelės šulinėlį su NTC.

- (8) Paruoštą plokštelę kruopščiai uždengti optinės plokštelės dengimo plėvele.
- (9) Paruoštą optinę plokštelę centrifuguoti 3 500 aps./min 1 min.

#### **Alelių nustatymo protokolo kūrimas**

- (10) Aktyvuoti programą *Sequence Detection Systems version 2.3 Software (SDS 2.3)*.
- (11) Pasirinkti *File>New Plate Wizard*.
- (12) Etape *Select assay type* pasirinkti *Allelic discrimination (AD)*, spausti *Next*.
- (13) Etape *Enter Plate Information* skiltyje *Plate Type* pasirinkti *96 wells*, skiltyje *Create Document From* pasirinkti *Blank Document*, laukelyje *Barcode* įvesti plokštelės pavadinimą, pavyzdžiui, 2010-X-20-Y-M170-LIT-179-267, spausti *Next*.
- (14) Etape *Enter Samples and Detectors to Use in Plate* skiltyje *Samples to Use in Plate* įvesti visus gDNR mėginių, k+, k0 ir NTC pavadinimus, skiltyje *Markers to Use in Plate* spausti *Detector Manager* iš esamo sąrašo pasirinkti Y-M170, spausti *Copy to Plate Document*, pasirinkus visus žymenis spausti *Done*, spausti *Next*.
- (15) Etape *Assign Samples and Markers to Wells* užpildyti virtualią plokštelės reprezentaciją, susiejant suvestus gDNR mėginius, teigiamą, neigiamą kontroles, NTC su detektoriumi ir jų padėtimi plokštelėje. Skiltyje *Plate Layout* pasirinkti laukelius su pirmu gDNR mėginiu, skiltyje *Add Samples and Markers* sąrašė *Markers in selected Wells* pažymėti Y-M170 ir detektoriaus stulpelyje *Task* pasirinkti *Unknown*;
- (16) 15 etapą kartoti su likusiais gDNR mėginiais, k+, k0, NTC – detektoriaus stulpelyje *Task* pasirinkti *NTC*, užpildžius protokolą spausti *Finish*.
- (17) Išsaugoti sukurtą protokolą pasirenkant *File > Save (D:\ (Barcode) scan.sds)*, iškelti protokolo informaciją pasirenkant *File > Export (D:\Temp (Barcode) scan.txt)*.
- (18) Programoje *SDS 2.3* pasirinkti *File > New Plate Wizard*.
- (19) Etape *Select Assay Type* pasirinkti *Standard Curve (AQ)*, spausti *Next*.



- (20) Etape *Enter Plate Information* skiltyje *Plate Type* pasirinkti *96 wells*, skiltyje *Create Document From* pasirinkti *Blank Document*, laukelyje *Barcode* įvesti plokštelės pavadinimą, pavyzdžiui, 2010-X-20-Y-M170-LIT-179-267, spausti *Next*.
- (21) Programoje *SDS 2.3* išsaugoti sukurtą protokolą pasirenkant *File > Save* (D:\\ (Barcode) txn.sds), iškelti protokolo informaciją pasirenkant *File > Export* (D:\\Temp (Barcode) txn.txt).
- (22) Programa *MS Office Excel* atsidaryti skurtus dokumentus „Temp (Barcode) scan.txt“ ir „Temp (Barcode) txn.txt“.
- (23) Perkelti laukelių B12-B107 informaciją iš dokumento „temp (Barcode) scan.txt“ į laukelį B7 dokumente „Temp (Barcode) txn.txt“, išsaugoti abu pakeistu dokumentus spaudžiant *File > Save*.

### **Gausinimo RL-PGR sistema protokolo kūrimas**

- (24) Programoje *SDS 2.3* pasirinkti *File > Open* iš esamo sąrašo pasirinkti dokumentą „Temp (Barcode) txn.sds“.
- (25) Etape *Enter Samples and Detectors to Use in Plate* skiltyje *Samples to Use in Plate* pasirinkti *Existing Plate* iš esamo sąrašo pasirinkti dokumentą „Temp (Barcode) txn.txt“, skiltyje *Markers to Use in Plate* spausti *Detector Manager*, iš esamo sąrašo pasirinkti VIC ir FAM dažus, spausti *Copy to Plate Document*, pasirinkus visus žymenis spausti *Done*, spausti *Next*.
- (26) Etape *Assign Samples and Markers to Wells* užpildyti virtualią plokštelės reprezentaciją, susiejant suvestus gDNR mėginius, teigiamą, neigiamą kontroles su detektoriumi ir jų padėtimi plokštelėje. Skiltyje *Plate Layout* pažymėti visus šulinėlius su gDNR mėginiais, k+, k0 skiltyje *Add Samples and Markers* lentelėje *Markers in selected Wells* pažymėti abu dažus (VIC ir FAM), detektoriaus stulpelyje *Task* pasirinkti *Unknown*.
- (27) Etape *Assign Samples and Markers to Wells* užpildyti virtualią plokštelės reprezentaciją, susiejant suvestus NTC su detektoriumi ir jų padėtimi plokštelėje. Skiltyje *Plate Layout* pažymėti visus šulinėlius su NTC skiltyje *Add Samples and Markers* lentelėje *Markers in selected Wells*

pažymėti abu dažus (VIC ir FAM), detektoriaus stulpelyje *Task* pasirinkti *NTC*, spausti *Finish*.

- (28) Pasirinkti skiltį *Instrument > Thermal Cycler > Standard > Thermal Profile* (4 lentelė).

4 lentelė. RL-PGR sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų sk.
Pirminė denatūracija	95	10:00	1
Denatūracija	95	00:15	50
Pradmenų, zondų prijungimas/sintezė	60	01:00	

- (29) Laukelyje *Volume* ( $\mu$ l) įrašyti 10, laukelyje *Cycles* įrašyti 50.

- (30) Išsaugoti protokolą spaudžiant *File > Save*.

### **Optinės plokštelės nuskaitymas iki gausinimo (bazinės emisijos nustatymas)**

- (31) Programoje *SDS 2.3* pasirinkti *File > Open*, pasirinkti dokumentą „(Barcode) scan.sds“.
- (32) Pasirinkti skiltį *Instrument > Plate Read* spausti *Connect*, spausti *Open/Close* ir laukti kol optinės plokštelės platforma sustos plokštelės įstatymo pozicijoje.
- (33) Įstatyti paruoštą optinę plokštelę į RL-PGR sistemą.
- (34) Spausti *Open/Close* ir laukti kol optinės plokštelės platforma grįš į pradinę padėtį sistemoje.
- (35) Spausti *Pre-read*.
- (36) Po skenavimo uždaryti informacinę lentelę, nurodančia, kad skenavimas įvykdytas sėkmingai, spaudžiant *OK*. Skenavimo data ir laikas įrašomas laukelyje *Date Collection Stamp*.
- (37) Optinės plokštelės platforma sustos plokštelės įstatymo pozicijoje.

### **Gausinimas 7900HT Fast Real-Time PCR System**

- (38) Programoje *SDS 2.3* pasirinkti laukelį „(Barcode) txn.sds“.
- (39) Pasirinkti skiltį *Instrument > Real Time*, spausti *Open/Close*, laukti, kol optinės plokštelės platforma grįš į pradinę padėtį sistemoje.
- (40) Spausti *Start Run*.
- (41) Pasibaigus reakcijai spausti *Analysis*.

(42) Išsaugoti duomenis spaudžiant *File > Save*.

**Optinės plokštelės nuskaitymas po gausinimo (pagausintų mėginių emisijos nustatymas)**

(43) Programoje *SDS 2.3* pasirinkti dokumentą „(Barcode) scan.sds“.

(44) Spausti skiltį *Instrument > Plate Read > Connect*.

(45) Spausti *Post-read*.

(46) Po skenavimo uždaryti informacinę lentelę, nurodančia, kad skenavimas įvykdytas sėkmingai, spaudžiant *OK*. Skenavimo data ir laikas įrašomas laukelyje *Date Collection Stamp*.

(47) Išsaugoti duomenis spaudžiant *File > Save*.

(48) Pašalinti optinę plokštelę iš sistemos pasirenkant skiltį *Instrument > Plate Read*, spausti *Open/Close*.

**Y chromosomos trumpų tandeminių kartotinių sekų genotipavimas *Applied Biosystems® AmpFlSTR® Yfiler™* rinkiniu**

Protokolas skirtas 17 Y chromosomos trumpų tandeminių kartotinių sekų (TTKS) žymenų genotipavimui *Applied Biosystems® AmpFlSTR® Yfiler™* rinkiniu vykdant kapiliarinę elektroforezę *ABI PRISM™ 3130xl* genetiniu analizatoriumi.

Darbą atliko dokt. Ingrida Uktverytė Populiacinės genetikos laboratorijoje, Rusijos medicinos mokslų akademijos Medicininės genetikos tyrimų centre RAMN (Maskva, Rusijos Federacija) nuo 2010 m. spalio 18 d. iki 2010 m. lapkričio 19 d. Darbui vadovavo dr. Oleg Balanovsky, prof. Elena Balanovska. (\*) etapai atlikti Populiacinės genetikos laboratorijos tyrėjos Svetlanos Frolovos. Optimizuotas protokolas įdiegtas VU MF ŽMGK dokt. Ingridos Uktverytės.

Protokolas parengtas remiantis:

- Y chromosomos 17 TTKS žymenų nustatymo rinkinio vartotojo vadovu (angl. *Applied Biosystems® AmpFlSTR® Yfiler™ PCR Amplification Kit, User Guide, Publication Number 4358101 Rev. H*).

### Reagentai ir cheminės medžiagos:

- **10× genetinio analizatoriaus buferis** (angl. *3130 Running Buffer (10×) with EDTA*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **Hi-Di™ formamidas** (angl. *Hi-Di™ Formamide*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **Molekulinis ilgio standartas** (angl. *GeneScan™ 500 LIZ® Size Standard*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **POP-7™ polimeras** (angl. *POP-7™ Performance Optimized Polymer*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **Y chromosomos 17 TTKS žymenų nustatymo rinkinys** (angl. *Applied Biosystems® AmpFlSTR® Yfiler™ kit*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV):
  - **PGR mišinys** (angl. *AmpFlSTR® Yfiler™ PCR reaction Mix*);
  - **Pradmenų rinkinys** (angl. *AmpFlSTR® Yfiler™ Primer Set*);
  - **DNR polimerazė** (angl. *AmpFlSTR® kit AmpliTaq Gold® DNA Polymerase*);
  - **Y chromosomos TTKS žymenų alelių rinkinys** (angl. *AmpFlSTR® Yfiler® Allelic Ladder*).

### Įranga ir laboratoriniai reikmenys:

- **96 šulinėlių optinė sekoskaitos plokštelė** (angl. *MicroAmp™ Optical 96-Well Reaction Plate*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **96 šulinėlių optinė sekoskaitos plokštelės dengimo guma** (angl. *Plate Septa 96-Well*, gamintojas *Applied Biosystems® Life Technologies Corp.*, JAV);
- **Centrifuga** (96 šulinėlių plokštelėms);

- **Genetinius analizatorius**, 16 kapiliarų (angl. *ABI PRISM™ 3130xl Genetic Analyzer*, gamintojas *Hitachi High-Technologies Corp.*, Japonija);
- **Latekso pirštinės**;
- **Mikrocentrifuga** (0,2 ml, 0,5 ml, 1,5 ml);
- **Mikropipetės**, vienkanalės ir daugiakanalės (0,2–2,5 µl, 0,5–10 µl, 2–20 µl, 10–100 µl);
- **PGR mikromėgintuvėliai**, juostelė po 8 (angl. *Tubes 8-strips*);
- **Purtyklė**;
- **Sterilūs antgaliai** (10 µl, 20 µl, 100 µl);
- **Sterilūs mėgintuvėliai** (0,5 ml, 1,5 ml);
- **Termocikleris**.

**Darbo eiga:**

- (1) Paruošti 17 TTKS žymenų gausinimo mišinį (5 lentelė) n+1 reakcijoms (n – tiriamų mėginių, neigiamos kontrolės skaičius).

5 lentelė. PGR mišinys

Reagentai	Tūris (µl)/ reakcijai
PGR mišinys (angl. <i>AmpFISTR® Yfiler™ PCR reaction Mix</i> )	4,6
Pradmenų rinkinys (angl. <i>AmpFISTR® Yfiler™ Primer Set</i> )	2,5
DNR polimerazė (angl. <i>AmpFISTR® kit AmpliTaq Gold® DNA Polymerase</i> )	0,4
<b>Iš viso</b>	7,5

- (2) Įpilti po 7,5 µl 17 TTKS žymenų gausinimo mišinio (5 lentelė) į PGR mikromėgintuvėlius.
- (3) Į kiekvieną PGR mikromėgintuvėlį įpilti po 2 µl gDNR mėginio (konc. 2 ng/µl).
- (4) Į vieną PGR mikromėgintuvėlį įpilti 2 µl ddH<sub>2</sub>O.
- (5) PGR mikromėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (6) Paruoštus mišinius inkubuoti termocikleryje naudojant iš anksto paruoštą programą (6 lentelė).

6 lentelė. PGR sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	95	11:00	1
Denatūracija	95	1:00	30
Pradmenų prijungimas	61	1:00	
Sintezė	72	1:00	
Galutinė sintezė	60	80:00	1

- (7) \*Paruošti kapiliarinės elektroforezės mišinį (7 lentelė) n+1 reakcijoms (n – tiriamų mėginių, neigiamos kontrolės, Y chromosomos TTKS žymenų alelių rinkinio skaičius).

7 lentelė. Kapiliarinės elektroforezės mišinys

Reagentai	Tūris (µl)/ reakcijai
Hi-Di™ formamidas (angl. <i>Hi-Di™ Formamide</i> )	8,7
Molekulinis ilgio standartas (angl. <i>GeneScan™ 500 LIZ® Size Standard</i> )	0,3
<i>Iš viso</i>	9,0

- (8) \*Įpilti kapiliarinės elektroforezės mišinio (7 lentelė) į kiekvieną optinės sekoskaitos plokštelės šulinėlį, skirtą tiriamiems mėginiams, neigiamai kontrolei ir Y chromosomos TTKS žymenų alelių rinkiniui.
- (9) \*Įpilti po 0,5 µl kiekvieno pagausinto mėginio, neigiamos kontrolės ir 1 µl Y chromosomos TTKS žymenų alelių rinkiniui.
- (10) \*Paruoštą optinę sekoskaitos plokštelę centrifuguoti 1 500 aps./min 1 min.
- (11) \*Įjungti kompiuterį ir genetinį analizatorių.
- (12) \*Pakeisti, jei reikia, vandens ir sekoskaitos buferio indus atitinkamai nauju ddH<sub>2</sub>O ir 1× genetinio analizatoriaus buferiu, įsitikinti užtektinu POP-7™ polimero tūriu.
- (13) \*Aktyvuoti programą *Data Collection Software v3.0 (DCSv3.0)*.
- (14) \*Kairėje programos lango pusėje pasirinkti *Plate Manager*, spausti *New*.
- (15) \*Etapė *New Plate Dialog* laukelyje *Name* įvesti kapiliarinės elektroforezės pavadinimą, laukelyje *Application* pasirinkti pritaikymą fragmentų ilgiams nustatyti, laukelyje *Plate type* pasirinkti 96-Well,

laukelyje *Owner* įvesti genetinio analizatoriaus savininko pavadinimą, laukelyje *Operator* įvesti tyrėjo pavadinimą, spausti *Ok*.

- (16) \*Etapė *Plate Editor* stulpelyje *Sample name* įvesti tiriamų mėginių, neigiamos kontrolės, Y chromosomos TTKS žymenų alelių rinkinio pavadinimus atsižvelgiant į jų vietą optinėje sekoskaitos plokštelėje, stulpelyje *Sample type*: pasirinkti *Sample* arba *Allelic ladder*, stulpelyje *Size standard* pasirinkti molekulinį ilgio standartą, atitinkantį 500 np pažymėtą LIZ dažų, stulpelyje *Panel* pasirinkti tyrimo panelę, stulpelyje *Analysis method* pasirinkti Y chromosomos 17 TTKS žymenų nustatymo rinkinio rezultatams analizuoti pritaikytą metodą, stulpelyje *Result* pasirinkti kapiliarinės elektroforezės rezultatų saugojimo vietą, stulpelyje *Instrument protocol* pasirinkti genetinio analizatoriaus kapiliarinės elektroforezės protokolą, skirtą Y chromosomos 17 TTKS žymenų analizei, spausti *Ok*.
- (17) \*Optinė sekoskaitos plokštelė įtvirtinama į dviejų dalių adapterį.
- (18) \*Paspaudus *Tray* mygtuką genetinio analizatoriaus korpuse optinės sekoskaitos plokštelės platforma nukreipiama į plokštelės įstatymo į genetinį analizatorių poziciją.
- (19) \*Atidaryti genetinio analizatoriaus duris, įstatyti adapteryje įtvirtintą plokštelę į skirtą vietą platformoje.
- (20) \*Uždaryti genetinio analizatoriaus duris ir sulaukti platformos grįžimo į pradinę padėtį.
- (21) \*Programos *DCSv3.0* kairėje lango pusėje spausti *View Run*, spausti *Find all* ir pateiktame sąrašė pažymėti sukurtą kapiliarinės elektroforezės protokolą, spausti dešiniu klavišu ant dešinėje pusėje pavaizduotos plokštelės, susiejant sukurtą protokolą su įstatyta paruošta optine sekoskaitos plokštele.
- (22) \*Spausti *Run*, pateiktoje informacinėje lentelėje spausti *Ok* patvirtinant kapiliarinės elektroforezės vykdymą.
- (23) \*Pasibaigus kapiliarinei elektroforezei, aktyvuoti programą *GeneMapper<sup>®</sup> Software v4.0 (GMSv4.0)*.

- (24) \*Įkelti gautus kapiliarinės elektroforezės duomenis spaudžiant *File > Add Samples to Project*, pateiktame lange pasirinkti sukurtą katalogą su gautais duomenimis, spausti *Add to List > Add*.
- (25) \*Atlikti įkeltų mėginių analizę spaudžiant piktogramą *Analyze*, pateiktame lange laukelyje *Project Name* įrašyti projekto pavadinimą.
- (26) \*Pažymėti neigiamą kontrolę iš pateikto sąrašo ir spausti piktogramą *Size Match Editor*, įvertinti molekulinio ilgio standarto kokybę ir priskirtus ilgius, esant būtinybei atlikti taisymus, spausti *Save*.
- (27) \*Pažymėtos neigiamos kontrolės su įvertintu molekulinio ilgio standartu iš pateikto sąrašo peržiūrai spausti piktogramą *Display plots*, nenustačius Y chromosomos 17 TTKS žymenų fragmentų, atlikti likusių mėginių analizę.
- (28) \*26–27 etapas kartoti su likusiais mėginiais ir Y chromosomos TTKS žymenų alelių rinkiniu.
- (29) \* Iškelti išanalizuotų Y chromosomos 17 TTKS žymenų ilgių ir alelių rezultatus spaudžiant piktogramą *Genotype Table*, pasirenkant *File > Export Table*.

### **Mitochondrijų genomo sekoskaitos *Illumina GAI, GAI* genetiniu analizatoriumi protokolas**

Protokolas skirtas vienu metu nustatyti daugelio tiriamųjų viso mitochondrijų genomo (mtDNR) nukleotidų seką, atrenkant tikslines sekas hibridizacijos būdu, naudojant naujos kartos *Illumina Genome Analyzer II* genetinį analizatorių.

Darbą atlikto dokt. Ingrida Uktverytė Makso Planko evoliucinės antropologijos institute (Leipcigas, Vokietija) nuo 2011 m. rugsėjo 4 d. iki 2011 m. spalio 1 d. Darbui vadovavo prof. Mark Stoneking. (\*) etapai atlikti Makso Planko evoliucinės antropologijos instituto darbuotojų.



Protokolas parengtas remiantis:

- Makso Planko evoliucinės antropologijos instituto parengtu mtDNR sekoskaitos protokolu (angl. *Mitochondrial DNA (mtDNA) Sequencing Library Preparation.doc*);
- Bioanalizatoriaus rinkinio vartotojo vadovu (angl. *Agilent DNA 1000 Kit Guide, Manual P/N G2938-90014*);
- Klasterių formavimo sistemos vartotojo vadovu (angl. *Cluster Station User Guice, Part#15018818Rev.D*);
- Maricic, T., Whitten, M., and Pääbo, S. (2010). Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 5, e14004;
- Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010, pdb prot5448;
- Meyer, M., Stenzel, U., and Hofreiter, M. (2008). Parallel tagged sequencing on the 454 platform. *Nat Protoc* 3, 267-278.

**Reagentai ir cheminės medžiagos:**

- **1× BWT buferis** (angl. *1× BWT buffer*, ruošiamas laboratorijoje, sudėtis: 5 mM Tris-HCl, pH = 8.0; 0.5 mM EDTA; 1 M NaCl, 0,05 % Tween 20);
- **10× Tango buferis** (angl. *Tango Buffer 10×*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **2× BWT buferis** (angl. *2× BWT buffer*, ruošiamas laboratorijoje, sudėtis: 10 mM Tris-HCl, pH = 8,0; 1 mM EDTA; 2 M NaCl, 0,1 % Tween 20);
- **2× Phusion® PGR mišinys** kartu su DMSO (angl. *Phusion® Hot Start Flex 2× Master Mix*, gamintojas *New England Biolabs, Inc., JAV*);
- **3×/6× DNR įleidimo dažas** (angl. *DNA Loading dye, 3×/ 6×*; gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **99,96 % etanolis** (angl. *Ethanol Absolute*, gamintojas *Sigma-Aldrich Co., JAV*);

- **Agarozės (0,8 %; 2 %) gelis** (100 ml), siaurais šulinėliais (ruošiamas laboratorijoje, sudėtis: agarozės 0,8 g arba 2 g, 1× TBE buferio 100 ml, 2 μl etidžio bromido, 10 mg/ml);
- **Blokuojantys zondai** (angl. *Blocking Oligos, 1–10*);
- **Bst DNR polimerazė kartu su 10× ThermoPol® reakcijos buferiu** ilgiems fragmentams gausinti (angl. *Bst DNA Polymerase, Large Fragment; supplied with 10× ThermoPol® Reaction Buffer*, gamintojas *New England Biolabs, Inc., JAV*);
- **Dalelės, padengtos streptavidinu** (angl. *Dynabeads® M-270 Streptavidin*, 10 mg/ml, gamintojas *Invitrogen™ Life Technologies Corp., JAV*);
- **dATP**, deoksiadenozintrifosfatas (angl. *dATP, 100 mM Solution*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **DNR molekulinis ilgio ir masės standartas 50 np**, DNR 0,5 mg/ml (angl. *GeneRuler™ 50 bp DNA Ladder*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **DNR molekulinis ilgio ir masės standartas 100 np**, DNR 0,5 mg/ml (angl. *GeneRuler™ 100 bp DNA Ladder*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **DNR molekulinis ilgio ir masės standartas 500–10 000 np**, DNR 20 mg/ml (angl. *FastRuler™ High Range DNA Ladder, 500-10 000 bp*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **dNTP mišinys**, deoksinukleotidtrifosfatų mišinys (angl. *dNTP Mix, 25 mM each*, gamintojas *Thermo Fisher Scientific Baltics UAB, JAV*);
- **EBT buferis** (angl. *EBT buffer*, ruošiamas laboratorijoje, sudėtis: 10 mM Tris-Cl, pH = 8,0; 0,05 % Tween 20);
- Fragmentų galų bukinimo rinkinys (angl. *Quick Blunting™ Kit*, gamintojas *New England Biolabs, Inc., JAV*):
  - **dNTF mišinys**, 1 nM (angl. *Deoxynucleotide Solution Mix, 1 mM*);
  - **Galų bukinimo reakcijos buferis 10×** (angl. *10× Blunting Buffer*);
  - **Galų bukinimo fermentų mišinys** (angl. *Blunting Enzyme Mix*);

- **gDNR**, genominės DNR mėginys (konc. 10 ng/μl, tūris 80–90 μl, DNR kiekis ~850 ng);
- Hibridizacijos rinkinys (angl. *Agilent Oligo aCGH Hybridization Kit (25)*), gamintojas *Agilent Technologies, Inc.*, JAV):
  - **2× Hibridizacijos tirpalas** (angl. *2× Oligo aCGH Hybridization Solution*);
  - **10× Blokuojantis veiksnys** (angl. *10× Blocking Agent*);
- **H<sub>2</sub>O**, tinkamas HPLC;
- **HWT buferis** (angl. *HWT buffer*, ruošiamas laboratorijoje, sudėtis: 2,5 mM MgCl<sub>2</sub>; 50 mM KCl; 15 mM Tris-HCl, pH = 8,0; 0,1% Tween-20);
- Ligavimo rinkinys (angl. *Quick Ligation™ Kit*, gamintojas *New England Biolabs, Inc.*, JAV):
  - **2× ligavimo reakcijos buferis** (angl. *Quick Ligation Reaction Buffer 2×*);
  - **T4 DNR ligazė** (angl. *Quick T4 DNA Ligase*);
- **Magnetinės dalelės** (angl. *Agencourt® AMPure® XP*, gamintojas *Beckman Coulter, Inc.*, JAV);
- **Oligonukleotidai** (angl. *Oligonucleotides*, gamintojas *Sigma-Aldrich Co. LLC*, JAV);
- PGR mišinys ilgų fragmentų gausinimui (angl. *Expand Long Range dNTPack*, gamintojas *Roche Applied Science*, JAV):
  - **5× Reakcijos buferis be MgCl<sub>2</sub>** (angl. *Expand Long Range Buffer without MgCl<sub>2</sub>*);
  - **5× Reakcijos buferis su MgCl<sub>2</sub>** (angl. *Expand Long Range Buffer with 12.5 mM MgCl<sub>2</sub>*);
  - **DMSO, 100 %** (angl. *DMSO 100 %*);
  - **dNTP mišinys**, kiekvieno 10 mM (angl. *PCR Nucleotide Mix, each 10 mM*);
  - **Fermentų mišinys** (angl. *Expand Long Range Enzyme Mix*);
  - **MgCl<sub>2</sub> tirpalas**, 25 mM (angl. *MgCl<sub>2</sub> Stock Solution, 25 mM*);

- **PGR produktų valymo kolonėlės** (angl. *MinElute PCR Purification Kit*, *Qiaquick PCR Purification Kit*, gamintojas *QIAGEN*, JAV);
- **RL-PGR DNR standartas**, realaus laiko PGR DNR standartas (angl. *qPCR Standard*, ruošiamas laboratorijoje);
- **SYBR Green kiekybinės RL-PGR rinkinys** (angl. *DyNAmo Flash SYBR Green qPCR Kit*, gamintojas *New England Biolabs, Inc.*, JAV):
  - **ROX dažas** (angl. *ROX Reference Dye*);
  - **Kiekybinės RL-PGR mišinys** (angl. *SYBR Green qPCR Master Mix 2×*);
- **T4 DNR ligazė**, 5 U/μl, kartu su 10× T4 DNR ligazės buferiu ir 50 % PEG-4000 tirpalu (angl. *T4 DNA Ligase, 5 U/μl; supplied with 10× T4 DNA Ligase Buffer, 50 % PEG-4000 Solution*, gamintojas *Thermo Fisher Scientific Baltics UAB*, JAV);
- **T4 DNR polimerazė**, 5 U/μl (angl. *T4 DNA polymerase, 5 U/μl*, gamintojas *Thermo Fisher Scientific Baltics UAB*, JAV);
- **T4 PNK**, T4 polinukleotido kinazė 10 U/μl (angl. *T4 polynucleotide kinase, 10 U/μl*, gamintojas *Thermo Fisher Scientific Baltics UAB*, JAV);
- **TE buferis** (angl. *TE buffer*, ruošiamas laboratorijoje, sudėtis: 1 mM EDTA, pH = 8,0; 10 mM Tris-Cl, pH = 8,0; dH<sub>2</sub>O);
- **Teigiama DNR kontrolė**, 200–300 np ilgio PGR produktai, 200–500 ng (angl. *Positive Control DNA, 200–300 bp fragment, generated via PCR using unmodified primers and polymerase with terminal transferase activity, 200–500 ng*);
- **TET buferis** (angl. *TET buffer*, ruošiamas laboratorijoje, sudėtis: TE buferis, 0,05 % Tween 20);
- **Tween<sup>®</sup> 20** (angl. *TWEEN<sup>®</sup> 20*, gamintojas *Sigma-Aldrich Co.*, JAV).

