

VILNIAUS UNIVERSITETAS

INGRIDA DOMARKIENĖ

LIETUVOS POPULIACIJOS GENETINĖS STRUKTŪROS TYRIMAS  
REMIANTIS VIENO NUKLEOTIDO POLIMORFIZMŲ ASOCIACIJOS SU  
LIGA ANALIZE

Daktaro disertacija

Biomedicininiai mokslai, medicina (06 B)

Vilnius, 2014 metai

Disertacija rengta 2010–2014 metais Vilniaus universitete

**Mokslinis vadovas:**

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

# TURINYS

SANTRUMPŲ SĄRAŠAS.....	5
ĮVADAS .....	8
1. LITERATŪROS APŽVALGA .....	15
1.1. Daugiaveiksnių ligų genetinė architektūra ir genomika koronarinės širdies ligos pavyzdžiu .....	15
1.2. Daugiaveiksnių ligų genetinės architektūros nustatymas .....	28
1.2.1. Nuo genų kandidatų iki naujos kartos sekoskaitos. <i>GWAS</i> .....	30
1.2.2. Specifinių populiacijų tyrimo svarba.....	37
1.2.3. Genetinės KŠL architektūros tyrimai Lietuvoje .....	39
1.3. Genetinės įvairovės vertinimas ir ligos prognozė .....	41
2. TYRIMO METODIKA .....	45
2.1. Tyrimo strategija .....	45
2.2. Tyrimo imtis.....	47
2.2.1. Tiriamųjų asmenų grupė plataus masto genotipavimui. Žinomų ligos riziką lemiančių VNP genetinio profilio nustatymas bendroje Lietuvos lietuvių populiacijoje .....	47
2.2.2. Tiriamųjų asmenų grupė plataus masto genotipavimui. Bandomoji asociacijų analizė naujų genetinių sričių paieškai .....	49
2.3. Tyrimo eiga ir metodai.....	52
2.3.1. DNR išskyrimas ir DNR koncentracijos bei švarumo nustatymas.....	53
2.3.2. Genotipavimas <i>Illumina Infinium® HD</i> VNP lustais .....	53
2.3.3. Polimerazės grandininė reakcija (PGR) ir sekoskaita .....	54
2.3.4. Statistinė ir bioinformacinė duomenų ir rezultatų analizė.....	54
3. TYRIMO REZULTATAI.....	61
3.1. Genų ir VNP, susijusių su ligos genetinė architektūra, atranka.....	61
3.2. Genetinės struktūros ir įvairovės nustatymas pagal pasirinktus VNP .....	66
3.2.1. Vidupopuliacinė genetinė struktūra ir įvairovė pagal tiriamus VNP.....	66
3.2.2. Tarpopuliacinė genetinė struktūra ir įvairovė pagal tiriamus VNP.....	86
3.3. Naujų genetinių sričių, susijusių su ligos genetinė architektūra, paieška .....	97
3.3.1. Šeimų viso genomo žymenų asociacijos tyrimas .....	97
3.3.2. Potencialių genų kandidatų patvirtinimo tyrimas .....	102

4. TYRIMO REZULTATŲ APTARIMAS.....	105
4.1. Genetinė įvairovė, alelių dažnių, haplotipų skirtumai populiacijos viduje ir tarp populiacijų .....	105
4.2. Genetinės struktūros ir įvairovės pagal tirtus VNP reikšmė bendroje Lietuvos lietuvių populiacijoje .....	112
4.3. Asociacijų rezultatai ir jų reikšmė .....	113
LITERATŪROS SĄRAŠAS .....	122
PADĖKA .....	138
PRIEDAI.....	139
I PRIEDAS. DNR išskyrimo ir DNR koncentracijos bei švarumo nustatymo protokolai.....	139
II PRIEDAS. Genotipavimo <i>Illumina Infinium</i> <sup>®</sup> <i>HD</i> VNP lustais protokolas .....	148
III PRIEDAS. Polimerazės grandininės reakcijos (PGR) ir sekoskaitos protokolas .....	164
IV PRIEDAS. Darbo su <i>GenomeStudio</i> <sup>™</sup> aprašas .....	175
APIE AUTOREJĮ .....	182

## SANTRUMPŲ SĄRAŠAS

CA – Kanada (pagal tarptautinį ISO standartą 3166-1)

CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai (JAV)

CH – Šveicarija (pagal tarptautinį ISO standartą 3166-1)

CRB – C reaktyvusis baltymas

$D'$  – sankibos pusiausvyros koeficientas, santykinis nuokrypis ( $D_{AB} = p_{AB} - p_{A}p_{B}$ ,  $D' = D/D_{max}$ , kai  $D > 0$ ;  $D' = D/D_{min}$ , kai  $D < 0$ ; A ir B skirtingos genetinės sritys, p – dažnis) nuo nustatyto ir tikėtino haplotipų dažnio

ddH<sub>2</sub>O – didejonizuotas vanduo

DE – Vokietija (pagal tarptautinį ISO standartą 3166-1)

denom df – laisvės laipsnių vardiklis lyginant dvi dispersijas F testu (angl. *denominator degree of freedom*)

df – laisvės laipsniai, naudojami statistikoje (angl. *degree of freedom (df)*)

DNR – deoksiribonukleorūgštis

DTL – didelio tankio lipoproteinas

F – dviejų nepriklausomų imčių dispersijų palyginimo kriterijus

FI – Suomija (pagal tarptautinį ISO standartą 3166-1)

FIN – suomiaai Suomijoje

GBR – britai Anglijoje ir Škotijoje

gDNR – genomine DNR

JAV – Jungtinės Amerikos Valstijos

JP – Japonija (pagal tarptautinį ISO standartą 3166-1)

KMI – kūno masės indeksas

KŠL – koronarinė širdies liga

LOD – statistinis sankibos įvertis  $Z$ , dešimtainis šansų logaritmas, kai įvertinama sankibos įvykio tikimybė (angl. *logarithm of odds*)

LSMU – Lietuvos sveikatos mokslų universitetas

LT – Lietuva (pagal tarptautinį ISO standartą 3166-1)

LTU – tiriama bendra Lietuvos lietuvių populiacija, asmenys iš šešių etnolingvistinių grupių

MAF – retojo alelio dažnis (angl. *minor allele frequency*)

MF – Medicinos fakultetas

MI – miokardo infarktas

MTL – mažo tankio lipoproteinas

NTS – netransliuojama geno sritis

num df – laisvės laipsnių skaitiklis lyginant dvi dispersijas F testu (angl. *numerator degree of freedom*)

$p$  – tikimybė pagrįstai priimti arba atmesti hipotezę vadovaujantis reikšmingumo lygmeniu

PGR – polimerazės grandininė reakcija

PI – pasikliauties intervalas, naudojamas statistikoje (angl. CI – *confidence interval*)

RA – rizikos alelis

RAD – rizikos alelio dažnis

RS – rizikos santykis

RT-PGR – realaus laiko PGR

SG – Singapūras (pagal tarptautinį ISO standartą 3166-1)

SPC – sveikatos priežiūros centras

ŠS – šansų santykis

$t$  – dviejų nepriklausomų imčių vidurkių palyginimo kriterijus

TSI – toskaniečiai Italijoje

TW – Taivanas (pagal tarptautinį ISO standartą 3166-1)

UK – Jungtinė Karalystė (pagal tarptautinį ISO standartą 3166-1)

US – Jungtinės Amerikos Valstijos (pagal tarptautinį ISO standartą 3166-1)

VNP – vieno nukleotido polimorfizmas (-ai) (angl. *SNP – single nucleotide polymorphism*)

VŠĮ VUL SK MGC – Viešoji įstaiga Vilniaus universiteto ligoninės Santariškių klinikų Medicininės genetikos centras

VU – Vilniaus universitetas

ŽMGK – Žmogaus ir medicininės genetikos katedra

$\chi^2$  – Pirsono arba Fišerio (pagal poreikį) chi kvadratu kriterijus nepriklausomumo hipotezei tikrinti

## ĮVADAS

### **Tiriamos problemos pagrindimas, darbo aktualumas, naujumas, reikšmė**

Disertacijos pavadinimas „Lietuvos populiacijos genetinės struktūros tyrimas remiantis vieno nukleotido polimorfizmų asociacijos su liga analize“ yra sudėtinis, apimantis net tris svarbias sritis, kurios kiekviena savaip vertos atskiro veikalo. Tai populiacijos genetinė struktūra, genetiniai asociacijų tyrimai ir liga. Disertacijoje šios trys sritys sujungiamos, kad būtų nustatyti svarbūs, išskirtiniai, palyginti su kitomis populiacijomis, ir reikšmingi sveikatai mūsų – Lietuvos populiacijos – genominiai pasirinktos ligos (koronarinės širdies ligos – KŠL) žymenys, kuriuos būtų galima taikyti praktinėje medicinoje ir toliau plėtoti išsamius mokslinius tyrimus. Koronarinė širdies liga yra pasirinkta kaip daugiaveiksnės ligos modelis, tad taikant panašią tyrimo strategiją būtų galima atlikti populiacijos genetinio profilio ir indėlio į ligos genetinę architektūrą įvertinimą kone kiekvienos daugiaveiksnės ligos kontekste.

Genetinė populiacijos struktūra ir įvairovė, genetinis profilis yra **genetinės ligos architektūros** tyrimų objektai. Genetinė ligos architektūra žvelgiant holistiškai yra genotipo ir fenotipo ryšio struktūra. Ji apima fenotipinės įvairovės genetinę prigimtį – genetines sritis ir jose esančius alelius, bei nagrinėja jų charakteristiką – alelių skaičių genetinėse srityse, poveikio fenotipui stiprumą, pleiotropijos tendencijas, adityvumą, dominavimą, epistazę, epigenetinius poveikius ir t. t. Genetinė ligos architektūra taip pat gali būti alelių pasiskirstymo laike ir erdvėje apibūdinimas, t. y. evoliucinė ir biogeografinė fenotipinių skirtumų prigimtis remiantis alelių įvairovės tendencijomis. Apibūdinti genetinę ligos architektūrą reiškia išsiaiškinti genetines fenotipo įvairovės dedamąsias, t. y. kaip molekulinę struktūrinių ir funkcinių vienetų įvairovė įgauna konkretų fenotipinį požymį ir jam būdingą variaciją. Tai buvo ir yra viena iš išskirtinių bei pamatinių genetikos mokslo problemų. **Genetinis profilis** ir jo nustatymas yra viena iš genetinės ligos architektūros dedamųjų. Siaurąja prasme gali būti



suprantamas kaip tiesioginis, techninis alelio(-ių) ir (ar) genotipo(-ų), jų derinio nustatymas genotipuojant. Plačiaja prasme genetinis profilis – tai nustatytų būdingų alelių ir (ar) genotipų bei jų derinių dažnio ir įvairovės analizė. Genetinis profilis gali būti tiriamas siekiant identifikuoti asmenį, rūšį, įvertinti giminybę, genetinį panašumą ir skirtumą arba siekiant nustatyti ligą lemiančių žinomų ir (ar) nežinomų genetinių veiksnių kokybinę ir (ar) kiekybinę sudėtį. Dažniausiai ligos genetinio profilio sąvoka yra suprantama ne mendelinių, o daugiaveiksnių ligų kontekste, todėl yra tariama, kad genetinį profilį sudaro **polinkį** ligai lemiantys genetiniai veiksniai.

Šiuo metu jau yra atlikta daug plataus masto genomo žymenų **asociacijų tyrimų** bendroje žmonių populiacijoje daugiaveiksnių ligų ir požymių genomikos srityje. Tarp tokių įsimintinų genomikos mokslo projektų kaip Žmogaus genomo projektas, HapMap projektas, 1000 genomų projektas, genetiniai asociacijų tyrimai yra vieni iš svarbiausių, padarę ne vieną svarbų atradimą daugiaveiksnių ligų genetinėje architektūroje. Tačiau dauguma šių tyrimų buvo atliekami vadinamosiose mišriose populiacijose (angl. *admixture*). Tai leido nustatyti, kas yra bendra žmonių populiacijai tiriamų požymių genetikoje. Tačiau ar nėra praleidžiama tai, kas yra specifiška atskiroms populiacijoms? Daugėja publikacijų, akcentuojančių atskirų populiacijų tyrimo poreikį genetinių ligų sampratai gilinti. Be to, remiantis asociacijos tyrimų duomenimis pagal tam tikrus genomo žymenis siūloma didelė įvairovė komercinių testų kai kurių daugiaveiksnių ligų prognozei. Tačiau genomo asociacijų tyrimų rezultatai nėra iki galo suprasti ir siekiant juos pritaikyti medicinos praktikai dar būtina atlikti papildomus tyrimus gilinant populiacijų genomikos žinias ir tobulinti rizikos skaičiavimo algoritmus. Todėl šiuos komercinius testus naudoti klinikinėje praktikoje nerekomenduojama.

Koronarinė širdies liga (KŠL) yra daugiaveiksnė. Sergamumas, mirštamumas ir gydymo išlaidos tiek Lietuvoje, tiek pasaulyje kelia specialistų susirūpinimą. Yra žinoma daug KŠL lemiančių įvairių rizikos veiksnių, tarp jų svarbų vaidmenį šios ligos etiopatogenezėje atlieka genominiai veiksniai, kurių

veikimo mechanizmai vis dar yra neaiškūs. Taigi papildomų tyrimų poreikis išlieka.

Kalbant apie KŠL plataus masto genomo tyrimus Lietuvos populiacijos kontekste, tokių mūsų šalyje nerasime. Šis darbas yra pirmasis tokio pobūdžio. Koronarinės širdies ligos genetiniai tyrimai Lietuvoje yra pavieniai ir fragmentiški, trūksta sistemiškesnio genetinės architektūros įvertinimo. Šio darbo pavyzdžiu galėtų būti atliekami ir kitų daugiaveiksnių ligų genetinės architektūros tyrimai.

Darbas yra aktualus ne tik tuo, kad padės pagrindą tolesniems daugiaveiksnių ligų architektūros tyrimams Lietuvoje ir pasaulyje, tačiau gali būti pritaikomas ir praktikoje, pradedant KŠL rizikos vertinimo algoritmo, baigiant KŠL diagnostikos, prognozės vertinimo ir prevencijos sistemos kūrimu. Šis tyrimas yra reikšmingas savo naujumu ne tik fundamentaliajam, bet ir taikomajam mokslui.

### **Tikslas, uždaviniai, ginamieji teiginiai**

Šio tyrimo **tikslas** – daugiaveiksnės ligos genetinės architektūros, naudojant dažnus vieno nukleotido polimorfizmo (VNP) genomo žymenis, analizė bendroje Lietuvos populiacijoje.

Tiksliui įgyvendinti iškelti tyrimo **uždaviniai**:

1. Tyrimo atlikimui pasirinkti daugiaveiksne ligą kaip modelį.
2. Vadovaujantis literatūros duomenimis ir analizuojant duomenų bazes sudaryti žinomų, su daugiaveiksne liga asocijuotų bendrų VNP rizikos alelių ir siejamų genų sąrašą pasirinktos daugiaveiksnės ligos genetinei įvairovei nustatyti ir genetinei architektūrai tirti bendroje Lietuvos lietuvių populiacijoje.
3. Atlikti atrinktų žinomų VNP rizikos alelių genetinės įvairovės analizę bendroje Lietuvos lietuvių populiacijoje, vidupopuliacinį ir tarppopuliacinį palyginimus:

- a. Nustatyti dažnų VNP, siejamų su liga, genetinę struktūrą ir įvairovę panaudojant plataus masto genotipavimo duomenis;
  - b. Atlikti dažnų rizikos alelių bei jų derinių paplitimo ir paveldimo polinkio pagal daugiaveiksnį modelį ypatumų analizę.
4. Nustatyti naujus daugiaveiksnės ligos rizikos alelius, atliekant bandomąjį plataus masto genomo asociacijų su liga tyrimą Lietuvos populiacijoje:
- a. Atlikti plataus masto genomo asociacijų analizę triadų, kuriose yra sergantis probandas, grupėje;
  - b. Nustačius potencialius genus kandidatus, atlikti jų patvirtinimą – papildomą konkrečių genetinių sričių asociacijos analizę atvejo ir kontrolės tiriamųjų grupėje.

Viena iš svarbiausių genetikoje yra genetinės architektūros mįslė, kurią pasaulio mokslas bando įminti. Siekiant nustatyti naujus genetinės architektūros vienetus, neišvengiamas yra ir jau žinomų genetinių vienetų tyrimas tiek bendroje žmonių populiacijoje, tiek etninėse populiacijose. Šiam tikslui pasiekti šio tyrimo metu sugalvota strategija, darbo modelis, kuris gali būti taikomas daugumai kitų daugiaveiksnų ligų tirti. Modelio atraminiai taškai yra žinomų genetinių sričių, siejamų su liga, analizė bendroje Lietuvos lietuvių populiacijoje ir naujų su liga siejamų genetinių sričių paieška. Šiuo būdu siekiama prisidėti prie daugiaveiksnų ligų genetinės architektūros žinių.

Atlikus šį tyrimą suformuluoti **ginamieji teiginiai**:

1. Pagal tirtus genetinius koronarinės širdies ligos rizikos žymenis (VNP) bendra Lietuvos lietuvių populiacija nuo kitų populiacijų, su kuriomis palyginta, statistiškai reikšmingai skyrėsi skirtingais variantais ir jų alelių dažnių deriniais, tuo įrodant, kad skirtingoms populiacijoms yra būdingi tiek bendri, tiek unikalūs genetinės architektūros vienetai.
2. Bendros Lietuvos lietuvių populiacijos genetinę struktūrą ir įvairovę pagal tirtus genetinius koronarinės širdies ligos rizikos žymenis (VNP)

galima orientuoti gradientiškai Europos Šiaurės–Pietų kryptimi, tai paaiškinant geografinių platumų lemiamą klimato ir ekologijos įtaką genofondui evoliucijos eigoje.

3. Pagal tirtų genetinių koronarinės širdies ligos žymenų (VNP) rizikos alelių skaičių, tenkantį asmeniui, tirtoje bendroje Lietuvos lietuvių populiacijoje daugiausia asmenų turi vidutinę riziką susirgti koronarine širdies liga.
4. *ITPR2* ir *FBXL17* genai yra galimi genai kandidatai, dalyvaujantys KŠL ir aterosklerozės patogenezėje.

### **Darbo aprobacija**

Darbo rezultatai yra paskelbti recenzuojamuose mokslo leidiniuose, pristatyti Lietuvos ir užsienio mokslinėse konferencijose.

### **Publikacijos disertacijos tema (2)**

1. **Pepalytė I**, Kučinskienė ZA, Grigalionienė K, Petrulionienė Ž, Dženkevičiūtė V, Bagdonaitė L, Kučinskas V. Genetic variants that participate in oxidation processes and/or oxidative stress and are associated with atherosclerosis. *European Medical, Health and Pharmaceutical Journal*. 2012; 3: 13–16. ISSN 1804-5804.
2. **Domarkienė I**, Pranculis A, Germanas S, Jakaitienė A, Vitkus D, Dženkevičiūtė V, Kučinskienė Z, Kučinskas V. *RTN4* and *FBXL17* Genes are Associated with Coronary Heart Disease in Genome-Wide Association Analysis of Lithuanian Families. *Balkan Journal of Molecular Genetics*. 2013; 16(2): 17–22. doi: 10.2478/bjmg-2013-0026.