#### Įranga ir laboratoriniai reikmenys:

- **96 šulinėlių plokštelių dengimo plėvelė** (angl. *MicroAmp<sup>™</sup> Optical Adhesive Film*, gamintojas *Applied Biosystems<sup>®</sup> Life Technologies Corp.*, JAV);

- **Centrifuga** (96 šulinėlių plokštelėms);
- **Gelių analizavimo sistema** (angl. *Herolab E.A.S.Y 442 K*, gamintojas *Herolab GmbH Laborgeraete*, Vokietija);
- **Genetinis analizatorius** (angl. *Genome Analyzer II*, gamintojas *Illumina, Inc.*, JAV);
- **Horizontalios elektroforezės sistema su maitinimo šaltiniu** (angl. *Midicell<sup>®</sup> EC350 Electrophoretic Gel System*, gamintojas *E-C Apparatus Corp.*, JAV);
- **Latekso ir nitrilo pirštinės**;
- **Magnetinis stovas 96 šulinėlių plokštelėms** (angl. *SPRI Plate 96R-Ring Super Magnet Plate* gamintojas *Agencourt<sup>®</sup>, Beckman Coulter, Inc.*, JAV);
- **Mikrocentrifuga** (0,2 ml, 0,5 ml, 1,5 ml, 2 ml);
- **Mikropipetės**, vienkanalės ir daugiakanalės (0,2–2,5 µl, 2–20 µl, 10–100 µl, 20–200 µl, 100–1000 µl);
- **PGR plokštelės** (96 šulinėlių, 0,2 ml);
- **PGR plokštelių dangteliai**, juostelės po 8;
- **Purtyklė** (angl. *Vortex-Genie 2, Model G560*, gamintojas *Scientific Industries Inc.*, JAV);
- **Reagentų lovelis**, daugiakanalėms pipetėms (angl. *Tip-Tube*, gamintojas *Eppendorf AG*, Vokietija);
- **RL-PGR sistema**, realaus laiko PGR sistema (angl. *Mx3005P QPCR System*, gamintojas *Stratagene, Agilent Technologies, Inc.*, JAV);
- **Skaidymo ultragarsu sistema su vėsinimu** (angl. *DNA shearing device, Bioruptor UCD-200*, gamintojas *Diagenode Inc.*, JAV);
- **Spektrofotometras** (angl. *NanoDrop<sup>®</sup> ND-1000 Spectrophotometer*, gamintojas *NanoDrop Technologies Inc.*, JAV);
- **Sterilūs antgaliai** (2,5 µl, 20 µl, 100 µl, 200 µl, 1000 µl);
- **Sterilūs mėgintuvėliai** (0,5 ml);
- **Sterilūs PGR mėgintuvėliai**, juostelės po 8 (0,2 ml);
- **Sterilūs optiniai dangteliai**, juostelės po 8;

- Šaldymo stovėlis arba grūstas ledas;
- Termocikleris (angl. *DNA Engine Tetrad<sup>®</sup> 2*, gamintojas *Bio-Rad Laboratories, Inc.* JAV).

**Darbo eiga:**

**\*mtDNR ilgų fragmentų gausinimas**

- (1) Paruošti ilgų mtDNR fragmentų gausinimo mišinį (8 lentelė), pradmuo P1 (H2986, L12279, produkto ilgis 7 276 np), pradmuo P2 (H12336, L2594, produkto ilgis 9 742 np).

8 lentelė. Ilgų mtDNR fragmentų gausinimo mišinys

Reagentai	Tūris (μl)/ reakcijai
dH <sub>2</sub> O	29,3
5× Reakcijos buferis su MgCl <sub>2</sub> (angl. <i>Buffer with 12.5 mM MgCl<sub>2</sub></i> )	10,0
dNTF mišinys (angl. <i>PCR Nucleotide Mix, each 10 mM</i> )	2,5
P1/P2 H pradmuo	3,0
P1/P2 L pradmuo	3,0
Fermentų mišinys (angl. <i>Enzyme Mix</i> )	0,7
<b>Iš viso</b>	48,5

- (2) Paruoštą mišinį išpilstyti po 48,5 μl į kiekvieną 0,2 ml tūrio mėgintuvėlį su 1,5 μl mėginio.
- (3) Mėgintuvėlius su mišiniu purtyti 2 s ir centrifuguoti 5 s.
- (4) Paruoštus mišinius inkubuoti termocikleryje naudojant iš anksto paruoštą programą (9 lentelė).

9 lentelė. mtDNR ilgų fragmentų gausinimo sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	92	02:00	1
Denatūracija	92	00:10	10
Pradmenų prijungimas	57	00:15	
Sintezė	68	08:00	
Denatūracija	92	00:10	27
Pradmenų prijungimas	57	00:15	
Sintezė	68	08:00 (+ 00:20 ciklui)	
Galutinė sintezė	68	07:00	1
Mėginių laikymas	7	∞	1

- (5) Atlikti pagausintos mtDNR įvertinimą (33–38 etapai) 0,8 % agarozės gelyje, naudojant 4 μl pagausinto mtDNR produkto, sumaišyto su 2 μl

3× įleidimo dažo, 4 μl DNR molekulinio ilgio ir masės standarto 500–10 000 bp.

- (6) Į du naujus 1,5 ml tūrio mėgintuvėlius sujungti atitinkamai P1 ir P2 mtDNR pagausintų produktų visus tūrius.
- (7) Atlikti P1 ir P2 mtDNR pagausintų produktų valymą magnetinėmis dalelėmis (52–68 etapai).
- (8) Įvertinti P1 ir P2 mtDNR pagausintų produktų koncentraciją spektrofotometru (žr. Genominės DNR koncentracijos ir švarumo matavimas).
- (9) Į naują 1,5 ml tūrio mėgintuvėlį sujungti paruoštus P1 ir P2 mtDNR produktus iki galutinio 3 μg kiekio, po 1,5 μg atitinkamai P1 ir P2 produkto.

#### **\*Ilgų mtDNR fragmentų skaidymas ultragarsu**

- (10) Įpilti 80 μl paruošto P1 ir P2 mtDNR produktų mišinio (kiekis 3 μg) į naują 0,5 ml tūrio mėgintuvėlį.
- (11) Perkelti 0,5 ml mėgintuvėlį su mišiniu į įjungtą ir paruoštą skaidymo ultragarsu sistemą (angl. *Bioruptor UCD-200*), įsitikinus užtektinu grūsto ledo kiekiu vėsavimo sistemoje.
- (12) Nustatyti skaidymo parametrus: galingumas 200 W (padėtis *HIGH*), ultragarso veiklos/pauzės trukmė po 30 s, ciklo trukmė 5 min.
- (13) Pasibaigus ciklui mėgintuvėlį centrifuguoti 7 000 aps./min 5 s.
- (14) 12–13 etapus kartoti vieną kartą (iš viso du ciklai).
- (15) Fragmentų ilgių pasiskirstymas tarp 400–800 np, vidurkis 600 np.
- (16) Atlikti skaidytų mtDNR produktų įvertinimą 2 % agarozės gelyje (33–38 etapai).

#### **\*Skaidytų mtDNR produktų galų bukinimas**

- (17) Paruošti skaidytų mtDNR produktų galų bukinimo mišinį 1,5 ml tūrio mėgintuvėlyje (10 lentelė).

10 lentelė. Skaidytos mtDNR fragmentų galų bukinimo mišinys

Reagentai	Tūris (μl)/ reakcijai
Galų bukinimo reakcijos buferis 10× (angl. <i>10× Blunting Buffer</i> )	10,0
dNTF mišinys, 1 nM (angl. <i>Deoxynucleotide Solution Mix, 1 mM</i> )	10,0
Galų bukinimo fermentų mišinys (angl. <i>Blunting Enzyme Mix</i> )	4,0
Skaidytas mtDNR produktas (< 20 μg)	76,0
<i>Iš viso</i>	100,0

- (18) Mėgintuvėlius su mišiniu purtyti 2 s ir centrifuguoti 5 s.
- (19) Paruoštus mėgintuvėlius su mišiniu inkubuoti kambario (20–25 °C) temperatūroje 30 min.
- (20) Atlikti reakcijos valymą PGR produktų valymo kolonėlėmis (angl. *MinElute PCR Purification Kit; Qiaquick PCR Purification Kit*) pagal gamintojo pateiktą protokolą.

**\*Adapterių prijungimas prie skaidytų mtDNR produktų**

- (21) Paruošti adapterių prijungimo prie skaidytos mtDNR fragmentų mišinį (11 lentelė) 1,5 ml tūrio mėgintuvėlyje.

11 lentelė. Adapterių prijungimo prie skaidytos mtDNR fragmentų galų mišinys

Reagentai	Tūris (μl)/ reakcijai
2× ligavimo reakcijos buferis (angl. <i>Quick Ligation Reaction Buffer</i> )	20,0
50 mM adapteriai (Sol-MP)	1,0
T4 DNR ligazė (angl. <i>Quick T4 DNA Ligase</i> )	4,0
Skaidytas mtDNR produktas po galų bukinimo reakcijos (< 5 μg)	15,0
<i>Iš viso</i>	40,0

- (22) Mėgintuvėlius su mišiniu purtyti ir centrifuguoti 5 s.
- (23) Paruoštus mišinius inkubuoti kambario (20–25 °C) temperatūroje 15 min.
- (24) Atlikti reakcijos valymą PGR produktų valymo kolonėlėmis *MinElute PCR Purification Kit* arba *Qiaquick PCR Purification Kit* pagal gamintojo pateiktą protokolą.
- (25) Įvertinti paruoštų mtDNR produktų koncentraciją spektrofotometru (žr. Genominės DNR koncentracijos ir švarumo matavimas).
- (26) Laikyti –20 °C temperatūroje iki naudojimo.

**gDNR skaidymas ultragarsu**



- (27) Įpilti po 85 µl gDNR (10 ng/µl, kiekis ~850 ng) į naujus 0,5 ml tūrio mėgintuvėlius.
- (28) Įstatyti į adapterį dvylika 0,5 ml mėgintuvėlių su gDNR.
- (29) Adapterį įstatyti į įjungtą ir paruoštą skaidymo ultragarsu sistemą (angl. *Bioruptor UCD-200*), įsitikinus užtektinu grūsto ledo kiekiu vėsavimo sistemoje.
- (30) Nustatyti skaidymo parametrus: galingumas 200 W (padėtis *HIGH*), ultragarso veiklos/pauzės trukmė po 30 s, ciklo trukmė 7 min.
- (31) Pasibaigus ciklui mėgintuvėlį centrifuguoti 7 000 aps./min 5 s.
- (32) 28–29 etapus kartoti keturis kartus (iš viso penki ciklai).
- (33) Fragmentų ilgių pasiskirstymas turi būti tarp 100–500 np, vidurkis 200 np.

#### **Skaidytos gDNR įvertinimas**

- (34) Paruoštą 2 % agarozės gelį įleisti į horizontalios elektroforezės vonelę, pripildytą 1× TBE buferio.
- (35) 10 µl skaidytos gDNR tirpalo sumaišyti su 5 µl 3× įleidimo dažų.
- (36) 15 µl mišinio įleisti į suformuotus šulinėlius.
- (37) Iš abiejų pusių skaidytos gDNR ir 3× įleidimo dažų mišinio į šulinėlius įleisti po 3,5 µl DNR molekulinio ilgio ir masės standarto 100 np.
- (38) Elektroforezę vykdyti 25 min, įtampa 100 V.
- (39) Genominės DNR mėginių kokybę įvertinti naudojant gelių analizavimo sistemą esant 302 nm bangos ilgiui (ultravioletas, UV).

#### **Skaidytos gDNR 96 šulinėlių plokštelės paruošimas**

- (40) Įpilti į atskirus plokštelės šulinėlius po 20 µl kiekvieno tiriamojo skaidytos gDNR mėginio, nustačius didelį mėginio kiekį.
- (41) Įpilti į atskirus plokštelės šulinėlius po 25 µl kiekvieno tiriamojo skaidytos gDNR mėginio, nustačius mažą mėginio kiekį.
- (42) Įpilti 20 µl teigiamos DNR kontrolės į vieną plokštelės šulinėlį.
- (43) Įpilti 20 µl H<sub>2</sub>O, neigiamos kontrolės, į vieną plokštelės šulinėlį.

#### **Skaidytų gDNR galų bukinimas**

- (44) Paruošti skaidytos gDNR galų bukinimo mišinį reagentų lovelyje (12 lentelė).

12 lentelė. Skaidytos gDNR fragmentų galų bukinimo mišinys

Reagentai	Tūris (µl)/ reakcijai
dH <sub>2</sub> O	6,6
10× Tango buferis (angl. <i>Tango Buffer</i> )	4,0
BSA (10 mg/ml)	1,0
dNTP mišinys, 2,5 mM (angl. <i>dNTP Mix</i> )	1,6
dATP, 10 mM (angl. <i>dATP Solution</i> )	4,0
T4 DNR polimerazė, 5 U/µl (angl. <i>T4 DNA polymerase</i> )	0,8
T4 PNK, 10 U/µl (angl. <i>T4 polynucleotide kinase</i> )	2,0
<b>Iš viso</b>	20,0

- (45) Įpilti po 20 µl paruošto mišinio į kiekvieną gDNR plokštelės šulinėlį.
- (46) Paruoštą plokštelę kruopščiai uždengti dengimo plėvele.
- (47) Plokštelę purtyti 2 s ir centrifuguoti 1000 g 30 s.
- (48) Paruoštą plokštelę su mišiniu inkubuoti termocikleryje 12 °C temperatūroje 15 min, vėliau 25 °C temperatūroje 15 min, po inkubacijos mišinius nedelsiant valyti magnetinėmis dalelėmis (53–69 etapai) arba laikyti 4 °C temperatūroje.

#### Magnetinių dalelių aktyvumo patikrinimas

- (49) Įpilti 1 µl Tween<sup>®</sup>20 į 2 ml magnetinių dalelių tirpalą iki galutinės 0,05 % koncentracijos.
- (50) Įpilti 35 µl DNR molekulinio ilgio ir masės standarto 50 np į 245 µl dH<sub>2</sub>O.
- (51) Paruoštų magnetinių dalelių įpilti skirtingomis dalimis (1,0; 1,4; 1,8; 2,2; 2,6; 3,0) į 40 µl skiesto DNR molekulinio ilgio ir masės standarto 50 np.
- (52) Atlikti įvertinimą 2 % agarozės gelyje (34–39 etapai).

#### Mišinio valymas magnetinėmis dalelėmis

- (53) Iš naujo paruošti 70 % etanolio.
- (54) Magnetines daleles supurtyti ir palikti sušilti tamsoje iki kambario (20–25 °C) temperatūros.
- (55) Į kiekvieną plokštelės šulinėlį įpilti 88 µl (220 % mėginio tūrio) gerai supurtytų magnetinių dalelių.
- (56) Plokštelę kruopščiai uždengti dengimo plėvele.

- (57) Paruoštą plokštelę purtyti 10 s ir palikti kambario (20–25 °C) temperatūroje 5 min.
- (58) Plokštelę centrifuguoti 2000 g 15 s.
- (59) Plokštelę su mišiniu perkelti į magnetinį stovą, laukti mažiausiai 3–5 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant ant šulinėlių sienelių susirinkusių dalelių ir pašalinti.
- (60) Pakartotinai išsiurbti tirpalą neliečiant ant šulinėlių sienelių susirinkusių dalelių, ir pašalinti.
- (61) Į magnetiniame stove įstatytos plokštelės kiekvieną šulinėlį įpilti po 150 µl naujai paruošto 70 % etanolio, laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, etanolį, neliečiant ant šulinėlių sienelių susirinkusių dalelių, išsiurbti ir pašalinti.
- (62) 61 etapą kartoti dar kartą.
- (63) Pakartotinai išsiurbti likusį etanolį neliečiant ant šulinėlių sienelių susirinkusių dalelių, ir pašalinti.
- (64) Mėgintuvėlius su magnetinėmis dalelėmis atidaryti ir palikti džiūti kambario (20–25 °C) temperatūroje 20–30 min.
- (65) Į kiekvieną plokštelės šulinėlį su išdžiovintomis magnetinėmis dalelėmis įpilti 22 µl EBT buferio.
- (66) Plokštelę kruopščiai uždengti dengimo plėvele.
- (67) Paruoštą plokštelę purtyti 10 s purtykle ir palikti kambario (20–25 °C) temperatūroje 1 min.
- (68) Plokštelę centrifuguoti 2000 g 15 s.
- (69) Plokštelę su buferio ir dalelių mišiniu perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, buferį išsiurbti neliečiant ant šulinėlių sienelių susirinkusių dalelių ir perkelti į naują 96 šulinėlių plokštelę.

#### **Adapterių MP-P5 ir MP-P7 paruošimas**

- (70) Įpilti 1× TE buferio iki adapterių (13 lentelė) galutinės 500 µM koncentracijos.

13 lentelė. MP-P5 ir MP-P7 adapteriai

Adapteris	Nukleotidų seka
Sol-MP-P5	A*C*A*C*TCTTTCCTACACGACGCTCTCCG*A*T*C*T
Sol-MP-P7	Biotinas- G*T*G*A*CTGGAGTTCAGACGTGTGCTCTTCCG*A*T*C*T
Sol-MP-P5P7-comp	A*G*A*T*CGGAA*G*A*G*C

(71) Paruošti 10× Hibridizacijos buferį, įpilant 100 µl 5 M NaCl į 900 µl 1× TE buferį.

(72) Paruošti atskirai MP-P5 ir MP-P7 adapteriams hibridizacijos mišinį (14 lentelė).

14 lentelė. MP-P5 ir MP-P7 hibridizacijos mišinys

Reagentai	Tūris (µl)/ reakcijai
Sol-MP-P5/ Sol-MP-P7 (500 µM)	20,0
Sol-MP-P5P7-comp (500 µM)	20,0
10× Hibridizacijos buferis	5,0
1× TE buferis	5,0
<i>Iš viso</i>	50,0

(73) Paruoštą mišinį inkubuoti termocikleryje 95 °C temperatūroje 1 min, toliau mažinant temperatūrą 0,1 °C/s greičiu iki galutinės 14 °C temperatūros.

(74) Po hibridizacijos sumaišyti MP-P5 ir MP-P7 adapterius santykiu 1:1 iki galutinės kiekvieno adapterio 100 µM koncentracijos, mėgintuvėlį pažymėti Sol-MP.

### Adapterių prijungimas prie gDNR

(75) Paruošti adapterių prijungimo reakcijos mišinį (15 lentelė).

15 lentelė. Adapterių prijungimo prie gDNR mišinys

Reagentai	Tūris (µl)/ reakcijai	Galutinė konc.
dH <sub>2</sub> O	10,0	
10× T4 ligazės buferis (angl. <i>T4 Ligase buffer</i> )	4,0	1×
PEG-4000 (50 %)	4,0	16 %
T4 DNR ligazė, 5 U/µl (angl. <i>T4 Ligase</i> )	1,0	0,125 U/µl
<i>Iš viso</i>	19,0	

(76) Įpilti po 19 µl paruošto mišinio į kiekvieną plokštelės su gDNR, teigiama, neigiama kontrolėmis šulinėlių.

- (77) Įpilti po 1 µl paruošto Sol-MP adapterių mišinio į kiekvieną plokštelės su gDNR, teigiama, neigiama kontrolėmis šulinėlių.
- (78) Paruoštą plokštelę kruopščiai uždengti dengimo plėvele.
- (79) Paruoštą plokštelę purtyti 10 s ir inkubuoti termocikleryje 22 °C temperatūroje 60 min.
- (80) Atlikti reakcijos mišinio valymą magnetinėmis dalelėmis, santykis 1:2,2 (53–69 etapai).

### Prijungtų adapterių galų pildymo reakcija

- (81) Paruošti adapterių galų pildymo reakcijos mišinį (16 lentelė).

16 lentelė. Adapterių galų pildymo reakcijos mišinys

Reagentai	Tūris (µl)/ reakcija	Galutinė konc.
dH <sub>2</sub> O	10,5	
10× ThermoPol <sup>®</sup> reakcijos buferis (angl. <i>10× ThermoPol<sup>®</sup> Reaction Buffer</i> )	4,0	1×
dNTF mišinys, 2,5 mM (angl. <i>dNTP Mix</i> )	4,0	250 µM kiekvieno
<i>Bst</i> DNR polimerazė, 8U/µl (angl. <i>Bst DNA Polymerase, Large Fragment</i> )	1,5	0,3 U/ µl
<b>Iš viso</b>	20,0	

- (82) Įpilti po 20 µl paruošto reakcijos mišinio į kiekvieną plokštelės su gDNR, teigiama, neigiama kontrolėmis šulinėlių.
- (83) Paruoštą plokštelę kruopščiai uždengti dengimo plėvele.
- (84) Paruoštą plokštelę purtyti 10 s ir inkubuoti termocikleryje 22 °C temperatūroje 60 min.
- (85) Atlikti reakcijos mišinio valymą magnetinėmis dalelėmis, santykis 1:2,2 (53–69 etapai).
- (86) Atlikti adapterių prijungimo ir pildymo reakcijų efektyvumo įvertinimą (34–39 etapai) 2 % agarozės gelyje.

### Kiekybinė RL-PGR

- (87) Skiesti aštuonis gDNR mėginius (šulinėlius, patenkančius į plokštelės įstrižainę nuo A1 iki H8) ir neigiamą kontrolę iš paruoštos plokštelės, įpilant 1 µl gDNR į naujus 0,2 ml mėgintuvėlius su 9 µl TE buferio (skiedimas 1:9).
- (88) Paruoštus mėgintuvėlius su mišiniu purtyti 10 s ir centrifuguoti 5s.

- (89) Skiesti paruoštus gDNR mėginius (87 etapas) ir neigiamą kontrolę, įpilant 1 µl gDNR į naujus 0,2 ml mėgintuvėlius su 9 µl TE buferio (galutinis skiedimas 1:99).
- (90) Paruoštus mėgintuvėlius su mišiniu purtyti 10 s ir centrifuguoti 5s.
- (91) Skiesti koncentruotą RL-PGR DNR standartą (konc. 10<sup>9</sup>) (17 lentelė)
- 17 lentelė. Kiekybinės RL-PGR DNR standartų skiedimai

DNR standarto numeris	Skiedimas (iš konc. DNR 10 <sup>9</sup> )	Komponentai	Galutinė koncentracija
1	1:9	1 µl konc. DNR + 9 µl TE buferis	10 <sup>8</sup>
2	1:99	1 µl DNR standarto 1 + 9 µl TE buferis	10 <sup>7</sup>
3	1:999	1 µl DNR standarto 2 + 9 µl TE buferis	10 <sup>6</sup>
4	1:9 999	1 µl DNR standarto 3 + 9 µl TE buferis	10 <sup>5</sup>
5	1:99 999	1 µl DNR standarto 4 + 9 µl TE buferis	10 <sup>4</sup>
6	1:999 999	1 µl DNR standarto 5 + 9 µl TE buferis	10 <sup>3</sup>

- (92) Paruošti kiekybinės RL-PGR mišinį, atsižvelgiant į visų mėginių, DNR standartų, neigiamos kontrolės trijų kartotinių skaičių 1,5 ml tūrio mėgintuvėliuose (18 lentelė).

18 lentelė. Kiekybinės RL-PGR mišinys

Reagentai	Tūris (µl)/ reakcijai	Galutinė konc.
dH <sub>2</sub> O	7,0	
Pradmuo F (Sol_iPCR-P5) 10 µM	1,0	500 nM
Pradmuo R (Sol_iPCR-P7) 10 µM	1,0	500 nM
2× Kiekybinės RT-PGR mišinys (angl. <i>SYBR Green qPCR Master Mix</i> )	10,0	1×
<i>Iš viso</i>	19,0	

- (93) Įpilti po 19 µl paruošto RL-PGR mišinio (18 lentelė) į kiekvieną 96 šulinėlių optinės plokštelės šulinėlį, kur į kiekvieną bus įpilta trys kartotiniai mėginio, neigiamos kontrolės, DNR standarto.
- (94) Įpilti tik po 19 µl paruošto RL-PGR mišinio (18 lentelė) į kiekvieną 96 šulinėlių optinės plokštelės šulinėlį su neigiama RL-PGR kontrole – NTC (angl. *No Template Control*).
- (95) Įpilti po 1 µl (trys kartotiniai) paruoštų gDNR mėginių, neigiamos kontrolės, RL-PGR DNR standarto.
- (96) Paruoštą optinę plokštelę kruopščiai uždengti optiniais dangteliais.