### **Žodiniai pranešimai disertacijos tema (3)**

1. **Pepalytė I**, Pranculis A, Dženkevičiūtė V, Bagdonaitė L, Kučinskas V, Kučinskienė ZA. Defining coronary heart disease genomics in the Lithuanian population. 11-asis Baltijos šalių laboratorinės medicinos kongresas, 2012, Vilnius, Lietuva.

2. Pranculis A, **Pepalytė I**, Arasimavičius J, Kučinskienė ZA, Kučinskas V. Pharmacogenomics of CVD treatment in patients from the Lithuanian population. Tarptautinė konferencija “Evolutionary medicine: new solutions for old problems”, 2012, Vilnius, Lietuva.
3. **Domarkienė I**, Pranculis A, Germanas Š, Jakaitienė A, Dženkevičiūtė V, Kučinskas V, Kučinskienė ZA. Are we ready to incorporate genes in risk estimation? Lithuanian way towards genomics of CHD. Šeštasis Baltijos šalių aterosklerozės kongresas, 2013, Ryga, Latvija.

### **Stendiniai pranešimai ir tezės, pateiktos konferencijose (7)**

1. **Domarkienė I**, Pranculis A, Germanas Š, Jakaitienė A, Dženkevičiūtė V, Kučinskas V, Kučinskienė ZA. Lithuanian way towards genomics of coronary heart disease. Šeštasis baltijos šalių aterosklerozės kongresas, 2013, Ryga, Latvija.
2. Pranculis A\*, **Pepalytė I\***, Dženkevičiūtė V, Bagdonaitė L, Kučinskas V, Kučinskienė ZA. A family based genomics and pharmacogenomics study of coronary artery disease in the patient group from the Lithuanian population. 11-asis Baltijos šalių laboratorinės medicinos kongresas, 2012, Vilnius, Lietuva.  
\*Autorių indėlis vienodas.
3. Kučinskas V, **Pepalytė I**, Pranculis A, Bagdonaitė L, Dženkevičiūtė V, Kučinskienė ZA. Coronary artery disease genomics – CNV findings in the group of patients from Lithuania. Amerikos žmogaus genetikos draugijos konferencija, 2012, San Franciskas, JAV.
4. Pranculis A, **Pepalytė I**, Arasimavičius J, Kučinskienė ZA, Kučinskas V. Pharmacogenomics of cardiovascular disease treatment in subjects from the Lithuanian population. Amerikos žmogaus genetikos draugijos konferencija, 2012, San Franciskas, JAV.
5. **Pepalytė I**, Uktverytė I, Dirsė V, Kučinskas V. Characteristics of the genomic structural variation in the Lithuanian population. Tarptautinė

konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, 2012, Vilnius, Lietuva.

6. Pranculis A, **Pepalytė I**, Arasimavičius J, Kučinskienė ZA, Kučinskas V. Pharmacogenomics of CVD treatment in patients from the Lithuanian population. Tarptautinė konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, 2012, Vilnius, Lietuva.
7. **Pepalytė I**, Grigalionienė K, Petrulionienė Ž, Bagdonaitė L, Kučinskienė ZA, Kučinskas V. Common genetic variants in *ITGA2*, *IL1B*, *ALOX5AP*, *OR13G1*, *MMP9* genes that participate in oxidative stress and influence atherosclerosis phenotype formation. Europos žmogaus genetikos draugijos konferencija, 2011, Amsterdamas, Nyderlandai.

## 1. LITERATŪROS APŽVALGA

### 1.1. Daugiaveiksnių ligų genetinė architektūra ir genomika koronarinės širdies ligos pavyzdžiu

„Biologijoje viskas prasmę įgyja tik evoliucijos šviesoje“

Teodosij Dobžanskij

#### **Daugiaveiksnės ligos samprata**

Mokslinėje literatūroje teigiama, kad, skirtingai nei monogeninių arba dar vadinamų mendelinių ligų atveju, kurias dažniausiai lemia vieno geno mutacijos, daugiaveiksnės ligos yra genetinių ir aplinkos sąveikos išdava. Tai Alzheimerio (*Alzheimer*) liga, skleroderma, astma, Parkinsono (*Parkinson*) liga, išsėtinė sklerozė, osteoporozė, jungiamojo audinio, inkstų, autoimuninės ligos ir daugelis kitų (*Hunter DJ, 2005*), iš jų aterosklerozė ir koronarinė širdies liga (KŠL). Daugiaveiksnės ligos nors ir būdingos šeimoms, nepaklūsta nė vienam žinomam mendeliniam paveldėjimo dėsniumui, be to, daugiaveiksnės ligos yra dažnesnės negu retos mendelinės. Tačiau tiriant daugiaveiksnes ligas reikia nepamiršti reiškinių, kurie lemia nukrypimus nuo Mendelio dėsnų: sumažėjęs penetrantiškumas, raiškos įvairavimas, fenotipo apibrėžtis (tą pačią ligą gali lemti ne vienas ir tas pats genas), poligeniniai požymiai, geno ir geno sąveika, geno ir aplinkos sąveika (*Hern LM, Bidichandani SI, 2004*). Daugiaveiksnių ligų atveju genai lemia ne pačią ligą, o **polinkį sirgti liga**, ligos išsivystymas priklauso nuo aplinkos, kurioje gyvenama, ir nuo gyvenimo būdo. Tačiau nėra taip, kad genai arba aplinka veiktų izoliuotai. Genai, lemiantys polinkį, gali būti specifiskai modifikuoti aplinkos, nes genų produktai sąveikauja su aplinkos poveikio produktais molekulinio lygmeniu (*Dempfle A et al., 2008*). Daugiaveiksniškumas šių ligų atveju atskleidžiamas pirmiausia per klinikinės raiškos įvairovę – skirtingas ligos pradžios amžius, skirtinga ligos trukmė, skirtingas atsakas į gydymą, daugybė simptomų, kurie yra labai įvairūs ir skiriasi tarp asmenų. Antra, nors

genetinių veiksnių poveikis tiriamam požymiui yra akivaizdus, tačiau pasireiškimui būtinas ne vieno, o daugelio genų, kurių kiekvieno poveikis vidutiniškas arba mažas, veikimas (kiekvieno) atskirai arba kartu. Veikiančiuose genuose gali būti daug retų ar dažnų / bendrų variantų (angl. *common variants*), kurie ir daro įtaką požymio pasireiškimui. Trečia, veikiant genetiniams veiksniams, kartu veikia ir negenetiniai veiksniai, pavyzdžiui, mityba. Tad tuos pačius požymius dažniausiai vienu metu veikia skirtingos prigimties veiksniai. Tai lemia kiekvieno ligos požymio unikalumą ir sudėtingumą (Rodén D, 2009).

Daugiaveiksnės ligos – **sudėtingos, nes yra kompleksinės**, t. y. dėl daugelio jas lemiančių veiksnių (kurių kiekvieno indėlis į ligą yra vidutinis arba mažas) nustatyti ligos etiopatogenezės mechanizmą yra sunkiau, o diagnostikos, gydymo, prognozės ir prevencijos sistemos sukūrimas tampa iššūkiu.

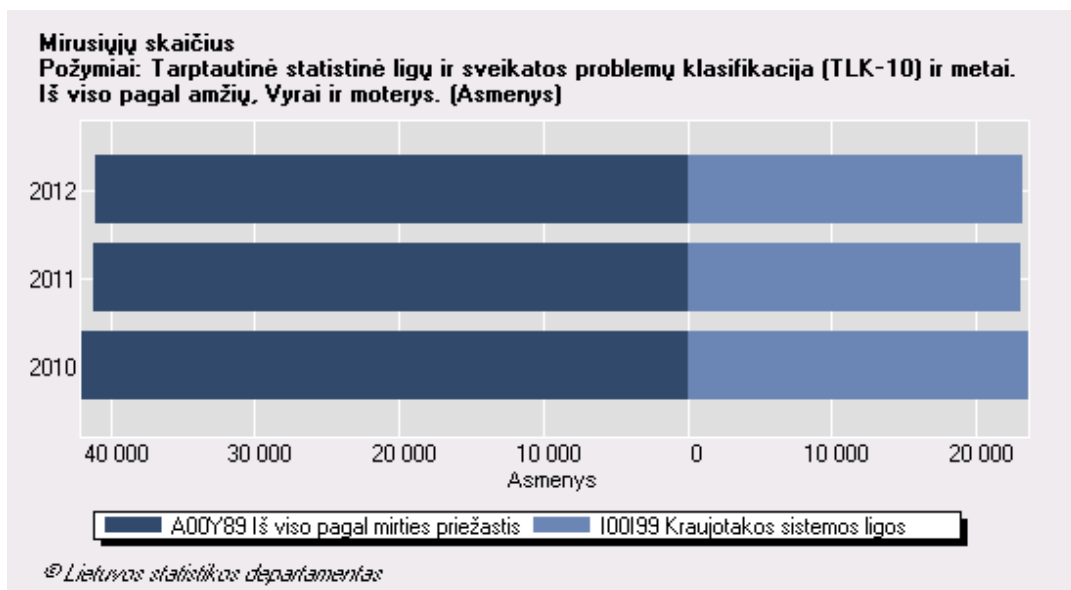
### **Koronarinė širdies liga – daugiaveiksnė**

Koronarinė širdies liga yra lemiamą daugelio veiksnių. Pamatinė KŠL lemianti patologija yra aterosklerozė, prasidedanti dar vaikystėje, vėliau ji komplikuojasi trombo susidarymu, kuris trukdo kraujo tiekimui į širdį ir sukelia išemiją bei miokardo infarktą (MI) (NCBI duomenų bazės enciklopedija *MedlinePlus*). Dėl svaraus aterosklerozės indėlio KŠL išsivystymui kartais yra sudėtinga išgryninti rizikos veiksnius, kurie šią ligą lemia tiesiogiai, o kurie netiesiogiai, pirmiausia veikdami aterosklerozės procesą.

Rizikos veiksnius būtina nustatyti, rasti tinkamus jų žymenis, nes rizikos veiksniai yra kritiniai nulemiant, žmogus sirgs ar ne. Rizikos veiksniai daro įtaką ne tik ligos vystymuisi, bet ir jos mastui. Pažymėtina, kad Lietuvos statistikos departamento analizės duomenimis (pateikti Higienos instituto Sveikatos informacijos centro duomenys), Lietuvoje, 2010–2012 metais kraujotakos sistemos ligos (atskirai KŠL statistinių duomenų nėra) sudarė



daugiau negu pusę visų vyrų ir moterų mirčių atvejų (**1.1 paveikslas**). Deja, ši tendencija nesikeičia. Remiantis 2012 metų Europos kardiologų draugijos parengtomis širdies ir kraujagyslių ligų prevencijos klinikinėje praktikoje rekomendacijomis, Lietuva kartu su Armėnija, Azerbaidžanu, Baltarusija, Bulgarija, Gruzija, Kazachstanu, Kirgizija, Latvija, Makedonija, Moldova, Rusija, Ukraina ir Uzbekija nurodoma kaip šalis, kuriai būdinga didelė rizika susirgti širdies ir kraujagyslių ligomis (*Perk J et al., 2012*).



**1.1 paveikslas.** 2010–2012 metais mirusiųjų dėl kraujotakos sistemos ligų proporcija visų mirusiųjų grupėje (vyrų ir moterų kartu). A00Y89 – kodas, iš viso pagal mirties priežastis mirusieji; I00I99 – kodas, mirusieji nuo kraujotakos sistemos ligų.

Sergamumas ir mirštamumas nuo KŠL pasaulyje, šios ligos masiškumo poveikis ekonomikai yra svarūs argumentai kurti naujas ligos diagnostikos, gydymo ir prevencijos strategijas. Norint keisti sergamumo ir mirtingumo statistiką, būtina paveikti rizikos veiksnių pasireiškimą (jeigu tai įmanoma). Todėl privalu nustatyti ir tirti rizikos veiksnius. Jau minėta, kad KŠL yra daugiaveiksni, sudėtinga liga, kurios veiksnius nėra paprasta identifikuoti, tačiau yra žinomi tiek **rizikos veiksniai**, lemiantys ligos išsivystymą, tiek šių veiksnių žymenys. Rizikos veiksnius (ir jų žymenis) galima suskirstyti į grupes pagal koreguojamumą: **modifikuojami ir**

**nemodifikuojami**; bei pagal naujumą: **tradiciniai ir netradiciniai** rizikos veiksniai.

Veiksniams, kurių įtaką KŠL pasireiškimui galima pakeisti, priklauso modifikuojami rizikos veiksniai: rūkymas, dislipidemija, cukrinis diabetas, nutukimas, netinkama mityba, trombogeniniai veiksniai, fizinis pasyvumas, perteklinis alkoholio vartojimas, metabolinis sindromas, stresas. Veiksniams, kurių įtakos KŠL pasireiškimui negalima pakeisti (tačiau galbūt ateityje bus galima), priklauso nemodifikuojami rizikos veiksniai, kuriais laikytini amžius (vyrams rizikos amžius yra  $\geq 45$  metai, moterims  $\geq 55$ ), lytis (moterims dėl estrogenų rizika yra mažesnė, tačiau tai priklauso nuo amžiaus), asmeninė ir šeimos KŠL istorija (esant KŠL įvykių šeimoje, rizika yra beveik du kartus didesnė), etninė priklausomybė (Afrikos ir Azijos protėvių populiacijos turi didesnę riziką susirgti širdies ir kraujagyslių ligomis), **genetiniai** veiksniai (*Boudi B, 2013*). Minėtų modifikuojamų veiksmų pasireiškimo įvairovės, nemodifikuojamų veiksmų – lyties, asmeninės ir šeimos KŠL istorijos bei etninės priklausomybės prigimtis taip pat yra nulemta genetinių veiksmų. Taigi genetinių veiksmų svarba yra nepaneigiama. Kalbant apie modifikuojamus ir nemodifikuojamus rizikos veiksmus būtina pabrėžti, kad riba tarp jų yra reliatyvi ir gali keistis. Galbūt netolimoje ateityje bus galima paveikti su lytimi susijusius ar genetinius rizikos veiksmus. Puikus pavyzdys – kaip nustačius genetinius ligos išsivystymo mechanizmus yra gydoma fenilketonurija, įgimta hipotirozė, mukopolisacharidozė, leukemijos ar Diušeno raumenų distrofija ir t. t. Tokių ligų, anksčiau laikytų nepagydomomis, tačiau dabar jau gydomų, sąrašas vis ilgėja. Viskas priklauso nuo to, kiek mes žinome apie ligą, ją lemiančius rizikos veiksmus bei patogenezės mechanizmus ir kiek tas žinias gebame pritaikyti. Genetiniai tyrimai gali prisidėti prie žinių gilinimo, taikymo ir sergamumo situacijos gerinimo.

Koronarinę širdies ligą lemiantys rizikos veiksniai yra tiriami jau daugelį metų, nemažai jų nustatyta dar praėjusiame amžiuje, tačiau vystantis

mokslui, nustatoma ir sužinoma vis naujų. Pagal naujumą rizikos veiksniai yra skiriami į tradicinius ir netradicinius. Tačiau, prognozuojant širdies ir kraujagyslių ligas, kol kas labiausiai pasitarnauja tradiciniai rizikos veiksniai, nes naujai nustatomų žymenų poveikis rizikai yra vidutinis. Šiuo metu geriausiai 10 metų rizikos prognozę aprašo amžius, lytis, etninė priklausomybė, bendras cholesterolio kiekis, DTL-cholesterolis, kraujospūdis, cukrinis diabetas, momentinis rūkymo statusas. Informatyviausi žymenys, geriausiai prognozuojantys mirštamumą: B-tipo natriurezinis peptidas (BNP), C reaktyvusis baltymas (CRB), homocisteinas, reninas, albumino ir kreatinino kiekių santykis šlapime. Informatyviausi žymenys, geriausiai prognozuojantys širdies ir kraujagyslių ligos pasireiškimo atvejus, – albumino ir kreatinino kiekių santykis šlapime ir BNP (*Boudi B, 2013*). Pastaruoju metu vis daugėja nustatomų naujų rizikos veiksnių ir netradicinių žymenų.

### **Genetikos ir genomikos vaidmuo koronarinės širdies ligos pavyzdžiu**

Koronarinė širdies liga, taip pat ir aterosklerozė, kaip minėta, yra daugiaveiksnės ligos, vadinasi, pasireiškia dėl aplinkos ir genetinių veiksnių sąveikos. Buvo manoma, kad apie 50 % polinkio sirgti KŠL yra dėl genetinių veiksnių polimorfizmų, veikiančių tiek žinomus, tiek nežinomus rizikos veiksnius, pavyzdžiui, cholesterolio kiekį ar kraujospūdį (*Chan L, Boerwinkle E, 1994*). Tačiau plataus masto genomo asociacijų tyrimais didelėse populiacijų grupėse nustatyta, kad genetiniais veiksniais gali būti paaiškinama nuo 20 % iki 50 % daugiaveiksnės ligos požymio įvairovės (*Visscher PM et al., 2012 (1)*). Yra teigiama, kad dominuojantys priešlaikinės aterosklerozės rizikos veiksniai yra genetiniai (*Genest JJ et al., 1992*). Verta paminėti, kad JAV daugiau negu vienas iš trijų asmenų serga viena ar daugiau širdies ir kraujagyslių ligų, o beveik pusė šių asmenų yra jaunesni negu 60 metų amžiaus (*Go AS et al, 2013*). Taigi ankstyvas ligos pradžios amžius rodo, kad nemažą įtaką sergamumui gali turėti genetiniai veiksniai.

Kai kurių mokslininkų teigimu, aterosklerozė ir jos pasekmės yra išvengiamos, gali būti sumažintos ir netgi visiškai pašalintos (*Wald NJ, Law*

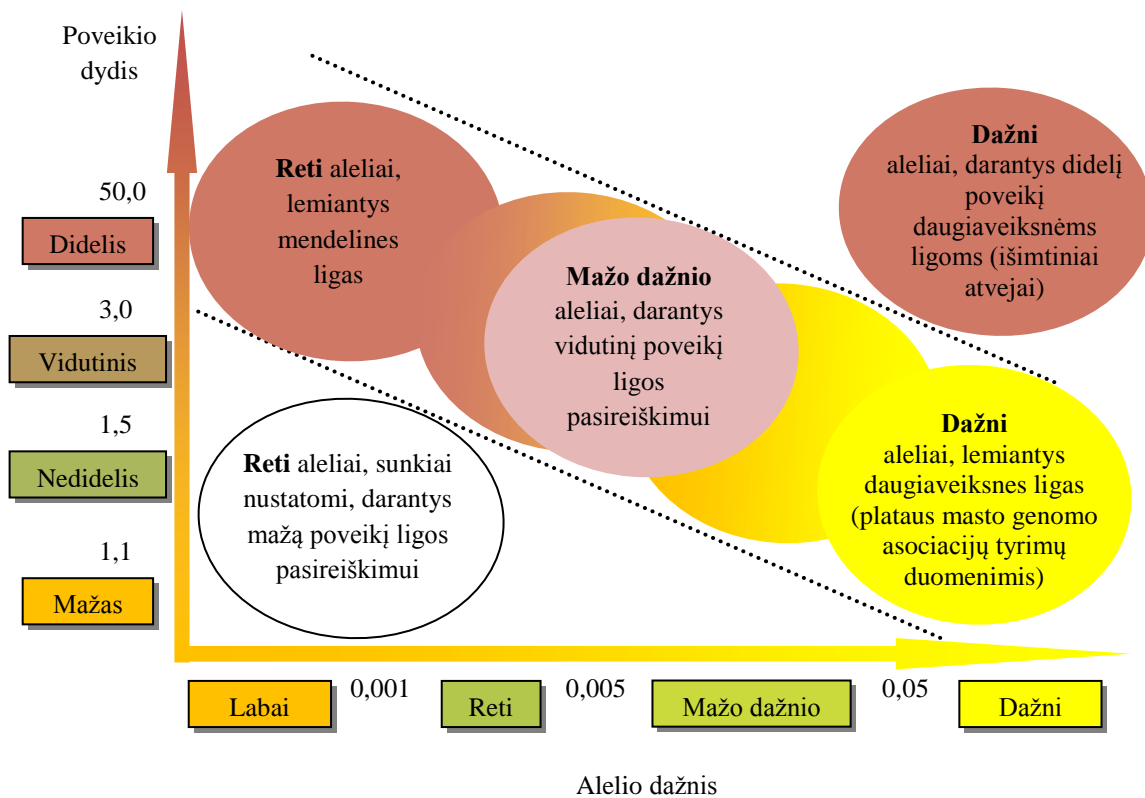
MR, 2003). Tam reikia apibrėžti ne tik aplinkos, bet ir genetinius bei genominius veiksnius. Genų ir individualių genetinių variantų, lemiančių koronarinę aterosklerozę, nustatymas sukurs individualizuotos medicinos ir optimalios ligos prevencijos bei gydymo prielaidas (Roberts R, Stewart AFR, 2006). Taigi svarbiausias širdies ir kraujagyslių ligų genomikos ir genetikos uždavinys – nustatyti tuos genomo variantus, kurie turi įtakos individualiam ligos pasireiškimui.

### Genetinė ligos architektūra

Genetinė ligos architektūra abstrakčiai yra genotipo ir fenotipo ryšio struktūra. Ja bandoma paaiškinti ligos fenotipinės įvairovės genetinę prigimtį. Genetinės ligos yra monogeninės (mendelinės), poligeninės, daugiaveiksnės. Yra žinoma daugiau negu 2000 mendelinių ligų, kurias lemia retos didelį poveikį ligos pasireiškimui darančios mutacijos (jos taip pat yra genomo variantai), kurių dažnis populiacijoje yra mažesnis nei 1 %, nes evoliucijos metu jas veikia neigiama gamtinė atranka. Kad ir kaip būtų, retų genomo variantų poveikis daugiaveiksnėms ligoms ir požymiams nėra iki galo suprastas. Teigiama, kad daugiaveiksnius požymius lemia  **tiek dažni, tiek reti variantai kartu su aplinkos veiksniais**, o dėl vėlyvo ligos pasireiškimo reti variantai išvengia neigiamos gamtinės atrankos (Reich DE, Lander ES, 2001).

Kalbant apie daugiaveiksnių ligų genetinę architektūrą KŠL atveju, vis daugėja nustatomų KŠL lemiančių genų, kuriuose randami tiek reti, tiek dažni / bendri variantai, kurių dažnis  $\geq 5\%$  (Wang WY et al., 2005). Pavyzdys galėtų būti mažo tankio lipoproteinų (MTL) receptorių koduojantis *LDLR* genas, kurio reti variantai lemia šeiminę hipercholesterolemiją, o dažni variantai, nustatyti asociacijų tyrimais, yra siejami su padidėjusiu MTL-cholesterolio kiekiu ir MI rizika bendroje populiacijoje (Willer CJ et al., 2008). Dažnų variantų poveikis ligos pasireiškimui neretai yra mažesnis negu retų variantų, o šansų santykio (ŠS) reikšmės yra apie du, tačiau bendrapopuliacinę ligos riziką gali paveikti labiau dėl didesnio variantų dažnio, palyginti su retais variantais, lemiančiais mendelines ligas. Reti variantai yra nustatomi naujos

kartos viso genomo sekoskaitos metodais, o dažni – taip pat ir plataus masto genomo asociacijų tyrimams naudojamais metodais. Dažni variantai dar vadinami **bendrais**, nes jų dažnis yra didelis ir jie būdingi didesnei populiacijos asmenų daliai, o reti variantai paprastai yra unikalūs, pasitaiko pavieniems asmenims (Arnett DK, 2012). Nustatomų variantų dažnio ir poveikio fenotipui ryšys vaizduojamas **1.2 paveiksle**. Išskyrus kelias išimtis, dažni ( $\geq 0,05$ ) variantai daro mažą poveikį fenotipui, reti ( $\leq 0,005$ ) – didelį. Mažo dažnio (0,005–0,05), tačiau ne reti variantai daro vidutinį poveikį (Manolio TA et al., 2009).



**1.2 paveikslas.** Genetiniai variantai pagal rizikos alelio dažnį ir genetinio poveikio dydį (šansų santykio reikšmes) (Manolio TA et al., 2009)

Yra genetinio poveikio skirtumas tarp retų didelę riziką susirgti lemiančių ir dažnų mažą riziką lemiančių variantų. Tai **kiekybinis**, ne kokybinis skirtumas. Manoma, kad genetinės įvairovės įtaka genų raiškai yra **daugiapakopė**. Genų

raišką reguliuoja daug elementų, o genetinė įvairovė tame gene galbūt keičia raiškos reguliavimą. Vadovaujantis šiuo modeliu, kiekvienoje genetinėje srityje yra daug variantų, kurie gali sudaryti vieną ar daugiau haplotipų blokų, esančių šalia paveikto DNR nuorašo. Taigi, ne vienas haplotipas, o jų rinkinys, ligos riziką didina arba mažina. Variantas, kuris daro didžiausią įtaką rizikai susirgti populiacijoje (alelio dažnio ir santykinės rizikos derinys), tikėtina, bus atrastas pirmas iš visos genetinės srities, kuri yra ligos rizikos veiksnys. Tolesnis atrasto varianto genetinės srities, kuriai jis priklauso, nagrinėjimas atveria galimybę nustatyti kitus, mažesnę įtaką rizikai turinčius alelius (*Hardy J, Singleton A, 2009*).

Šalia dažnų ir retų variantų aiškėja ir **somatinių mutacijų poveikis** dažnai stebimiems požymiams. Somatinės mutacijos dažniausiai yra siejamos su vėžio vystymusi. Vis dėlto manoma, kad somatinės mutacijos gali daryti įtaką ir paplitusioms ligoms, bet kitu, kancerogenezei nebūdingu, mechanizmu (*Roden D, 2009*).

Dažni, reti ir somatiniai variantai turi bendrus veikimo mechanizmų scenarijus. Pirmia, jie gali veikti **struktūriniu lygmeniu**, kai variantai yra aminorūgštis koduojančiose genų sekose ar splaisingo srityse ir lemia koduojamų baltymų struktūros pokyčius, tokiu būdu veikdami baltymų funkciją. Antra, minėtieji variantai gali veikti **reguliaciniu lygmeniu**, būdami reguliaciniuose genomo regionuose gali lemti genų raiškos valdymo pokyčius. Galiausiai yra manoma, kad rezultatas, t. y. daugiaveiksnė liga, yra suminis daugelio variantų (dažnų, retų, somatinių), esančių daugelyje skirtingų genomo sričių, veikimo padarinys, kai variantų deriniai reguliuoja biocheminius požymius (metabolitų perteklių, genų raišką ir hormonų kiekius), kurie ir lemia tam tikrą fenotipą (*Gibson G, 2012*).

Yra palaikoma ir plėtojama hipotezė apie **genų ir baltymų tinklus**, t. y. suminį daugelio variantų derinių veikimą. Daugelio ligų atveju skirtingos genetinės sritys dalyvauja bendrame patogenezės kelyje. Vadinasi,

nustačius vieną konkretų rizikos alelį konkrečioje gentinėje srityje, atsiveria galimybės aptikti ir kitas riziką lemiančias sritis bei mechanizmus (*Hardy J, Singleton A, 2009*). Pasitelkiant cirkuliuojančius leukocitus buvo pademonstruota genų raiška kartu (angl. *co-expression*), kuri priklausė nuo serumo metabolitų koncentracijos, tuo pagrindžiant molekulinių tinklų tarpusavio sąveikas, veikiant aplinkos veiksniams, hipotezę (*Inouye M et al., 2010*). Tikimasi, kad gilesnės žinios apie **epigenomą**, ypač genetinės variacijos ir genų raiškos kontekste, padės susieti aplinkos ir biologinių kelių daromą poveikį ligų patogenezei (*Hardy J, Singleton A, 2009*). Žmogaus ir kitų individų požymių įvarovei bei ligų patogenezei neabejotinai svarbūs yra ir **genomo sekų kopijų skaičiaus persitvarkymai** (KSP) (angl. *copy number variation, CNV*). Tai naujas genomo tyrimų objektas, tačiau KSP, kaip ir dažni bei reti vieno nukleotido variantai, gali daryti tiek didelį, tiek mažą poveikį ligos rizikai. Šiuo metu KSP tyrimai intensyviai taikomi aiškinantis daugiaveiksnių ligų patologiją ir šalia epigenominių tyrimų atradimų turėtų suteikti daugiau fundamentalių genetinės daugiaveiksnių ligų architektūros žinių (*Roden D, 2009*).

Išsiaiškinę konkrečius veikimo principus, galėsime nuspėti ne tik ligos pasireiškimą, tačiau ir greitį, koku ji vystysis, ligos mastą ir atsako į gydymą specifiškumą. Tad genetinės KŠL architektūros tyrimų dar laukia daug uždavinių: identifikuoti naujus ligos riziką lemiančius žymenis populiaciniu lygmeniu; nustatyti naujus genus ir jų veikimo kelius, lemiančius žmogaus patofiziologiją; nustatyti priežastingumą epidemiologiniuose tyrimuose; atrasti vaistų taikinius, padėsiančius optimizuoti gydymą (*Roden D, 2009*). Plataus masto genomo asociacijų tyrimais gauti duomenys ir besivystantys sekoskaitos metodai yra kelias į ligos genetinių priežasčių analizę.

### **Evoliuciniai genetinės KŠL architektūros modeliai**

Genetinė KŠL architektūra neatsiejama nuo genetinių polimorfizmų, darančių poveikį KŠL rizikai, alelių dažnių ir poveikio dydžio pasiskirstymo bei genetinio veikimo pobūdžio (adityvus, dominantinis ir (ar)

epistazinis, pleiotropinis). Svarbu išsiaiškinti genetinę architektūrą, siekiant įvertinti ligos riziką ir ją prognozuoti (*Wray NR et al., 2008*). Bet kurie evoliuciniai pokyčiai – mutacijos, genų dreifas ir gamtinė atranka – gali paveikti genetinę architektūrą. Evoliucinė genetika padeda paaiškinti rizikos alelių, lemiančių KŠL, skaičiaus ir dažnio paplitimo tendencijas (*Pritchard JK, 2001*).

Dažnu atveju rūšis yra apibrėžiama kaip grupė individų, kurie savo sandaros ypatybėmis prisitaikę gyventi tam tikromis aplinkos sąlygomis, skiriasi nuo kitų rūšių, gamina į save panašius ir laisvai kryžminasi tarpusavyje. Tai biologinė rūšies koncepcija. Yra ir kitų sampratų, tačiau būtent šiuo atveju rūšys yra traktuojamos ne kaip individai, o kaip populiacijos, pabrėžiamas rūšies kaip reprodukcinio, genetinio ir ekologinio vieneto egzistavimas (pagal *Rakauskas R, 2001*). Žmogus taip pat yra rūšis. Jam netaikomos biologinės išlygos, todėl gyvensenos pobūdžio ir kokybės, sykiu ir sveikatos, aspektai yra susieti su šiuo rūšies apibrėžime skambančiu teiginiu „prisitaikiusi gyventi tam tikromis aplinkos sąlygomis“. Prisitaikymas yra evoliucinis procesas. Žvelgiant į žmogų (rūšį) kaip į genetinį vienetą, veikiamą laiko ir aplinkos, evoliucijos procesas tampa akivaizdus, kaip ir jo įtaka genofondui.

Evoliucinės biologijos svarbą aiškinantis ligos priežastį pabrėžė *G. C. Williams* ir *R. M. Nesse* (1991), tai vadinamoji darvinistinė medicina. Gamtinė atranka veikia tiek mendelines, tiek daugiaveiksnes ligas lemiančius genetinius veiksnius (*Nielsen R et al., 2007*). Evoliucinio konteksto svarba priklauso nuo tiriamos ligos, tačiau KŠL atveju evoliucinė genetika yra ypač svarbi.

Daugiaveiksnių ligų atveju, skirtingai nei mendelinių, genetiniai **variantai, lemiantys polinkį ligai, manoma, yra ir reti, ir dažni (bendri)** (*Bodmer W, Bonilla C, 2008*). Ši hipotezė jungia žemiau pateikiamas dažnų ir retų variantų hipotezes, aiškinančias atitinkamų variantų ryšį su



daugiaveiksnėmis ligomis. Plataus masto genomo asociacijų tyrimais nustatyta dešimtys riziką lemiančių dažnų (bendrų) alelių (aptariama 2.1 skyrelyje), kurių poveikis yra mažas, tačiau dažnis bendroje populiacijoje yra didelis. Tai bandoma paaiškinti „**dažna liga – dažni variantai**“ (angl. *Common Disease/Common Variants*) hipoteze (Reich DE, Lander ES, 2001). Evoliuciškai tai arba neutralūs, arba pranašumą, palyginti su kitais individais, prisitaikant aplinkoje praeityje suteikę aleliai. Gyvenimo būdo pokyčiai, kultūra ir gyvenamoji aplinka žmonėms palikus Afriką lėmė specifinius atrankos spaudimo mechanizmus. Protėviniai aleliai, kurie tiko protėvių aplinkai, neatitinka šių dienų aplinkos, todėl tai vadinama **protėvinio polinkio** (angl. *ancestral-susceptibility*) modeliu daugiaveiksnių ligų atveju (Di Rienzo A, Hudson RR, 2005). „**Reti variantai – dažna liga**“ hipotezė teigia, kad polinkį ligai lemia daug retų variantų skirtingose genetinėse srityse, kurių kiekvienas daro vidutinišką poveikį ligos pasireiškimui. Šie aleliai ne tokie žalingi, kaip mendelinių ligų atveju, todėl nėra eliminuojami gamtinės atrankos (Bodmer W, Bonilla C, 2008). Yra pasiūlyta ir „**netikrų asociacijų**“ (angl. *synthetic association*) hipotezė, teigianti, kad dažni VNP yra asocijuojami su daugiaveiksne liga ar požymiu tik tada, kai genetinėje srityje yra nepusiausviroje sankiboje su vienu ar daugiau retų variantų, kurie turi didesnę poveikį fenotipui. Daugiaveiksnių ligų atveju retų variantų nustatymas yra ankstyvoje tyrimų stadijoje, kadangi šiuo metu atliekami plataus masto genomo asociacijų tyrimai neleidžia jų nustatyti, o naujos kartos viso genomo sekoskaitos tyrimai, įgalinantys juos nustatyti, yra dar tik prasidėję (Dickson SP et al., 2010; Wang K et al., 2010).

Protėvinio polinkio modelio atveju, kadaise apsaugoję nuo ligos aleliai, dabar tapo nebetinkami. Tai lemia kelios priežastys. Pirmiausia – atranka palaiko variantus, kurie didina reprodukcinę sėkmę, net jei tie variantai lemia didesnę riziką susirgti vyresniame amžiuje (Williams GC, 1957). Antra – atsilikimas laike, kai aplinkos pokyčiai vyksta greičiau negu prisitaikymas. Trečia – gamtinė atranka nesukuria kitų apsauginių variantų. Galiausiai,

atranka yra stochastinis procesas, visiškai atsitiktinai išstumiantis nuo ligos saugančius alelius arba palaikantis rizikos alelius, tai lemia padidėjusį polinkį ligai (*Houle D, 1992*). Koronarinės širdies ligos atveju genetinius ir evoliucinius ligos tyrimus sunkina sudėtinis ligos fenotipas ir ligą lemiančių veiksnių kiekybė. Tačiau epidemiologiniais tyrimais buvo nustatyti biologiniai keliai ir veiksniai, svarbūs KŠL patogenezei. Tai aterosklerozę lemiantys biologiniai keliai – kraujospūdžio reguliavimo, lipoproteinų ir gliukozės metabolizmo, krešėjimo ir uždegimo procesų, kurie gali būti veikiami gamtinės atrankos gyvenamomis klimato ir mitybos sąlygomis bei šeiminingo atsako į patogeną ypatybių (*Ding K, Kullo IJ, 2009*).

Hipertenzijos atveju skirtingą polinkį jai lėmė istorinis prisitaikymas prie klimatinių sąlygų (*Hancock AM et al., 2008*) – natrij tausojantis mechanizmas užtikrino išgyvenamumą karštoje ir drėgnoje aplinkoje Afrikoje, tačiau vidutinio klimato zonose jis lemia hipertenziją (*Ding K, Kullo IJ, 2009*). Nustatyti genetiniai natrio apytakos dėsniumai – natrij tausojantys aleliai – yra dažnesni Afrikos populiacijose, palyginti su Šiaurės Europos populiacijomis. Tai turi reikšmės hipertenzijos diagnostikai ir gydymui geografiškai ir etniškai išnagrinėtose ir aprašytose populiacijose visame pasaulyje (*Nakajima T et al., 2004*).

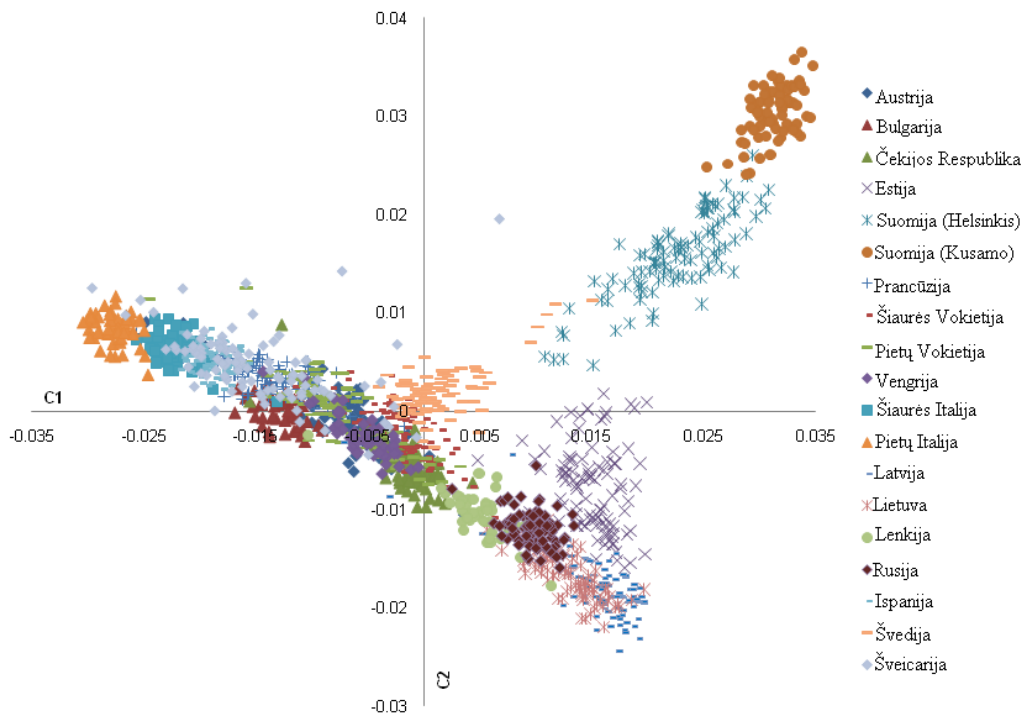
Lipoproteinų metabolizmas evoliuciškai galėjo būti paveiktas jau minėtų klimato ir mitybos sąlygų pokyčių. Didesni cholesterolio kiekiai serume teikė privalumą evoliucionuojant žmogaus ūgiui ir masei, nes cholesterolis dalyvauja steroidinių hormonų sintezėje. Todėl *APOE* geno  $\epsilon 4$  alelis, lėmęs geresnį prisitaikymą anuomet, šiomis maisto pertekliaus išsivysčiusiose šalyse sąlygomis tapo lemiančiu KŠL riziką (*Stengard JH et al., 1995*). Plėtojant biologinių metabolizmo kelių, susijusių su KŠL, evoliucijos temą būtina paminėti metabolinį sindromą, kuris apima nutukimą (padidėjusi juosmens apimtis), aterogeninę dislipidemiją, hipertenziją, sutrikusį gliukozės metabolizmą. Šie veiksniai lemia aterosklerozės ir KŠL riziką (*Dandona P et al., 2005*). Evoliuciškai polinkis į nutukimą ir diabetą tam

tikrose žmonių populiacijose aiškinamas „taupaus geno“ hipoteze (angl. „*thrifty-gene*“). Teigiama, kad tam tikri genetiniai variantai atsirado siekiant maksimizuoti metabolinį efektyvumą, lipidų atsargas ir maisto ieškojimo elgseną. Šiais laikais tokie genetiniai variantai lemia polinkį sirgti ligomis, kurias lemia perteklinis maisto vartojimas (*Neel JV, 1962*). Nesenų tyrimų išvadomis, klimatiniai pokyčiai galėjo būti svarbūs koreguojant genetinę įvairovę, susijusią su metaboliniais keliais (*Hancock AM et al., 2008*).