- (97) Paruoštą plokštelę purtyti 10 s ir centrifuguoti 1000 g 1 min.
- (98) Įjungti RL-PGR sistemą, aktyvuoti programą *Mx3005P Real-Time PCR System Software (Mx3005P Software)*.
- (99) Programoje *Mx3005P Software* pasirinkti *Instrument > Lamp On* 20 min iki RL-PGR pradžios.
- (100) Spausti *File > New*.
- (101) Etape *New Experiment Options* skiltyje *Real-time* pasirinkti *SYBR Green (with Dissociation Curve)*, spausti *OK*.
- (102) Etape *Setup* skiltyje *Plate Setup* užpildyti paruoštos optinės plokštelės protokolą.
- (103) Pažymėti virtualioje plokštelėje laukelius (šulinėlius) su tiriamais mėginiais, dešinėje programos lango pusėje laukelyje *Well Type* pasirinkti *Unknown*, skiltyje *Collect Fluorescence data* pažymėti *SYBR* ir *ROX*, laukelyje *Reference Dye* pasirinkti *ROX*.
- (104) Pažymėti virtualioje plokštelėje laukelius (šulinėlius) su neigiama RL-PGR kontrole, dešinėje programos lango pusėje laukelyje *Well Type* pasirinkti *NTC*, skiltyje *Collect Fluorescence data* pažymėti *SYBR* ir *ROX*, laukelyje *Reference Dye* pasirinkti *ROX*.
- (105) Virtualioje plokštelės reprezentacijoje pažymėti laukelius (šulinėlius) su RL-PGR DNR standartu, dešinėje programos lango pusėje laukelyje *Well Type* pasirinkti *Standard*, skiltyje *Collect Fluorescence data* pažymėti *SYBR* ir *ROX*, laukelyje *Reference Dye* pasirinkti *ROX*, laukelyje *Standard Quantity* pasirinkti *ROX* ir įrašyti kopijų skaičių  $10^8$ , pasirinkti *Auto-decrement 10×*. Virtualioje plokštelės reprezentacijoje pažymėjus kito DNR standarto trijų kartotinių laukelius, priskiriami nustatyti parametrai, sumažinant kopijų skaičių iki  $10^7$ .
- (106) Skiltyje *Identify Replicates* laukelyje *Replicate Symbol* pasirinkti 1, spausti mygtuką *Auto-Increment* ir virtualioje plokštelės reprezentacijoje pažymėti pirmus kartotinius laukelius (šulinėlius), kuriems bus priskirtas kartotinio numeris 1, etapą kartoti su visais likusiais gDNR mėginio, DNR standarto, neigiamos RL-PGR kontrolės kartotiniais.

- (107) Užpildžius plokštelės protokolą, spausti *Next*.
- (108) Etape *Setup* skiltyje *Thermal Profile Setup* užpildyti RL-PGR sąlygas (19 lentelė).

19 lentelė. Kiekybinės RL-PGR sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	95	15:00	1
Denatūracija	94	00:30	40
Pradmenų prijungimas	58	00:30	
Sintezė	72	01:00	

- (109) Pažymėti disociacijos etapą ir programos dešinėje skiltyje *Selection* spausti *Delete*.
- (110) Įstatyti paruoštą optinę plokštelę į RL-PGR sistemą, uždaryti duris.
- (111) Spausti programos dešinėje *Start Run*, langelyje *Run Status* spausti *Start*.
- (112) Pasibaigus reakcijai, etape *Analysis*, įvertinti gautus rezultatus.
- (113) Įvertinti unikalių DNR sekų prijungimui reikalingų ciklų skaičių, nustatant  $C_t$  vertę prieš RL-PGR reakcijos plato arba plato stadijoje.
- (114) Atlikti kiekybinės RL-PGR pagausintų gDNR mėginių, DNR standarto, neigiamos RL-PGR kontrolės (skiedimai 1:9, 1:99) tikslinių produktų įvertinimą (34–39 etapai) 2 % agarozės gelyje, 5  $\mu$ l mėginio tirpalo sumaišyti su 1  $\mu$ l 6 $\times$  įleidimo dažų, 3  $\mu$ l DNR molekulinio ilgio ir masės standarto 50 np.

### Unikalių DNR sekų prijungimas prie gDNR

- (115) Paruošti dviejų unikalių DNR sekų prijungimo prie gDNR reakcijos mišinį (20 lentelė).

20 lentelė. Unikalių DNR sekų prijungimo prie gDNR mišinys

Reagentai	Tūris ( $\mu$ l)/ reakcijai	Galutinė konc.
dH <sub>2</sub> O	14,5	
2 $\times$ <i>Phusion</i> <sup>®</sup> PGR mišinys (angl. <i>Phusion</i> <sup>®</sup> <i>Hot Start Flex Master Mix</i> )	25,0	1 $\times$
<i>Iš viso</i>	39,5	

- (116) Įpilti po 39,5  $\mu$ l paruošto unikalių sekų prijungimo mišinio (20 lentelė) į kiekvieną 96 šulinėlių plokštelės šulinėlį.



- (117) Įpilti po 1,5 μl P5 unikalios DNR sekos adapterio į kiekvieną 96 šulinėlių plokštelės šulinėlį.
- (118) Įpilti po 1,5 μl P7 unikalios DNR sekos adapterio į kiekvieną 96 šulinėlių plokštelės šulinėlį.
- (119) Paruoštą plokštelę purtyti 10 s ir centrifuguoti 1000 g 1 min.
- (120) Paruoštą plokštelę inkubuoti termocikleryje pagal iš anksto paruoštą programą (21 lentelė), ciklų skaičių nustačius iš RL-PGR (113 etapas).

21 lentelė. Unikalių DNR sekų prijungimo reakcijos sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	98	00:30	1
Denatūracija	98	00:10	16/17 (113 etapas)
Pradmenų prijungimas	60	00:20	
Sintezė	72	00:20	
Galutinė sintezė	72	10:00	1

- (121) Atlikti dviejų unikalių DNR sekų prijungimo reakcijos įvertinimą (34–39 etapai) 2 % agarozės gelyje, 5 μl gDNR tirpalo sumaišyti su 1 μl 6× įleidimo dažų, 2,5 μl DNR molekulinio ilgio ir masės standarto 50 np.
- (122) Atlikti reakcijos mišinio valymą magnetinėmis dalelėmis, santykis 1:2,2 (53–69 etapai), 64 etape įpilti 25 μl EBT buferio.
- (123) Įvertinti gDNR mėginių koncentraciją spektrofotometru (žr. Genominės DNR koncentracijos ir švarumo matavimas).

#### **gDNR mėginių, žymėtų unikalėmis DNR sekomis, sujungimas**

- (124) Įpilti kiekvieno mėginio iš paruoštos plokštelės tūrį (μl) su galutine 200 ng gDNR mase į naują sterilų 0,5 ml tūrio sumažintos adhezijos mėgintuvėlį.

$$(125) V_{\text{Hibridizacijai}} (\mu\text{l}) = 2\,000 \text{ ng} / \frac{\sum_{i=1}^n m_i}{\sum_{i=1}^n v_i}$$

#### **Paruoštų ilgų mtDNR produktų prijungimas prie dalelių, padengtų streptavidinu**

- (126) Paruošti mtDNR fragmentų mišinį (22 lentelė) prijungti prie dalelių, padengtų streptavidinu.

22 lentelė. mtDNR fragmentų mišinys prijungti prie dalelių, padengtų streptavidinu

Reagentai	Tūris (µl)/ reakcijai	Galutinė konc.
mtDNR fragmentų mišinys (~139 ng/µl)	5,0	~700 ng
2× BWT buferis	10,0	1×
dH <sub>2</sub> O	5,0	
<i>Iš viso</i>	20,0	

- (127) Paruoštą mišinį inkubuoti termostate 95 °C temperatūroje 1 min, po inkubavimo nedelsiant perkelti į šaldymo bloką arba grūstą ledą.
- (128) Įpilti 7 µl dalelių, padengtų streptavidinu, į naują 0,5 ml tūrio mėgintuvėlį
- (129) Įpilti 200 µl 1× BWT buferio.
- (130) Mėgintuvėlį su mišinių purtyti 5 s, centrifuguoti 1 s, perkelti mėgintuvėlį į magnetinį stovą, laukti mažiausiai 1 min, kol tirpalas pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (131) Įpilti 200 µl TET buferio.
- (132) 130 etapą kartoti vieną kartą.
- (133) Įpilti 20 µl denatūruoto mtDNR fragmentų mišinio į magnetiniame stove įstatytą mėgintuvėlį su dalelėmis, dengtomis streptavidinu.
- (134) Mėgintuvėlį su mišiniu perkelti iš magnetinio stovo, mišinį sumaišyti pipetuojuant, inkubuoti kambario (20–25 °C) temperatūroje 20 min.
- (135) Mėgintuvėlį su mišiniu perkelti į magnetinį stovą, laukti mažiausiai 1 min, kol tirpalas pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (136) Įpilti 200 µl 1× BWT buferio.
- (137) 130 etapą kartoti vieną kartą.
- (138) 136–137 etapas kartoti vieną kartą.
- (139) Įpilti 50 µl TET buferio, mėgintuvėlį su mišiniu purtyti 5 s, centrifuguoti 1 s, laikyti 4 °C temperatūroje iki naudojimo.

### Hibridizacija

- (140) Įvertinti gDNR bibliotekos, sudarytos iš sujungtų gDNR mėginių, koncentraciją spektrofotometru (žr. Genominės DNR koncentracijos ir švarumo matavimas).

$$(141) \text{ Faktorius} = \left( \frac{2\,000\text{ ng}}{\text{Konc. (126 etapas)}} \right) / 17,8$$

(142) Paruošti hibridizacijos mišinį (23 lentelė) naujuose 0,5 ml tūrio mėgintuvėliuose kiekvienai gDNR bibliotekai.

23 lentelė. Hibridizacijos mišinys

Reagentai	Tūris (μl)/ reakcijai	Tūris (μl)/ reakcijai
		Faktorius (141 etapas): 3
BO1 (angl. <i>Blocking Oligos 1</i> )	0,5	1,5
BO2	0,5	1,5
BO3	0,5	1,5
BO4	0,5	1,5
BO5	0,5	1,5
BO6	0,5	1,5
BO7	0,5	1,5
BO8	0,5	1,5
BO9	0,5	1,5
BO10	0,5	1,5
10× blokuojantis veiksnys (angl. <i>Blocking Agent</i> )	5,2	15,6
2× hibridizacijos tirpalas (angl. <i>Hybridization Solution</i> )	26,0	78,0
gDNR biblioteka (2 μg)	17,8	53,4
<i>Iš viso</i>	54,0	162,0

(143) Paruoštą mėgintuvėlį su mišiniu inkubuoti termostate 95 °C temperatūroje 3 min.

(144) Mišinį perkelti į naują 1,5 ml tūrio mėgintuvėlį, inkubuoti termostate 37 °C temperatūroje 30 min.

(145) Mėgintuvėlį su mišiniu centrifuguoti 1 s.

(146) Mėgintuvėlį su prie dalelių prijungtais mtDNR fragmentais ir TET buferio mišiniu perkelti į magnetinį stovą, laukti mažiausiai 1 min, kol tirpalas pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.

(147) Įpilti visą tūrį hibridizacijos mišinio (142 etapas), mėgintuvėlį purtyti 1 s, centrifuguoti 1 s.

(148) Mėgintuvėlius kruopščiai uždaryti, apvynioti parafino juostelėmis.

(149) Paruoštus mėgintuvėlius inkubuoti iš anksto įjungtame termostate 65 °C temperatūroje 48 val.

#### **Mišinio valymas po hibridizacijos**

(150) HWT buferį inkubuoti 60 °C temperatūroje 30 min iki naudojimo.

- (151) Mėgintuvėlius su hibridizacijos mišiniu perkelti į magnetinį stovą, laukti mažiausiai 1 min, kol tirpalas pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (152) Įpilti 200 μl 1× BWT buferio.
- (153) Mėgintuvėlį su mišiniu purtyti 5 s, centrifuguoti 1 s, perkelti mėgintuvėlį į magnetinį stovą, laukti mažiausiai 1 min, kol tirpalas pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (154) 152–153 etapus kartoti tris kartus.
- (155) Įpilti 200 μl 60 °C temperatūros 1× HWT buferio (150 etapas).
- (156) Mėgintuvėlį su mišiniu perkelti iš magnetinio stovo į termostatą ir inkubuoti 60 °C temperatūroje 2 min.
- (157) Įpilti 200 μl 1× BWT buferio.
- (158) 153 etapą kartoti vieną kartą.
- (159) Įpilti 100 μl TET buferio ir visą mišinio tūrį įpilti į naują 0,5 ml sumažintos adhezijos tūrio mėgintuvėlį.
- (160) 153 etapą kartoti vieną kartą.
- (161) Įpilti 30 μl EBT buferio, inkubuoti termostate 95 °C temperatūroje 3 min, po inkubavimo nedelsiant perkelti į šaldymo bloką arba grūstą ledą.
- (162) Visą mišinio tūrį įpilti į naują 0,5 ml sumažintos adhezijos tūrio mėgintuvėlį.

### **Kiekybinė RL-PGR**

- (163) Atlikti visų gDNR bibliotekų kiekybinę RL-PGR, naudojant pradmenį F (Sol\_bridge\_P5) 10 μM, pradmenį R (Sol\_bridge\_P7) 10 μM (87–112 etapas).
- (164) Atlikti kiekybinės RL-PGR pagausintų gDNR mėginių, DNR standarto, neigiamos RL-PGR kontrolės (skiedimai 1:9, 1:99) tikslinių produktų įvertinimą (34–39 etapai) 2 % agarozės gelyje (114 etapas).
- (165) Ciklų skaičius =  $\log_2 \frac{2 \times 10^{12}}{\text{konc.}(148 \text{ etapas}) \times 10}$

### **gDNR bibliotekos gausinimas**

- (166) Paruošti gDNR bibliotekos gausinimo reakcijos mišinį (24 lentelė).

24 lentelė. gDNR bibliotekos gausinimo mišinys

Reagentai	Tūris (μl)/ reakcijai	Galutinė konc.
dH <sub>2</sub> O	12,5	
Pradmuo F (Sol_bridge_P5) 10 μM	1,25	250 nM
Pradmuo R (Sol_bridge_P7) 10 μM	1,25	250 nM
2× <i>Phusion</i> <sup>®</sup> PGR mišinys (angl. <i>Phusion</i> <sup>®</sup> <i>Hot Start Flex Master Mix</i> )	25,0	1×
<b><i>Iš viso</i></b>	40,0	

- (167) Įpilti po 40 μl paruošto gDNR gausinimo mišinio (24 lentelė).
- (168) Įpilti 10 μl paruoštos gDNR bibliotekos į kiekvieną mėgintuvėlį su mišiniu.
- (169) Mėgintuvėlius su mišiniu purtyti 10 s ir centrifuguoti 5 s.
- (170) Paruoštą mėgintuvėlį su mišiniu inkubuoti termocikleryje pagal iš anksto paruoštą programą (25 lentelė), nustačius iš RL-PGR įvertintą ciklų skaičių (165 etapas).

25 lentelė. gDNR bibliotekos gausinimo sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	98	00:30	1
Denatūracija	98	00:10	12/13 (165 etapas)
Pradmenų prijungimas	60	00:30	
Sintezė	72	00:30	
Galutinė sintezė	72	05:00	1
Mėginių laikymas	4	∞	1

- (171) Atlikti reakcijos mišinio valymą magnetinėmis dalelėmis, santykis 1:1,4 (53–69 etapai), 65 etape įpilti 22 μl EBT buferio.
- (172) Atlikti pagausintos gDNR bibliotekos įvertinimą bioanalizatoriumi pagal gamintojo pateiktas rekomendacijas (angl. *Agilent DNA 1000 Kit Guide*), nustatyti paruoštos gDNR bibliotekos molinę koncentraciją.
- (173) Atlikti paruoštos gDNR bibliotekos skiedimą Tris-Cl 10 mM, pH = 8,5; 0,1 % *Tween*<sup>®</sup> 20 iki galutinės 2 nM koncentracijos.
- (174) Tęsti protokolą pagal *Illumina* Klasterių formavimo sistemos vartotojo vadovą.

## **Žmogaus genominės DNR vieno nukleotido polimorfizmų genotipavimas *Illumina Infinium*<sup>®</sup> *HD* technologijos lustais**

Protokolas skirtas *Infinium*<sup>®</sup> *HD* technologija genotipuoti vieno nukleotido polimorfizmus ant lustų, naudojant *Illumina HiScan*<sup>™</sup>*SQ* sistemą. Protokolas įdiegtas VU MF ŽMGK dokt. Ingridos Domarkienės, dokt. Ingridos Uktverytės, dr. Laimos Ambrozaitytės.

Protokolas parengtas pagal gamintojo pateiktas rekomendacijas:

- Naudotojo vadovas (angl. *Infinium*<sup>®</sup> *HD Assay Ultra Manual Experienced User Card*);
- Išplėstinio protokolo vadovas (angl. *Infinium*<sup>®</sup> *HD Assay Ultra Protocol Guide*);
- Laboratorijos paruošimo ir procedūrų aprašas (angl. *Illumina Infinium*<sup>®</sup> *Assay Lab Set Up and Procedures*).

Protokolo vykdymo metu pildoma gamintojo rekomenduojama protokolo eigos forma:

- Protokolo eigos vykdymo forma (angl. *Infinium*<sup>®</sup> *HD Assay Ultra Lab tracking Form*).

**Dirbant laikomasi geros laboratorinės praktikos, laboratorijos priežiūros, parengimo darbui laboratorijoje ir protokolų reikalavimų.**

### **Reagentai ir cheminės medžiagos:**

- **2-propanolis (izopropanolis)**, 100 % (angl. *2-propanol, iso-propyl alcohol*, gamintojas *Sigma-Aldrich, Co.*, JAV);
- **Baliklio tirpalas**, 99,6 % natrio hipochlorito tirpalas (Kometa Pelėsių valiklis, gamintojas *Pufas Werk KG*, Vokietija);
- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **Detergento tirpalas**, 1 % (angl. *Alconox*<sup>®</sup> *Powdered Precision Cleaner*, gamintojas *Alconox, Inc.*, JAV);
- **Etanolis**, 100 % (angl. *Absolut Ethanol*, gamintojas *Sigma-Aldrich, Co.*, JAV);
- **FA/ EDTA tirpalas** (95 % formamidas, 1 mM EDTA);
- **gDNR**, genominės DNR mėginys (konc. 50 ng/μl, tūris 4 μl, DNR kiekis 200 ng);
- *Infinium*<sup>®</sup> *HD Assay Kit WGS-Pre LVI* rinkinys:

- **MA1 reagentas**, daugiau nei vieno mėginio gausinimo mišinys 1 (angl. *Multi-Sample Amplification 1 Mix*, gamintojas *Illumina Inc.*, JAV);
- **MA2 reagentas**, daugiau nei vieno mėginio gausinimo mišinys 2, pradmenys (angl. *Multi-Sample Amplification 2 Mix*, gamintojas *Illumina Inc.*, JAV);
- **MSM reagentas**, pagrindinis daugiau nei vieno mėginio gausinimo mišinys (angl. *Multi-Sample Amplification Master Mix*, gamintojas *Illumina Inc.*, JAV);
- *Infinium<sup>®</sup> HD Assay Kit WG-Post 1 LV1* rinkinys:
  - **ATM reagentas**, dvispalvis reagentų mišinys, fluoroforo žymeį sustiprinti (angl. *Anti-Stain Two-Color Master Mix*, gamintojas *Illumina Inc.*, JAV);
  - **FMS reagentas**, fragmentavimo tirpalas (angl. *Fragmentation Solution*, gamintojas *Illumina Inc.*, JAV);
  - **STM reagentas**, pagrindinis dvispalvis dažymo reagentų mišinys (žymėjimui fluoroforais) (angl. *Superior Two-Color Master Mix*, gamintojas *Illumina Inc.*, JAV);
  - **TEM reagentas**, dvispalvis vieno nukleotido pratęsimo reakcijos reagentų mišinys (angl. *Two-Color Extension Master Mix*, gamintojas *Illumina Inc.*, JAV);
  - **XC1 reagentas**, VNP lustų dažų tirpalas 1, blokuojantis (angl. *Xstain BeadChip Solution 1*, gamintojas *Illumina Inc.*, JAV);
  - **XC2 reagentas**, VNP lustų dažų tirpalas 2, buferio keitimui (angl. *Xstain BeadChip Solution 2*, gamintojas *Illumina Inc.*, JAV);
- *Infinium<sup>®</sup> Assay Kit Post 2 LMV* rinkinys:
  - **RA1 reagentas**, resuspendacijos, hibridizacijos ir plovimo (stipraus surišimo) tirpalas (angl. *Resuspension, hybridization and wash solution*, gamintojas *Illumina, Inc.*, JAV);
- *Infinium<sup>®</sup> Assay Kit Single Post 3 LV* rinkinys:
  - **PB2 reagentas**, drėgmę palaikantis hibridizacijos buferis (angl. *Humidifying buffer used during hybridization*, gamintojas *Illumina, Inc.*, JAV);
  - **PM1 reagentas**, nusodinimo tirpalas (angl. *Precipitation solution*, gamintojas *Illumina, Inc.*, JAV);
- *Infinium<sup>®</sup> Assay Kit Post 4 LV* rinkinys:

- **PB1 reagentas**, paruošiamasis VNP lustų hibridizacijos reagentas, silpno surišimo plovimui (angl. *Reagent used to prepare BeadChips for hybridization*, gamintojas *Illumina Inc.*, JAV);
- **XC3 reagentas**, VNP lustų dažų tirpalas 3 silpno surišimo plovimui (angl. *Xstain BeadChip Solution 3*, gamintojas *Illumina Inc.*, JAV);
- **XC4 reagentas**, VNP lustų dažų tirpalas 4 dažų apsaugojimui (angl. *Xstain BeadChip Solution 4*, gamintojas *Illumina Inc.*, JAV);
- **NaOH**, 0,1 N (ruošiamas laboratorijoje);
- **VNP lustai**, *Infinium® HD* technologija (angl. *BeadChip Array 12-sample HD DNA Analysis BC HumanOmniExpress-12v1-1*, gamintojas *Illumina Inc.*, JAV).

#### Įranga ir laboratoriniai reikmenys:

- **Beplaušės servetėlės** (angl. *lint-free wipes*);
- **Beplaušiai sugertukai** (angl. *lint-free absorbent pad*);
- **Centrifuga** 15 ml mėgintuvėliams (angl. *Hettich Zentrifugen Universal 32 R*, gamintojas Hettich Zentrifugen, Vokietija);
- **Daugiavietis VNP lustų talpiklis** (angl. *Multi-Sample BeadChip Alignment Fixture*, gamintojas *Illumina Inc.*, JAV);
- **Folijos užlydytuvas**, plokštelių dengiamosios folijos užlydytuvas (angl. *Combi Thermo-Sealer ALPS™ 25 (Heat Sealer)*, gamintojas Thermo Fisher Scientific, JK);
- **Hibridizacijos kameros** (angl. *Hybridization chambers*, gamintojas *Illumina Inc.*, JAV);
- **Hibridizacijos kameros guminis tarpiklis** (angl. *Hybridization chamber gasket*, gamintojas *Illumina Inc.*, JAV);
- **Hibridizacijos kameros įdėklai** (angl. *Hybridization chamber inserts*, gamintojas *Illumina Inc.*, JAV);
- **Hibridizacijos krosnis**, hibridizacijos krosnis su lustų hibridizacijos padėklų, modelis 5521 (angl. *Illumina® Hybridization Oven, Model 5521 with BeadChip Hyb Wheel*, gamintojas *Boehel Scientific*, JAV);
- **HiScan™SQ sistema, genetinis analizatorius** (angl. *HiScan™SQ system*, gamintojas *Illumina Inc.*, JAV);



- **Laikiklis ant talpiklio** (angl. *Multi-Sample BeadChip Alignment Fixture Alignment Bar*, gamintojas *Illumina Inc.*, JAV);
- **Mėgintuvėlių stovai**;
- **Mentelė** gnybtukams nuimti (angl. *dismantling tool*, gamintojas *Illumina Inc.*, JAV);
- **Mikropipetės**, vienkanalės ir daugiakanalės (2–20 µl, 10–100 µl, 20–200 µl, 30–300 µl, 100–1000 µl);
- **Mikroplokštelių inkubatorius Hybex<sup>®</sup>** (angl. *Hybex<sup>®</sup> Microsample Incubator (Heat block)*, gamintojas *SciGene Corp.*, JAV);
- **MSA3 plokštelė**, 96 šulinėlių polipropileno 0,8 ml mikrotitro plokštelė (MIDI) (angl. *MSA3 plate*, gamintojas *Thermo Scientific*, JK);
- **MSA3 plokštelės 96 šulinėlių dengiamoji guma** (angl. *96-well cap mat*, gamintojas *Thermo Scientific*, JK);
- **MSA3 plokštelės dengiamoji folija** (angl. *Heat Sealing Foil Sheets*, gamintojas *Thermo Scientific*, JK);
- **Nešiojamas termometras OMEGA<sup>®</sup>**, nešiojamas RTD termometras 869C (angl. *OMEGA<sup>®</sup> 869C Handheld RTD (resistance temperature detector) Thermometer*, *OMEGA<sup>®</sup>*, gamintojas *OMEGA Engineering Inc.*, JAV);
- **Pincetas** stiklelių perkėlimui (angl. *Self-locking tweezers 6 1/2"*, gamintojas *Aven Inc.*, JAV);
- **Plastikiniai antgaliai su filtrais** (20 µl, 100 µl, 200 µl, 1000 µl);
- **Plastikiniai loveliai**, reagentų lovelis daugiakanalėms pipetėms (angl. *Tip-tubes*, gamintojas *Eppendorf AG*, Vokietija);
- **Plokštelių centrifuga** (angl. *Heraeus Multifuge 3L*, gamintojas *Thermo Electron Corp.*, Vokietija);
- **Plokštelių purtyklė**, didelio greičio 230V skaitmeninė plokštelių purtyklė (angl. *Illumina<sup>®</sup> High Speed Microplate Shaker Digital 230V*, gamintojas *Henry Troemer, LLC*, JAV);
- **Pratekamosios kameros**, Te-Flow<sup>™</sup> pratekamosios kameros su juodais rėmeliais, tarpinėmis, dengiamaisiais stiklais ir gnybtukais (angl. *Te-Flow<sup>™</sup> Flow-Through Chambers with black frames, spacers, glass backplates, and clamps*, gamintojas *Illumina Inc.*, JAV);

- **Pratekamųjų kamerų laikiklis**, 48 vietų (angl. *Tecan Te-Flow™ Module Chamber Rack 48 Pos.*, gamintojas *Tecan Schweiz AG*, Šveicarija);
- **Sterilūs mikromėgintuvėliai** (0,2 ml, 0,5 ml, 1,5 ml);
- **Stikliniai ir plastikiniai indai** plovimui (angl. *Wash dish*, gamintojas *Illumina Inc.*, JAV);
- **Suspaustas oras** nupūsti dulkiems nuo dengiamųjų stiklų (angl. *Air pressure cleaner (Laboratory air gun)*, gamintojas *UAB ACME® Europe*, Lietuva);
- **Termostatas su vandens cirkuliacijos sistema**, pažangus programuojamas temperatūros reguliatorius su cirkuliacijos sistema (angl. *Advanced Programmable Controller Refrigerated/ Heated Circulating Baths*, gamintojas *VWR® International, LLC*, JAV);
- **Vakuuminė džiovyklė**, vakuuminė džiovyklė „Space Saver“ (angl. *Vakuum Desiccator "Space Saver"*, gamintojas *Scienceware® Bel-ART products*, JAV);
- **VNP lustų laikiklis** plovimui po dažymo, plastmasinis (angl. *Staining rack*, gamintojas *Illumina Inc.*, JAV);
- **VNP lustų laikiklis** plovimui prieš dažymą, metalinis (angl. *Wash rack*, gamintojas *Illumina Inc.*, JAV);
- **Žirklys** tarpinėms perkirpti (gamintojas *Illumina Inc.*, JAV).