Krešėjimo mechanizmo genetinės įvairovės pokyčiai lėmė didesnę išgyvenamumą ankstesnėje aplinkoje. Galbūt tie pokyčiai buvo hemostazės sustiprinimas, apsaugantis nuo nukraujavimo susižeidus, medžiojant ar po gimdymo (*Zivelin A et al., 1997; Dahlback B, 2008*).

Atrankos spaudimas neabejotinai veikė šeimininko atsako į patogeną mechanizmus, kartu ir juose dalyvaujančius genetinius veiksnius. Padidėjęs imuninis atsakas, išsivystęs kovoje su patogenais, gali padidinti KŠL riziką. Uždegimo procesų įtaka aterosklerozei yra įrodyta, taip pat pasiūlyti genai kandidatai, koduojantys imuniniame atsake dalyvaujančius baltymus (*Ding K, Kullo IJ, 2009*).

Svarbu pažymėti, kad evoliucinis aspektas aiškina ligos genetinę architektūrą tiek laiko, tiek erdvės atžvilgiu. Kiekviena populiacija geografiniame žemėlapyje užima tik tam tikrą vietą. Bėgant laikui žemėlapis kito. Tačiau nustatyta, kad ir genetiniu požiūriu Europos gyventojų populiacijos, sykiu ir Lietuvos, išsidėsto pagal dabartinį geografinį žemėlapi, (žr. **1.3 paveikslą**).



**1.3 paveikslas.** Europos populiacijų genetinė struktūra (*Nelis M et al., 2009*). Sutartiniais ženklais, remiantis pagrindinės genetinės komponentės analize, pavaizduotos Europos populiacijos.

Kiekvienos šalies, tautos konkreti geografinė padėtis ir jos sutapimas su genetinės struktūros pasiskirstymu tik patvirtina *E. Mayr* suformuluotą teiginį, kad rūšis, kaip uždara genetinė sistema, pasižymi genofondo specifiškumu, turėtų pasižymėti ir ekologinės nišos specifiškumu (*Rakauskas R, 2001*). Todėl aiškinantis ligų genetinę architektūrą, svarbu tirti vietines populiacijas, nes jose glūdi genetiniai skirtumai nuo kitų populiacijų, nulemti genetinės įvairovės veikiant aplinkai. Šie skirtumai nulemia ir ligos architektūros skirtumus, kuriuos svarbu išsiaiškinti.

## 1.2. Daugiaveiksnių ligų genetinės architektūros nustatymas

„*Naujas mokslas – dėl genomo asociacijų tyrimų*“

Peter M. Visscher

Dvidešimt metų genetinė sankiba kartu su poziciniu klonavimu buvo du nepakeičiami genetinės analizės būdai, leidę tiesiogiai nustatyti

mutacijas ir genus lemiančius monogenines ligas. Taip buvo išaiškintos cistinės fibrozės, Hantingtono (*Huntington*) ir daugelio kitų mendelinių ligų priežastys. Šiais metodais buvo nustatytos mutacijos, keičiančios aminorūgščių seką baltyme ir gerokai didinančios ligos riziką. Taip susiformavo molekulinės genetinės patologijos sampratos pamatai.

Suvokus, jog šie metodai nėra tinkami daugiaveiksnių ligų genetinių priežasčių tyrimui, buvo pasiūlyta ir ištobulinta genomo asociacijos tyrimų metodika. Šalia nuodugnaus genų kandidatų asociacijos su liga tyrimo, grįsto išankstine hipoteze apie genetinio veiksnio reikšmę ligos patogenezei, išsirutuliojo plataus masto genomo asociacijos su tiriamu požymiu tyrimas, nereikalaujantis išankstinės hipotezės (*Hunter DJ, Kraft P, 2007*). Pirmąjį sėkmingą plataus masto genomo tyrimą 2005 metais išspausdino „*Science*“ žurnalas, kai buvo nustatyti VNP, dažniau pasitaikę senatvinę geltonosios dėmės degeneraciją (angl. *age-related macular degeneration*) turinčių pacientų grupėje, palyginti su kontroliniais asmenimis (*Klein RJ et al., 2005*). Nuo to laiko, genomo asociacijos tyrimų daugėjo, gerėjo ekonominės ir analitinės atlikimo sąlygos. Nustatyta daugybė statistiškai reikšmingų asociacijų tarp specifinių chromosominių sričių ir daugiaveiksnių žmogaus ligų. Dabar nustatyti dažnus, mažą riziką lemiančius variantus (kurių dažnis populiacijoje  $\geq 5\%$ , šansų santykis 1,2–5,0 intervale), darančius įtaką daugiaveiksnių ligų vystymuisi, tapo laboratorinė rutina (*Manolio TA et al., 2008*). Genomo asociacijų tyrimai leido aptikti genetines sritis, kuriose yra dažni (bendri) ligos riziką lemiantys variantai. Šių tyrimų rezultatai padėjo suformuluoti „dažna liga – dažnas variantas“ hipotezę bei jai alternatyvią – „reti variantai – dažna liga“ hipotezę. Šiandien šias dvi hipotezes jungia hipotezė, kuri teigia, kad daugiaveiksnes ligas lemia tiek dažni, tiek reti variantai. Tačiau plataus masto genomo asociacijų tyrimas nėra tinkamas nei tiesiogiai nustatyti ligos genus, nei retų variantų sritis. Tai yra naujos kartos sekoskaitos ir genų kandidatų tyrimų uždavinys (*Roden D, 2009*).

### 1.2.1. Nuo genų kandidatų iki naujos kartos sekoskaitos. GWAS

Nustatant DNR sekos variantus, lemiančius polinkį sirgti liga, istoriškai yra susiklostę du pagrindiniai metodai: sankibos ir asociacijų analizės, kuri skirstoma į genų kandidatų asociacijos ir plataus masto genomo asociacijų tyrimus.

#### Genetinės sankibos analizė

Genetinės sankibos analizė kildinama iš mendelinės genetikos, kurios vienas iš pagrindinių principų yra nepriklausomas požymių paveldėjimas. Nukrypimai nuo šio dėsnio, kai genetinis žymuo ir ligą lemiantis tiriamas požymis paveldimi kartu, tarsi įrodo, kad ligą lemianti sritis ir žymuo yra fiziškai netoli vienas kito, todėl vadinami „sukibusiais“ arba vertinami kaip nerekombinuojantys (arba rekombinacija reta). Sankibos analizė remiasi šeimų, kuriose nustatomi nepriklausomi mejozės įvykiai (informatyvios mejozės) ir įvertinama kosegregacija, tyrimais. Svarbu pasirinkti analizei tinkamus genetinius žymenis, kurie turi būti išsidėstę po genomą. Klasikiniu atveju sankiba statistiškai yra vertinama LOD įverčiu ( $Z$ ). Tai šansų logaritmas, kai įvertinama sankibos įvykio tikimybė. LOD esant  $> 3$  vertinama kaip sankiba,  $< (-2)$  vertinama kaip nesankiba. Reikšmės, esančios tarp  $(-2)$  ir  $3$ , yra neinformatyvios. Esant įvairioms tyrimo strategijoms ir galimybėms, galima pasirinkti tinkamiausią kompiuterinę programą sankibos analizei atlikti. Sankibos analizė leidžia skenuoti genomą 20 Mb segmentais (pagal *Strachan T, Read A, 2011*). Šis metodas skirtas ligos genams nustatyti, juo buvo aptiktas ne vienas mendelinių ligų genas. Tačiau metodas turi trūkumų. Visų pirma, tai jautrumas klaidoms, skaičiavimo apribojimai atsižvelgiant į šeimų struktūrą, genetinių sričių heterogeniškumo įtaką, antra – ribota skiriamoji geba, griežtas parametrų poreikis (reikia nurodyti paveldėjimo tipą, alelinių genų dažnius, kiekvieno genotipo penetrantiškumą) ir trečia – ligos, nepaklūstančios mendelinio paveldėjimo dėsningumams, negali būti nustatytos sankibos analizės metodu (pagal *Strachan T, Read A, 2011*). Daugiaveiksnių ligų atveju

tinkamesni yra genų kandidatų ir plataus masto viso genomo asociacijų tyrimai.

### **Genų kandidatų asociacijos tyrimas**

Istoriškai genotipavimo našumo ir ekonominiais sumetimais asociacijų tyrimai buvo ribotos apimties. Buvo tiriama keletas variantų siaurame genetiniame regione, susitelkiant į specifinį geną. Tai genų kandidatų tyrimai. Jie tebėra populiarūs. Šis tyrimas yra priklausomas nuo hipotezės, t. y. genai pasirenkami tyrimui ne atsitiktinai, o turint tam tikrą hipotezę. Hipotezė gali būti formuluojama keliais pagrindais: pagal įtariamą geno funkciją lemiant tiriamą požymį; pagal sankibos analize nustatytą geno priklausomybę genetinei sričiai; pagal geno raiškos ar proteomikos tyrimų rezultatus. Tai vadinamoji **tiesioginė asociacijų analizė**. Šios analizės metu dėmesys telkiamas į vieno nukleotido polimorfizmus (VNP). Šio metodo pranašumas yra ekonomiškasis ir detalus kiekvieno tiriamo geno sekos ir jos variantų ištyrimas. Šių tyrimų rezultatai yra aiškūs ir nesudėtingai interpretuojami dėl nustatomos tiesioginės asociacijos. Be to, tolesnį šių rezultatų patvirtinimą nesudėtinga perkelti į funkcinį lygmenį. Pagrindinis metodo trūkumas yra išankstinė hipotezė, kurios formulavimas priklauso nuo sukauptų mokslo duomenų kokybės ir aktualumo. Pasenusių ir (ar) klaidingų, netikslių duomenų iš literatūros, duomenų bazių ar kt. naudojimas hipotezei formuluoti gali paveikti gaunamą rezultatą, netinkamai pakreipti tyrimo eigą. Tačiau agnostiškai, t. y. nepriklausomai nuo hipotezės, atrasti su tiriamu požymiu susietą naują genetinę sritį (geną) šiuo metodu yra neįmanoma (Rodén D, 2009).

### **Plataus masto genomo asociacijų tyrimas**

Kitas būdas nustatyti genetinius variantus, susijusius su liga, yra viso genomo žymenų asociacija su požymiu ar liga. Plataus masto genomo asociacijų tyrimas (angl. *genome-wide association study*, *GWAS*), pasiūlytas 1996 metais, apima didelę grupę dažnų genomo variantų. Asociacijų tyrimai yra skirti aptikti vidutinį poveikį požymiui darantiems variantams, kurių negali

nustatyti sankibos analizė (*Spielman RS, Ewens WJ, 1996*). Šio tyrimo metu tiriamas genomo žymenų (dažniausiai VNP, kurie pasklidę tolygiai po genomą) sąsaja su požymiu ar liga. Tai **netiesioginis asociacijos** nustatymo būdas, kuriam nereikalinga išankstinė hipotezė. Netiesioginė asociacija remiasi reta rekombinacija tarp arti esančių VNP. Taigi VNP tam tikruose chromosomų regionuose telkiasi į neatsitiktinai susijusius nepusiausviroi sankibos blokus (angl. *LD blocks*) (*Daly MJ et al., 2001*), kurie gali būti fiziškai pertraukti. Nustačius pirmąjį, efektyviausiai haplotipų bloką reprezentuojantį VNP alelį, kuris dar vadinamas žyminiu (angl. *tag SNP*), galima nustatyti ir kitus blokui priklausančius alelius (*Johnson GC et al., 2001*). Tad netiesioginio asociacijos tyrimo tikslas yra nustatyti nepusiausviroi sankibos VNP rinkinį, kuris yra statistiškai reikšmingai susijęs su tiriamu požymiu ar liga. Tyrimas gali būti atliekamas tiek šeimų, tiek atvejo ir kontrolės grupėse. Sergančių arba turinčių tam tikrą požymį asmenų grupėje žyminio VNP alelių dažniai skirsis nuo kontrolinių (nesergančių arba neturinčių požymio). Šio metodo pranašumas – analizė be išankstinės hipotezės. Trūkumas tas, kad nė vienas genas ar genetinė sritis nėra nuodugnai tiriami, todėl kai kurios asociacijos gali likti nenustatytos.

Atliekant plataus masto genomo asociacijos tyrimą keliami tam tikri reikalavimai. Pirmia, turi būti tinkamai pasirinktas požymis ar liga ir didelė tiriamųjų imtis. Antra, atlikus tiriamųjų genotipavimą ir gavus žaliavinius duomenis būtina atlikti kokybės kontrolę ir duomenų valymą šalinant blogos kokybės duomenis ir išskirtis. Trečia, tik atlikus kokybės kontrolę ir tinkamai pakoregavus žaliavinius duomenis atliekama asociacijų analizė, kuri yra atvaizduojama *Manhattan* grafiku. Ketvirta, pagal pasirinktą (pageidautina  $1-5 \times 10^{-8}$ ) reikšmingumo lygmenį nustatomi ir atrenkami reikšmingai asocijuoti variantai tolesnei analizei nepriklausomoje imtyje asociacijai patvirtinti. Penkta, priklausomai nuo to, ar asociacija atkartojama, toliau atliekamas naujas papildomas nepriklausomas tyrimas (naujai genotipuojant atskirą nepriklausomą imtį) siekiant nustatyti, ar atsikartos tų pačių variantų



asociacijos. Atkartojus asociaciją, toliau atliekama duomenų bazių (angl. *data mining*) analizė, nustatant genetinę sritį ir jai priklausančius kitus VNP. Galiausiai, siekiant nustatyti kritinius biologiniam efektui variantus, atliekamas tikslus genetinės srities kartografavimas (angl. *fine mapping*), jį derinant su giliais pakartotinės sekoskaitos (angl. *deep-resequencing*) metodais (Hardy J, Singleton A, 2009).

Atsiradus galimybei (techniniu ir ekonominiu požiūriais) genotipuoti milijonus VNP, iškilo poreikis tobulinti bioinformatikos metodus. Duomenų kiekiai tapo neaprėpiami skaičiuoti įprastais metodais, todėl atsirado naujos statistinės problemos, kurias reikėjo spręsti, rezultatų interpretacija tapo sudėtingesnė. Plataus masto genomo asociacijų tyrimas yra jautrus populiacijos stratifikacijai, trūkstamiems genotipams (angl. *missing genotypes*), nukrypimams nuo Hardžio ir Vainbergo pusiausvyros. Tai sprendžiama atliekant žaliavinių duomenų kokybės kontrolę ir valymą. Tikrinant daugelio žymenų asociaciją, dėl daugybinio palyginimo didėja klaidingai teigiamų rezultatų tikimybė, todėl yra griežtas reikšmingumo lygmens ( $\alpha$ ) reikalavimas. Kuo  $\alpha$  mažesnis, tuo reikšmingos asociacijos nustatymo galia mažesnė (ypač mažos imties atveju), todėl reikalingos kuo didesnės imtys (Lewis CM, Knight J, 2012).

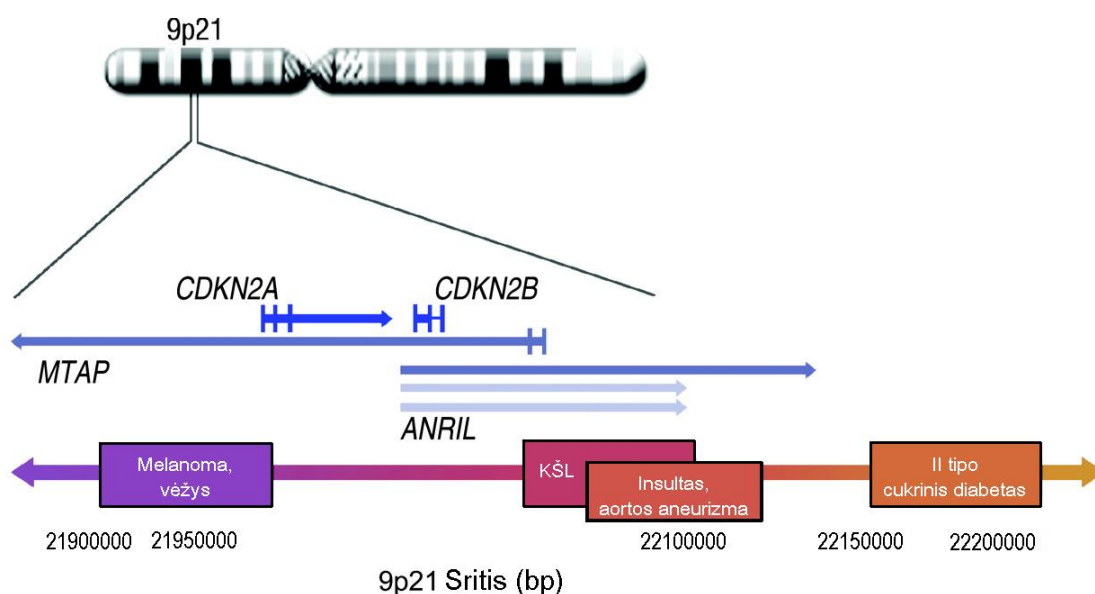
Tiek genų kandidatų, tiek plataus masto genomo asociacijų tyrimai turi savo privalumų ir trūkumų. Jų palyginimas pateikiamas **1.1 lentelėje**. Tinkamai pritaikyti abu metodai gali būti pravartūs naujiems atradimams.

**1.1 lentelė.** Genų kandidatų ir plataus masto genomo asociacijos tyrimų palyginimas (lentelė adaptuota pagal *Roden D, 2009*)

	<b>Genų kandidatų tyrimas</b>	<b>Plataus masto genomo tyrimas</b>
<b>Pobūdis</b>	Mažas VNP kiekis gene kandidate	100000–1000000 ir daugiau VNP, išsidėsčiusių genome
<b>Imtis</b>	Didelė	Labai didelė
<b>Kaina</b>	Santykinai nedidelė	Santykinai didesnė
<b>Jautrumas daugybiniam palyginimui</b>	Mažiau jautrus, variantai turi didesnę pirminę tikimybę būti asocijuoti	Labai jautrus, reikalingos didelės imtys pasiekti reikšmingas asociacijų vertes
<b>Klaidingai teigiamų rezultatų rizika</b>	Santykinai maža	Santykinai didesnė, reikalauja kruopščios kokybės kontrolės ir griežto reikšmingumo lygmens
<b>Stratifikacijos nustatymas</b>	Nenustatoma	Nustatoma
<b>Naujų sričių nustatymas</b>	Nenustatomos, tikrinama išankstinė hipotezė	Nustatomos, išankstinės hipotezės nėra

Iki 2012 metų plataus masto genomo asociacijų tyrimais nustatyta daugiau negu 2000 genetinių sričių, siejamų su daugiaveikniais požymiais ir (ar) ligomis (*Visscher PM et al., 2012 (1)*). Koronarinės širdies ligos atveju nustatytos 46-ios genetinės sritys ir 104 nepriklausomi variantai, patvirtinantys kitų tyrėjų grupių gautus rezultatus. Genų tinklų analizė nustatė 5 potencialius biologinius kelius, lemiančius KŠL patogenezę, iš jų keturi susiję su lipidų metabolizmu, vienas – su uždegiminiu atsaku (*Deloukas P, The CARDIoGRAMplusC4D Consortium et al., 2013*). Pirmoji ir bene nuodugniausiai tiriama genetinė sritis, susijusi su KŠL, yra **9p21** (*McPherson R et al., 2007*). Šios srities asociacija su KŠL buvo patvirtinta nepriklausomais tyrimais ne viena kartą. Nustatyta, kad ši sritis lemia KŠL riziką nepriklausomai nuo kitų ligą lemiančių žinomų veiksnių (*Roberts R, Stewart AF, 2012*). Tiriant 9p21 sritį, nenustatyta nė vieno baltymą koduojančio geno. Sritis yra sudėtingos sandaros (**1.4 paveikslas**). Ją sudaro *CDKN2B-AS1*

(priešpasmės *CDKN2B RNR 1* genas), nekoduojanti RNR, dar žinoma kaip *ANRIL*. *CDKN2B-AS1* yra greta *CDKN2A* geno ir sutampa su *CDKN2B* genu. Pastarieji koduoja nuo ciklinų priklausančių kinazių inhibitorius. Šie genai pelių tyrimuose yra siejami su sparčiu auglių vystymusi (*Krimpenfort P et al., 2007*). Šioje srityje taip yra *MTAP* – metiltioadenozino fosforilazės genas. Tiriant *CDKN2B*, *CDKN2A* ir *MTAP* genų raiškos dėsningumus žmogaus pažeistose kraujagyslių ląstelėse, buvo aptikta gausi šių baltymų gamyba kraujagyslių sienelėse, tuo patvirtinant šiuos genus kaip potencialius kandidatus aterogenezėje (tačiau nepriklausomai nuo 9p21) (*Holdt LM et al., 2011*). *ANRIL* nuorašai yra būdingi audiniams ir ląstelėms, siejamiems su ateroskleroze. Manoma, kad *ANRIL* modifikuoja *CDKN2B* raišką ir tokiu būdu yra skatinama aterosklerozei būdingų ląstelių tipų proliferacija (*Jarinova O et al., 2009*).



**1.4 paveikslas.** 9p21 genetinė sritis ir regionai, susiję su skirtingais požymiais (*Zeller T et al., 2012*). KŠL – koronarinė širdies liga, bp – bazių poros, *CDKN2A* – nuo ciklinų priklausomą kinazės inhibitorių 2A koduojantis genas, *CDKN2B* – nuo ciklinų priklausomą kinazės inhibitorių 2B koduojantis genas, *MTAP* – metiltioadenozino fosforilazę koduojantis genas, *ANRIL* (*CDKN2B-AS1*) – *CDKN2B* priešpasmės RNR 1 genas.

Genetinės srities 9p21 tyrimas yra puikus pavyzdys, kokius struktūrinius ir funkcinius atradimus genome leidžia daryti plataus masto genomo asociacijų tyrimai. Šiuo ir kitais plataus masto genomo asociacijų tyrimų atradimais esame vienu žingsniu arčiau daugiaveiksnių ligų ir požymių genetinės architektūros suvokimo. Tačiau akivaizdu, kad dalis paveldimumo yra trūkstama (angl. *missing heritability*) (Gibson G, 2012). Todėl dar daug reikia nustatyti ir išsiaiškinti. Artimiausi taikiniai – retų genomo variantų įvairovės ir reikšmės bei atskirų populiacijų genetinės įvairovės tyrimai daugiaveiksnių ligų kontekste.

### **Sekoskaita ir jos perspektyva**

Tiek genų kandidatų tyrimai, tiek besivystančių didelio našumo sekoskaitos metodų rezultatai atskleidžia retų, didelę riziką lemiančių genomo variantų sistemingo nustatymo poreikį. Genų kandidatų tyrimai pademonstravo gilią pakartotinės sekoskaitos galią nustatant retų variantų reikšmę fenotipui. Pavyzdžiui, pacientų grupėje, pasižyminčioje mažais DTL-cholesterolio kiekiais kraujo plazmoje, buvo tiriami genai, koduojantys cholesterolį metabolizuojančius baltymus (Cohen JC et al., 2004). Buvo nustatyta, kad reti genomo variantai šiuose genuose yra dažnesni pacientų grupėje, palyginti su kontrolinių asmenų grupe.

Vis dar nėra iki galo aišku, kuri hipotezė („dažna liga – dažni variantai“, „reti variantai – dažna liga“, „netikrų asociacijų“, „dažnų ir retų variantų“) geriausiai aprašo daugiaveiksnes ligas ir požymius. Todėl būtina toliau tirti dažnus ir retus variantus bei ieškoti veikimo tinklų ir sąveikų (Gibson G, 2012).

Šiuo metu yra sulygstama dėl to, kas yra nustatyta daugiaveiksnių požymių ir ligų genetikos ir genomikos tyrimais:

1. Daugiaveiksnių požymių įvairovę nulemia **daug** genetinių sričių.
2. Nustatytose riziką lemiančiose genetinėse srityse yra daug su liga asocijuotų alelių, kurių **dažnis įvairuoja**.

3. Galima **pleiotropija**, t. y. tie patys variantai yra asocijuoti su daugeliu požymių (*Sivakumaran S et al., 2011*).
4. Variantai, susiję su daugiaveiksne liga ar požymiais vienoje etninėje populiacijoje, yra susiję ta pačia liga ar požymiu kitose populiacijose.
5. Nėra aiškių įrodymų, pagrindžiančių „netikrų asociacijų“ hipotezę. Vadovaujamesi tuo, kad dauguma priežastinių variantų, segreguojančių populiacijose, yra reti (neprieštarauja evoliucinei teorijai). Reti variantai paaikškina genetinės įvairovės dalį, tačiau genetinė įvairovė yra alelių dažnių, poveikio dydžio ir funkcinių ypatybių derinys, darantis įtaką tinkamumui (angl. *fitness*) (*Visscher PM et al., 2012 (2)*).
6. Vertinant visus VNP vienu metu galima nustatyti didelę dalį suminės genetinės įvairovės, be to, genuose ar šalia jų esantys VNP paaikškina daugiau įvairovės, negu VNP, esantys tarpgeninėse srityse (*Yang J et al., 2011*).

Vystantis didelio našumo naujos kartos genomo analizės metodams, viso genomo sekoskaita tampa prieinama tiek didelės apimties moksliniams tyrimams, tiek diagnostinei praktikai. Šie darbai nulems tolesnius atradimus tiriant mendelinių ir daugiaveiksnių ligų genetinę architektūrą – nežinomų genų ir priežastinių mutacijų, biologinių kelių ir mechanizmų nustatymą, naujas biologines išvalgas (*Visscher PM et al., 2012 (1)*).

### 1.2.2. Specifinių populiacijų tyrimo svarba

Dideliuose miestuose, ypač šalių sostinėse, matomas gyventojų heterogeniškumas. Todėl egzistuoja ir natūralūs skirtumai, susiję su ligą lemiančiais rizikos veiksniais ir pačiu ligos paplitimu, gyvenimo trukme bei sveikatos priežiūros paslaugomis. Šie klausimai neišvengiamai susiję su genetiniais ir aplinkos veiksniais (gyvensena remiantis kultūriniais ypatumais) (*Roden D, 2009*).

Terminas **etniškumas** yra kilęs iš graikų kalbos, *ethnos*, reiškiantis liaudį ar tautą. Tai „daugiaveidė“ savybė grupės, kuriai pagal tam tikras

bendras charakteristikas, apimančias geografines ir protėvines ištakas, o ypač kultūrinės tradicijas ir kalbą, priklauso (ar priskiriami) žmonės (*Česnys G, 2009; Bhopal R, 2004*). Kad ir kaip ten būtų, etniškumo koncepcija sveikatos priežiūros ir visuomenės sveikatos sistemoje vis dar yra prieštaringa, kelia daugybę apibrėžties ir etinių klausimų. Dėl tų pačių priežasčių paskelbtuose moksliniuose tyrimuose sudėtinga atrasti tikrai unikalių populiacijų, kadangi neretai už termino „juodaodžių populiacija“ ar „azijiečių populiacija“ slepiasi subpopuliacijos, apie kurias neužsimenama (*Roden D, 2009*). Rasė ar etniškumas – tai socialinis biologinis, sociokultūrinis, psichologinis ir elgsenos kontekstus apimantis konstruktas, kuriame nėra vietos genetinei apibrėžčiai (*Česnys G, 2009; Cooper RS et al., 2003*). Priešingai rasei, genetinė kilmė (angl. *genetic ancestry*) priklauso genetinei koncepcijai, kuri nagrinėja genomo įvairovės architektūrą tarp populiacijų (*Kittles RA, Weiss KM, 2003*). Visa genetinė įvairovė prasideda lokaliai nuo vienos mutacijos individe ir todėl visos naujos mutacijos pradžioje yra geografiškai nulemtos (*Dries DL, 2009*). Tačiau dauguma klinikinių tyrimų yra atlikta baltųjų populiacijoje, o rezultatai generalizuoti ir ekstrapolijuojami visiems pacientams bendrai, neatsižvelgiant į jų kilmę. Trūksta žinių apie daugiaveiksnes ligas apskritai, jau nekalbant apie kiekvienos populiacijos genetinę įvairovę šių ligų kontekste. Koronarinės širdies ligos atveju esama užmojo tirti milijonus genetinių žymenų, netgi visą genomo seką, didelėse keleto kruopščiai atrinktų populiacijų imtyse, siekiant nustatyti aterosklerozės polinkį lemiančius genus. Paskui būtų atliekamas pakartotinis tyrimas panašiose ir skirtingose etninėse populiacijose, kas leistų nustatyti bendrus ir unikalius populiacijoms polinkį lemiančius genus. Tolesnis būtų aterosklerozės pasekmių – KŠL ir MI, analogiškas tyrimas papildant imtį atitinkamais fenotipais. Viso genomo tyrimai populiacijose pagal specifinius aterosklerozės riziką lemiančius veiksnius (hipertenziją, nutukimą ar cukrinį diabetą) leis įvertinti kiekvieno veiksnio poveikį atskirai bei jų derinių poveikį bendrai. Tik tokie tyrimai atvers galimybes tikslesniam genų–genų ir aplinkos sąveikos įvertinimui (*Roden D, 2009; Henn BM et al., 2010*). Vienas iš tokių

pavyzdinių projektų yra Jungtinėje Karalystėje vykdomas jungtinis mokslininkų projektas *UK10K* (<http://www.uk10k.org/>).

### 1.2.3. Genetinės KŠL architektūros tyrimai Lietuvoje

Apskritai aptariant žmogaus genetikos ir genomikos populiacinius tyrimus Lietuvoje, svarbiausi tyrimai ir atradimai yra paskelbti prof. habil. dr. V. Kučinsko monografijoje „Genomo įvairovė: Lietuviai Europoje“, išleistoje 2004 metais. Tai pagrindas tolesniems genetiniams ir genominiams Lietuvos populiacijos tyrimams. Jungtinio, VU mokslininkų 2004–2006 metais vykdyto, „Žmogaus genomo įvairovės ypatumų nulemti aterosklerozės patogenezės mechanizmai“ (akronimas *ATHEROGEN*) projekto metu buvo atliekami genų kandidatų asociacijos su ateroskleroze tyrimai, o vėliau – plataus masto genomo asociacijų tyrimai. *ATHEROGEN* projekto metu buvo tiriama *APOE* (rs426358 ir rs7412), *LIPC* (rs180058), *PONI* (c.632A>G ir c.221T>A) ir *IL-6* (-174G>C) genų VNP žymenų asociacija su ateroskleroze. Į tyrimą įtrauktos 65 šeimos, genotipavimas atliktas restrikcijos endonukleazėmis, asociacijai nustatyti naudotas nepusiausviro perdavimo testas (angl. *transmission disequilibrium test (TDT)*). Reikšmingos asociacijos su paveldimu polinkiu sirgti ateroskleroze nenustatyta, tačiau autoriai pabrėžia, kad genų žymenų asociacijos su polinkiu sirgti ateroskleroze atmesti negalima ir būtini tolesni išsamesni tyrimai (*Grigalionienė K et al., 2007(2); (1)*). *ATHEROGEN* projekto metu, atlikus bioinformacinę duomenų bazių analizę, pagal pasirinktą metodiką (*Timinskas A et al., 2007*) buvo sudarytas aterogenezės genų kandidatų sąrašas ir atrinkta 150 VNP 89-iuose genuose kandidatuose genetiniam APEX lustui sukurti. Vėliau, naudojant APEX technologija sukurtus lustus, genotipuotos 28 šeimos (pacientai, sergantys ateroskleroze, ir jų tėvai arba sibsai). Atlikus *TDT* analizę nustatyti šeši VNP *ITGA2*, *IL1B*, *ALOX5A*, *OR13G1*, *MMP9* ir *NFKB1* genuose, statistiškai reikšmingai asocijuoti su aterosklerozės fenotipu. Šie genai vienaip ar kitaip susiję su oksidacijos procesais ląstelėse ir neprieštarauja teiginiui, kad oksidaciniai procesai yra svarbūs aterosklerozės vystymuisi (*Pepalyte I et al., 2012*).

Vilniaus universiteto Medicinos fakultete 2010–2011 metais vykdytas projektas „Lietuvos populiacijos genomine įvairovė ir koronarinės širdies ligos individualaus gydymo galimybės“ (akronimas LIGENKOR) (*Domarkienė I et al., 2013*) ir 2011–2015 metais vykdomas projektas „Lietuvos populiacijos genetinė įvairovė ir sandaros kitimai, susiję su evoliucija ir dažniausiai paplitusiomis ligomis“ (akronimas LITGEN). Pastarųjų projektų rezultatai aptariami šiame darbe.

Lietuvos sveikatos mokslų universiteto Kardiologijos institute 2012–2014 metais vykdomo projekto „Genetinių ir gyvenimo veiksnių ryšiai, vertinant lėtinių ligų riziką ir jos pokyčius nuo vaikystės“ tyrėjai nagrinėjo tuos pačius *APOE* geno variantus, kaip ir ATHEROGEN projekto vykdytojai, tačiau tyrimo pobūdis kitoks. Tikslas buvo nustatyti *APOE* VNP (rs429358 ir rs7412) genotipų ir kitų rizikos veiksnių ryšį su lipidų kiekiais kraujyje 996 asmenų iš penkių Lietuvos savivaldybių. Genotipavus RT-PGR metodu ir atlikus daugialypės tiesinės regresijos analizę buvo nustatyta, kad *APOE* tiriamų VNP genotipai, sočiųjų riebalų rūgščių suvartojimas ir nutukimas yra susiję su tiriamųjų grupės lipidų kiekiais kraujyje (*Petkevičienė J et al., 2012*). Vėliau tiriant didesnę imtį (1030) asmenų, šalia pakartotinai tiriamų *APOE* geno variantų (rs429358 ir rs7412) buvo vertinama ir *SCARB1* (rs5888) bei *PPARα* (rs1800206) variantų genetinė įvairovė bei ryšys su dislipidemija. Buvo nustatyti su lytimi susiję skirtumai. Vyrų grupėje *APOE* ketvirtas ir *PPARα* CG genotipai koreliavo su aterogeniniu lipidų profiliu, o *SCARB1* CT genotipas demonstravo apsauginį poveikį. Moterų grupėje *APOE* antro alelio nešiotojoms buvo nustatyti mažiausi MTL-cholesterolio kiekio šansai (*Smalinskienė A et al., 2013*). Toliau tęsiant tyrimus aterogeneze srityje, 1976 kontrolinių asmenų ir 463 asmenų, patyrusių MI, grupėse buvo tiriama *SCARB1* geno jau minėto VNP rs5888 ir plazmos lipidų profilio asociacija su KŠL. Buvo nustatyta asociacija priklausomai nuo amžiaus ir lyties. Kontrolinėje grupėje *SCARB1* rs5888 C/T genotipas vyresnių vyrų grupėje buvo susietas su priešaterogeniniu lipidų profiliu, o T/T genotipas – jaunų



moterų grupėje. *SCARB1* rs5888 TT genotipas vyresnių vyrų grupėje taip pat buvo siejamas su sumažėjusiais MI šansais (*Stanislovaitienė D et al., 2013*).

Širdies ir kraujagyslių ligų morfogenezės tyrimo duomenimis, pacientams, kuriems diagnozuotas kairiojo skilvelio remodeliavimas, *AGT* geno M235T variantas kartu su priekine infarkto sienele, padidėjusiu leukocitų kiekiu, sumažėjusiu kairiojo skilvelio išilginiu pokyčiu (angl. *left ventricular longitudinal strain*) gali padėti prognozuoti kairiojo skilvelio remodeliavimą įvykus MI (*Žaliaduonytė-Pekšienė D et al., 2013*).

Iki šiol, KŠL genetiniai tyrimai Lietuvoje buvo atliekami gana fragmentiškai, tiriant pavienius genus įvairiose tiriamųjų grupėse. Siekiant kurti ir diegti efektyvesnę diagnostiką, prognozavimą, gydymą ir prevenciją Lietuvoje, reikalingi detalesni populiacijos lygmens genetiniai tyrimai, kurie sudarys prielaidas genetinės KŠL architektūros žinių plėtojimui ir pritaikymui.

### 1.3. Genetinės įvairovės vertinimas ir ligos prognozė

*„Prognozė yra labai sudėtinga, ypač ateities“*

Niels Bohr

Vertinant genetinį polinkį sirgti KŠL, DNR testai bus kliniškai vertingi tik tuomet, kai jų informatyvumas **pasiuks ir viršys** tradicinių rizikos veiksnių (amžius, lytis, lipidų profilis, kraujospūdis, KMI, šeimos istorija, rūkymas), šiuo metu naudojamų prognozės algoritmuose, informatyvumą (*Humphries SE et al., 2010*). 2012 metų Europos kardiologų draugijos parengtose širdies ir kraujagyslių ligų prevencijos klinikinėje praktikoje rekomendacijose (*Perk J et al., 2012*) ir Amerikos širdies draugijos parengtoje 2013 metų Širdies ligos ir smūgio statistikos ataskaitoje (*Go AS et al., 2013*) įdiegti genetinius tyrimus į praktiką (ligos rizikai nustatyti) dar nėra rekomenduojama.