***Eiga:***

**gDNR gausinimas (MSA3 plokštelės paruošimas) pre-PGR patalpoje**

- (1) Atsižvelgiant į tiriamų DNR mėginių skaičių į naują MSA3 plokštelę atitinkamai išpilstyti po 20 µl MA1 reagento.
- (2) Į atitinkamus (pagal pildomą protokolą) MSA3 plokštelės šulinėlius įpilti po 4 µl gDNR (~ 200 ng).
- (3) Į kiekvieną šulinėlį su MA1 ir gDNR įpilti po 4 µl 0,1 N NaOH.
- (4) Uždengti plokštelę dengiamąja guma, purtyti plokštelių purtyklėje 1600 aps./min 1 min, centrifuguoti 280 g 1 min, inkubuoti plokštelę kambario (20–25 °C) temperatūroje 10 min.
- (5) Į kiekvieną šulinėlį su mėginiais įpilti po 34 µl MA2 reagento;
- (6) Į kiekvieną šulinėlį su mėginiais įpilti po 38 µl MSM reagento;

- (7) Uždengus plokštelę dengiamąja guma, purtyti plokštelių purtyklėje 1600 aps./min 1 min, centrifuguoti 280 g 1 min, inkubuoti plokštelę hibridizacijos krosnyje 37 °C temperatūroje 20–24 val.

#### **gDNR fermentinis skaidymas post-PGR patalpoje**

- (8) Išėmus iš hibridizacijos krosnies MSA3 plokštelę centrifuguoti 50 g 1 min.
- (9) Į kiekvieną šulinėlį su mėginiu įpilti po 25 µl FMS reagento, uždengti plokštelę dengiamąja guma.
- (10) Purtyti plokštelę 1600 aps./min 1 min, centrifuguoti 50 g 1 min.
- (11) Inkubuoti plokštelę mikroplokštelių inkubatoriuje Hybex<sup>®</sup> 37 °C temperatūroje 1 val.

#### **gDNR nusodinimas**

- (12) Į kiekvieną šulinėlį su mėginiu įpilti po 50 µl PM1 reagento, uždengti plokštelę dengiamąja guma.
- (13) Purtyti plokštelę 1600 aps./min 1 min.
- (14) Inkubuoti plokštelę mikroplokštelių inkubatoriuje Hybex<sup>®</sup> 37 °C temperatūroje 5 min.
- (15) Po inkubacijos plokštelę centrifuguoti 50 g 1 min.
- (16) Į kiekvieną šulinėlį su mėginiu įpilti po 155 µl 100 % 2-propanolio, uždengti plokštelę **nauja** dengiamąja guma, ~10 kartų apversti plokštelę.
- (17) Inkubuoti plokštelę 4 °C temperatūroje 30 min.
- (18) Po inkubacijos MSA3 plokštelę centrifuguoti 4 °C temperatūroje 3000 g 20–30 min.

**NB!** Po centrifugavimo plokštelės šulinėlių su mėginiais dugne turi būti matomas melsvos spalvos precipitatas.

- (19) Išėmus plokštelę iš centrifugos, nedelsiant nuimti dengiamąją gumą ir staigiu judesiu apversti plokštelę ant beplaušių sugertukų.
- (20) Neleidžiant viršnuosėdiniui skysčiui pakliūti į kitus šulinėlius, stipriai daužyti plokštelę ant absorbuojančių beplaušių sugertukų apie 1 min.
- (21) Apverstą neuždengtą plokštelę ant mėgintuvėlių stovo džiovinti kambario (20–25 °C) temperatūroje 1 val., neapvertus plokštelės įsitikinti, kad precipitatas išdžiūvo.

#### **gDNR resuspendacija**

- (22) Į MSA3 plokštelę su precipitatu įpilti po 23 µl atšildyto RA1 reagento.

- (23) Užlydyti dengiamosios folijos lapą ant plokštelės naudojant folijos užlydytuvą. Užlydytuvą prispaudus prie folijos laikyti ~1,5–5 s.
- (24) Plokštelę inkubuoti hibridizacijos krosnyje 48 °C temperatūroje 1 val.
- (25) Po inkubacijos MSA3 plokštelę purtyti 1800 aps./min 1 min, trumpai centrifuguoti 280 g.

### **gDNR hibridizacija ant VNP lustų (tipas 12x1)**

#### **Hibridizacijos kameros paruošimas**

- (26) Surinkti hibridizacijos kamerą – į hibridizacijos kameros dugną tinkamai įdėti guminį tarpiklį.
- (27) Į hibridizacijos kameros dugne esančius rezervuarus įpilti po 400 µl PB2 reagento.
- (28) Uždengti hibridizacijos kameros dangčiu.

#### **gDNR hibridizacija ant lustų**

- (29) Mėginius MSA3 plokštelėje denatūruoti mikroplokštelių inkubatoriuje Hybex® 95 °C temperatūroje 20 min.
- (30) Išimti VNP lustus ir neatidarius pakuočių ir leisti sušilti iki kambario (20–25 °C) temperatūros.
- (31) MSA3 plokštelę perkelti iš mikroplokštelių inkubatoriaus ant stalo ir palikti atvėsti iki kambario (20–25 °C) temperatūros 30 min, trumpai centrifuguoti 280 g.
- (32) Išpakuoti VNP lustus ir sudėti į hibridizacijos kameros įdėklus, kad VNP lusto brūkšninis kodas būtų arčiau tyrėjo.
- (33) Į kiekvieną lusto laukelį įpilti po 15 µl paruošto DNR mėginio, mėginių vietą VNP luste ir VNP lusto brūkšninį kodą pažymėti protokole.
- (34) Nedelsiant VNP lustus su hibridizacijos kameros įdėklais tinkamai įdėti į hibridizacijos kamerą ir atsargiai uždengti hibridizacijos kameros dangčiu.
- (35) Dviem judesiais užspausti hibridizacijos kameros užraktus (po du priešingose pusėse, įstrižai).
- (36) Hibridizacijos kamerą padėti ant hibridizacijos krosnies padėklo, kad kameros užraktai būtų nukreipti į kairę ir dešinę krosnies sienas, hibridizacijos kameros logotipas turi būti nukreiptas į hibridizacijos krosnies viršų.
- (37) Inkubuoti įjungus 5 vienetų hibridizacijos krosnies padėklo siūbavimo greitį 48 °C temperatūroje 16–24 val.

### **XC4 reagento paruošimas**

- (38) Į XC4 butelį įpilti 330 ml 100 % etanolio. Stipriai purtyti rankos judesiu 15 s. Esant būtinybei purtyti plokštelių purtyklėje 1625 aps./min.
- (39) Palikti nusistovėti kambario (20–25 °C) temperatūroje 24 val.  
**NB!** XC4 yra daugkartinio naudojimo, galiojantis dvi savaites po pagaminimo ir tinkamas 24 lustams padengti.

### **VNP lustų plovimas (parengimas dažymui)**

- (40) Hibridizacijos kamerą išimti iš hibridizacijos krosnies, palikti atvėsti iki kambario (20–25 °C) temperatūros 25–35 min.
- (41) Įpilti po ~200 ml PB1 reagento į du stiklinius indus, VNP lustų laikiklį pamerkti į pirmąjį stiklinį indą su PB1 reagentu.
- (42) Į daugiavietį VNP lustų talpiklį įpilti ~150 ml PB1 reagento, DNR lustai turi būti visiškai apsemti.
- (43) Paruošti dengiamuosius stiklus:
  - pirmą kartą ir kas septynis naudojimo kartus plauti stiklus 0,525 % baliklio tirpalu pagal Laboratorijos paruošimo ir procedūrų aprašą;
  - kiekvieną kartą plauti 1 % detergento tirpalu (angl. *Alconox*) pagal Laboratorijos paruošimo ir procedūrų aprašą.
- (44) Įjungti termostatą ir vandens cirkuliacijos sistemą leidžiant nusistovėti 44 °C temperatūrai.
- (45) Atidaryti hibridizacijos kamerą ir išimti įdėklus su VNP lustais.
- (46) Iš įdėklo išimti po vieną VNP lustą, laikant nuožulniai brūkšniniu kodu arčiau tyrėjo, vienu lėtu įstrižu tęstiniu judesiu nulupti VNP lusto plėvelę (angl. *IntelliHyb Seal*) neliečiant lusto paviršiaus.
- (47) VNP lustą palengva įstatyti į metalinį DNR lustų laikiklį, panardintą pirmame inde su PB1 reagentu.
- (48) 46–47 etapus kartoti su likusiais VNP lustais.
- (49) DNR lustų laikiklį aktyviai, bet švelniai kilnoti kertant skysčio paviršių, ~1 min.
- (50) Perkelti laikiklį su VNP lustais ir panardinti į antrą indą su PB1 reagentu, aktyviai, bet švelniai kilnoti kertant skysčio paviršių ~1 min.

### **Pratekamųjų kamerų surinkimas**

- (51) Juodus rėmelius įdėti į skirtas vietas daugiaviečiame VNP lustų talpiklyje.

- (52) VNP lustus, nukreiptais brūkšniniais kodais į tyrėją, panardinti ir įstatyti į rėmelius, įsitikinti, kad VNP lustai visiškai panirę PB1 reagente.
- (53) Permatomas plastikines tarpines atskirti nuo baltų apsauginių plėvelių ir tinkamai uždėti ant panardintų VNP lustų.
- (54) Tinkamai uždėti talpiklio laikiklį ant talpiklio.
- (55) Suspaustu oru nupūsti dulkes nuo detergentu nuplautų ir 70 % etanoliu nuvalytų dengiamųjų stiklų.
- (56) Nedelsiant, nukreipus stiklo nuožulniąją vietą į apačią ir tyrėją, stiklus panardinti į PB1 reagento pripildytą talpiklį, įstatyti ant VNP lustų su plastikinėmis tarpinėmis.
- (57) Kiekvienos surinktos pratekamosios kameros abiejuose galuose, 5 mm atstumu nuo viršaus ir 5 mm atstumu nuo reagentų rezervuaro siauriausios dalies, užspausti gnybtukus.
- (58) Nukirpti išsikišusias plastikinių tarpinių dalis abiejuose surinktos pratekamosios kameros pusėse.
- (59) Surinktas pratekamasias kameras dėti į pratekamųjų kamerų laikiklį, įsitikinus, kad viskas paruošta dažymo procedūrai.
- (60) Hibridizacijos kamerą nedelsiant plauti ddH<sub>2</sub>O.

#### **Pasiruošimas VNP lustų dažymui**

- (61) Vartant pratekamų kamerų laikiklį, pašalinti iš sistemos burbulus.
- (62) Tinkamos 44 °C temperatūros įvertinimui į pratekamųjų kamerų laikiklį įstatyti nešiojamą termometrą OMEGA, įsitikinus esama temperatūra keliose vietose.
- (63) Atšildytus iki kambario (20–25 °C) temperatūros reagentus centrifuguoti 3000 g 1 min.
- (64) Pakartotinai supurtyti XC4 reagento butelį.

#### **Vienos bazės pratęsimo reakcija**

- (65) Į kiekvieną pratekamąją kamerą įpilti po 150 µl RA1 reagento, inkubuoti 30 s.
- (66) 65 etapą kartoti penkis kartus.
- (67) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC1 reagento, inkubuoti 10 min.
- (68) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC2 reagento, inkubuoti 10 min.

- (69) Į kiekvieną pratekamąją kamerą įpilti po 200 µl TEM reagento, inkubuoti 15 min.
- (70) Į kiekvieną pratekamąją kamerą įpilti po 450 µl FA/ EDTA reagento, inkubuoti 1 min.
- (71) 70 etapą kartoti vieną kartą ir inkubuoti 5 min.
- (72) Termostato ir vandens cirkuliacijos sistemoje nustatyti temperatūrą nurodytą ant STM reagento buteliuko, o jei nenurodyta, nustatyti 37 °C temperatūrą.
- (73) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (74) 73 etapą kartoti vieną kartą.
- (75) Laukti, iki pasiekama ant STM reagento pakuotės nurodyta temperatūra.

#### **VNP lustų dažymas**

- (76) Į kiekvieną pratekamąją kamerą įpilti po 250 µl STM reagento, inkubuoti 10 min.
- (77) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (78) 77 etapą kartoti vieną kartą, inkubuoti 5 min.
- (79) Į kiekvieną pratekamąją kamerą įpilti po 250 µl ATM reagento, inkubuoti 10 min.
- (80) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (81) 80 etapą kartoti vieną kartą, inkubuoti 5 min.
- (82) Į kiekvieną pratekamąją kamerą įpilti po 250 µl STM reagento, inkubuoti 10 min.
- (83) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (84) 83 etapą kartoti vieną kartą, inkubuoti 5 min.
- (85) Į kiekvieną pratekamąją kamerą įpilti po 250 µl ATM reagento, inkubuoti 10 min.
- (86) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (87) 86 etapą kartoti vieną kartą, inkubuoti 5 min.
- (88) Į kiekvieną pratekamąją kamerą įpilti po 250 µl STM reagento, inkubuoti 10 min.
- (89) Į kiekvieną pratekamąją kamerą įpilti po 450 µl XC3 reagento, inkubuoti 1 min.
- (90) 89 etapą kartoti vieną kartą, inkubuoti 5 min.
- (91) Nedelsiant išimti pratekamasias kameras iš pratekamųjų kamerų laikiklio, padėti horizontaliai ant darbaltalio.

### **HiScan<sup>TM</sup>SQ sistemos paruošimas**

- (92) Įjungti kompiuterį ir *HiScan<sup>TM</sup>SQ* prietaisą
- (93) Aktyvuoti programą *iScan Control Software (iScan CS)*.

**NB!** Iš anksto programa *DMAP Client* atsisiųsti kiekvienam VNP lustui (pagal kiekvieną unikalų brūkšninį kodą) unikalų dekodavimo (angl. *Decode*) dokumentą.

### **VNP lustų plovimas ir padengimas**

- (94) Tinkamai panardinti plastikinį VNP lustų plovimui po dažymo laikiklį į plastikinį plovimo indą pripildytą ~310 ml PB1 reagento.
- (95) Atsargiai išardyti pratekamašias kameras, mentele atkabinant gnybtukus, neliečiant VNP lusto mėginių zonų nuimti dengiamuosius stiklus ir plastikines tarpines.
- (96) Nedelsiant į PB1 tirpale panardintą plastikinį laikiklį įstatyti VNP lustą brūkšniniu kodu nuo tyrėjo, įsitikinti, kad lustai visiškai panardinti PB1 reagente.
- (97) Lėtai kertant PB1 reagento paviršių kilnoti plastikinį laikiklį su VNP lustais aukštyn ir žemyn 10 kartų, palikti mirkti 5 min.
- (98) Pakartotinai supurtyti XC4 reagentą.
- (99) Į plastikinį plovimo indą įpilti ~310 ml XC4 reagento, neleidžiant inde stovėti ilgiau nei 10 min.
- (100) Plastikinį VNP lustų plovimui po dažymo laikiklį tinkamai panardinti į indą su XC4 reagentu.
- (101) Lėtai kertant XC4 reagento paviršių kilnoti plastikinį laikiklį su VNP lustais aukštyn ir žemyn 10 kartų, palikti mirkti 5 min.
- (102) Lėtai ištraukti plastikinį laikiklį su VNP lustais iš XC4 reagento indo.
- (103) Laikiklį su į viršų nukreiptais VNP lustų brūkšniniais kodais padėti ant mėgintuvėlių stovo.
- (104) Atsargiai pincetu iš plastikinio laikiklio po vieną ištraukti VNP lustus ir padėti ant mėgintuvėlių stovo brūkšniniais kodais į viršų.
- (105) VNP lustus džiovinti vakuuminėje džiovyklėje 50–55 min esant 508 mm Hg (0,68 bar).
- (106) Išdžiovintų VNP lustų apatinėse pusėse esantį XC4 reagento perteklių atsargiai nuvalyti beplaušėmis servetėlėmis, suvilgytomis 70 % etanoliu.



(107) Visus VNP lustų paruošimui naudotus laboratorinius reikmenis nedelsiant užmerkti ir gerai išplauti ddH<sub>2</sub>O.

### **VNP lustų skenavimas *Illumina iScan* skeneriu *HiScan<sup>TM</sup>SQ* sistema**

(108) Programoje *iScan CS* spausti *Start*.

(109) Atsidarius prietaiso durelėms, į atitinkamas platformos vietas sudėti paruoštus VNP lustus.

(110) Paspausti platformos šonuose juodus mygtukus, skirtus VNP lustų fiksatorių padėčiai keisti, įsitikinti, kad VNP lustai įsistato į taisyklingą skenavimui padėtį.

(111) Spausti *Next*, laukti VNP lustų brūkšninių kodų nuskaitymo užsidarius prietaiso durelėms.

(112) Pasirodžius lentelei (3×4) su kiekvieno VNP lusto duomenimis (brūkšninis kodas, lusto tipas, skenavimo protokolas ir skenavimo metu sukurtų dokumentų formatas), jei reikia atlikti nustatymų korekcijas spaudžiant *Settings* mygtuką.

(113) Po lentele esančiuose laukeliuose nurodyti .dmap dokumentų katalogo vietą kompiuteryje (**Input path:** D:/DMAPs) ir skenavimo metu sukurtų dokumentų vietą kompiuteryje (**Output path:** D:/Illumina/LITGEN Scans).

(114) Spausti *Scan*.

### **Žmogaus branduolio genomo visos koduojančios dalies nukleotidų sekoskaitos protokolas**

Protokolas skirtas nustatyti žmogaus branduolio genomo visos koduojančios dalies (1,5 % genomo), egzomo, nukleotidų seką naujos kartos masiniu lygiagrečiu sekoskaitos metodu naudojant 5500 SOLiD<sup>TM</sup> genetinį analizatorių. Optimizuotą protokolą VU MF ŽMGK įdiegė dokt. Ingrida Uktverytė, dr. Laima Ambrozaitytė, dokt. Ingrida Domarkienė.

Optimizuotas protokolas parengtas remiantis:

- DNR fragmentų bibliotekos ruošimo vartotojo vadovu (angl. *Fragment Library Preparation 5500 Series SOLiD<sup>TM</sup> Systems User Guide, Publication Part Number 4460960 Rev. A*);

- Tikslinių DNR fragmentų atrinkimo vartotojo vadovu (angl. *TargetSeq™ Exome Enrichment System User Guide, Publication Part Number MAN0004396*);
- Dalelių emulsijos sudarymo sistemos vartotojo vadovu (angl. *Applied Biosystems SOLiD® EZ Bead™ Emulsifier User Guide, Publication Part Number 4441486 Rev. E*);
- Dalelių gausinimo sistemos vartotojo vadovu (angl. *Applied Biosystems SOLiD® EZ Bead™ Amplifier User Guide, Publication Part Number 4443494 Rev. E*);
- Dalelių praturtinimo sistemos vartotojo vadovu (angl. *Applied Biosystems SOLiD® EZ Bead™ Enricher User Guide, Publication Part Number 4443496 Rev. E*);
- *5500 SOLiD™* genetinio analizatoriaus vartotojo vadovu (angl. *5500 Genetic Analysis System User Guide, Publication Part Number 4456991 Rev. L*).

#### **Reagentai ir cheminės medžiagos:**

- **1 mg/ml žmogaus DNR Cot-1** (angl. *Human Cot-1 DNA®*, gamintojas *Invitrogen™ Life Technologies Corp., JAV*);
- **1× TBE buferis** (89 mM Tris, 89 mM boro rūgštis, 2 mM EDTA, pH (10× TBE) = 8,3, gamintojas *Thermo Fisher Scientific Baltics UAB, Lietuva*);
- **1× TE buferis (0,1 mM EDTA)**, 1× TE buferis su sumažinta EDTA koncentracija, (angl. *SOLiD™ Buffer Kit 1× Low TE Buffer*, gamintojas *Applied Biosystems®, Life Technologies Corp., JAV*);
- **1× TEX buferis** (angl. *SOLiD™ Buffer Kit 1× TEX Buffer*, gamintojas *Applied Biosystems®, Life Technologies Corp., JAV*);
- **3× DNR įleidimo dažas** (angl. *DNA Loading dye, 3×/ 6×*; gamintojas *UAB Thermo Fisher Scientific Baltics, JAV*);
- **10× Sekoskaitos buferis Fv2** (angl. *SOLiD® Buffer F, Double V2 10×*, gamintojas *Applied Biosystems®, Life Technologies Corp., JAV*);

- **99,96 % etanolis** (angl. *Ethanol Absolute*, gamintojas *Sigma-Aldrich Co.*, JAV);
- **Agarozės (2 %) gelis** (100 ml), plačiais šulinėliais (ruošiamas laboratorijoje, sudėtis: agarozė 2 g, 1× TBE buferis 100 ml, 1,8 μl etidžio bromido 10 mg/ml);
- **Butanolis** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Buthanol*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Blokuojantys zondai 1-16** (angl. *TargetSeq™ Blockers 1-16*, 0.2 mM, gamintojas *Life Technologies Corp.*, JAV);
- **Blokuojantys zondai P1** (angl. *TargetSeq™ Blockers P1*, 0.2 mM, gamintojas *Life Technologies Corp.*, JAV);
- **Dalelės padengtos streptavidinu** (angl. *Dynabeads® M-270 Streptavidin*, 10 mg/ml, gamintojas *Invitrogen™ Life Technologies Corp.*, JAV);
- **ddH<sub>2</sub>O** – didejonizuotas vanduo;
- **DNR molekulinis ilgio ir masės standartas 100 np**, DNR 0,5 mg/ml (angl. *GeneRuler™ 100 bp DNA Ladder*, gamintojas *UAB Thermo Fisher Scientific Baltics*, JAV);
- **Egzomo zondu mišinys** (angl. *TargetSeq™ Exome Probe Pool, 4rxn*, gamintojas *Life Technologies Corp.*, JAV);
- Emulsijos alyvos rinkinys (angl. *SOLiD™ EZ Bead™ Emulsifier Oil Kit*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Alyva** (angl. *SOLiD™ EZ Bead™ Emulsifier Oil Kit – Oil*);
  - **Stabilizatorius 1** (angl. *SOLiD™ EZ Bead™ Emulsifier Oil Kit – Stabilizer 1*);
  - **Stabilizatorius 2** (angl. *SOLiD™ EZ Bead™ Emulsifier Oil Kit – Stabilizer 2*);
- Emulsijos reagentų rinkinys 1 (angl. *SOLiD™ EZ Bead™ Emulsifier E120 Reagents Kit – Box1*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):

- **Emulsijos mišinys vandens pagrindu** (angl. *SOLiD™ EZ Bead™ Emulsifier E120 – Aqueous Master Mix*);
- **P1 dalelės** (angl. *SOLiD™ EZ Bead™ Emulsifier E120 – P1 Beads*);
- Emulsijos reagentų rinkinys 2 (angl. *SOLiD™ EZ Bead™ Emulsifier E120 Reagents Kit – Box2*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **BB tirpalas** (angl. *SOLiD™ EZ Bead™ Emulsifier E120 – BB Solution*);
  - **P1 reagentas** (angl. *SOLiD™ EZ Bead™ Emulsifier E120 – P1 Reagent*);
  - **P2 reagentas** (angl. *SOLiD™ EZ Bead™ Emulsifier E120 – P2 Reagent*);
- **gDNR**, genominės DNR mėginys (konc. 100 ng/μl, tūris 20–30 μl, DNR kiekis 2–3 μg);
- Gausinimo rinkinys (angl. *5500 SOLiD™ Fragment Library Amplification Module*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Gausinimo mišinys** (angl. *Platinum® PCR, Amplification Mix*);
- gDNR bibliotekos ruošimo fermentų rinkinys (angl. *5500 SOLiD™ Fragment Library Enzyme Module*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **5× reakcijos buferis** (angl. *5× Reaction Buffer*);
  - **Skaidymo buferis** (angl. *Shear Buffer*);
  - **10 mM dATP** – deoksiadenozino trifosfatas (angl. *dATP*);
  - **10 mM dNTP** – deksinukleozido trifosfatai (angl. *dNTP*);
  - **T4 DNR ligazė, 5U/μl** (angl. *T4 DNA Ligase*);
  - **DNR fragmentų galų bukinimo fermentas 1** (angl. *End Polishing E1*);
  - **DNR fragmentų galų bukinimo fermentas 2** (angl. *End Polishing E2*);
  - **Adenoziną prijungiantis fermentas I** (angl. *A-tailing Enzyme I*);
- gDNR bibliotekos unikalios DNR sekos adapteriai (angl. *5500 SOLiD™ Fragment Library Barcode Adaptors 1-16*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):

- **50 µM unikalios DNR sekos adapteriai** (angl. *Barcode-T-001-016*);
- **50 µM bendras DNR sekos adapteris** (angl. *P1-T-Adapter*);
- **50 µM pradmuo 1** – gDNR bibliotekos PGR pradmuo 1 (angl. *Library PCR Primer 1*);
- **50 µM pradmuo 2** – gDNR bibliotekos PGR pradmuo 2 (angl. *Library PCR Primer 2*);
- Hibridizacijos ir plovimo rinkinys (angl. *TargetSeq™ Hybridization and Wash Kit*, gamintojas *Life Technologies Corp.*, JAV):
  - **2,5× magnetinių dalelių plovimo tirpalas** (angl. *TargetSeq™ Magnetic bead Wash 2.5×*);
  - **10× plovimo tirpalas A** (angl. *TargetSeq™ Wash Solution A 10×*);
  - **10× plovimo tirpalas B** (angl. *TargetSeq™ Wash Solution B 10×*);
  - **10× plovimo tirpalas C** (angl. *TargetSeq™ Wash Solution C 10×*);
  - **Hibridizaciją skatinantis tirpalas B** (angl. *TargetSeq™ Hyb Enhancer B*);
  - **Hibridizacijos tirpalas A** (angl. *TargetSeq™ Hyb Solution A*);
  - **10× surišimo tirpalas** (angl. *TargetSeq™ Stringent Wash Solution 10×*);
- **Magnetinės dalelės** (angl. *Agencourt® AMPure® XP*, gamintojas *Beckman Coulter, Inc.*, JAV);
- Praturtinimo buferių rinkinys 1 (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Box1*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Plovimo buferis** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Rinse Buffer*);
  - **Hibridizacijos buferis** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Hyb Buffer*);
  - **CS tirpalas 2** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – CS Solution 2*);
  - **CS tirpalas 3** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – CS Solution 3*);

- Praturtinimo buferių rinkinys 2 (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Box2*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Sugavimo buferis** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Capture Buffer*);
  - **Išlaisvinimo buferis** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – Release Buffer*);
  - **CS tirpalas 1** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – CS Solution 1*);
  - **CS tirpalo 1 priedas** (angl. *SOLiD™ EZ Bead™ Enricher Buffer Kit – CS Solution 1 Additive*);
- Praturtinimo reagentų rinkinys 1 (angl. *SOLiD™ EZ Bead™ Enricher E120 – Box1*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Sugavimo dalelės** (angl. *SOLiD™ EZ Bead™ Enricher E120 – Capture Beads*);
- Praturtinimo reagentų rinkinys 2 (angl. *SOLiD™ EZ Bead™ Enricher E120 – Box2*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Sugavimo zondai** (angl. *SOLiD™ EZ Bead™ Enricher E120 – Capture Oligos*);
- Prieš suleidimą būtinų reagentų rinkinys (angl. *SOLiD® Pre Deposition Plus Kit Box1*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Dalelių jungiklio tirpalas** (angl. *SOLiD™ Bead Linker*);
  - **10× transferazės buferis** (angl. *SOLiD™ 10× Transferase Buffer*);
  - **10× kobalto chloridas** (angl. *SOLiD™ 10× Cobalt Chloride*);
  - **Terminalinė transferazė, 20U/μl** (angl. *SOLiD® Terminal Transferase*);
- DNR standartas, RL-PGR DNR standartas (angl. *SOLiD™ Library qPCR Standard*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **5 nM DNR standartas** (angl. *SOLiD™ Library qPCR Standard, 5nM*);

- RL-PGR rinkinys, realaus laiko PGR rinkinys (angl. *SOLiD™ Library TaqMan® qPCR Module*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **ROX dažas** (angl. *ROX Reference Dye, 500 µl*);
  - **RL-PGR mišinys**, realaus laiko PGR mišinys (angl. *SOLiD™ Library qPCR Mix, 250 rxn*);
  - **20× DNR pradmenys ir zondai** (angl. *Ac00010015\_a1 TaqMan Assay, 20×*);
- **Sekoskaitos buferis A** (angl. *SOLiD™ Buffer A Kit, Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Sekoskaitos buferis B** (angl. *SOLiD™ Buffer B Kit, Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Sekoskaitos buferis C** (angl. *SOLiD™ Buffer C Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Sekoskaitos buferis D** (angl. *SOLiD™ Buffer D, Double (reset)*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Sekoskaitos buferis E** (angl. *SOLiD™ Buffer E Kit, Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Sekoskaitos buferis G** (angl. *SOLiD™ Buffer G, Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- Sekoskaitos DNR pradmenų rinkinys 1 (angl. *SOLiD™ FWD1 SP KIT*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV ):
  - **Sekoskaitos DNR pradmenų rinkinys** (angl. *3 SOLiD™ FWD1 Seq. Primers*);
  - **Sekoskaitos unikalių DNR sekų pradmenų rinkinys** (angl. *7 SOLiD™ BA Primers*);
  - **Sekoskaitos kokybės kontrolės rinkinys** (angl. *4 SOLiD™ Sequencing QC*);
  - **Mėgintuvėliai sekoskaitos reagentų maišymui** (angl. *10a/b SOLiD™ Mixing Tubes*);

- Sekoskaitos DNR pradmenų rinkinys 2 (angl. *SOLiD™ FWD2 SP KIT*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV ):
  - **Sekoskaitos pradmenų rinkinys** (angl. 2 *SOLiD™ FWD2 Seq. Primers*);
  - **Sekoskaitos kokybės kontrolės rinkinys** (angl. 4 *SOLiD™ Sequencing QC*);
- Sekoskaitos reagentų rinkinys (angl. *SOLiD® FWD SR S75 Kit, Single*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV ):
  - **Sekoskaitos reagentų rinkinys** (angl. 6a/b *SOLiD® FWD Seq S75 Reagents, Single*);
- **Suleidimo buferis** (angl. *SOLiD™ FlowChip Deposition Buffer 1*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Tiesioginis sekoskaitos buferis** (angl. *SOLiD™ FWD Buffer, Double*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV).

#### **Įranga ir laboratoriniai reikmenys:**

- **96 šulinėlių optinė plokštelė** (angl. *ABI PRISM® 96-Well Optical Reaction Plate with Barcode*, gamintojas *Applied Biosystems®* *Life Technologies Corp.*, JAV);
- **96 šulinėlių optinės plokštelės optinė dengimo plėvelė** (angl. *MicroAmp™ Optical Adhesive Film*, gamintojas *Applied Biosystems®* *Life Technologies Corp.*, JAV);
- **Centrifuga** (15 ml);
- **Centrifuga** (96 šulinėlių plokštelėms);
- **Dalelių emulsijos sudarymo sistema** (angl. *SOLiD™ EZ Bead™ Emulsifier, Bead Emulsification Station*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Dalelių gausinimo sistema** (angl. *SOLiD™ EZ Bead™ Amplifier, Bead Thermocycling Station*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);



- **Dalelių praturtinimo sistema** (angl. *SOLiD™ EZ Bead™ Enricher, Bead Enrichment Station*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Didelių tūrių skysčių automatinė pipetė** (angl. *Eppendorf EasyPet 4420*, gamintojas *Eppendorf AG*, Vokietija);
- Emulsijos ruošimo reikmenų rinkinys (angl. *SOLiD™ EZ Bead™ Emulsifier E-Large Accessories Pack*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Emulsijos indas** (angl. *SOLiD™ EZ Bead™ Emulsifier E-Large Jar Pack*);
  - **Emulsijos vandeninės fazės indas** (angl. *SOLiD™ EZ Bead™ Emulsifier E-Large AQ Tube Pack*);
  - **Emulsijos ruošimo vamzdelis**;
- **Emulsijos alyvos ruošimo rinkinys** (angl. *SOLiD™ EZ Bead™ Emulsifier Oil Accessories Kit*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Gausinimo maišelis** (angl. *SOLiD™ EZ Bead™ Amplifier E-Large Accessories Kit*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Gausinimo maišelio laikiklis**;
- **Gelių analizavimo sistema** (angl. *Herolab E.A.S.Y 442 K*, gamintojas *Herolab GmbH Laborgeraete*, Vokietija);
- **Genetinis analizatorius** (angl. *5500 SOLiD™ Sequencer*, gamintojas *Hitachi High-Technologies Corp.*, Japonija);
- **Horizontalios elektroforezės sistema su maitinimo šaltiniu** (angl. *Midicell® EC350 Electrophoretic Gel System*, gamintojas *E-C Apparatus Corp.*, JAV);
- **Latekso ir nitrilo pirštinės**;
- **Liofilizatorius** (angl. *Eppendorf Concentrator 5301*, gamintojas *Eppendorf AG*, Vokietija);

- **Magnetinis stovas 2 ml tūrio mėgintuvėliams** (angl. *DynaMag™-2*, gamintojas *Life Technologies Corp.*, JAV);
- **Mėgintuvėliai sekoskaitos reagentų maišymui** (angl. *SOLiD™ Mixing Strip Tube*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- **Mikrocentrifuga** (0,2 ml, 0,5 ml, 1,5 ml);
- **Mikropipetės**, vienkanalės (0,2–2,5 µl, 2–20 µl, 10–100 µl, 20–200 µl, 100–1000 µl);
- **Parafino juostelės** (angl. *Parafilm® M*, gamintojas *Bemis Company, Inc.*, JAV);
- **Pildymo sistema** (angl. *SOLiD™ EZ Bead™ Fill Station*, gamintojas *UniSlide® Assemblies, Velmex, Inc.*, JAV);
- **Piltuvėlis** (angl. *SOLiD™ EZ Bead™ Amplifier Funnel Pack*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV);
- Praturtinimo reikmenų rinkinys (angl. *SOLiD™ EZ Bead™ Enricher Accessories Kit*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV):
  - **Praturtinimo kolonėlė** (angl. *SOLiD™ EZ Bead™ Enricher Column, 2 Blue-green*);
  - **Emulsijos indas** (angl. *SOLiD™ EZ Bead™ Enricher Input Tube, 8 Orange*);
  - **Nepraturtintų dalelių indas** (angl. *SOLiD™ EZ Bead™ Enricher Unenriched Bead Bottle, 4 Black*);
  - **Išlaisvinimo mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher Release Tube, 11 Purple*);
  - **Kokybės kontrolės mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher QC Output Tube, 5 Yellow*);
  - **Sugavimo zondų mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher Capture Oligo Tube, F Lime-green*);

- **Dalelių išėigos mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher Bead Output Tube, 3 Red*);
- **Hibridizacijos mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher Hyb Tube, 10 Tan*);
- **Filtrai** (angl. *Acrodisc® Syringe Filter 0,45 µm Supor® Membrane Low Protein Binding Non-Pyrogenic, sterile, 7 Blue*, gamintojas *Pall Corp., JAV*);
- **Sugavimo dalelių mėgintuvėlis** (angl. *SOLiD™ EZ Bead™ Enricher Capture Bead Tube, 9 Pink*);
- **Purtyklė** (angl. *Vortex-Genie 2, Model G560*, gamintojas *Scientific Industries Inc., JAV*);
- **RL-PGR sistema**, realaus laiko PGR sistema (angl. *7900HT Fast Real-Time PCR System*, gamintojas *Applied Biosystems® Life Technologies Corp., JAV*);
- **Serologinė 25 ml ir (ar) 10 ml tūrio pipetė** (angl. *SpaceSaver™ 25 ml Serological Pippet*, gamintojas *BD, JAV*);
- **Skaidymo ir 1,5 ml tūrio mikromėgintuvėlių laikikliai** (angl. *MicroTUBE holder 6 mm TUBE, MicroTUBE holder 1,5 ml Microfuge Tube*, gamintojas *Covaris, Inc, JAV*);
- **Skaidymo mikromėgintuvėlis** (angl. *microTUBE AFA Fiber Pre-Slit Snap Cap 6\*16 mm*, gamintojas *Covaris, Inc, JAV*);
- **Skaidymo ultragarsu sistema su vėsinimu** (angl. *Covaris™ S220*, gamintojas *Covaris, Inc, JAV*; *VWR™ MM7*, gamintojas *VWR International, JAV*);
- **Sterilūs 1 ml tūrio švirkštai** (angl. *BD Plastipak™ 1 ml Syringe*, gamintojas *BD, JAV*);
- **Sterilūs 3 ml tūrio švirkštai** (angl. *BD 3 ml Syringe*, gamintojas *BD, JAV*);
- **Sterilūs antgaliai** (10 µl, 20 µl, 100 µl, 200 µl, 1000 µl);
- **Sterilūs mėgintuvėliai** (0,2 ml, 0,5 ml, 1,5 ml, 50 ml);

- **Sumažintos adhezijos mėgintuvėliai** (angl. *DNA LoBind Tubes, 1,5 ml*, gamintojas *Eppendorf AG*, Vokietija);
- **Svarstyklės** (angl. *Kern Pls 310-3F*, gamintojas *Kern&Sohn GmbH*, Vokietija);
- **Termocikleris** (angl. *Eppendorf Mastercycler*, gamintojas *Eppendorf AG*, Vokietija);
- **Termopurtyklė 1,5 ml tūrio mėgintuvėliams** (angl. *Eppendorf Thermomixer comfort*, gamintojas *Eppendorf AG*, Vokietija);
- **Termostatas** (angl. *Amersham Hybridization Oven / Shaker*, gamintojas *GE Healthcare, General Electric Corp.*, Jungtinė Karalystė);
- **Tėkmės lustas** (angl. *SOLiD™ 6-Lane FlowChip*, gamintojas *Applied Biosystems®*, *Life Technologies Corp.*, JAV).

**Darbo eiga:**

**gDNR skaidymas ultragarsu**

- (1) Paruošti 12 gDNR mėginių skaidymo mišinius 1,5 ml tūrio sumažintos adhezijos mėgintuvėliuose (26 lentelė).

26 lentelė. gDNR mėginių skaidymo mišinys

Reagentai	Tūris (µl)/ mėginiui
gDNR (100 ng/µl; 2–3 µg)	20–30
1× TE buferis (0,1 mM EDTA), (angl. <i>1× Low TE buffer</i> )	88,8–98,8
Skaidymo buferis (angl. <i>Shear buffer</i> )	1,2
<i>Iš viso</i>	120,0

- (2) Paruoštus mišinius supurtyti ir centrifuguoti 5 s.
- (3) Mišinius perkelti į skaidymo mikromėgintuvėlius.
- (4) Skaidymo mikromėgintuvėlį įtvirtinti į skaidymo mikromėgintuvėlių laikiklį.
- (5) Laikiklį su skaidymo mikromėgintuvėliu perkelti į iš anksto įjungtą ir paruoštą skaidymo ultragarsu sistemą, įsitikinus, kad vėsinimo blokas pasiekęs nustatytą 3 °C temperatūrą.
- (6) Programoje *SonoLab™ v7.1.0.5* pasirinkti iš anksto paruoštą programą (27 lentelė).

27 lentelė. Skaidymo programa

Galia (W) (angl. <i>Peak Power</i> )	Impulsų užpildos faktorius (%) (angl. <i>Duty Factor</i> )	Impulsų vora/ciklui (angl. <i>Cycle/Burst</i> )	Ciklo trukmė (s)	Ciklų skaičius
175	10	100	60	6

- (7) Laikiklį su skaidymo mikromėgintuvėliu pašalinti iš skaidymo ultragarsu sistemos.
- (8) Skaidymo mikromėgintuvėlį pašalinti iš laikiklio.
- (9) Mišinį perkelti į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (10) 4–9 etapus kartoti visiems gDNR mėginiams.

#### **gDNR įvertinimas**

- (11) Paruoštą 2 % agarozės gelį įleisti į horizontalios elektroforezės vonelę, pripildytą 1× TBE buferio.
- (12) 2 μl skaidytos gDNR tirpalo sumaišyti su 1 μl 3× įleidimo dažų.
- (13) 3 μl mišinio įleisti į suformuotus šulinėlius.
- (14) Iš abiejų pusių skaidytos gDNR ir 3× įleidimo dažo mišinio į šulinėlius įleisti po 1 μl molekulinio DNR ilgio ir masės standarto 100 np.
- (15) Elektroforezę vykdyti 25 min, įtampa 100 V.
- (16) gDNR mėginių kokybę įvertinti naudojant gelių analizavimo sistemą esant 302 nm bangos ilgiui (ultravioletas, UV).

#### **Skaidytos gDNR galų bukinimas**

- (17) Paruošti skaidytos gDNR galų bukinimo mišinį 1,5 ml tūrio sumažintos adhezijos mėgintuvėlyje (28 lentelė).

28 lentelė. Skaidytos gDNR galų bukinimo mišinys

Reagentai	Tūris (μl)/ reakcijai
5× reakcijos buferis (angl. <i>5× Reaction buffer</i> )	40
10 mM dNTP (angl. <i>dNTP</i> )	8,0
DNR fragmentų galų bukinimo fermentas 1 (angl. <i>End Polishing E1</i> )	8,0
DNR fragmentų galų bukinimo fermentas 2 (angl. <i>End Polishing E2</i> )	10,0
ddH <sub>2</sub> O	14,0
<b>Iš viso</b>	<b>80,0</b>

- (18) Paruoštą mišinį išpilstyti po 80  $\mu\text{l}$  į kiekvieną 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį su 120  $\mu\text{l}$  mėginio.
- (19) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (20) Paruoštus mišinius inkubuoti kambario (20–25 °C) temperatūroje 30 min.

**Norimo ilgio fragmentų atrinkimas ir mišinio valymas magnetinėmis dalelėmis**

- (21) Iš naujo paruošti 30 ml 70 % etanolio.
- (22) Magnetines daleles supurtyti ir palikti sušilti tamsoje iki kambario (20–25 °C) temperatūros.
- (23) Į kiekvieną mėgintuvėlį įpilti 100  $\mu\text{l}$  (50 % mėginio tūrio) gerai supurtytų magnetinių dalelių.
- (24) Paruoštą mišinį purtyti 10 s purtykle, centrifuguoti 5 s ir palikti kambario (20–25 °C) temperatūroje 5 min.
- (25) Mėgintuvėlius su mišiniu perkelti į magnetinį stovą, laukiama mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių ir perkelti į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (26) Į kiekvieną mėgintuvėlį įpilti 60  $\mu\text{l}$  (30 % mėginio tūrio) gerai supurtytų magnetinių dalelių.
- (27) Paruoštą mišinį purtyti 10 s purtykle, centrifuguoti 5 s ir palikti kambario (20–25 °C) temperatūroje 5 min.
- (28) Mėgintuvėlius su mišiniu perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių ir pašalinti.
- (29) Į kiekvieną mėgintuvėlį įpilti po 200  $\mu\text{l}$  naujai paruošto 70 % etanolio.
- (30) Mėgintuvėlius su magnetinių dalelių ir etanolio mišiniu supurtyti, centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, neliečiant magnetinio stovo pusėje susirinkusių dalelių, etanolį išsiurbti ir pašalinti.
- (31) 29–30 etapus kartoti dar du kartus.

- (32) Mėgintuvėlius su magnetinėmis dalelėmis centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, neliečiant magnetinio stovo pusėje susirinkusių dalelių pašalinti likusį etanolį.
- (33) Mėgintuvėlius su magnetinėmis dalelėmis atidaryti ir palikti džiūti 28 °C temperatūroje 5–10 min.
- (34) Į mėgintuvėlį su išdžiovintomis magnetinėmis dalelėmis įpilti 36 µl 1× TE buferio (0,1 mM EDTA) arba ddH<sub>2</sub>O.
- (35) Mėgintuvėlius su mišiniu purtyti 10 s, centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, neliečiant magnetinio stovo pusėje susirinkusių dalelių tirpalą išsiurbti ir perkelti į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (36) Atlikti gDNR įvertinimą (11–16 etapai) 2 % agarozės gelyje.

#### **Deoksiadenozino monofosfato (dAMP) prijungimas prie gDNR fragmentų galų**

- (37) Paruošti mišinį dAMP prijungimui prie gDNR fragmentų galų 1,5 ml tūrio sumažintos adhezijos mėgintuvėlyje (29 lentelė).

29 lentelė. dAMP prijungimo prie gDNR fragmentų galų mišinys

<b>Reagentai</b>	<b>Tūris (µl)/ reakcijai</b>
5× reakcijos buferis (angl. 5× <i>Reaction buffer</i> )	10
10 mM dATP (angl. <i>dATP</i> )	1,0
dAMP prijungiantis fermentas I (angl. <i>A-tailing Enzyme I</i> )	5,0
<b><i>Iš viso</i></b>	16,0

- (38) Paruoštą mišinį išpilstyti po 16 µl į kiekvieną 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį su 34 µl mėginio.
- (39) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (40) Paruoštus mišinius inkubuoti termostate 68 °C temperatūroje 30 min, po inkubacijos mišinius atvėsinti iki kambario (20–25 °C) temperatūros.

#### **Unikalių ir bendrų adapterių prijungimas**

- (41) Paruošti mišinį, išskyrus unikalius adapterius, unikalių ir bendrų adapterių prijungimui prie gDNR fragmentų galų 1,5 ml tūrio sumažintos adhezijos mėgintuvėlyje (30 lentelė).

30 lentelė. Unikalių ir bendrų adapterių prijungimo prie gDNR fragmentų galų mišinys

Reagentai	Tūris (μl)/ reakcijai
5× reakcijos buferis (angl. <i>5× Reaction buffer</i> )	10
50 μM bendras adapteris (angl. <i>P1-T-Adapter</i> )	1,1/1,6
50 μM unikalios DNR sekos adapteriai (angl. <i>Barcode-T-001-016</i> )	<b>1,1/1,6</b>
T4 DNR ligazė, 5U/μl (angl. <i>T4 DNA Ligase</i> )	6,5
10 mM dNTP (angl. <i>dNTP</i> )	1,2
ddH <sub>2</sub> O	5,0
<i>Iš viso</i>	13,9/13,4

- (42) Bendrų ir unikalių adapterių tūris priklauso nuo pradinio gDNR kiekio, esant 2 μg įpilti po 1,1 μl unikalių ir bendrų adapterių vienam mėginiui, esant 3 μg įpilti po 1,6 μl unikalių ir bendrų adapterių vienam mėginiui.
- (43) Paruoštą mišinį priklausomai nuo pradinio gDNR kiekio (2 μg arba 3 μg) išpilstyti po 13,9 μl arba 13,4 μl į kiekvieną 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį su 50 μl mėginio.
- (44) Į kiekvieną mėgintuvėlį su skirtingu mėginiu priklausomai nuo pradinio gDNR kiekio (2 μg arba 3 μg) įpilti po 1,1 μl arba 1,6 μl unikalios DNR sekos adapterių.
- (45) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (46) Paruoštus mišinius inkubuoti termostate 20 °C temperatūroje 30 min, vėliau 72 °C temperatūroje 20 min, po inkubacijos mišinius atvėsinti iki 4 °C temperatūros.
- (47) Po inkubavimo atlikti mišinio valymą magnetinėmis dalelėmis (26–35 etapai), naudojant 26 etape 39 μl gerai supurtytų magnetinių dalelių ir 34 etape 22 μl 1× TE buferio (0,1 mM EDTA) arba ddH<sub>2</sub>O.

### **gDNR gausinimas**

- (48) Paruošti mišinį gDNR gausinimui 1,5 ml tūrio sumažintos adhezijos mėgintuvėlyje (31 lentelė).



31 lentelė. gDNR gausinimo mišinys

Reagentai	Tūris (μl)/ reakcijai
Gausinimo mišinys (angl. <i>Platinum<sup>®</sup> PCR, Amplification Mix</i> )	100
50 μM pradmuo 1 (angl. <i>Library PCR Primer 1</i> )	2,5
50 μM pradmuo 2 (angl. <i>Library PCR Primer 2</i> )	2,5
<b><i>Iš viso</i></b>	105

- (49) Paruoštą mišinį išpilstyti po 105 μl į kiekvieną 0,2 ml tūrio mėgintuvėlį su 20 μl mėginio.
- (50) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (51) Paruoštus mišinius inkubuoti termocikleryje naudojant iš anksto paruoštą programą (32 lentelė).

32 lentelė. gDNR gausinimo sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	95	05:00	1
Denatūracija	95	00:15	6
Pradmenų prijungimas	62	00:15	
Sintezė	70	01:00	
Galutinė sintezė	70	05:00	1

- (52) Po gDNR gausinimo atlikti mišinio valymą magnetinėmis dalelėmis (26–35 etapai), naudojant 26 etape 187,5 μl gerai supurtytų magnetinių dalelių ir 34 etape 30 μl 1× TE buferio (0,1 mM EDTA) arba ddH<sub>2</sub>O.
- (53) Atlikti gDNR įvertinimą (11–16 etapai) 2 % agarozės gelyje.

#### **gDNR koncentracijos nustatymas kiekybinės RL-PGR metodu**

- (54) Paruošti kiekvieno gDNR mėginio 20 000 kartų skiedimą. Ruošti 20 kartų skiedimą (įpilti 2 μl gDNR į naują mėgintuvėlį su 18 μl ddH<sub>2</sub>O), toliau ruošti 1 000 kartų skiedimą (įpilti 1 μl 20 kartų skiestos gDNR į naują mėgintuvėlį su 999 μl ddH<sub>2</sub>O).
- (55) Paruošti RL-PGR DNR standarto skiedimus (33 lentelė).

33 lentelė. Kiekybinės RL-PGR DNR standartų skiedimai

DNR standarto numeris	Skiedimas (iš 5 nM konc. DNR)	Komponentai	Galutinė koncentracija
1	1:49	1 µl konc. DNR + 49 µl ddH <sub>2</sub> O	100 pM
2	1:499	5 µl DNR standarto 1 + 45 µl ddH <sub>2</sub> O	10 pM
3	1:4 999	5 µl DNR standarto 2 + 45 µl ddH <sub>2</sub> O	1 pM
4	1:49 999	5 µl DNR standarto 3 + 45 µl ddH <sub>2</sub> O	0,1 pM
5	1:499 999	5 µl DNR standarto 4 + 45 µl ddH <sub>2</sub> O	0,01 pM

(56) Paruošti kiekybinės RL-PGR mišinį, atsižvelgiant į visų mėginių, DNR standartų, neigiamos kontrolės trijų kartotinių skaičių 1,5 ml tūrio mėgintuvėliuose (34 lentelė).

34 lentelė. Kiekybinės RL-PGR mišinys

Reagentai	Tūris (µl)/ reakcijai
RL-PGR mišinys (angl. <i>SOLiD™ Library qPCR Mix</i> )	10,0
20× DNR pradmenys ir zondai (angl. <i>Ac00010015_a1 TaqMan Assay</i> )	1,0
ROX dažas (angl. <i>ROX Reference Dye</i> )	0,4
ddH <sub>2</sub> O	3,6
<b>Iš viso</b>	15,0

- (57) Paruoštą mišinį išpilstyti po 15 µl į kiekvieną 96 šulinėlių optinės plokštelės šulinėlį, kur į kiekvieną bus įpilta trys kartotiniai 5 µl mėginio, neigiamos kontrolės, DNR standarto.
- (58) Paruoštą 96 šulinėlių optinę plokštelę kruopščiai uždengti optine dengimo plėvele ir centrifuguoti 1500 aps./min 1 min.
- (59) Įjungti RL-PGR sistemą, aktyvuoti programą *SDS 2.3*.
- (60) Pasirinkti *File > New Plate Wizard*;
- (61) Etape *Select Assay Type* skiltyje *Quantification* pasirinkti *Standard Curve (AQ)*, spausti *Next*.
- (62) Etape *Enter Plate Information* skiltyje *Plate Type* pasirinkti *96 well Clear Plate*, skiltyje *Create Document From* pasirinkti *Blank Document*, spausti *Next*.
- (63) Etape *Enter Samples and Detectors to Use in Plate* skiltyje *Samples to Use in Plate* suvesti visus skirtingus mėginius, skiltyje *For Markers to*

*Use in Plate* spausti *Existing Plate* pasirinkti anksčiau sukurtą dokumentą ir spausti *Open*, spausti *Next*;

- (64) Etape *Assign Samples and Markers to Wells* užpildyti virtualią plokštelės reprezentaciją, susiejant suvestus gDNR mėginius, DNR standartus, neigiamą kontrolę su detektoriumi ir jų padėtimi plokštelėje. Skiltyje *Plate Layout* pažymėti visus laukelius su trimis kartotiniais kiekvieno skirtingo gDNR mėginio, kiekvieno skirtingo DNR standarto, neigiamos kontrolės, skiltyje *Detectors in Selected Wells* pažymėti detektorių FGFR3 kartu su reporteriniu dažu (FAM), detektoriaus *Task* stulpelyje pasirinkti atitinkamai *Unknown* (gDNR mėginiui), *Standard* (DNR standartui, įvedant koncentraciją), *NTC* (neigiamai kontrolei); skiltyje *Samples in Selected Wells* pažymėti konkretų mėginį, šį žingsnį kartoti kiekvienam skirtingam mėginiui, spausti *Finish*.
- (65) Spausti *Save*, įvesti pavadinimą ir išsaugoti sukurtą plokštelės protokolą.
- (66) Pasirinkti paskyrą *Instrument > Thermal Cyclers > Standard > Thermal Profile* (35 lentelė).