Perspektyva nustatyti naujus rizikos žymenis, kuriuos būtų galima panaudoti daugiaveiksnių dažnų ligų prognozei, buvo akstinas atlikti genetinių asociacijų tyrimus. Nors vis dar esame pradinėse prognozės algoritmų, apimančių genetinius veiksnius, kūrimo stadijose, ieškoma papildomų rizikos veiksnių ir kuriami modeliai, vertinantys daugelį DNR sekos variantų, aplinkos bei klinikinius rizikos veiksnius. Per 15 metų pasiektas reikšmingas postūmis prognozuojant KŠL riziką, kai iš plataus masto genomo asociacijų tyrimų nustatyti VNP rizikos vertinimo modeliuose pridėti prie tradicinių rizikos veiksnių. Šis indėlis gali būti ypač naudingas identifikuojant žemą prognozavimo indeksą turinčius asmenis, kurie iš tiesų priklauso didelės rizikos susirgti KŠL grupei, kai vien tik tradiciniai rizikos veiksniai to neaprašo (*Bolton JL et al., 2013*).

Iki šiol sukurtas ne vienas KŠL rizikos vertinimo algoritmas. Klasikinis yra *Framingham* rizikos vertinimo algoritmas (*Wilson PW et al., 1998*), tačiau jo tikslumas įvairiose populiacijose skiriasi, o Europos šalyse rizika dažnai pervertinama (*Eichler K et al., 2007*). Todėl sukurta alternatyvių modelių, apimančių papildomus rizikos veiksnius, pavyzdžiui, socialinį ekonominį rodiklį (*Hippisley-Cox J et al., 2007; Woodward M et al., 2007*), šeimos KŠL istoriją (*Martin C et al., 2008*). Tačiau jų galia tiksliai aprašyti riziką vis dėlto išlieka ribota. Pirmiau yra minėta šeimos ligos istorija, kuri yra KŠL rizikos veiksnys, svarbus sudarant prognozę. Taip pat yra žinomas faktas, kad 50 % KŠL įvairovės gali būti paaiškinama genetiniais veiksniais. Tačiau nepaisant genetinių veiksnių akivaizdumo, sukurti gerą prognozinį modelį integruojant genetinius veiksnius yra sudėtinga. Tikėtina, kad viena iš priežasčių – maži individualių VNP poveikiai ligos pasireiškimui, palyginti su tradicinių rizikos veiksnių. Ne mažiau svarbu yra pasirinkti **tinkamus** VNP ir patį prognozavimo modelį (*Humphries SE et al., 2010*).

Jau atsiranda modelių, prognozuojančių KŠL atvejus remiantis VNP, susietais su KŠL (*Bolton JL et al., 2013*). Genetiniu rizikos vertinimo modeliu apskaičiuota santykinė rizika gaunama panaši į tradicinių rizikos

veiksmių modelių apskaičiuotą riziką. Gaunama genetinė komponentė, kuri yra nepriklausoma nuo šeimos ligos istorijos (pastaroji ilgą laiką genetinės rizikos vertinimo algoritmuose buvo naudojama kaip pakaitalas). *E. Tikkanen* ir kolegų 2013 m. darbe 28 į algoritmą integruoti VNP pagerino prognozavimą lyginant su tradicinių rizikos veiksmių prognozės modeliais. Teigiama, kad genetinis testavimas ypač naudingas asmenų, kuriems būdinga vidutinė ligos rizika (t. y. 10 metų rizika yra 10–20 %), grupėje, nes iki šiol taikyti modeliai ir metodikos orientavosi į didelės rizikos grupę (t. y. 10 metų rizika >20%) (*Tikkanen E et al., 2013*). Aprašytasis modelis integruojant genetinius veiksmius – tai Suomijos populiacijoje atliktas tyrimas. Jungtinės Karalystės mokslininkai taip pat tobulina metodus ieškodami tinkamiausio rizikos prognozavimo algoritmo vietinei populiacijai (*Dent T, 2009*). Šie pavyzdžiai įrodo vietinių populiacijų genetinio KŠL tyrimo pranašumą. Dėl ligos mastų ir gydymo išlaidų yra iškilusi būtinybė atrasti tinkamiausią, optimalų KŠL rizikos vertinimo modelį. Bet ar jis gali būti universalus?.. Jau buvo pademonstruotas daugelio (pvz.: *Framingham, SCORE*) naudojamų tradicinių rizikos prognozės algoritmų neuniversalumas (*Eichler K et al., 2007; Perk J et al., 2012*). Netikslus rizikos įvertinimas lemia klaidingą asmenų priskyrimą rizikos grupei, iš ko išplaukia netinkamas asmenų, iš tiesų turinčių didelę ligos riziką, gydymas ir motyvacija arba asmenų, turinčių mažesnę riziką, neadekvatus gydymas (tiek vaistų, tiek kainos prasme). Be to, jei modelis yra tendencingas, jis gali padidinti sveikatos būklės netolygumą, pavyzdžiui, sistemingai neidentifikuoti (nuvertinti) socialiniu ekonominiu požiūriu jautrių grupių asmenų KŠL riziką (*Dent T, 2009*). Sveikatos priežiūros specialistams taip pat didėja poreikis išmanyti skirtingų tautų genetinę prigimtį spartėjančio globalizacijos proceso akivaizdoje.

### **Nuo tautos prie asmens**

Asmeninė medicina – nauja ligų prevencijos ir gydymo paradigma. Rutininis rizikos alelių genotipavimas, būtinas asmeninei medicinai, pareikalaus pokyčių visuomenėje dar anksčiau, nei bus atliekamas

visuotinai. Greičiausiai technologijos ir jų prieinamumas atlikti plataus masto genomo tyrimus visuotinai aplenks teisinius, socialinius ir etinius klausimus, kurie turės būti išspręsti, pavyzdžiui, kaip saugoti genetinius duomenis, kad asmenims nekiltų diskriminacijos pavojus gyvenimo kokybės ir sveikatos draudimo plotmėje (pagal *Roden D, 2009*). Jau dabar viso genomo sekoskaita didelio našumo sistemomis yra palyginti greitai, įperkama bei vis tobulinama ir jau netoli laikas, kai galėsime įdiegti į diagnostinę praktiką. Visuotinai tiriant jau žinomus rizikos alelius, atsivers galimybės nustatyti ir naujus genetinius rizikos veiksnius, pagilinsiančius daugiaveiksnes ligas lemiančių biologinių kelių suvokimą, ir taikinius naujoms terapijoms. Taigi technologijos poreikis ir nauda turėtų būti stimulas, reikalingas naujų terapijų, genetinio visuotinio tyrimo ir prevencijos vystymuisi. Ilgas kelias turės būti nueitas, tačiau vieną dieną tai paskatins gydymo ir prevencijos strategijų permainas.

Taigi akivaizdu, kad, plėtojantis technologijoms ir galimybėms jas eksploatuoti, buvo atlikta daug didelės apimties tyrimų, padariusių proveržį nustatant daugiaveiksnių ligų genetinę architektūrą. Iškelta ne viena hipotezė, tačiau sudėtinga iki galo kiekvieną pagrįsti. Todėl siekiant atradimus pritaikyti praktikoje reikia atlikti tolesnius fundamentinius tiek dažnų, tiek retų genomo variantų tyrimus atskirose populiacijose.

## 2. TYRIMO METODIKA

### 2.1. Tyrimo strategija

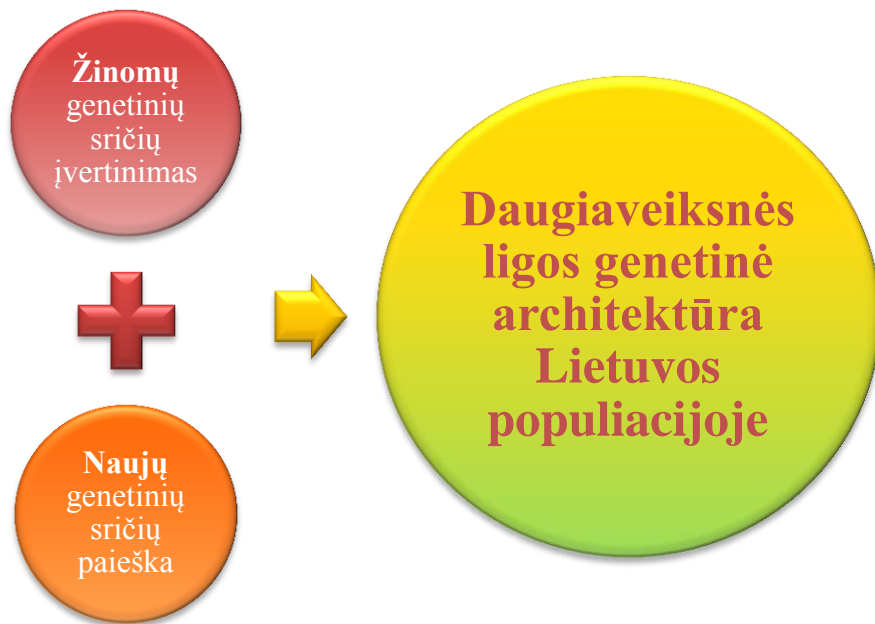
Disertacijos tema suformuluota iš Lietuvos valstybinio mokslo ir studijų fondo finansuojamo 2004–2006 metais VU vykdyto jungtinio mokslininkų projekto „Žmogaus genomo įvairovės ypatumų nulemti aterosklerozės patogenezės mechanizmai“ (akronimas ATHEROGEN; vad. prof. habil. dr. Z. A. Kučinskienė) genetinių tyrimų rezultatų.

Atsiradus galimybių atlikti platesnės apimties genomo tyrimus naujos kartos genetinėmis metodais, darbas aterosklerozės tematika buvo tęsiamas 2010–2011 metais VU MF vykdamas Lietuvos mokslo tarybos finansuojamą projektą „Lietuvos populiacijos genomine įvairove ir koronarines širdies ligos individualaus gydymo galimybes“ (akronimas LIGENKOR; vad. prof. habil. dr. Z. A. Kučinskienė).

Didžioji dalis disertacijos tyrimų yra atlikta kaip dalis 2011–2015 metais VU MF vykdomo projekto „Lietuvos populiacijos genetinė įvairovė ir sandaros kitimai, susiję su evoliucija ir dažniausiai paplitusiomis ligomis“, kuris finansuojamas kaip Visuotinės dotacijos priemonė Europos socialinio fondo ir Lietuvos Respublikos bendrojo finansavimo lėšomis (akronimas LITGEN; vad. prof. habil. dr. V. Kučinskas).

Visiems tyrimams disertacijos tema, atliktiems vykdamas pirmiau minėtus projektus, yra gauti Vilniaus regioninio biomedicininių tyrimų etikos komiteto (VRBTEK) leidimai. Projektui ATHEROGEN buvo suteiktas vietinis leidimas VU MF ŽMGK vykdomiems moksliniams tyrimams. Projektui LIGENKOR 2010 m. lapkričio 5 d. suteiktas VRBTEK pritarimas leidimo papildymui Nr. 158200-11-255-067LP2. Projektui LITGEN 2011 m. gegužės 3 d. suteiktas VRBTEK leidimas Nr. 158200-05-329-79. Tyrimams atlikti yra gautos pasirašytos asmenų, dalyvaujančių atitinkamame projekte, informuoto sutikimo formos.

Galutinė disertacijos tyrimo strategija ir darbo modelis pateikiama **2.1 paveiksle**. Tyrimą sudaro dvi dalys – naujų genetinių sričių, siejamų su liga, paieška pacientų šeimų grupėje ir žinomų genetinių sričių, siejamų su liga, analizė bendroje Lietuvos lietuvių populiacijoje.



**2.1 paveikslas.** Genetinės daugiaveiksnės ligos architektūros tyrimo strategija

**Žinomų genetinių sričių**, siejamų su KŠL, analizė bendroje Lietuvos lietuvių populiacijoje susideda iš kelių dalių. Pirmiausia, remiantis literatūra ir duomenų bazių analize yra sudaromas VNP ir reikšmingai asocijuotų su KŠL genų sąrašas. Antra, siekiant nustatyti atrinktų VNP alelius atliekamas genotipavimas. Trečia, bendroje Lietuvos lietuvių populiacijoje nustatomi atrinktų VNP alelių ir genotipų dažniai, atliekama vidupopuliacinė ir tarppopuliacinė lyginamoji analizė, nustatomas genetinis profilis, jo įvairovė bendroje Lietuvos lietuvių populiacijoje.

**Naujų genetinių sričių** paieška vykdoma atliekant bandomąjį plataus masto genomo žymenų, VNP, asociacijos su KŠL fenotipu, šeimų (triadų) tyrimą. Remiantis statistine analize, nustačius reikšmingiausias genetines sritis, t. y. potencialius genus kandidatus, asocijuotus su KŠL,

hipotezė tikrinama genetinei asociacijai patvirtinti ar paneigti. Atliekama atvejo ir kontrolės analizė pagal nustatytus potencialius genus kandidatus, t. y. padidinama tiriamų pacientų imtis ir papildomai į tyrimą įtraukiami kontroliniai asmenys iš bendros Lietuvos lietuvių populiacijos.

## **2.2. Tyrimo imtis**

Tyrimas yra sudėtinis, sudarytas iš dviejų dalių. Pirmoji dalis – atrinktų žinomų VNP rizikos alelių analizė **bendroje Lietuvos lietuvių populiacijoje**. Antroji – naujų daugiaveiksnės ligos rizikos alelių paieška Lietuvos populiacijoje atliekant **šeimų** bei **atvejo ir kontrolės** asociacijų tyrimus. Atitinkamai pagal tai sudarytos tiriamųjų imtys. Tyrimui naudota DNR, išskirta iš veninio kraujo (2.3.1 skyrelyje aprašomais metodais) VU MF Žmogaus ir medicininės genetikos katedroje.

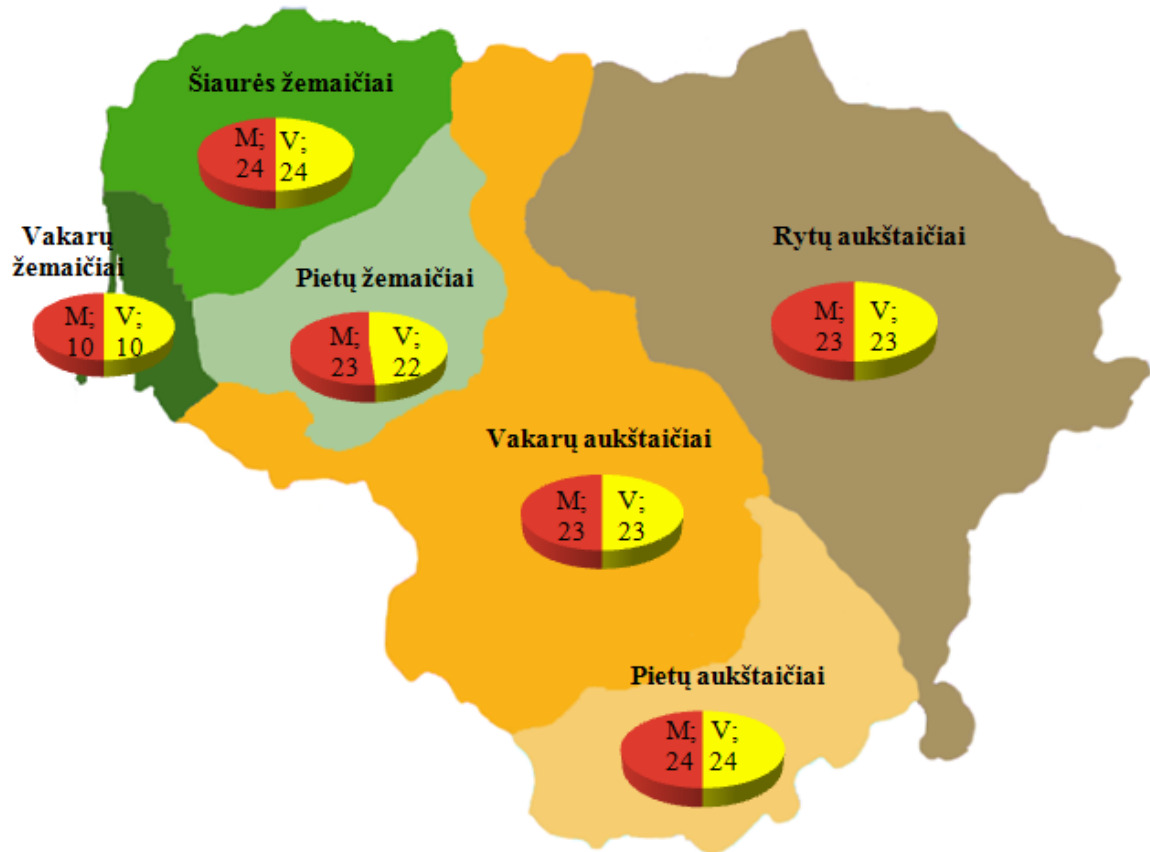
### **2.2.1. Tiriamųjų asmenų grupė plataus masto genotipavimui.**

#### **Žinomų ligos riziką lemiančių VNP genetinio profilio nustatymas bendroje Lietuvos lietuvių populiacijoje**

Tiriamųjų asmenų imtis sudaryta 2011–2013 metais, vykdant LITGEN projektą. Genetinei vidupopuliacinei ir tarppopuliacinei analizei labiausiai tinkamas yra Lietuvos suskirstymas į dvi grupes – aukštaičius ir žemaičius, kuriuos pagal tarmes galima suskirstyti dar smulčiau į šešias etnolingvistines grupes: rytų, vakarų, pietų aukštaičius ir šiaurės, pietų, vakarų žemaičius (**2.2 paveikslas**) (*Girdenis A, Zinkevičius Z, 1966*). Pagal šį suskirstymą kraujo ėminiai buvo surinkti iš:

- ✓ Viečiūnų, Alytaus, Druskininkų, Lazdijų, Elektrėnų ir Lazdijų sveikatos priežiūros centrų (SPC), esančių Pietų Aukštaitijoje;
- ✓ Birštono, Šakių, Marijampolės, Kėdainių, Birštono ir Šiaulių SPC Vakarų Aukštaitijoje;
- ✓ Širvintų, Ignalinos, Panevėžio, Rokiškio ir Molėtų SPC Rytų Aukštaitijoje;
- ✓ Raseinių, Tauragės, Šilalės, Varnių, Kelmės SPC Pietų Žemaitijoje;

- ✓ Šilutės, Plikių ir Priekulės SPC Vakarų Žemaitijoje;
- ✓ Mažeikių, Naujosios Akmenės, Skuodo, Palangos, Telšių ir Gargždų SPC Šiaurės Žemaitijoje.



**2.2 paveikslas.** Bendrą Lietuvos lietuvių populiaciją reprezentuojanti tiriamųjų imtis. Skritulinėse diagramose nurodytas moterų (M) ir vyrų (V) skaičius. Žemėlapyje Lietuva suskirstyta į šešias etnolingvistines grupes (pagal A. Girdenio ir Z. Zinkevičiaus tarmių klasifikaciją). Žemėlapio šabloną parengė VU MF ŽMGK darbuotojai (pagal *Morkūnas K, 1977, žemėlapių Nr.1*).