35 lentelė. Kiekybinės RL-PGR gausinimo sąlygos

<b>Etapas</b>	<b>Temperatūra (°C)</b>	<b>Trukmė (min)</b>	<b>Ciklų skaičius</b>
Pirminė denatūracija	95	02:00	1
Denatūracija	95	00:15	40
Pradmenų, zondų prijungimas/sintezė	60	01:00	

- (67) Laukelyje *Volume (µl)* įrašyti 20.
- (68) Pasirinkti paskyrą *Instrument > Real Time > Connect to Instrument*, toliau spausti *Open/Close* ir laukti optinės plokštelės platformos.
- (69) 96 šulinėlių optinę plokštelę, uždengus apsauginiu plokštelės dangalu, įstatyti į RL-PGR sistemos platformą.
- (70) Spausti *Open/Close* ir laukti, kol optinės plokštelės platforma grįš į pradinę padėtį RL-PGR sistemoje.
- (71) Spausti *Start Run* ir laukti, kol bus vykdoma kiekybinė RL-PGR, kurios rezultatus galima stebėti realiu laiku.

- (72) Pasibaigus reakcijai spausti *Analyze* ir įvertinti gautus rezultatus, kuriuos išsaugoti pasirenkant *File > Save*.
- (73) Gautus rezultatus iškelti .txt dokumento formatu, pasirenkant *File > Export Table*.
- (74) Remiantis RL-PGR metodu nustatytais DNR mėginių koncentracijomis, apskaičiuoti kiekvieno gDNR mėginio tūrį, reikalingą sujungti po keturis gDNR mėginius.
- (75) 
$$V (\mu l) = 125 \text{ ng} / \frac{(konc.(pM) \times 20\,000) \times (660 \times \text{fragmento ilgis (bp)})}{1 \times 10^9}$$

### **Hibridizacija**

- (76) gDNR mėginius sujungti po keturis, įpilant kiekvieno mėginio protokolo 75 etape apskaičiuotą tūrį, naujuose trijuose 1,5 ml tūrio sumažintos adhezijos mėgintuvėliuose pagal unikalių DNR sekų adapterių spalvinį žymėjimą.
- (77) Žmogaus DNR Cot-1 (1 mg/ml) skiesti iki 0,1 mg/ml, paruošiant bendrą 20 µl tūrį.
- (78) Į tris mėgintuvėlius su sujungtais gDNR mėginiais įpilti 5 µl žmogaus DNR Cot-1 (0,1 mg/ml), po 5 µl blokuojančių zondų 1–16 ir blokuojančių zondų P1.
- (79) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (80) Mišinius mėgintuvėliuose džiovinti liofilizatoriuje 60 °C temperatūroje apie 30–40 min.
- (81) Į mėgintuvėlius su išdžiovintu mišiniu įpilti 7,5 µl hibridizacijos tirpalo A ir 3 µl hibridizaciją skatinančio tirpalo B.
- (82) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (83) Inkubuoti termostate 95 °C temperatūroje 10 min.
- (84) Mėgintuvėlius su mišiniu centrifuguoti 5 s didžiausiu greičiu.
- (85) Į mėgintuvėlius su mišiniu įpilti po 4,5 µl egzomo zondų mišinio.
- (86) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.

- (87) Mėgintuvėlius kruopščiai uždaryti, apvynioti parafino juostelėmis ir inkubuoti termocikleryje 47 °C temperatūroje, nustačius termociklerio dangtelio 57 °C temperatūrą, apie 72 val.

### Mėginių valymas po hibridizacijos

- (88) Valymo tirpalus skiesti iki 1× tirpalų (36 lentelė).

36 lentelė. Valymo tirpalai

Reagentai	Konc. tirpalo tūris (μl)	ddH <sub>2</sub> O tūris (μl)	1× tirpalo tūris (μl)
2,5× magnetinių dalelių plovimo tirpalas (angl. <i>Magnetic bead Wash 2.5×</i> )	640	960	1600
10× plovimo tirpalas A (angl. <i>Wash Solution A 10×</i> )	100	900	1000
10× plovimo tirpalas B (angl. <i>Wash Solution B 10×</i> )	70	630	700
10× plovimo tirpalas C (angl. <i>Wash Solution C 10×</i> )	70	630	700
10× surišimo tirpalas (angl. <i>Stringent Wash Solution 10×</i> )	130	1170	1300

- (89) 300 μl 1× plovimo tirpalą A ir 1600 μl 1× surišimo tirpalą inkubuoti 47 °C temperatūroje 2 val. iki naudojimo.
- (90) Daleles padengtas streptavidinu laikyti kambario (20–25 °C) temperatūroje 30 min iki naudojimo.
- (91) Į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį įpilti 300 μl dalelių padengtų streptavidinu (kiekvienam sujungtam gDNR mišiniui po 100 μl).
- (92) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (93) Įpilti 600 μl 1× magnetinių dalelių plovimo tirpalo.
- (94) Mišinį mėgintuvėlyje purtyti 10 s, centrifuguoti 5 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (95) 93–94 etapus kartoti vieną kartą.
- (96) Įpilti 300 μl 1× magnetinių dalelių plovimo tirpalo, mėgintuvėlį su mišiniu supurtyti ir centrifuguoti 5 s.

- (97) Įpilti po 100 µl dalelių, padengtų streptavidinu, ir 1× magnetinių dalelių plovimo tirpalo į du naujus 1,5 ml tūrio sumažintos adhezijos mėgintuvėlius.
- (98) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (99) Į kiekvieną atskirą mėgintuvėlį su dalelėmis, padengtomis streptavidinu, perkelti po 15 µl hibridizacijos mišinio (87 etapas).
- (100) Mišinį mėgintuvėliuose supurtyti ir centrifuguoti 5 s.
- (101) Mėgintuvėlius su mišiniu inkubuoti 47 °C temperatūroje 45 min, kas 15 min supurtyti 3 s ir centrifuguoti 5 s.
- (102) Po inkubavimo į kiekvieną mėgintuvėlį įpilti po 100 µl 47 °C temperatūros 1× plovimo tirpalo A.
- (103) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (104) Mėgintuvėlius perkelti iš magnetinio stovo ir į kiekvieną mėgintuvėlį įpilti po 200 µl 47 °C temperatūros 1× surišimo tirpalo. Atlikti greitai, kad mišinys neatvėstų.
- (105) Mėgintuvėlius su mišiniu supurtyti, centrifuguoti 2 s ir inkubuoti termostate 47 °C temperatūroje 5 min.
- (106) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (107) 104–106 etapus kartoti vieną kartą.
- (108) Į kiekvieną mėgintuvėlį įpilti po 200 µl 1× plovimo tirpalo A.
- (109) Mėgintuvėlius purtyti 2 min, centrifuguoti 5 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (110) Į kiekvieną mėgintuvėlį įpilti po 200 µl 1× plovimo tirpalo B.

- (111) Mėgintuvėlius purtyti 1 min, centrifuguoti 5 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (112) Į kiekvieną mėgintuvėlį įpilti po 200 µl 1× plovimo tirpalo C.
- (113) Mėgintuvėlius purtyti 30 s, centrifuguoti 5 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (114) Į kiekvieną mėgintuvėlį įpilti po 15 µl ddH<sub>2</sub>O.

**gDNR, prijungtos prie streptavidinu padengtų dalelių, gausinimas**

- (115) Paruošti gDNR prijungtos prie streptavidinu padengtų dalelių gausinimo mišinį 1,5 ml tūrio sumažintos adhezijos mėgintuvėlyje (37 lentelė).

37 lentelė. Prijungtos gDNR gausinimo mišinys

Reagentai	Tūris (µl)/ reakcijai
Gausinimo mišinys (angl. <i>Platinum<sup>®</sup> PCR, Amplification Mix</i> )	200
50 µM pradmuo 1 (angl. <i>Library PCR Primer 1</i> )	5
50 µM pradmuo 2 (angl. <i>Library PCR Primer 2</i> )	5
ddH <sub>2</sub> O	25
gDNR prijungta prie padengtų streptavidinu dalelių	15
<b><i>Iš viso</i></b>	250

- (116) Paruoštą mišinį išpilstyti po 125 µl į du 0,2 ml tūrio mėgintuvėlius kiekvienam mėginiui.
- (117) Mėgintuvėlius su mišiniu supurtyti ir centrifuguoti 5 s.
- (118) Paruoštus mišinius inkubuoti termocikleryje naudojant iš anksto paruoštą programą (38 lentelė).

38 lentelė. gDNR, prijungtos prie streptavidinu padengtų dalelių, gausinimo sąlygos

Etapas	Temperatūra (°C)	Trukmė (min)	Ciklų skaičius
Pirminė denatūracija	95	05:00	1
Denatūracija	95	00:15	10
Pradmenų prijungimas	62	00:15	
Sintezė	70	01:00	
Galutinė sintezė	70	05:00	1

- (119) Po inkubavimo kiekvieno mėginio mišinius iš dviejų 0,2 ml tūrio mėgintuvėlių sujungti į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (120) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių ir perkelti į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (121) Atlikti mišinio valymą magnetinėmis dalelėmis (26–35 etapai), naudojant 26 etape 375 µl gerai supurtytų magnetinių dalelių ir 34 etape 50 µl 1× TE buferio (0,1 mM EDTA) arba ddH<sub>2</sub>O.
- (122) Atlikti gDNR koncentracijos nustatymą kiekybinės RL-PGR metodu (54–74 etapai), mėginius skiesti 2 000 kartų.
- (123) 
$$V (\mu l) = \frac{500 \text{ pM}}{(\text{konc. (pM)} \times 2\,000)} \times 16,643$$

**Dalelių emulsijos sudarymo sistema: alyvos ruošimas**

- (124) 50 ml tūrio mėgintuvėlyje paruošti alyvos mišinį (39 lentelė), reagentus pilti pagal nurodytą tvarką. Stabilizatorių 1 matuoti steriliu 3 ml tūrio švirkštu, stabilizatorių 2 – steriliu 1 ml tūrio švirkštu.

39 lentelė. Alyvos mišinys

Reagentai	Tūris (ml)/ reakcijai
Stabilizatorius 1 (angl. <i>Stabilizer 1</i> )	1,8
Stabilizatorius 2 (angl. <i>Stabilizer 2</i> )	0,4
Alyva (angl. <i>Oil</i> )	37,8
<b><i>Iš viso</i></b>	40

- (125) Mišinį 50 ml tūrio mėgintuvėlyje susidarius verpetui purtyti dar 15 s purtykle, toliau mišinį intensyviai purtyti rankos judesiu 15 s.
- (126) Mišinį perpilti į emulsijos alyvos ruošimo reikmenų rinkinyje esantį alyvos indą, neužsuktu dangteliu laikyti tamsoje 30 min, kad pasišalintų oras.

**Dalelių emulsijos sudarymo sistema: sistemos paruošimas**

- (127) Įjungti dalelių emulsijos sudarymo sistemą ir pasirinkti programą E120.



(128) Emulsijos ruošimo vamzdelį įstatyti į dalelių emulsijos sudarymo sistemos pompos kasetę, laikant vamzdelio galą su fiksatoriumi dešinėje ir įterpiančiam vamzdelio vidurinius 2 ir 3 fiksatorius nuo dešinio krašto į atitinkamai pompos kasetės briauną su veržimo mechanizmu (dešinė) ir be veržimo mechanizmo (kairė).

(129) Paruoštą pompos kasetę įstatyti į dalelių emulsijos sudarymo sistemą.

**Dalelių emulsijos sudarymo sistema: P1 dalelių ruošimas**

(130) P1 daleles supurtyti, centrifuguoti 5 s ir įpilti po 1430 µl P1 dalelių į du naujus 1,5 ml tūrio sumažintos adhezijos mėgintuvėlius.

(131) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.

(132) Į kiekvieną mėgintuvėlį įpilti po 1430 µl atšildyto BB tirpalo ir mėgintuvėlius su mišiniu gerai supurtyti, kad dalelės tolygiai pasiskirstytų po tirpalą.

(133) Mėgintuvėlį su dalelėmis įtvirtinti į 1,5 ml tūrio mikromėgintuvėlių laikiklį.

(134) Laikiklį perkelti į iš anksto įjungtą ir paruoštą skaidymo ultragarsu sistemą, įsitikinus, kad vėsinimo blokas pasiekęs nustatytą 3 °C temperatūrą.

(135) Programoje *SonoLab™ v7.1.0.5* pasirinkti iš anksto paruoštą programą (40 lentelė).

40 lentelė. Dalelių skaidymo programa

<b>Galia (W)</b> (angl. <i>Peak Power</i> )	<b>Impulsų užpildos faktorius (%)</b> (angl. <i>Duty Factor</i> )	<b>Impulsų vora/ciklui</b> (angl. <i>Cycle/Burst</i> )	<b>Ciklo trukmė (s)</b>	<b>Ciklų skaičius</b>
175	1	50	5	1
175	5	100	60	1

(136) 133–135 etapas kartoti su kitu mėgintuvėliu.

(137) Mėgintuvėlius su mišiniu centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.

- (138) Į kiekvieną mėgintuvėlį įpilama po 715 µl 1× TEX buferio ir mėgintuvėliai su mišiniu gerai supurtomi, kad dalelės tolygiai pasiskirstytų po tirpalą.
- (139) Vieno mėgintuvėlio P1 dalelės su 715 µl 1× TEX buferiu perkeliama į kitą mėgintuvėlį su P1 dalelėmis.
- (140) Mėgintuvėlius su mišiniu centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (141) Į mėgintuvėlį įpilti 1430 µl 1× TEX buferio ir mėgintuvėlį su mišiniu gerai supurtyti, kad dalelės tolygiai pasiskirstytų po tirpalą.

**Dalelių emulsijos sudarymo sistema: alyvos perkėlimas į sistemą**

- (142) Įpilti 67,9 g alyvos į emulsijos ruošimo indą, užsukti dangtelį ir indą įtvirtinti į dalelių emulsijos sudarymo sistemoje skirtą vietą (dešinėje).
- (143) Prijungti paruošto emulsijos ruošimo vamzdelio dešinį galą su fiksatoriumi prie įtvirtinto indo dangtelio.

**Dalelių emulsijos sudarymo sistema: mišinio vandens pagrindu ruošimas**

- (144) Įpilti kiekvieno iš trijų sujungtų gDNR mėginių į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį pagal 123 etape apskaičiuotą tūrį.
- (145) Įpilti 21 µl atitirpinto P1 reagento į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį su 189 µl 1× TE buferio (0,1 mM EDTA).
- (146) Emulsijos vandeningos fazės inde, laikomame ant ledo, paruošti mišinį (41 lentelė), reagentus pilant pagal nurodytą tvarką. Emulsijos mišinį vandens pagrindu pilti naudojant didelių tūrių skysčių automatinę pipetę kartu su serologinėmis pipetėmis.

41 lentelė. Emulsijos mišinys vandens pagrindu

Reagentai	Tūris (µl)/ reakcijai
Emulsijos mišinys vandens pagrindu (angl. <i>Aqueous Master Mix</i> )	48000-n
Paruoštas P1 reagentas (angl. <i>P1 Reagent</i> );	200
P2 reagentas (angl. <i>P2 Reagent</i> );	300
gDNR mišinys (144 etapas)	n
<i>Iš viso:</i>	48500

- (147) Indą su mišiniu užsukti, sumaišyti rankos judesiu ir perkelti atgal ant ledo.

- (148) Mėgintuvėlį su paruoštomis P1 dalelėmis supurtyti ir centrifuguoti 5 s.
- (149) Mėgintuvėlį su dalelėmis įtvirtinti į 1,5 ml tūrio mikromėgintuvėlių laikiklį.
- (150) Laikiklį perkelti į iš anksto įjungtą ir paruoštą skaidymo ultragarsu sistemą, įsitikinus, kad vėsinimo blokas pasiekęs nustatytą 3 °C temperatūrą.
- (151) Programoje *SonoLab™ v7.1.0.5* pasirinkti iš anksto paruoštą programą (42 lentelė).

42 lentelė. Dalelių skaidymo programa

Galia (W) (angl. <i>Peak Power</i> )	Impulsų užpildos faktorius (%) (angl. <i>Duty Factor</i> )	Impulsų vora/ciklui (angl. <i>Cycle/Burst</i> )	Ciklo trukmė (s)	Ciklų skaičius
200	1,1	100	5	1
275	5	100	30	1
200	1,1	100	5	1
275	5	100	30	1
200	1,1	100	5	1
275	5	100	30	1

- (152) Mėgintuvėlį su P1 dalelėmis centrifuguoti 1 s ir visą mišinį skubiai įpilti į paruoštą emulsijos mišinį vandens pagrindu.
- (153) Indą užsukti, mišinį sumaišyti rankos judesiu ir įtvirtinti į dalelių emulsijos sudarymo sistemoje skirtą vietą (kairėje).
- (154) Prijungti paruošto emulsijos ruošimo vamzdelio kairįjį galą be fiksatoriaus prie įtvirtinto indo dangtelio.
- (155) Aktyvuoti pasirinktą E120 programą (trukmė 15 min).

#### **Dalelių gausinimo sistema: gausinimo maišelio paruošimas**

- (156) Naują gausinimo maišelį su atsuktais abiem kamšteliais įstatyti į pildymo sistemą.
- (157) Naudojant piltuvėlį įpilti 10 ml paruoštos alyvos (124–126 etapai).
- (158) Užsukti kamštelių ir gausinimo maišelio vidinį paviršių gerai įtrinti alyva, jos perteklių, atsukus kamštelių ir apvertus gausinimo maišelį, išpilti ant vienkartinųjų servetėlių.

#### **Dalelių gausinimo sistema: gausinimo maišelio pripildymas**

- (159) Gausinimo maišelį perkelti į pildymo sistemą ir įpilti visą susidariusią emulsiją (155 etapas).
- (160) Pildymo sistema iš gausinimo maišelio išspausti kiek įmanoma daugiau oro ir užsukti kamšteliu.
- (161) Gausinimo maišelį įtvirtinti į plastikinį stovėlį su magnetais ir įstatyti į dalelių gausinimo sistemą.
- (162) Uždaryti sistemos duris ir aktyvuoti iš anksto pasirinktą E120 programą (trukmė 3 h 36 min).

### **Dalelių praturtinimo sistema: sistemos paruošimas ir aktyvavimas**

- (163) Dalelių praturtinimo sistemą, jei reikia, išvalyti.
- (164) Į sugavimo zondų mėgintuvėlį įpilti 1 ml sugavimo buferio ir 28 µl atšildytų sugavimo zondų.
- (165) Mišinį mėgintuvėlyje sumaišyti rankos judesiu.
- (166) CS tirpalo 1 priedo visą tūrį supilti į CS tirpalą 1.
- (167) Sugavimo dalelių mėgintuvėlio dangtelį persukti ant sugavimo dalelių indo.
- (168) Paruoštus praturtinimo buferių rinkinius 1 ir 2, butanolį ir praturtinimo reikmenų rinkinio sudedamąsias dalis, išskyrus emulsijos indą, įtvirtinti į dalelių praturtinimo sistemą ir sujungti su atitinkamais vamzdeliais pagal spalvinį žymėjimą.
- (169) Išvalyti atliekų indą.
- (170) Įpilti emulsiją iš gausinimo maišelio į emulsijos indą ir nustatyti emulsijos masę sveriant.
- (171) Emulsijos indą įtvirtinti į dalelių praturtinimo sistemą, prijungti atitinkamos spalvos vamzdelį.
- (172) Sandariai uždaryti dalelių praturtinimo sistemos duris ir aktyvuoti iš anksto pasirinktą E120 programą, įvedant nustatytą emulsijos svorį (trukmė 4 val. 4 min).
- (173) Atlikti dalelių praturtinimo sistemos valymo procedūras pagal gamintojo pateiktas rekomendacijas.

### **Praturtintų dalelių koncentravimas**

- (174) Įpilti 1 ml dalelių tirpalo iš dalelių išeigos mėgintuvėlio į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį.
- (175) Mėgintuvėlį su dalelių tirpalu perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (176) 174–175 etapus kartoti iki dalelių išeigos mėgintuvėlyje liks 1 ml tirpalo.
- (177) Įpylus paskutinį 1 ml tirpalą, 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį perkelti iš magnetinio stovo, supurtyti ir centrifuguoti 2 s.

### **gDNR mėginių fragmentų 3' galų modifikavimas**

- (178) Paruošti 5 ml terminalinės transferazės mišinio (43 lentelė).

43 lentelė. Terminalinės transferazės mišinys

<b>Reagentai</b>	<b>Tūris (μl)</b>
10× terminalinės transferazės buferis (angl. <i>10× Terminal Transferase Buffer</i> )	550
10× kobalto chloridas (angl. <i>10× Cobalt Chloride</i> )	550
ddH <sub>2</sub> O	3900
<i>Iš viso</i>	5000

- (179) Įpilti 10 μl 50 mM dalelių jungiklio tirpalo (angl. *Bead Linker*) į naują 1,5 ml tūrio sumažintos adhezijos mėgintuvėlį su 490 μl 1× TE buferio (0,1 mM EDTA).
- (180) Perkelti mėgintuvėlį su koncentruotomis dalelėmis į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (181) Įpilti 600 μl terminalinės transferazės mišinio į mėgintuvėlį su dalelėmis.
- (182) Mėgintuvėlį su mišiniu supurtyti ir centrifuguoti 2 s, perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (183) Įpilti 600 μl terminalinės transferazės mišinio į mėgintuvėlį su dalelėmis.
- (184) Mėgintuvėlį su mišiniu supurtyti ir centrifuguoti 2 s.
- (185) Įpilti po 150 μl terminalinės transferazės mišinio su dalelėmis į tris naujus 1,5 ml tūrio sumažintos adhezijos mėgintuvėlius (iš viso keturi mėgintuvėliai).

- (186) Perkelti mėgintuvėlius į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (187) Į kiekvieną mėgintuvėlį įpilti po 890 μl terminalinės transferazės mišinio.
- (188) Į kiekvieną mėgintuvėlį įpilti po 100 μl paruošto 1 mM dalelių jungiklio tirpalo.
- (189) Mėgintuvėlį su dalelėmis įtvirtinti į 1,5 ml tūrio mikromėgintuvėlių laikiklį.
- (190) Laikiklį perkelti į iš anksto įjungtą ir paruoštą skaidymo ultragarsu sistemą, įsitikinus, kad vėsinimo blokas pasiekęs nustatytą 3 °C temperatūrą.
- (191) Programoje *SonoLab*<sup>TM</sup> v7.1.0.5 pasirinkti iš anksto paruoštą programą (42 lentelė).
- (192) Mėgintuvėlį su mišiniu centrifuguoti 1 s (dalelės neturi nusėsti mėgintuvėlio dugne ar šonuose) ir įpilti 10 μl 20 U/μl terminalinės transferazės.
- (193) Mėgintuvėlį su mišiniu akimirksniu supurtyti ir centrifuguoti 1 s (dalelės neturi nusėsti mėgintuvėlio dugne ar šonuose) ir perkelti į termopurtyklę.
- (194) 189–193 etapus kartoti su kitais trimis mėgintuvėliais.
- (195) Mėgintuvėlius su mišiniu inkubuoti termopurtyklėje 37 °C temperatūroje 1000 aps./min 150 min.

#### **Tėkmės lusto kamerų pripildymas paruoštomis gDNR fragmentų bibliotekomis**

- (196) Mėgintuvėlius su mišiniu centrifuguoti 1 s (dalelės neturi nusėsti mėgintuvėlio dugne ar šonuose), du mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (197) Iš likusių dviejų mėgintuvėlių perpilti visą kiekvieno mišinio tūrį atitinkamai į du mėgintuvėlius, esančius magnetiniame stove.
- (198) Laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.

- (199) Į kiekvieną mėgintuvėlį įpilti po 500 µl 1× TEX buferio, akimirksniu supurtyti ir centrifuguoti 1 s (dalelės neturi susirinkti mėgintuvėlio dugne ar šonuose).
- (200) Vieno mėgintuvėlio visą mišinio tūrį perkelti į kitą mėgintuvėlį.
- (201) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (202) Įpilti 1000 µl 1× TEX buferio, mėgintuvėlį akimirksniu supurtyti ir centrifuguoti 1 s (dalelės neturi nusėsti mėgintuvėlio dugne ar šonuose).
- (203) Įpilti po 166 µl mišinio su dalelėmis į penkis naujus 1,5 ml tūrio sumažintos adhezijos mėgintuvėlius (iš viso šeši mėgintuvėliai).
- (204) Į kiekvieną mėgintuvėlį įpilti po 334 µl suleidimo buferio.
- (205) Tęsti protokolą su trimis mėgintuvėliais.
- (206) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (207) Į kiekvieną mėgintuvėlį įpilti po 400 µl suleidimo buferio.
- (208) Mėgintuvėlius su mišiniu akimirksniu supurtyti ir centrifuguoti 1 s.
- (209) 206–208 etapas kartoti dar du kartus.
- (210) Mėgintuvėlius perkelti į magnetinį stovą ir laukti mažiausiai 1 min, kol mišinys pasidarys skaidrus, tirpalą išsiurbti neliečiant magnetinio stovo pusėje susirinkusių dalelių.
- (211) Įpilti ant tėkmės lusto pakuotės nurodytą suleidimo buferio tūrį (~30 µl).
- (212) Mėgintuvėlius su mišiniu akimirksniu supurtyti ir centrifuguoti 1 s.
- (213) Mėgintuvėlį su dalelėmis įtvirtinti į 1,5 ml tūrio mikromėgintuvėlių laikiklį.
- (214) Laikiklį perkelti į iš anksto įjungtą ir paruoštą skaidymo ultragarsu sistemą, įsitikinus, kad vėsinimo blokas pasiekęs nustatytą 3 °C temperatūrą.
- (215) Programoje *SonoLab*<sup>TM</sup> v7.1.0.5 pasirinkti iš anksto paruoštą programą (44 lentelė).

44 lentelė. Dalelių skaidymo programa

<b>Galia (W)</b> (angl. <i>Peak Power</i> )	<b>Impulsų užpildos faktorius (%)</b> (angl. <i>Duty Factor</i> )	<b>Impulsų vora/ciklui</b> (angl. <i>Cycle/Burst</i> )	<b>Ciklo trukmė (s)</b>	<b>Ciklų skaičius</b>
200	2	100	5	3
275	5	100	30	1

- (216) Mėgintuvėlį su mišiniu akimirksniu centrifuguoti 1 s (dalelės neturi nusėsti mėgintuvėlio dugne ar šonuose).
- (217) Nedelsiant visą mišinio tūrį suleisti į paruoštą (nulupti lipnias plėveles šonuose) adapteryje įtvirtintą tėkmės lusto kamerą.
- (218) 213–217 etapus kartoti su likusiais dviem mėgintuvėliais.
- (219) 206–218 etapus kartoti su likusiais trimis mėgintuvėliais.
- (220) Paruoštą tėkmės lustą apvynioti parafino juostele ir inkubuoti termostate 37 °C temperatūroje 90 min.