Kaip matyti **2.2 paveiksle**, iš kiekvieno regiono surinkta po 45–48 mėginius, išskyrus, Vakarų Žemaitiją, kurioje surinkta 20 mėginių. Tačiau atsižvelgiant į tai, kad Vakarų Žemaitija yra mažiausia užimamu plotu, galima daryti prielaidą, kad kartu su kitų regionų imtimis, bendra Lietuvos lietuvių populiacija genetinei struktūrai tirti yra reprezentuojama tinkamai. Dėl gyventojų mobilumo tarp etnolingvistinių regionų yra sudėtinga įvertinti



lietuvių populiacijos struktūrą pagal etnolingvistines grupes (*Jakaitienė A, Kučinskas V, 2013*). Tačiau atkreiptinas dėmesys, kad šiame darbe buvo siekiama atrinkti tokius asmenis, kurių mažiausiai trys kartos gyveno tame pačiame regione. Todėl imtis reprezentuoja lietuvių populiacijos struktūrą pagal etnolingvistines grupes, o ne lietuvių, gyvenančių etnolingvistinėse teritorijose, struktūrą.

Tyrimui atrinkti 253 negiminingi asmenys iš visų šešių etnolingvistinių grupių, iš kurių 113 žemaičių (20 vakarų, 48 šiaurės, 45 pietų) ir 140 aukštaičių (46 rytų, 46 vakarų, 48 pietų) arba skirstant pagal lytį – 127 moterys ir 126 vyrai.

### **2.2.2. Tiriamųjų asmenų grupė plataus masto genotipavimui.**

#### **Bandomoji asociacijų analizė naujų genetinių sričių paieškai**

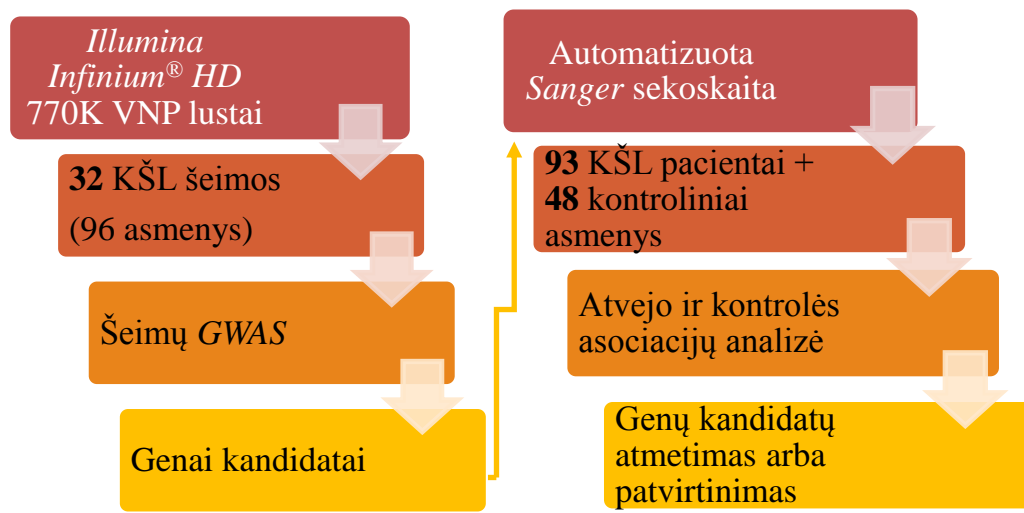
Asociacijų tyrimui atrinkti reikalingi asmenys, kurie turi tiriamą ligą ar požymį, taip pat tie, kurie tokio požymio neturi. Šiam tyrimui pasirinktas prieššlaikinės KŠL (angl. *premature coronary heart disease*) fenotipas (toliau KŠL).

Šią tyrimo dalį sudaro dar dvi dalys, nes plataus masto genomo asociacijų tyrimu nustatčius statistiškai reikšmingai asocijuotas genetines sritis reikia patvirtinti. Tad šiam tyrimui strategiškai buvo reikalingos kelios skirtingos imtys (**2.4 paveikslas**). Iš viso reikalingos ir atrinktos trys imtys: KŠL šeimų grupė, kurią sudaro tėvai ir vaikai (probandas yra vaikas, kuris serga KŠL); sergančių KŠL pavienių asmenų grupė ir kontroliniai bendros populiacijos asmenys, kurie neserga KŠL.

**KŠL šeimų grupė.** Šią grupę sudaro asmenys, sergantys KŠL (probandai), ir jų tėvai. Iš 32 probandų viena yra moteris, 31 vyras. 64 tėvų grupę sudarė 32 vyrai ir 32 moterys.

**KŠL sergančių pavienių asmenų grupė.** Šią grupę sudaro 32 KŠL sergantys asmenys iš šeimų grupės ir papildomai 61 KŠL sergantis asmuo. Iš viso 86 vyrai ir 7 moterys.

**Kontrolinių asmenų grupė.** Šią grupę sudaro 48 bendros Lietuvos lietuvių populiacijos asmenys (po 8 asmenis iš kiekvienos (iš šešių esamų) etnolingvistinės grupės), 24 vyrai ir 24 moterys, kurie neserga KŠL.



**2.4 paveikslas.** Naujų daugiaveiksnės ligos rizikos alelių paieškos modelio pavyzdys. Šio tyrimo metu naudotos imtys ir metodai, skirti naujų genetinių KŠL veiksmų paieškai. Geltona rodyklė vaizduoja eksperimentų seką. Šeimų GWAS – plataus masto genomo asociacijų analizė taikant nepusiausviro perdavimo testą šeimų imtyje, kurią sudaro tėvai ir vaikas, sergantis KŠL.

**Koronarinės širdies ligos fenotipas.** Asmenų klinikinis ištyrimas buvo atliekamas VU MF Kardiologijos ir angiologijos klinikoje. Biocheminis fenotipas tirtas VU MF Fiziologijos, biochemijos, mikrobiologijos ir laboratorinės medicinos katedros klinikinėje bazėje – VUL SK Laboratorinės medicinos centre. Klinikinis KŠL fenotipas buvo vertinamas atliekant antropometrinius matavimus, klinikius, instrumentinius ir laboratorinius biocheminius tyrimus. Renkant anamnezę taip pat vertinta informacija apie KŠL rizikos veiksmus (rūkymą, alkoholio vartojimą, mitybos bei fizinio aktyvumo įpročius ir kt.), kitas ligas ir vaistų vartojimą.

Asmenys buvo **įtraukti** į tyrimą, jei tenkino šiuos reikalavimus:

- ✓ vyrai nuo 35 iki 55 metų, moterys nuo 35 iki 65 metų;
- ✓ asmenys pirmą kartą gyvenime hospitalizuoti į intensyviosios kardiologijos skyrių, susirgę ūminiu koronariniu sindromu:
  - ✓ miokardo infarktu su arba be Q bangos (patvirtintu būdingais elektrokardiografiniais pokyčiais ir kardiосpecificinių žymenų padidėjimu);
  - ✓ nestabilią krūtinės angina (patvirtinta būdingais elektrokardiografiniais pokyčiais);
- ✓ yra pagrindas manyti, kad asmeniui praeityje nustatyta KŠL. Pagrindą įtarti KŠL sudaro šie faktai: įvykęs MI, atlikta perkutaninė koronarinė angioplastija, atliktas koronarinių arterijų šuntavimas ar koronarinė angiograma, kurioje matoma hemodinamiškai reikšminga stenozė.

Asmenys **neįtraukti** į tyrimą, jei:

- ✓ insulinu gydomas cukrinis diabetas;
- ✓ turi inkstų funkcijos nepakankamumą;
- ✓ serga III–IV funkcinės klasės širdies nepakankamumu;
- ✓ serga onkologinėmis ligomis (išskyrus odos bazaliomą);
- ✓ turi alkoholizmą ir kitas socialines priežastis, galinčias turėti įtakos tyrimui.

Anatominiai kraujagyslių pokyčiai patvirtinti neinvaziniais metodais vertinant aterosklerozinių plokštelių būklę, atliekant arterijų standumo ir endotelio funkcijos tyrimus. Laboratoriniais biocheminiais tyrimais buvo vertinami metaboliniai ir uždegiminiai žymenys, dalyvaujantys aterosklerozės procese: CRP, hemoglobinas A<sub>1C</sub>, lipoproteinas a, apolipoproteinai A1, B ir jų santykis (ApoB/A1), oksiduoti MTL, homocisteinas, interleukinas-6, gliukozės kiekis plazmoje, K<sup>+</sup>, Na<sup>+</sup>, šlapalas, kreatininas.

Molekuliniais genetiniais metodais, išskyrus tiriamųjų DNR, pagal pirmiau pateiktą tyrimo strategiją toliau kiekvienai sudarytai imčiai buvo atlikti tyrimai, kurie aprašyti toliau.

### **2.3. Tyrimo eiga ir metodai**

Siekiant darbo vientisumo, detalios tiriamojo darbo metodikos pateikiamos prieduose.

Dirbant laboratorijoje yra laikomasi geros laboratorinės praktikos, išorinio kokybės kontrolės vertinimo, laboratorijos priežiūros, parengimo darbui laboratorijoje ir protokolų reikalavimų.

**Pirma darbo dalis.** Siekiant nustatyti bendros Lietuvos lietuvių populiacijos genetinį KŠL žymenų profilį ir įvertinti jų genetinę įvairovę, 253 asmenims VU MF ŽMGK atliktas plataus masto genotipavimas *Illumina Infinium® HD* VNP lustais *HumanOmniExpress-12 v1.1* (719666 VNP) ir tolesnė bioinformacinė statistinė pirminių duomenų kokybės analizė *GenomeStudio v2011.1* programa (*Illumina Inc., US*) bei tolesni vidupopuliaciniai ir tarpopuliaciniai genomo įvairovės ir genomo analizės tyrimai.

**Antra darbo dalis.** Naujų genetinių KŠL veiksnių paieškai atlikta plataus masto genomo asociacijos analizė KŠL šeimose. Trisdešimt dviejų šeimų asmenims (n=96) Tartu universiteto Molekulinės ir ląstelės biologijos institute Estijoje atliktas plataus masto genotipavimas *Illumina Infinium® HD* VNP lustais *HumanOmniExpress-12 v1.0* (733202 VNP). Tolesnė bioinformacinė-statistinė pirminių duomenų kokybės analizė bei šeimų plataus masto asociacijos analizė atlikta VU MF ŽMGK *BC/Gene (Biocomputing Platforms Ltd., FI)* duomenų valdymo platforma.

Potencialiems genams kandidatams patvirtinti atlikta atvejo ir kontrolės asociacijos analizė. Devyniasdešimt trims KŠL sergantiems pavieniams asmenims ir 48 kontroliniams asmenims atlikta potencialių genų

kandidatų (*RTN4* ir *FBXL17*) sekoskaita ir nustatytų pokyčių atvejais atlikta asociacijų analizė bei tolesni statistiniai skaičiavimai.

### **2.3.1. DNR išskyrimas ir DNR koncentracijos bei švarumo nustatymas**

Atliekant visus genetinius tyrimus buvo naudota DNR, išskirta VU MF ŽMGK darbuotojų iš veninio kraujo ėminių. DNR buvo išskiriama dviem metodais:

fenolio ir chloroformo išskyrimo metodu pagal VŠĮ VUL SK MGC patvirtintą diagnostikos protokolą „DBR išskyrimas fenolio chloroformo metodu“ (Generalinio direktoriaus įsakymas Nr. 461);

automatizuota robotine sistema *TECAN Freedom EVO<sup>®</sup> 200* (gamintojas *Tecan Schweiz AG, CH*) magnetinių dalelių prijungimo prie DNR metodu pagal prietaiso gamintojo pateikiamus protokolus ir DNR išskyrimo rinkinio techninį aprašą (žr. literatūros sąrašą).

DNR koncentracija ir švarumas buvo nustatoma spektrofotometru *NanoDrop<sup>®</sup>*. DNR koncentracijos ir švarumo nustatymo protokolas parengtas pagal *NanoDrop<sup>®</sup>* prietaiso gamintojo pateikiamą vartotojo vadovą *NanoDrop<sup>®</sup> ND-1000 Spectrophotometer User Manual*.

Išsamūs DNR išskyrimo ir DNR koncentracijos bei švarumo nustatymo protokolai pateikiami **I priede**.

### **2.3.2. Genotipavimas *Illumina Infinium<sup>®</sup> HD VNP* lustais**

Protokolas skirtas *Infinium<sup>®</sup> HD* technologija genotipuoti VNP ant lustų, naudojant *Illumina HiScan<sup>™</sup> SQ* naujos kartos didelio našumo genomo analizės sistemą.

Protokolas parengtas pagal gamintojo pateiktas rekomendacijas: vartotojo vadovą (angl. *Infinium<sup>®</sup> HD Assay Ultra Manual Experienced User Card*); išplėstinį protokolo vadovą (angl. *Infinium<sup>®</sup> HD Assay Ultra Protocol Guide*); laboratorijos paruošimo ir procedūrų aprašą (angl. *Illumina Infinium<sup>®</sup>*

*Assay Lab Set Up and Procedures*). Protokolo vykdymo metu pildoma gamintojo rekomenduojama protokolo eigos forma (angl. *Infinium<sup>®</sup> HD Assay Ultra Lab tracking Form*).

Išsamus protokolas yra pateikiamas **II priede**. Genotipavimą atliko VU MF ŽMGK darbuotojai (tarp jų ir doktorantė).

Genotipavimas *Illumina Infinium<sup>®</sup> HD VNP* lustais *HumanOmniExpress-12 v1.0* Tartu universiteto Molekulinės ir ląstelės biologijos institute Estijoje vykdytas V. Soo ir doktorantės pagal gamintojo pateikiamą protokolą (*Infinium<sup>®</sup> Multi-Use Assay, Manual Protocol, 15013850 Rev. A*), beveik identišką protokolui, kuris naudotas genotipuojant *HumanOmniExpress-12 v1.1* lustais VU MF ŽMGK.

### **2.3.3. Polimerazės grandininė reakcija (PGR) ir sekoskaita**

Polimerazės grandininė reakcija buvo atliekama siekiant pagausinti bandomojo tyrimo (VNP genetinės asociacijos su KŠL) metu nustatytų potencialių genų kandidatų koduojančią DNR seką ir toliau atlikti sekoskaitą sekos pokyčiams nustatyti. Išsamus PGR ir sekoskaitos protokolas (apima PGR, elektroforezę PGR produktui patikrinti, PGR produkto valymą prieš sekoskaitos reakciją, sekoskaitos reakciją, sekoskaitos reakcijos produkto valymą ir kapiliarinę elektroforezę nukleotidų sekai nustatyti) bei potencialių genų kandidatų oligonukleotidinių pradmenų sekos pateikiamos **III priede**. Oligonukleotidiniai pradmenys sukurti naudojantis internetinėmis programomis *Primer 3 v4.0.0* (*Untergrasser A et al., 2012; Koressaar T et Remm M, 2007*) ir *Primer-BLAST* (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

### **2.3.4. Statistinė ir bioinformacinė duomenų ir rezultatų analizė**

Toliau aprašomuose skyreliuose nurodyti duomenų statistinės ir bioinformacinės geno analizės metodai suskirstyti pagal tyrimo strategiją, nurodytą **2.1 skyrelyje**. Aprašomų tapačių analizės metodų įvairovė ir

naudotos tos pačios programinės įrangos versijų netapatumas yra ketverių metų laikotarpiu vykdyto darbo pasekmė, tačiau rezultatų kokybei įtakos neturi.

#### **2.3.4.1. Ligos riziką lemiančių VNP genetinės įvairovės analizės bendroje Lietuvos lietuvių populiacijoje metodai**

##### **Genotipavimo *Illumina Infinium*<sup>®</sup> VNP lustais duomenų analizė ir kokybės procedūros *GenomeStudio v2011.1* programine įranga**

*GenomeStudio v2011.1* programinė įranga (*Illumina*<sup>®</sup> *GenomeStudio 2011 (c) Illumina, Inc. 2003–2011*) yra prieinama įsigijus *Illumina Inc., US* gamintojo (toliau *Illumina*) prietaisą *HiScanSQ*<sup>™</sup> ir suvedus pateikiamą specifinį genotipavimo ar kitam naudojamam moduliui skirtą vartotojo raktą iš 16 skaitmenų.

*GenomeStudio*<sup>™</sup> *Genotyping Module v1.9* yra skirtas analizuoti genotipavimo duomenims, gautiems atliekant genotipavimą *Illumina* gaminamų *Infinium*<sup>®</sup> ir *GoldenGate*<sup>®</sup> DNR lustais.

Genotipavimo duomenys, gauti *Infinium*<sup>®</sup> VNP lustais, yra analizuojami pagal *Illumina* pateikiamą aktualų *GenomeStudio*<sup>™</sup> Genotipavimo modulio vartotojo vadovą *Genome Studio*<sup>™</sup> *Genotyping Module v1.0 User Guide* ir atsižvelgiant į pateiktas pataisas *GenomeStudio*<sup>™</sup> *Software v2011.1 Release Notes* (prieiga internete <http://support.illumina.com/array/arraysoftware/genomestudio/downloads.ilmn>). Tyrimo projekto sukūrimo, duomenų kokybės kontrolės vertinimo ir įvesties bylų kūrimo trečių šalių programoms aprašai yra pateikiami **IV priede**.

##### **Kiti bioinformaciniai statistiniai metodai**

Hardžio ir Vainbergo pusiausvyros nustatymas, genotipų ir alelių dažniai, įvesties bylos *Haploview v4.2* programai atliktos viso genomo asociacijų analizės programiniu atviros prieigos įrankių paketu *PLINK v1.07* (*Purcell S et al., 2007*; prieiga internete <http://pngu.mgh.harvard.edu/purcell/plink/>).

**Hardžio ir Vainbergo pusiausvyra** nustatoma tiksliau testu (*Wigginton JE et al., 2005*) naudojant *PLINK v1.07* programos funkciją *-hardy*. Reikšmingumo lygmuo pasirinktas 0,001. Jei tiriamosios imties VNP aleliai ir genotipai nėra Hardžio ir Vainbergo pusiausvyroje,  $p \leq 0,001$ , tai gali būti ženklas apie prastą genotipavimo kokybę bei populiacijos stratifikaciją arba indikuoti kitus populiaciją veikiančius procesus – migraciją, neatsitiktinį partnerių pasirinkimą, mutagenezę, gamtinę atranką ir kt. Jei VNP genotipai ir alelių dažniai yra Hardžio ir Vainbergo pusiausvyroje, vadinasi, pasiskirsto pagal dėsnį:  $p^2 + 2pq + q^2 = 1$  ir  $p + q = 1$ , čia  $p$  – vieno alelio dažnis,  $q$  – alternatyvaus alelio dažnis,  $p^2$  – vieno alelio homozigotų dažnis,  $q^2$  – alternatyvaus alelio homozigotų dažnis,  $2pq$  – heterozigotų dažnis. **Genotipų ir alelių dažniai** nustatyti naudojant *PLINK v1.07* programos funkcijas *-freq* bei *-hardy*.