#### **Genetinio analizatoriaus paruošimas**

- (221) Įjungti genetinį analizatorių (angl. *5500 SOLiD™ Sequencer*) ir aktyvuoti programą *Instrument Control Software 1.2.1 (ICS 1.2.1)*.
- (222) Įjungti lempą, *ICS 1.2.1* programos viršuje spausti *Lamp > Turn On*, valandą iki sekoskaitos pradžios.
- (223) Pripildyti skysčių sistemos (angl. *System Fluid*) indą nauju ddH<sub>2</sub>O.
- (224) Ištuštinti, jeigu reikia, cheminių atliekų (angl. *Chemical Waste*) ir plovimo atliekų (angl. *Rinse Waste*) indus.
- (225) Skiesti 10× sekoskaitos buferį Fv2, atsargiai įpilant 200 ml 10× sekoskaitos buferio F į indą su 1800 ml ddH<sub>2</sub>O, palikti nusistovėti, nesudarant burbulų perpilti į genetiniam analizatoriui pritaikytą indą, įstatyti į genetiniame analizatoriuje skirtą vietą.
- (226) Nesudarant burbulų perpilti sekoskaitos buferį G į genetiniam analizatoriui pritaikytą indą ir įstatyti į genetiniame analizatoriuje skirtą vietą.
- (227) Atšildytų sekoskaitos buferių A, B ir E, susidedančių iš dviejų dalių, papildomą reagentų dalį supilti į pagrindinį kiekvieno buferio indą.



- (228) Paruoštus sekoskaitos tiesioginių buferių, A, B, C, D, E buferius įstatyti į genetinio analizatoriaus šaldymo bloke skirtą vietą.
- (229) Atšildyti, centrifuguoti 500 aps./min 1 min ir į genetinio analizatoriaus šaldymo bloke skirtą vietą įstatyti sekoskaitos reagentų rinkinį, pradmenų rinkinius 1 ir 2, mėgintuvėlius sekoskaitos reagentams maišyti.

### **Sekoskaitos protokolo kūrimas**

- (230) *ICS 1.2.1* programoje skiltyje *Samples* pasirinkti *Create a Sample*.
- (231) Laukelyje *Sample Name\** įvesti tėkmės lusto pirmos kameros mėginių pavadinimą, pavyzdžiui, „LITGEN\_20140101\_lane1“, skiltyje *Library Type* pažymėti *Fragment*, pažymėti laukelį *Barcoding*, laukelyje *Kit* pasirinkti *BC Kit Module 1–16*, skiltyje *Tag Selection* laukelyje *F3 Tag* pasirinkti *75BP*, laukelyje *BC Tag* pasirinkti *5BP*, protokolo apačioje pateiktų unikalių adapterių sąrašė pašalinti keturis gDNR mėginių ruošimo metu nenaudotus unikalius adapterius, pažymint eilutes su unikaliais adapteriais ir spaudžiant dešiniu klavišu *Delete Library(ies)*, likusiame sąrašė įrašyti mėginių, kuriems buvo prijungtas unikalus adapteris, pavadinimus, suvedus mėginius spausti *Done*.
- (232) 229 etapą kartoti su kitomis penkiomis tėkmės lusto kameromis, *ICS 1.2.1* programos kairėje juostoje matyti naujai kuriamų įrašų sąrašas.
- (233) Skiltyje *Run* susieti kiekvieną sukurtą sekoskaitos protokolą su tėkmės lusto kamera, tempiant protokolo pavadinimą iš sąrašo kairėje pusėje ant virtualiai pavaizduotos tėkmės lusto kameros.
- (234) 233 etapą kartoti su likusiais penkiais sekoskaitos protokolais ir virtualiai pavaizduotomis tėkmės lusto kameromis.
- (235) Spausti *Continue*.
- (236) Įvesti sekoskaitos projekto pavadinimą paskyroje *Run name*.
- (237) Spausti *Continue*.

### **Tėkmės lusto įstatymas ir sekoskaitos vykdymas**

- (238) Po inkubacijos, išardyti tėkmės lusto adapterį ir nuvalyti lusto kraštus beplauše servetėle suvilgyta ddH<sub>2</sub>O.
- (239) Paruoštą tėkmės lustą įstatyti į genetiniame analizatoriuje skirtą vietą.

- (240) Uždėti tēkmės lusto dangtelį ir spaudžiant tolygiai dviem rankomis vienu judesiu užsukti abu fiksatorius.
- (241) Virš tēkmės lusto uždėti apsaugą nuo skysčių (angl. *Splash Guard*).
- (242) Uždaryti genetinio analizatoriaus duris ir *ICS 1.2.1* programoje pažymėti *Door > Lock Door*.
- (243) Programoje *ICS 1.2.1* skiltyje *Run* pažymėti varnele *Loaded Flowchip*, esant tēkmės lustui, įstatytam į genetiniame analizatoriuje skirtą vietą.
- (244) Pasirinkti programą *Imager 1.2.1*.
- (245) Paskyroje *FlowChip FCI* laukelyje *Lane* pasirinkti tēkmės lusto kameros numerį, laukelyje *Panel* – tēkmės lusto kameros skilties numerį, spausti *Find Focus Range*.
- (246) Atsiradus informacinei lentelei su nustatyta fokusavimo z verte, kuri turėtų būti 5 500–7 000, esant tinkamai reikšmei spausti *Ok*.
- (247) Pasirinkti programą *ICS 1.2.1* ir spausti *Start Run*.

## 2 PRIEDAS. PAPILDOMA MEDŽIAGA

### Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis Y chromosomos haplogrupių ir haplotipų pasiskirstymu

1 lentelė. Šešių Lietuvos etnolingvistinių grupių geografinės koordinatės (dešimtainiai laipsniai) naudotos nustatyti genetinės įvairovės barjerus *Barrier v2.2* (Manni *et al.* 2004)

Etnolingvistinė grupė	Platuma	Ilguma	Artimiausias miestas
Rytų Aukštaitija	55,95	25,46	Rokiškis
Vakarų Aukštaitija	54,73	23,18	Kazlų Rūda
Pietų Aukštaitija	54,2	23,88	Seirijai
Šiaurės Žemaitija	56,16	21,76	Mosėdis
Pietų Žemaitija	55,31	22,72	Eržvilkas
Vakarų Žemaitija	55,53	21,35	Priekulė

2 lentelė. Šešių Lietuvos etnolingvistinių grupių geografinių atstumų (km) matrica, naudota vertinant koreliaciją tarp genetinių ir geografinių atstumų matricų (Mantelio kriterijus *R v3.0.3 (Ade4)*)

	RA	ŠŽ	PŽ	VŽ	VA	PA
RA	0	230,9	74,58	112,3	182,5	256,1
ŠŽ	230,9	0	261,4	186,1	197,9	184,6
PŽ	74,58	261,4	0	89,85	146,4	200,7
VŽ	112,3	186,1	89,85	0	70,85	143,8
VA	182,5	197,9	146,4	70,85	0	74,29
PA	256,1	184,6	200,7	143,8	74,29	0

3 lentelė. Y chromosomos 15 TTKS mutacijų dažnis, tirtų mejozių skaičius ir suteiktas svartinis koeficientas haplogrupių N1c1, R1a1a, R1a1alg filogenetinio medžio sudarymui *Network v4.6.1.2* ([www.fluxus-engineering.com](http://www.fluxus-engineering.com)) (Ballantyne *et al.* 2010)

Žymuo (DYS)	393	390	19	391	439	389I	392	389II
Kartotinio motyvo ilgis	4	4	4	4	4	4	3	4
Mutacijų skaičius	3	2	7	5	6	9	1	6
Mejozių skaičius	1750	1758	1756	1759	1736	1751	1728	1743
Svertinis koeficientas	7	8	4	6	5	2	9	5
Žymuo (DYS)	458	437	448	H4	456	438	635	Vid.
Kartotinio motyvo ilgis	4	4	6	4	4	5	4	
Mutacijų skaičius	14	2	0	5	8	1	6	5
Mejozių skaičius	1756	1760	1747	1755	1757	1751	1732	1749
Svertinis koeficientas	1	8	10	6	3	9	5	

4 lentelė. Dažniausių Y chromosomos haplogrupių TMRCA skaičiavimui naudoti išankstiniai parametrai programa *BATWING* (Wilson *et al.* 2003)

Mutacijų dažnis – $8,5 \cdot 10^{-3}$ mutacija/žymuo/karta (Ballantyne <i>et al.</i> 2010)					
Haplogrupė	Imtis	N prior <sup>1</sup>	mu prior <sup>2</sup>	alfa prior <sup>3</sup>	beta prior <sup>4</sup>
N1c1	119	gamma(1,0.001)	gamma(5,1749)	uniform(0,0.04)	uniform(0,1)
R1a1a	97	gamma(1,0.001)	gamma(5,1749)	uniform(0,0.04)	uniform(0,1)
R1a1a1g	25	gamma(1,0.0025)	gamma(5,1749)	uniform(0,0.04)	uniform(0,1)
Mutacijų dažnis – $6,9 \cdot 10^{-4}$ mutacija/žymuo/karta (Zhivotovsky <i>et al.</i> 2004)					
Haplogrupė	Imtis	N prior	mu prior	alfa prior	beta prior
N1c1	119	gamma(1,0.001)	gamma(1.47,2130)	uniform(0,0.04)	uniform(0,1)
R1a1a	97	gamma(1,0.001)	gamma(1.47,2130)	uniform(0,0.04)	uniform(0,1)
R1a1a1g	25	gamma(1,0.0025)	gamma(1.47,2130)	uniform(0,0.04)	uniform(0,1)
Mutacijų dažnis – $2 \cdot 10^{-3}$ mutacija/žymuo/karta (Luca <i>et al.</i> 2007)					
Haplogrupė	Imtis	N prior	mu prior	alfa prior	beta prior
N1c1	119	gamma(1,0.001)	gamma(2,1000)	uniform(0,0.04)	uniform(0,1)
R1a1a	97	gamma(1,0.001)	gamma(2,1000)	uniform(0,0.04)	uniform(0,1)
R1a1a1g	25	gamma(1,0.0025)	gamma(2,1000)	uniform(0,0.04)	uniform(0,1)

<sup>1</sup>Pradinis populiacijos dydis (*N prior*)

<sup>2</sup>Mutacijos dažnis, mutacija/žymuo/karta (*mu prior*)

<sup>3</sup>Populiacijos augimo greitis (*alfa prior*)

<sup>4</sup>Populiacijos augimo pradžia (*beta prior*)

5 lentelė. Y chromosomos haplogrupių pasiskirstymas Europos populiacijose

	R1a	R1b	N3	I1	J2	E1	G	Šaltinis
Latvija	0,389	0,097	0,416	0,071	0	0,009	0	(Lappalainen <i>et al.</i> 2008)
Estija	0,373	0,042	0,339	0,169	0,017	0,025	0	(Lappalainen <i>et al.</i> 2008)
Lenkija	0,57	0,116	0,037	0,173	0,025	0,005	0	(Kayser <i>et al.</i> 2005)
Baltarusija	0,456	0,044	0,088	0,176	0,044	0,044	0,015	(Kharkov <i>et al.</i> 2005)
R. Suomija	0,059	0,026	0,709	0,197	0	0	0	(Lappalainen <i>et al.</i> 2008)
V. Suomija	0,087	0,052	0,413	0,413	0	0,009	0	(Lappalainen <i>et al.</i> 2008)
Rusija	0,483	0,068	0,14	0,159	0,014	0,048	0,012	(Malyarchuk, Derenko 2008)
Lietuva	0,4164	0,044	0,406	0,0478	0,01365	0,0273	0,0102	Darbo metu gauti rezultatai
Ukraina	0,397	0,208	0,057	0,038	0,076	0	0	(Varzari <i>et al.</i> 2013)
Švedija	0,244	0,131	0,144	0,375	0	0,013	0	(Lappalainen <i>et al.</i> 2008)
Vokietija	0,179	0,389	0,016	0,236	0,04	0,062	0	(Kayser <i>et al.</i> 2005)
Moldavija	0,304	0,16	0,016	0,048	0,04	0,128	0,008	(Varzari <i>et al.</i> 2013)
Rumunija	0,2037	0,1296	0	0,037	0,055	0,074	0,055	(Varzari <i>et al.</i> 2013)
Norvegija	0,263	0,05	0,038	0,325	0,011	0,011	0	(Dupuy <i>et al.</i> 2006)
Graikija	0,113	0,133	0,0066	0,026	0,153	0,32	0,053	(King <i>et al.</i> 2011)
Turkija	0,055	0,1775	0,0475	0,025	0,2475	0,1075	0,115	(King <i>et al.</i> 2011)
Italija	0,12	0,233	0	0,0067	0,193	0	0,09	(Brisighelli <i>et al.</i> 2012)
Danija	0,165	0,361	0,005	0,387	0,026	0	0	(Sanchez <i>et al.</i> 2004)
Grenlandija	0,088	0,199	0	0,167	0,019	0	0	(Sanchez <i>et al.</i> 2004)
Čekija	0,342	0,28	0,016	0,051	0,035	0,008	0,051	(Luca <i>et al.</i> 2007)
Britanija	0,054	0,59	0,009	0,172	0,009	0,009	0,036	(King, Jobling 2009)
Kroatija	0,339	0,156	0	0,376	0,018	0,055	0,009	(Barac <i>et al.</i> 2003)

6 lentelė. Nei (D) atstumų matrica pagal Y chromosomos haplogrupių (E1, G, I1, J2, N3, R1a, R1b) pasiskirstymą Europos populiacijose

	Latvija	Estija	Lenkija	Baltarusija	R. Suomija	V. Suomija	Rusija	Lietuva	Ukraina	Švedija	Vokietija
Latvija	0	0,003395	0,032796	0,022759	0,036619	0,037339	0,016834	0,000835	0,025829	0,033708	0,058159
Estija	0,003395	0	0,023638	0,012501	0,040461	0,026169	0,009374	0,003673	0,022454	0,018834	0,048798
Lenkija	0,032796	0,023638	0	0,003885	0,128073	0,077964	0,003832	0,031576	0,010195	0,027803	0,041433
Baltarusija	0,022759	0,012501	0,003885	0	0,096403	0,053715	0,000900	0,021101	0,009209	0,017312	0,036530
R. Suomija	0,036619	0,040461	0,128073	0,096403	0	0,022714	0,089503	0,041407	0,106003	0,068989	0,113248
V. Suomija	0,037339	0,026169	0,077964	0,053715	0,022714	0	0,053315	0,043067	0,070074	0,018480	0,056871
Rusija	0,016834	0,009374	0,003832	0,000900	0,089503	0,053315	0	0,015547	0,009420	0,019378	0,039043
Lietuva	0,000835	0,003673	0,031576	0,021101	0,041407	0,043067	0,015547	0	0,026782	0,038076	0,066102
Ukraina	0,025829	0,022454	0,010195	0,009209	0,106003	0,070074	0,009420	0,026782	0	0,027337	0,021989
Švedija	0,033708	0,018834	0,027803	0,017312	0,068989	0,018480	0,019378	0,038076	0,027337	0	0,019735
Vokietija	0,058159	0,048798	0,041433	0,036530	0,113248	0,056871	0,039043	0,066102	0,021989	0,019735	0
Moldavija	0,032829	0,026009	0,017878	0,011239	0,105077	0,064886	0,012996	0,033013	0,005206	0,024890	0,018718
Rumunija	0,037783	0,029969	0,027635	0,016955	0,098834	0,059468	0,020361	0,038359	0,009309	0,024844	0,018664
Norvegija	0,039186	0,021810	0,020776	0,010904	0,088038	0,030958	0,015003	0,040996	0,022188	0,003250	0,022956
Graikija	0,066976	0,057111	0,062146	0,043454	0,118917	0,080816	0,048675	0,066124	0,035072	0,050173	0,035154
Turkija	0,061465	0,054055	0,064809	0,046109	0,102093	0,069609	0,052188	0,062974	0,030138	0,045583	0,029329
Italija	0,055093	0,049476	0,048755	0,037253	0,110480	0,073455	0,042236	0,057829	0,017637	0,039736	0,019895
Danija	0,070494	0,055085	0,047666	0,042461	0,116927	0,048067	0,046433	0,079822	0,035777	0,014032	0,004899
Grenlandija	0,047928	0,036965	0,040273	0,028153	0,090817	0,042543	0,032480	0,052836	0,019471	0,014834	0,007985
Čekija	0,035168	0,031549	0,016658	0,015990	0,114951	0,073236	0,016596	0,037916	0,002460	0,027568	0,013575
Britanija	0,094955	0,091646	0,086534	0,082770	0,143541	0,091088	0,084353	0,108044	0,049948	0,053730	0,010781
Kroatija	0,049468	0,031071	0,017304	0,013127	0,113147	0,044525	0,016859	0,052443	0,022644	0,005766	0,017909

7 lentelė. TĘSINYS. Nei (D) atstumų matrica pagal Y chromosomos haplogrupių (E1, G, I1, J2, N3, R1a, R1b) pasiskirstymą Europos populiacijose

	<b>Moldavija</b>	<b>Rumunija</b>	<b>Norvegija</b>	<b>Graikija</b>	<b>Turkija</b>	<b>Italija</b>	<b>Danija</b>	<b>Grenlandija</b>	<b>Čekija</b>	<b>Britanija</b>	<b>Kroatija</b>
<b>Latvija</b>	0,032829	0,037783	0,039186	0,066976	0,061465	0,055093	0,070494	0,047928	0,035168	0,094955	0,049468
<b>Estija</b>	0,026009	0,029969	0,021810	0,057111	0,054055	0,049476	0,055085	0,036965	0,031549	0,091646	0,031071
<b>Lenkija</b>	0,017878	0,027635	0,020776	0,062146	0,064809	0,048755	0,047666	0,040273	0,016658	0,086534	0,017304
<b>Baltarusija</b>	0,011239	0,016955	0,010904	0,043454	0,046109	0,037253	0,042461	0,028153	0,015990	0,082770	0,013127
<b>R. Suomija</b>	0,105077	0,098834	0,088038	0,118917	0,102093	0,110480	0,116927	0,090817	0,114951	0,143541	0,113147
<b>V. Suomija</b>	0,064886	0,059468	0,030958	0,080816	0,069609	0,073455	0,048067	0,042543	0,073236	0,091088	0,044525
<b>Rusija</b>	0,012996	0,020361	0,015003	0,048675	0,052188	0,042236	0,046433	0,032480	0,016596	0,084353	0,016859
<b>Lietuva</b>	0,033013	0,038359	0,040996	0,066124	0,062974	0,057829	0,079822	0,052836	0,037916	0,108044	0,052443
<b>Ukraina</b>	0,005206	0,009309	0,022188	0,035072	0,030138	0,017637	0,035777	0,019471	0,002460	0,049948	0,022644
<b>Švedija</b>	0,024890	0,024844	0,003250	0,050173	0,045583	0,039736	0,014032	0,014834	0,027568	0,053730	0,005766
<b>Vokietija</b>	0,018718	0,018664	0,022956	0,035154	0,029329	0,019895	0,004899	0,007985	0,013575	0,010781	0,017909
<b>Moldavija</b>	0	0,002716	0,017980	0,015474	0,020637	0,014935	0,033492	0,012866	0,005500	0,048013	0,019967
<b>Rumunija</b>	0,002716	0	0,017333	0,013084	0,011350	0,007407	0,031961	0,007266	0,007809	0,044040	0,023336
<b>Norvegija</b>	0,017980	0,017333	0	0,041444	0,039653	0,033836	0,019078	0,012941	0,023593	0,062186	0,003654
<b>Graikija</b>	0,015474	0,013084	0,041444	0	0,011452	0,019882	0,055052	0,024304	0,032687	0,061538	0,047630
<b>Turkija</b>	0,020637	0,011350	0,039653	0,011452	0	0,004090	0,045198	0,016199	0,026935	0,046187	0,049300
<b>Italija</b>	0,014935	0,007407	0,033836	0,019882	0,004090	0	0,034537	0,010734	0,013680	0,033353	0,040240
<b>Danija</b>	0,033492	0,031961	0,019078	0,055052	0,045198	0,034537	0	0,012808	0,026883	0,019109	0,013350
<b>Grenlandija</b>	0,012866	0,007266	0,012941	0,024304	0,016199	0,010734	0,012808	0	0,014208	0,025685	0,017945
<b>Čekija</b>	0,005500	0,007809	0,023593	0,032687	0,026935	0,013680	0,026883	0,014208	0	0,033317	0,021955
<b>Britanija</b>	0,048013	0,044040	0,062186	0,061538	0,046187	0,033353	0,019109	0,025685	0,033317	0	0,054925
<b>Kroatija</b>	0,019967	0,023336	0,003654	0,04763	0,0493	0,04024	0,01335	0,017945	0,021955	0,054925	0

**Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis mtDNR haplogrupių pasiskirstymu**

8 lentelė. Nei (D) atstumų tarp šešių Lietuvos etnolingvistinių grupių matrica. Aukščiau įstrižainės – Nei (D) atstumai (23 grupės). Žemiau įstrižainės – Nei (D) atstumai (69 grupės). Trumpiniai: ŠŽ – Šiaurės Žemaitija, PŽ – Pietų Žemaitija, VŽ – Vakarų Žemaitija, VA – Vakarų Aukštaitija, PA – Pietų Aukštaitija, RA – Rytų Aukštaitija

	<b>RA</b>	<b>PA</b>	<b>VA</b>	<b>ŠŽ</b>	<b>PŽ</b>	<b>VŽ</b>
<b>RA</b>	0	0,002739	0,001627	0,001231	0,002881	0,000936
<b>PA</b>	0,000784	0	0,001554	0,0027	0,00207	0,001566
<b>VA</b>	0,000577	0,000789	0	0,002281	0,001921	0,001386
<b>ŠŽ</b>	0,000353	0,000731	0,000635	0	0,003716	0,001123
<b>PŽ</b>	0,000885	0,000823	0,000779	0,000846	0	0,002387
<b>VŽ</b>	0,000592	0,000895	0,000695	0,000464	0,000881	0

9 lentelė. Porinių skirtumų atstumų tarp šešių Lietuvos etnolingvistinių grupių matrica. Aukščiau įstrižainės – porinių skirtumų atstumai, naudojant haplotipus (577–16 568 np). Žemiau įstrižainės – porinių skirtumų atstumai, naudojant haplotipus (16 568 np). Kursyvu  $p = 0,07$ , paryškinta ir kursyvu  $p = 0,06$ . Trumpiniai: ŠŽ – Šiaurės Žemaitija, PŽ – Pietų Žemaitija, VŽ – Vakarų Žemaitija, VA – Vakarų Aukštaitija, PA – Pietų Aukštaitija, RA – Rytų Aukštaitija

	<b>RA</b>	<b>PA</b>	<b>VA</b>	<b>ŠŽ</b>	<b>PŽ</b>	<b>VŽ</b>
<b>RA</b>	0	0,00242	-0,00546	0,00101	0,0035	-0,00012
<b>PA</b>	0,00094	0	-0,00661	0,00139	0,01087	-0,00849
<b>VA</b>	-0,00438	-0,00419	0	0,00198	-0,00017	-0,00237
<b>ŠŽ</b>	0,00128	0,00285	0,00439	0	<b>0,01244</b>	-0,00433
<b>PŽ</b>	0,00236	0,00942	-0,00057	0,00972	0	0,009
<b>VŽ</b>	-0,00147	-0,00732	-0,00001	-0,00325	0,00788	0

10 lentelė. Europos populiacijų mtDNR haplotipai (16 569 np), naudoti šio darbo metu vykdytoms MDS ir PK analizėms

Populiacija	Imtis	db <i>GenBank</i> numeris <sup>1</sup>	Šaltinis
Samiai	17	AY882379, DQ902708	(Achilli et al. 2005; Ingman, Gyllensten 2007)
Baltarusija	33	KC867103	(Kushniarevich et al. 2013)
Kazanės totoriai	73	GU122975	(Malyarchuk et al. 2010)
Lenkija	26	JX307099, JX128041, JX266260	(Mielnik-Sikorska et al. 2013)
Ukraina	7	JX307099, JX128041, JX266260	(Mielnik-Sikorska et al. 2013)
Rusija	21	JX307099, JX128041, JX266260	(Mielnik-Sikorska et al. 2013)
Slovakija	11	JX307099, JX128041, JX266260	(Mielnik-Sikorska et al. 2013)
Čekija	8	JX307099, JX128041, JX266260	(Mielnik-Sikorska et al. 2013)
Gruzija	28	HM852756	(Schonberg et al. 2011)
Turkija	29	HM852756	(Schonberg et al. 2011)
Armėnija	30	HM852756	(Schonberg et al. 2011)
Azerbaidžanas	30	HM852756	(Schonberg et al. 2011)
Iranas	30	HM852756	(Schonberg et al. 2011)
Aškenaziai <sup>2</sup>	72	KC878709, JX273243	(Schonberg et al. 2011)
Ispanija	11	AY882379	(Costa et al. 2013)
Italija	20	AY882379	(Achilli et al. 2005)
Sardinija	9	GQ129143	(Achilli et al. 2005)

<sup>1</sup>Duomenys prieinami db *NCBI DNA & RNA PopSet*

<sup>2</sup>Artimieji Rytai



11 lentelė. Porinių skirtumų (žemiau įstrižainės) ir  $\Phi_{ST}$  (aukščiau įstrižainės) atstumų matricos, pagal mtDNR haplotipai (577–16 023 np) pasiskirstymą Europos populiacijose. Trumpiniai: AKZ – aškenaziai, ARM – Armėnija, AZR – Azerbaidžanas, BUY – Baltarusija, CZH – Čekija, ESP – Ispanija, GEO – Gruzija, IRN – Iranas, ITA – Italija, LIT – Lietuva, PLN – Lenkija, RUS – Rusija, SAA – samiai, SLO – Slovakija, SRD – Sardinija, TAT – Totoriai, TUR – Turkija, UKR – Ukraina. Paryškinti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) įverčiai

	AKZ	ARM	AZR	BUY	CZH	ESP	GEO	IRN	ITA
AKZ	0	0,0002	0,0002	<b>0,01225</b>	0,00022	0,00021	0,0002	0,00133	0,00273
ARM	<b>0,38361</b>	0	0	<b>0,01236</b>	0	0	0	0,00115	0,00259
AZR	<b>0,36423</b>	0,00312	0	<b>0,01236</b>	0	0	0	0,00115	0,00259
BUY	<b>0,59565</b>	<b>0,26931</b>	<b>0,24261</b>	0	0,0136	0,01315	<b>0,01239</b>	<b>0,0135</b>	<b>0,01517</b>
CZH	<b>0,59314</b>	<b>0,13284</b>	<b>0,1354</b>	<b>0,44405</b>	0	0	0	0,00126	0,00284
ESP	<b>0,46992</b>	<b>0,15571</b>	<b>0,13799</b>	<b>0,4405</b>	<b>0,43708</b>	0	0	0,00122	0,00274
GEO	<b>0,36082</b>	-0,0069	0,00328	<b>0,24868</b>	<b>0,13948</b>	<b>0,13045</b>	0	0,00115	0,00259
IRN	<b>0,44062</b>	0,01242	0,0119	<b>0,23806</b>	<b>0,10874</b>	<b>0,19527</b>	<b>0,02597</b>	0	0,00375
ITA	<b>0,44296</b>	<b>0,18812</b>	<b>0,1741</b>	<b>0,45519</b>	<b>0,44039</b>	-0,02004	<b>0,16682</b>	<b>0,23161</b>	0
LIT	<b>0,39475</b>	<b>0,02431</b>	<b>0,06029</b>	<b>0,27173</b>	<b>0,05121</b>	<b>0,19551</b>	<b>0,04336</b>	<b>0,02932</b>	<b>0,22035</b>
PLN	<b>0,52225</b>	<b>0,1154</b>	<b>0,11778</b>	<b>0,33908</b>	0,03756	<b>0,30631</b>	<b>0,1276</b>	<b>0,07772</b>	<b>0,33917</b>
RUS	<b>0,61405</b>	<b>0,19384</b>	<b>0,20112</b>	<b>0,48309</b>	<b>0,0651</b>	<b>0,50884</b>	<b>0,20583</b>	<b>0,16924</b>	<b>0,49841</b>
SAA	<b>0,50439</b>	<b>0,13872</b>	<b>0,11078</b>	<b>0,37377</b>	<b>0,30169</b>	<b>0,16666</b>	<b>0,13456</b>	<b>0,11288</b>	<b>0,18456</b>
SLO	<b>0,60467</b>	<b>0,16367</b>	<b>0,16593</b>	<b>0,46706</b>	0,03741	<b>0,47725</b>	<b>0,17151</b>	<b>0,13913</b>	<b>0,47009</b>
SRD	<b>0,59411</b>	<b>0,3133</b>	<b>0,26542</b>	<b>0,57491</b>	<b>0,75291</b>	<b>0,13441</b>	<b>0,28425</b>	<b>0,33206</b>	<b>0,11124</b>
TAT	<b>0,3518</b>	0,00223	0,01102	<b>0,21615</b>	<b>0,092</b>	<b>0,13743</b>	0,00734	0,00244	<b>0,17063</b>
TUR	<b>0,40001</b>	0,00317	0,00986	<b>0,25891</b>	<b>0,10572</b>	<b>0,1494</b>	0,01179	0,00572	<b>0,18474</b>
UKR	<b>0,54835</b>	<b>0,10711</b>	<b>0,06681</b>	<b>0,34821</b>	<b>0,14995</b>	<b>0,31754</b>	<b>0,10803</b>	0,037	<b>0,35776</b>

12 lentelė. TĖSINYS. Porinių skirtumų (žemiau įstrižainės) ir  $\Phi_{ST}$  (aukščiau įstrižainės) atstumų matricos, pagal mtDNR haplotipai (577–16 023 np) pasiskirstymą Europos populiacijose. Trumpiniai: AKZ – aškenaziai, ARM – Armėnija, AZR – Azerbaidžanas, BUY – Baltarusija, CZH – Čekija, ESP – Ispanija, GEO – Gruzija, IRN – Iranas, ITA – Italija, LIT – Lietuva, PLN – Lenkija, RUS – Rusija, SAA – samiai, SLO – Slovakija, SRD – Sardinija, TAT – Totoriai, TUR – Turkija, UKR – Ukraina. Paryškinti statistiškai reikšmingi ( $p < 0,05$ ; 10 000 kėlinių) įverčiai

	LIT	PLN	RUS	SAA	SLO	SRD	TAT	TUR	UKR
AKZ	<b>0,00157</b>	<b>0,00768</b>	<b>0,00479</b>	<b>0,05448</b>	0,00021	<b>0,04843</b>	<b>0,00134</b>	<b>0,00381</b>	0,00022
ARM	0,0014	<b>0,00765</b>	<b>0,00469</b>	<b>0,05642</b>	0	<b>0,04977</b>	0,00116	0,00369	0
AZR	0,0014	<b>0,00765</b>	<b>0,00469</b>	<b>0,05642</b>	0	<b>0,04977</b>	0,00116	0,00369	0
BUY	<b>0,01326</b>	<b>0,02005</b>	<b>0,01723</b>	<b>0,06881</b>	0,01315	<b>0,06288</b>	<b>0,01321</b>	<b>0,01605</b>	0,01383
CZH	0,00152	0,00841	0,00515	<b>0,0641</b>	0	0,05654	0,00127	0,00405	0
ESP	0,00148	0,00813	0,00498	<b>0,0615</b>	0	<b>0,0541</b>	0,00123	0,00391	0
GEO	0,0014	<b>0,00767</b>	<b>0,0047</b>	<b>0,05666</b>	0	<b>0,04996</b>	0,00117	0,0037	0
IRN	<b>0,00251</b>	<b>0,0088</b>	<b>0,00586</b>	<b>0,0576</b>	0,00122	<b>0,05101</b>	<b>0,00229</b>	0,00484	0,00128
ITA	<b>0,00393</b>	<b>0,01039</b>	0,0074	<b>0,06068</b>	0,00274	<b>0,05385</b>	<b>0,00372</b>	0,00634	0,00289
LIT	0	<b>0,00878</b>	<b>0,00595</b>	<b>0,05439</b>	0,00148	<b>0,04885</b>	<b>0,0025</b>	<b>0,00497</b>	0,00155
PLN	<b>0,07788</b>	0	<b>0,01249</b>	<b>0,06479</b>	0,00813	<b>0,05843</b>	<b>0,00864</b>	<b>0,01137</b>	0,00855
RUS	<b>0,09034</b>	<b>0,04201</b>	0	<b>0,06263</b>	0,00498	<b>0,05595</b>	<b>0,00577</b>	<b>0,00844</b>	0,00524
SAA	<b>0,15383</b>	<b>0,20189</b>	<b>0,37653</b>	0	<b>0,0615</b>	<b>0,11466</b>	<b>0,05544</b>	<b>0,06032</b>	<b>0,0654</b>
SLO	<b>0,07535</b>	0,01402	-0,01279	<b>0,3382</b>	0	<b>0,0541</b>	0,00123	0,00391	0
SRD	<b>0,3129</b>	<b>0,44151</b>	<b>0,72082</b>	<b>0,33237</b>	<b>0,76362</b>	0	<b>0,04944</b>	<b>0,05383</b>	0,05779
TAT	<b>0,02239</b>	<b>0,0812</b>	<b>0,13773</b>	<b>0,09933</b>	<b>0,11586</b>	<b>0,25855</b>	0	<b>0,00478</b>	0,00129
TUR	<b>0,0183</b>	<b>0,08655</b>	<b>0,16777</b>	<b>0,11016</b>	<b>0,13736</b>	<b>0,30432</b>	0,00284	0	0,00411
UKR	<b>0,09891</b>	0,0336	<b>0,24015</b>	<b>0,14466</b>	<b>0,20313</b>	<b>0,54435</b>	<b>0,04779</b>	<b>0,07329</b>	0

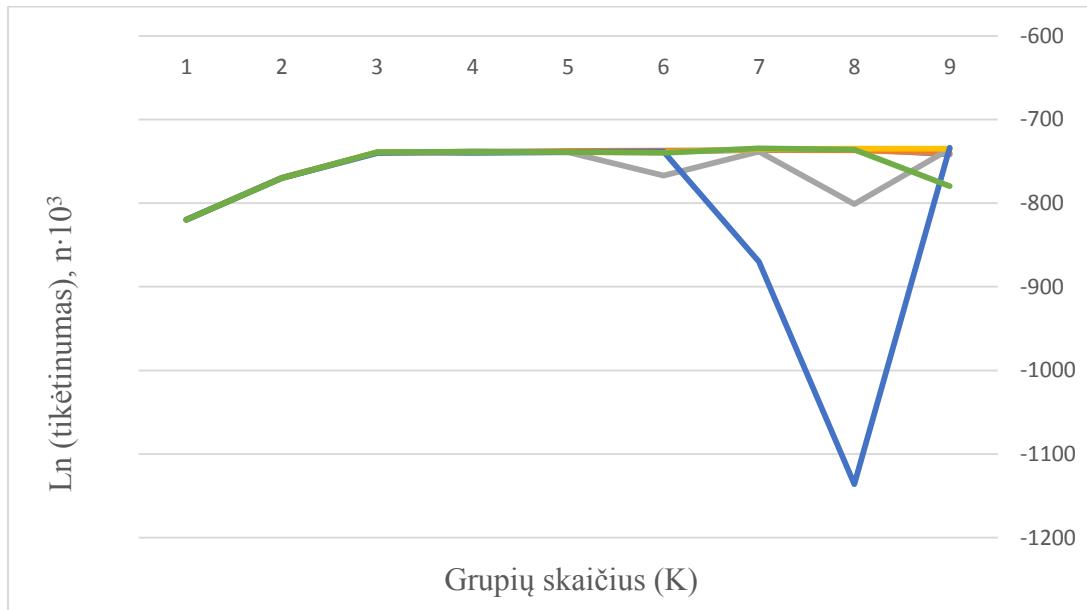
**Lietuvos populiacijos genetinė struktūra ir padėtis kitų populiacijų atžvilgiu, remiantis autosomų VNP pasiskirstymu**

13 lentelė. Naudoti plataus masto genotipavimo kokybės kontrolės įverčiai (*GenomeStudio v2011.1*)

<b>KOKYBĖS VERTINIMAS</b>		
<b>MĖGINIŲ</b>		
<b>Parametras</b>	<b>Angl. terminas</b>	<b>Pastabos</b>
Genotipų kokybės santykis	<i>Call rate</i>	Intervalas [0,97–100] (visi mėginiai); 0,97 priimtina, įvertinus kitus kriterijus, kurie atitinka geros kokybės reikalavimus
A ir B alelio intensyvumo 95 proc.	<i>p95Grn, p95Red</i>	Raudonas [3 874–9 931]; Žalias [3 002–7 751]
<i>GenCall</i> reikšmių 10 procentilė	<i>p10GC</i>	>0,7 (visi mėginiai); 0,4 LTG-1075 netinkamas
Paveldėjimo klaida	<i>P-P-C Error Rate</i>	Intervalas [0–0,07]
Mėginių grafikas	<i>Sample graph</i>	Visi mėginiai tinkami; LTG-1075 netinkamas
<b>GENOTIPUOTŲ VNP</b>		
Klasifikavimo dažnis	<i>Call frequency, Call Freq</i>	Intervalas [0,13–1]; Remiantis 1 % taisykle, visi, kurie <0,9845 yra pašalinami, pašalinti 10 262 VNP
Paveldėjimo klaida	<i>Heritability error</i>	Rep = 0 (tinkama); P-C kl.=0 (tinkama); P-P-C kl. yra, žiūrėti pagal poreikį
VNP klasterio kokybės įvertis	<i>GenTrain score</i>	Intervalas [0,35–0,98] (visi mėginiai)
Heterozigotų perteklius	<i>Het excess</i>	<(-0,3) – 31 VNP (tinkami); >0,2 – 25 VNP (tinkami); (X, XY, Y, 0 chromosomos vertintos atskirai)
VNP genotipų klasterių atsiskyrimas	<i>ClusterSep</i>	<0,27 pašalinti arba įvertinus palikti kaip tinkamus; 0,27<x<0,4 nustatyti 2 128 atvejai, peržiūrėjus, dalis koreguoti ir palikti, dalis anuliuoti
AB klasterio centro $\theta$ vertė	<i>AB T Mean</i>	<0,3 – 2 209 VNP (tinkami); <0,2 – 30 VNP (tinkami); >0,7 – 6 722 VNP (tinkami), >0,8 nėra
AB klasterio centro R vertė	<i>AB R Mean</i>	Intervalas [0,16–4,85]; <0,145 pašalinti
Retojo alelio dažnis	<i>Minor Freq</i>	MAF<0,1 autosomose – 140 839 VNP, MAF<0,001 nėra
X ir Y chr. vyrų grupėje		Y chr. – 1 372 VNP, pakoreguoti visi VNP, kurie netinkami; X chr. – 17 814 VNP, iš jų 14 036 VNP AB $\neq$ 0. Rekomenduojama analizei eliminuoti heterozigotines reikšmes. XY (PAR) – 463 VNP (tinkami)

14 lentelė. TĖSINYS. Naudoti plataus masto genotipavimo kokybės kontrolės įverčiai (*GenomeStudio v2011.1*)

GENOTIPAVIMO PROCESO		
Parametras	Angl. terminas	Pastabos
Dažymas	<i>Staining Control</i>	Intensyvumas:
		>3 000 (visi mėginiai); LTG-1075 netinkamas
		>2 000 (visi mėginiai); LTG-1075 netinkamas
Nukleotidinės bazės pratęsimas	<i>Extension Control</i>	Intensyvumas:
		>3 000 (visi mėginiai); LTG-1075 netinkamas
		>2 000 (visi mėginiai); LTG-1075 netinkamas
Taikinio pašalinimas	<i>Target Removal Control</i>	Intensyvumas:
		Vertinama tik <i>green</i> sistema
		<200 (visi mėginiai); LTG-1075 netinkamas
Hibridizacija	<i>Hybridization Control</i>	Intensyvumas:
		Vertinama tik <i>green</i> sistema
		J ~2 000; M ~3–7 000; Ž ~5–12 000; (visi mėginiai); LTG-1075 netinkamas
Atkūrimas	<i>Restoration Control</i>	Intensyvumas:
		Vertinama tik <i>green</i> sistema
		<500 (visi mėginiai); LTG-1075 netinkamas
Atitikimas, tikslumas	<i>Stringency Control</i>	Intensyvumas:
		PM > 6 000; MM < 1 000; visi mėginiai; LTG-1075 netinkamas
		Vertinama tik <i>red</i> sistema
Specifiškumas	<i>Non-Specific Binding Control</i>	Intensyvumas:
		<500 (visi mėginiai); LTG-1075 netinkamas
		<400 (visi mėginiai); LTG-1075 netinkamas
Polimorfiškumas	<i>Non-Polymorphic Control</i>	Intensyvumas:
		>3 000 (visi mėginiai); LTG-1388?; LTG-1075 netinkamas
		>3 000 (visi mėginiai); LTG-1075 netinkamas



1 pav. Lietuvos ir HapMap3 populiacijų struktūros modelių su skirtingu grupių skaičiumi ( $K = 1-9$ ) tikėtinumo pasiskirstymas

15 lentelė.  $F_{ST}$  atstumų matrica, pagal 106 545 autosomų VNP alelių pasiskirstymą (žemiau įstrižainės) tarp tirtų Lietuvos ir HapMap3 populiacijų. PK1 (PK2) vektorių dispersijos analizės p reikšmės (aukščiau įstrižainės). Paryškinti statistiškai reikšmingi įverčiai

	LIT	CEU	ASW	MKK	MEX	CHD	CHB	JPT	LWK	TSI	GIH	YRI
LIT	0	0	0	0	0	0	1,11E <sup>-16</sup>	2,22E <sup>-16</sup>	0	1,11E <sup>-16</sup>	2,22E <sup>-16</sup>	0
CEU	0,005	0	0	0	0	0	0	1,11E <sup>-16</sup>	0	0	0	0
ASW	0,061	0,058	0	0,0002 (PK2)	0 (PK2)	0	0	0	0	1,11E <sup>-16</sup>	0	0
MKK	0,066	0,062	0,013	0	0	0	0	1,11E <sup>-16</sup>	2,22E <sup>-16</sup>	0	0	0
MEX	0,029	0,025	0,058	0,063	0	0	0	0	0	9,61E <sup>-5</sup>	0,0107 (PK2)	0
CHD	0,08	0,079	0,088	0,092	0,054	0	0,155089	0,340182	0	0	0	0
CHB	0,079	0,078	0,088	0,091	0,053	0,001	0	0,00071 (PK2)	1,11E <sup>-16</sup>	0	0	0
JPT	0,081	0,08	0,089	0,093	0,054	0,008	0,007	0	0	0	0	0
LWK	0,089	0,086	0,008	0,013	0,082	0,108	0,107	0,109	0	0	1,11E <sup>-16</sup>	0
TSI	0,01	0,003	0,058	0,059	0,027	0,079	0,078	0,08	0,084	0	0	0
GIH	0,033	0,03	0,06	0,062	0,033	0,058	0,058	0,059	0,082	0,03	0	0
YRI	0,096	0,093	0,006	0,022	0,089	0,114	0,114	0,115	0,007	0,092	0,089	0

## Lietuvos populiacijos genetinė struktūra, remiantis egzomo variantų pasiskirstymu

16 lentelė. Egzomo sekoskaitos duomenų *LifeScope™ Genomic Analysis Software v2.5* programa pirminės, antrinės ir tretinės analizių nustatymai

<b>NUSTATYMAS</b>	<b>IVERTIS</b>	<b>PASTABOS</b>
<i>Analysis Assembly Name</i>	hg 19	(GRCh37 –2009-02)
<i>Annotation dbSNP File</i>	dbSNP_b132_00-All.vcf	Build132
<i>Analysis Regions</i>	TargetSeq_exome_named_targets_hg19.bed/ SureSelect_regions2.bed	
<i>Analysis Space</i>	Auto	Base/Color
<i>Analysis Platform</i>	SOLiD	
<b>SAET</b>		
<b>NUSTATYMAS</b>	<b>IVERTIS</b>	
<i>Update Quality Values</i>	True	
<i>Trusted Quality Values</i>	25	
<i>Support Votes</i>	3	
<i>Maximum Correction per Read</i>	0	
<i>K-mer size</i>	0	
<i>Genome Length</i>	30 000 000	
<i>Position of Error Inflation Point</i>	0	
<i>Disable Random Sampling for Large Data</i>	False	
<i>Trusted Frequency</i>	0	
<i>On Target Ratio</i>	0.5	
<i>Number of Recursive Runs</i>	1	
<b>Fragment Mapping</b>		
<b>NUSTATYMAS</b>	<b>IVERTIS</b>	
<i>Mapping QV Threshold</i>	0	
<i>Create Unmapped BAM Files</i>	False	
<i>Reference Weight</i>	8	
<i>Second Map Gapped Algorithm</i>	GLOBAL	
<i>Base Quality Filter Threshold</i>	10	
<i>Map in Base Space</i>	False	
<i>Add Color Sequence</i>	True	
<i>BAM soft clip</i>	False	
<b>BAM Stats</b>		
<b>NUSTATYMAS</b>	<b>IVERTIS</b>	
<i>Whether to combine Data from Both the Strands for Coverage in WIG format</i>	1	
<i>Bin Size for Coverage in WIG file format</i>	100	
<i>Maximum insert size</i>	100 000	
<i>Insert size bin</i>	100	
<i>Primary Alignments only for Coverage in WIG file format</i>	1	
<i>Maximum Coverage</i>	500 000	
<b>Small Indel</b>		
<b>NUSTATYMAS</b>	<b>IVERTIS</b>	
<i>Detail Level</i>	3	
<i>Zygoty Profile Name</i>	Max-mapping	
<i>Genomic Region</i>		
<i>Display Base Qvs</i>	False	

<i>Number Alignment per Pileup</i>	1 000
<i>Random Seed</i>	94404
<i>Min Num Evid</i>	2
<i>Max Num Evid</i>	-1
<i>ConsGroup</i>	1
<i>Max Reported Alignments</i>	-1
<i>Min Mapping Quality (MAPQ)</i>	8
<i>Min Best Mapping Quality</i>	10
<i>Min Anchor Mapping Quality</i>	-1
<i>Ungapped BAM Flag Filter</i>	ProperPair, Primary
<i>Gapped BAM Flag Filter</i>	Primary
<i>Edge Length Deletions</i>	0
<i>Edge Length Insertion</i>	0
<i>Perform Filtering</i>	True
<i>Indel Size Distribution Allowed</i>	Can-cluster
<i>Remove Singletons</i>	True
<i>Alignment Compatibility Filter</i>	1
<i>Max Coverage Ratio</i>	12
<i>Max Nonreds 4Fit</i>	2
<i>Min From End Position</i>	9.1
<i>Min Insertion Size</i>	0
<i>Min Deletion Size</i>	0
<i>Max Insertion Size</i>	1 000 000 000
<i>Max Deletion Size</i>	1 000 000 000
<b>Small Indel Annotation</b>	
<b>NUSTATYMAS</b>	<b>IVERTIS</b>
<i>Show Only Variants not in dbSNP</i>	False
<i>Show only coding variants</i>	False
<i>dbSNP concordance for dbSNP SNPs</i>	False
<i>dbSNP Indel border slack</i>	5
<i>Show Only Variants in dbSNP</i>	False
<i>dbSNP concordance for dbSNP Indels</i>	True
<i>Show only variant in genes</i>	False
<b>Enrichment</b>	
<b>NUSTATYMAS</b>	<b>IVERTIS</b>
<i>Extend Bases</i>	0
<i>Min Mapping Score</i>	8
<i>Minimum Target Overlap</i>	0.0001
<i>Minimum Target Overlap Reverse</i>	0.0001
<i>Summary Report</i>	True
<i>Target Coverage Stats</i>	True
<i>Coverage Frequency</i>	False
<i>Coverage Bedgraph</i>	False
<i>Genome Coverage Frequency</i>	False
<b>SNP Finding Annotation</b>	
<b>NUSTATYMAS</b>	<b>IVERTIS</b>
<i>Show Only Variants not in dbSNP</i>	False
<i>Show Only Coding Variants</i>	False
<i>dbSNP Concordance for dbSNP SNPs</i>	True
<i>dbSNP Indel Border Slack</i>	5
<i>Show Only Variants in dbSNP</i>	False



<i>dbSNP Concordance for dbSNP Indels</i>	False
<i>Show Only Variant in Genes</i>	False
<b>SNP Finding</b>	
<b>NUSTATYMAS</b>	<b>IVERTIS</b>
<i>Call Stringency</i>	Medium
<i>Skip High Coverage Positions (Het)</i>	False
<i>Minimum Mapping QV</i>	8
<i>Detect Adjacent SNPs</i>	False
<i>Polymorphism rate</i>	0.001
<i>Include Reads with unmapped Mate</i>	False
<i>Exclude Reads with Indels</i>	True
<i>Require Only Uniquely Mapped Reads</i>	False
<i>Ignore Reads with a Higher Mismatch Count to Alignment Length Ratio</i>	1.0
<i>Ignore Reads with a Lower Alignment Length to Read Length Ratio</i>	0.0
<i>Minimum Ratio of the Filtered Reads and Raw reads</i>	0
<i>Require alleles to be present in both strands</i>	False
<i>Minimum Base QV a Read for a Position</i>	28
<i>Minimum Color QV of a Read for a Position</i>	7
<i>Min Base QV of the non-reference allele of the position</i>	28
<i>Minimum Unique Start Positions of Less Common Allele</i>	0
<i>The Less Common Allele on Both Strands</i>	False
<i>Maximum Difference of Color QVs of the Most Common and Less Common Alleles</i>	99
<i>Minimum Average Color QV of Less Common Allele</i>	0
<i>Minimum Allele Ratio (Het)</i>	0.15
<i>Minimum Coverage (Het)</i>	2
<i>Minimum Unique Start Positions (Het)</i>	2
<i>Minimum non-Reference Color QV (Het)</i>	7
<i>Minimum non-Reference Base QV (Het)</i>	28
<i>Minimum Ratio of Valid Reads (Het)</i>	0.65
<i>Minimum Valid Tricolor Count (Het)</i>	2
<i>Minimum Coverage (Hom)</i>	1
<i>Minimum Count of the Non-Reference Allele (Hom)</i>	2
<i>Minimum Average non-Reference Base QV (Hom)</i>	28
<i>Minimum Average non-Reference Color QV (Hom)</i>	7
<i>Minimum Unique Start Position of the non-Reference allele (Hom)</i>	2
<i>Compress the Consensus File</i>	False
<i>Output Consensus File</i>	True
<i>Output FASTA File</i>	True

## APIE AUTORE

Vardas: Ingrida  
Pavardė: Uktverytė  
Gimimo data: 1985 08 29  
Darbovietės adresas: Santariškių g. 2, LT–08661 Vilnius  
Telefono nr.: +370 600 64870  
El. paštas: ingrida.uktveryte@mf.vu.lt

### ***Išsilavinimas:***

2008–2010 m. Vilniaus universitetas Gamtos mokslų fakultetas Genetikos studijų programa, biologijos magistro kvalifikacinis laipsnis.

2004–2008 m. Vilniaus universitetas Gamtos mokslų fakultetas Molekulinės biologijos studijų programa, biologijos bakalauro kvalifikacinis laipsnis.

2000–2004 m. Klaipėdos „Ažuolyno“ gimnazija.

1992–2000 m. Klaipėdos „Saulėtekio“ pagrindinė mokykla.

### ***Papildomi mokymai:***

2014 m. rugsėjo 8-11 d. kursai „Next-generation sequencing in a diagnostic setting“, Atėnai, Graikija.

2013 m. birželio 8–11 d. Europos žmogaus genetikos draugijos konferencija, Paryžius, Prancūzija.

2012 m. gruodžio 3–4 d. konferencija „Genomics in health and disease – Towards personal genomics“, Naijmegenas, Olandija.

2012 m. rugsėjo 18d. 4-asis Žmogaus identifikavimo molekulinų produktų vartotojų susitikimas, Ryga, Latvija.

2012 m. birželio 12–15 d. tarptautinė mokslinė konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, Vilnius, Lietuva.

2012 m. birželio 8–9 d. konferencija „Gene Forum 2012“, Tartu, Estija.

2012 m. gegužės 10–12 d. konferencija „11<sup>th</sup> Baltic Congress of Laboratory Medicine“, Vilnius, Lietuva.

2012 m. vasario 29 d. – kovo 1 d. kursai „5500 LifeScope Bioinformatics training“, Darmštadas, Vokietija.

2011 m. rugsėjo 5–31 d. mokslinė išvyka į Makso Planko evoliucinės antropologijos institutą, Leipcigas, Vokietija.

2011 m. birželio 13–17 d. mokslinė išvyka į Tartu universiteto Molekulinės ir ląstelės biologijos institutą, Tartu, Estija.

2011 m. gegužės 28–31 d. Europos žmogaus genetikos draugijos konferencija, Amsterdamas, Olandija.

2010 m. spalio 17 d. – lapkričio 20 d. mokslinė išvyka dalyvaujant Nacionalinės geografijos projekte Genografija į Rusijos medicinos mokslų akademijos Medicininės genetikos tyrimų centrą (RAMN), Maskva, Rusijos Federacija.

***Narystė profesinėse draugijose:***

Lietuvos žmogaus genetikos draugija,

Europos žmogaus genetikos draugija.