**Haplotipų analizė** atlikta programa *Haploview v4.2* (*Barrett JC et al., 2005*). Pagal tiriamos populiacijos genotipus ir pateiktas VNP grupes su VNP koordinatėmis genome tiriamai populiacijai ir jos grupėms pagal lytis ir etnolingvistines grupes sukonstruoti haplotipai, apskaičiuoti jų dažniai, įvertinta sankibos pusiausvyra, nustatyti haplotipų blokai. Parametrai *Haploview v4.2* programoje: Hardžio ir Vainbergo pusiausvyros mažiausia reikšmė 0,001; mažiausias genotipų skaičius 75 %; leistinas mendelinių klaidų skaičius 1; mažiausia retojo alelio dažnio (MAF) reikšmė 0,001. Haplotipai yra apskaičiuojami ir sudaromi pagal akceleruotą tikėtimumo maksimizavimo (angl. *expectation maximization* (EM)) algoritimą panašų į skaldymo / jungimo metodą (angl. *partition / ligation method*), aprašytą *Qin ZS et al., 2002*. Iš nežinomos fazės genotipų duomenų didžiausio tikėtimumo (angl. *maximum likelihood*) principu ypač tiksliai apskaičiuojami žinomos fazės haplotipų dažniai populiacijoje. Haplotipų blokai konstruojami pagal  $D'$  reikšmės (santykinis nepusiausvyros koeficientas) 95 % pasikliauties intervalą (*Gabriel SB et al., 2002*), pagal kuri kiekvienas palyginimas yra klasifikuojamas į grupes „stipri nepusiausvira sankiba“, „neinformatyvu“, „stipri



rekombinacija“. Haplotipų blokai sukuriami, kai daugiau negu 95 % informatyvių palyginimų priskiriami grupei „stipri nepusiausvira sankiba“, o į analizę įtraukiami tik tie VNP, kurių MAF yra didesnis negu 0,05.

Laisvos prieigos programiniu *R* paketu *v3.0.3* (*R Development Core Team, 2008*; prieiga internete <http://www.R-project.org>) atlikti skaičiavimai pateikiami **2.1 lentelėje**.

**2.1 lentelė.** *R* paketu *v3.0.3* atlikti statistiniai skaičiavimai

Statistika		Pastabos	<i>R</i> paketo komandinė eilutė
Aprašomoji statistika	Aritmetinis vidurkis	-	numSummary(RA\$apiskirst[, "RA.sk"], statistics=c("mean", "sd", "IQR", "quantiles", "cv"), quantiles=c(0,.25,.5,.75,1))
	Moda	-	
	Mediana	-	
	Kvartiliai	-	
$\chi^2$ kriterijus	Pirsono	-	library(abind, pos=4) .Table <- matrix(c(x1,x2,x3,y1,y2,y3), 2, 3, byrow=TRUE) rownames(.Table) <- c('1', '2') colnames(.Table) <- c('1', '2', '3') .Table # Counts .Test <- chisq.test(.Table, correct=FALSE).Test.Test\$expected # Expected Counts round(.Test\$residuals^2, 2) # Chi-square Components remove(.Test) fisher.test (.Table) remove(.Table)
	Fišerio	-	
Skirstinio normalumo testas	Grafinis būdas	Asmeniui tenkančių rizikos alelių (RA) skirstinio normalumo patikrinimas	with(RA, Hist(Rizikos_aleliu_skaiciu s, scale="density", breaks="Sturges", col="darkgray", ylim=c(0,0.08))) cur ve(dnorm(x, mean=mean(RA\$Riziko s_aleliu_skaicius), sd=sd(RA\$Riziko _aleliu_skaicius)), add=TRUE, col=" black", lwd=2))
<i>t</i> testas	Dviejų nepriklausomų imčių vidurkių palyginimas	Asmeniui tenkančių RA vidurkių palyginimas vyrų ir moterų, aukštųjų ir žemųjų grupėse	> t.test(RA.sk~Lytis, alternative='two.sided', conf.level=.95, var.equal=TRUE, + data=RA) >t.test(RA.sk~Etnolingvi stinė_grupė, alternative='two.sided', conf.level=.95, + var.equal=TRUE, data=RA)
<i>F</i> testas	Dviejų dispersijų palyginimas	Asmeniui tenkančių RA dispersijos palyginimas vyrų ir moterų, aukštųjų ir žemųjų grupėse	>var.test(RA.sk ~ Lytis, alternative='two.sided', conf.level=.95, data=RA\$apiskirstymas) > var.test(RA.sk ~ Etnogr, alternative='two.sided', conf.level=.95, + data=RA\$apiskirstymas)

Grafiniam rezultatų atvaizdavimui taip pat naudotas *R* paketas bei *Microsoft Office Excel 2007* programa.

### 2.3.4.2. Bandomosios asociacijų analizės, naujų genetinių sričių paieškos metodai

Šeimų plataus masto genotipavimo duomenys perkelti į duomenų valdymo platformą *BC/Gene* (*Biocomputing Platforms Ltd., FI*). Šioje platformoje integruotu programiniu paketu *PLINK v1.07* (*Purcell S et al., 2007*) atlikta pirminių **duomenų kokybės kontrolė**, kokybės kriterijai pateikiami **2.2 lentelėje**.

**2.2 lentelė.** *BC/Gene* integruotu *PLINK v1.7* programiniu paketu atliktos genotipavimo rezultatų kokybės analizės kriterijai

Kriterijus	Santrumpa programoje	Slenkstinė vertė	Pageidaujam a reikšmė
Retojo alelio dažnis	<i>MAF</i>	0,01	> 0,01
Trūkstamų genotipų pagal VNP dažnis (angl. <i>missingness per marker</i> )	<i>GENO</i>	0,1	< 0,1
Hardžio ir Vainbergo pusiausvyros reikšmingumo lygmuo	-	0,0005	> 0,0005
Trūkstamų genotipų pagal tiriamuosius (angl. <i>missingness per individual</i> ) dažnis	<i>MIND</i>	0,05	< 0,05

Naudojant funkciją *Family GWAS* po kokybės kontrolės likusiems tiriamųjų VNP atliktas **nepusiausviro perdavimo testas** (*Spielman RS, Ewens WJ, 1996*) VNP ir fenotipo asociacijai nustatyti. Daugybiniam palyginimui naudota pataisa – adaptyvi kėlinių (angl. *adaptive permutation*) procedūra, kur maksimalus kėlinių skaičius  $10^6$ . Asociacijai nustatytas reikšmingumo lygmuo yra  $10^{-4}$ .

Reikšmingai asocijuotiems VNP programiniu R paketu v2.15.3 (*R Development Core Team, 2008*) buvo apskaičiuota statistinė empirinė galia, ŠS ir 95 % pasikliauties intervalas. Šansų santykis išreikštas perduotų rizikos alelių skaičiaus santykiu su perduotų alternatyvių alelių skaičiumi. Pasikliauties intervalas apskaičiuotas pagal formulę  $(L, U) = (\exp\{T-1,96SE\}, \exp\{T+1,96SE\})$ , kur L ir U yra 95 % pasikliauties intervalo ribinės reikšmės, T – logaritmuota ŠS reikšmė, SE – standartinė logaritmuotų ŠS reikšmių paklaida, apskaičiuojama pagal formulę  $SE = \sqrt{(1/b + 1/c)}$ , kur b ir c atitinkamai perduotų alelių skaičiai. Statistinė empirinė galia apskaičiuota pagal informatyvius perduotus alelius naudojant binominio skirstinio aproksimaciją (*Germanas Š, Jakaitienė A, 2012*).

Tęsiant potencialių genų kandidatų patvirtinimo tyrimą, sekoskaitos metodu nustatytiems VNP atlikta atvejo ir kontrolės asociacijų analizė naudojant Pirsono arba Fišerio  $\chi^2$  kriterijų, apskaičiuotą R programiniu paketu v3.0.3 (*R Development Core Team, 2008*). Prieš atliekant asociacijų analizę, nustatytiems VNP atliktas Hardžio ir Vainbergo pusiausvyros testas ir alelių dažnio skaičiavimas naudojant skaičiuoklę internete (*Rodriguez S et al., 2009*; prieiga internete <http://www.oege.org/software/hardy-weinberg.html>). Pasirinktas reikšmingumo lygmuo 0,001. Genotipų dažniai apskaičiuoti tiesiogiai. Šansų santykis ir rizikos santykis su 95 % pasikliauties intervalu apskaičiuoti pagal *Altman DG, 1991* naudojantis *MedCalc v13.1.2* programine įranga (*MedCalc Software, Belgium*; prieiga internete <http://www.medcalc.org>).

### **2.3.4.3. Rezultatų analizei naudota programinė įranga ir duomenų bazės internete**

#### **Programinė įranga**

1. Oligonukleotidinių pradmenų PGR reakcijai kūrimo programinė įranga internete *Primer 3 v4.0.0* (*Untergrasser A et al., 2012; Koressaar T et Remm M, 2007*).

2. Oligonukleotidinių pradmenų PGR reakcijai kūrimo programinė įranga internete *Primer-BLAST* (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>).
3. Genomo analizės programinė įranga, Genotipavimo modulis *Illumina GenomeStudio™ v2011.1 Genotyping Module v1.9*.
4. Viso genomo asociacijų analizės programinis paketas *PLINK v1.07* (Purcell S et al., 2007; prieiga internete <http://pngu.mgh.harvard.edu/purcell/plink/>).
5. Statistinės ir grafinės duomenų analizės paketas *R v2.15.3 ir v3.0.3* (R Development Core Team, 2008; prieiga internete <http://www.r-project.org/>).
6. Statistinės ir grafinės duomenų analizės paketas *Microsoft Office Excel 2007*.
7. Haplotipų analizės programa *Haploview v4.2* (Barrett JC et al., 2005).
8. DNR ir baltymų sekų palyginimo įrankis internete *Clustal W2* (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).
9. Hardžio ir Vainbergo skaičiuoklė internete (Rodriguez S et al., 2009; prieiga internete <http://www.oege.org/software/hardy-weinberg.html>).
10. Statistinės duomenų analizės programinė įranga internete *MedCalc v13.1.2* (MedCalc Software, Belgium; prieiga internete <http://www.medcalc.org>).
11. Duomenų valdymo platforma *BC/Gene* (Biocomputing Platforms Ltd., FI).
12. Funkcinės baltymų sąveikos analizės programinė įranga internete *STRING v9.1* (Franceschini A et al., 2013; prieiga internete <http://string-db.org/>).

### **Duomenų bazės internete**

1. Baltymų duomenų bazė *UniProtKB/Swiss-Prot* <http://www.uniprot.org/>
2. Biotechnologinės informacijos nacionalinio centro – *NCBI* – duomenų bazės *Gene*, *dbSNP*, *PubMed*, *1000 Genomes*, *MedlinePlus* <http://www.ncbi.nlm.nih.gov/>
3. Genomo duomenų bazė *Ensembl* <http://www.ensembl.org/index.html>
4. Genų ontologijos duomenų bazė *The Gene Ontology (GO)* <http://www.geneontology.org/>

### 3. TYRIMO REZULTATAI

#### 3.1. Genų ir VNP, susijusių su ligos genetinė architektūra, atranka

Šiam tyrimui kaip daugiaveiksnių ligos modelis buvo pasirinkta KŠL. Jos genetinės architektūros tyrimo pavyzdžiu galima analogiškai tirti ir kitas daugiaveiksnes ligas bendroje populiacijoje. Genetinei struktūrai ir įvairovei nustatyti yra svarbu tinkamai pasirinkti ligą lemiančias genetines sritis ir genetinius variantus. Šiam tyrimui genai ir VNP buvo atrinkti remiantis naujausių mokslinių literatūros šaltinių analize: (1) *Correia M et al., 2013*; (2) *Deloukas et al., 2013*; (3) *Drenos et al., 2007*; (4) *Dumitrescu L et al., 2011*; (5) *Humphries SE et al., 2010*; (6) *Peloso GM et al., 2010*; (7) *Roberts R, Stewart AFR, 2012*. Išvardintieji tyrimai, nėra pirminiai, jie atkartoja ir patvirtina jau nustatytas asociacijas, taip pat apžvelgia ir nustato naujas statistiškai reikšmingas asociacijas didelėse imtyse. Todėl šiam tyrimui atlikti pasirinkti **tik** tie VNP, kurie buvo statistiškai reikšmingai susiję su KŠL ir ją lemiančiais veiksniais – lipidų lipoproteinograma bei MI, o VNP sąsaja atkartota daugiau nei viename nepriklausomame didelės imties plataus masto genomo tyrime. Tyrimui atrenkant VNP, kurie svarbūs lemiant KŠL, buvo analizuota ne tik literatūra, bet ir duomenų bazės *Ensembl*, *NCBI Gene*, *dbSNP*, *1000 Genomes*. **Lentelėje 3.1** pateikiamos atrinktos su liga siejamos 46 genetinės sritys ir (ar) genai bei jas reprezentuojantys 60 VNP, nurodyti ligos riziką lemiantys aleliai. Taip pat pateikiami literatūros šaltiniai, kuriais remtasi. Atrinktų VNP dažniai tiriamoje Lietuvos lietuvių populiacijoje pateikiami toliau.

#### 3.1 lentelė. Tyrimui atrinkti ligos riziką lemiantys VNP aleliai

Genai arba genetinės sritys	VNP	Rizikos alelis*	Šaltinis
<i>APOB</i>	rs1042031	A	<i>Drenos F et al., 2007</i>
<i>9p21</i>	rs10757274	G	<i>Correia M et al., 2013</i>
<i>RXRA</i>	rs11185660	A	<i>Peloso GM et al., 2010</i>
<i>LIPA</i>	rs11203042	A	<i>Deloukas P et al., 2013</i>
<i>PCSK9</i>	rs11206510	A	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>

<i>3.1 lentelės tęsinys</i>			
<b>Genai arba genetinės sritys</b>	<b>VNP</b>	<b>Rizikos alelis*</b>	<b>Šaltinis</b>
<b>ZC3HC1</b>	rs11556924	G	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>
<b>PCSK9</b>	rs11591147	A	<i>Dumitrescu L et al., 2011</i>
<b>GCKR</b>	rs1260326	A	<i>Dumitrescu L et al., 2011</i>
<b>LIPA</b>	rs1412444	A	<i>Roberts R, Stewart AFR, 2012</i>
<b>ABCC6</b>	rs150468	A	<i>Peloso GM et al., 2010</i>
<b>UBE2Z</b>	rs15563	G	<i>Deloukas P et al., 2013</i>
<b>MLXIPL</b>	rs17145738	A	<i>Dumitrescu L et al., 2011</i>
<b>SMAD3</b>	rs17228212	G	<i>Humphries SE et al., 2010</i>
<b>FADS1</b>	rs174547	A	<i>Dumitrescu L et al., 2011</i>
<b>MIA3</b>	rs17464857	A	<i>Deloukas P et al., 2013</i>
<b>MIA3</b>	rs17465637	C	<i>Roberts R, Stewart AFR, 2012; Humphries SE et al., 2010</i>
<b>FURIN/FES</b>	rs17514846	A	<i>Deloukas P et al., 2013</i>
<b>MTHFR</b>	rs1801133	A	<i>Drenos F et al., 2007</i>
<b>HDAC9</b>	rs2023938	G	<i>Deloukas P et al., 2013</i>
<b>SLC22A3/LPAL2/LPA</b>	rs2048327	G	<i>Deloukas P et al., 2013</i>
<b>TOMM40 (APOE/APOC1)</b>	rs2075650	G	<i>Deloukas P et al., 2013</i>
<b>LPL</b>	rs2197089	A	<i>Dumitrescu L et al., 2011</i>
<b>ZEB2</b>	rs2252641	G	<i>Deloukas P et al., 2013</i>
<b>HNFI1A/C12orf43</b>	rs2259816	A	<i>Humphries SE et al., 2010</i>
<b>LCAT</b>	rs2271293	A	<i>Dumitrescu L et al., 2011</i>
<b>MRAS</b>	rs2306374	G	<i>Roberts R, Stewart AFR, 2012</i>
<b>9p21</b>	rs2383206	G	<i>Preuss M et al., 2010; Correia M et al., 2013</i>
<b>9p21</b>	rs2383207	G	<i>Correia M et al., 2013</i>
<b>LIPC</b>	rs261332	A	<i>Dumitrescu L et al., 2011</i>
<b>USF1</b>	rs2774279	G	<i>Correia M et al., 2013</i>
<b>SH2B3</b>	rs3184504	A	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>
<b>CDKN2BAS1</b>	rs3217992	A	<i>Deloukas P et al., 2013</i>
<b>CETP</b>	rs3764261	A	<i>Dumitrescu L et al., 2011</i>
<b>PLG</b>	rs4252120	A	<i>Deloukas P et al., 2013</i>
<b>APOE/APOC4</b>	rs4420638	A	<i>Dumitrescu L et al., 2011</i>
<b>APOE</b>	rs445925	G	<i>Deloukas P et al., 2013</i>
<b>COL4A1/COL4A2</b>	rs4773144	G	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>
<b>LIPC</b>	rs4775041	C	<i>Dumitrescu L et al., 2011</i>
<b>LIPC</b>	rs4775065	A	<i>Peloso GM et al., 2010</i>
<b>IL6R</b>	rs4845625	A	<i>Deloukas P et al., 2013</i>
<b>CDKN2A/CDKN2B</b>	rs4977574	G	<i>Roberts R, Stewart AFR, 2012; Correia M et al., 2013</i>
<b>CXCL12</b>	rs501120	A	<i>Humphries SE et al., 2010; Deloukas P et al., 2013</i>
<b>APOB</b>	rs562338	A	<i>Dumitrescu L et al., 2011</i>

3.1 lentelės tęsinys			
Genai arba genetinės sritys	VNP	Rizikos alelis*	Šaltinis
<i>ABO</i>	rs579459	G	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>
<i>CETP</i>	rs5882	A	<i>Peloso GM et al., 2010</i>
<i>ITGB3</i>	rs5918	G	<i>Drenos F et al., 2007</i>
<i>SORT1</i>	rs599839	A	<i>Humphries SE et al., 2010; Dumitrescu L et al., 2011; Roberts R, Stewart AFR, 2012</i>
<i>CELSR2/SORT1</i>	rs646776	A	<i>Humphries SE et al., 2010; Dumitrescu L et al., 2011</i>
<i>PONI</i>	rs662	G	<i>Drenos F et al., 2007</i>
<i>LDLR</i>	rs688	A	<i>Correia M et al., 2013</i>
<i>C6orf105</i>	rs6903956	A	<i>Roberts R, Stewart AFR, 2012</i>
<i>APOB</i>	rs693	A	<i>Dumitrescu L et al., 2011</i>
<i>ADAMTS7</i>	rs7173743	A	<i>Deloukas P et al., 2013</i>
<i>SELP</i>	rs732314	A	<i>Peloso GM et al., 2010</i>
<i>GCKR</i>	rs780094	A	<i>Dumitrescu L et al., 2011</i>
<i>CUBN</i>	rs7893395	G	<i>Peloso GM et al., 2010</i>
<i>PHACTR1</i>	rs9369640	A	<i>Deloukas P et al., 2013</i>
<i>COL4A1/COL4A2</i>	rs9515203	A	<i>Deloukas P et al., 2013</i>
<i>PDGFD</i>	rs974819	A	<i>Deloukas P et al., 2013; Roberts R, Stewart AFR, 2012</i>
<i>MRAS</i>	rs9818870	A	<i>Humphries SE et al., 2010; Deloukas P et al., 2013</i>

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, \*pateikiami aleliai gali nesutapti su literatūroje nurodomais, nes remiamasi skirtinga DNR grandine (tiesiogine arba atvirkštine), kurios atžvilgiu yra nurodomas alelis. Skirtumas tik tekstinėje išraiškoje, tad faktinio skirtumo nėra.

Koronarinė širdies liga – tai daugialypė liga, kuri pasireiškia kaip aterosklerozės komplikacija. Aterosklerozės etiopatogenezės procese dalyvauja daug veiksnių, todėl nėra vieno biologinio kelio, kuris galėtų paaiškinti šios ligos vystymosi mechanizmą. Dėl šios priežasties, norint atlikti genetinių asociacijų tyrimą, yra ganėtinai sudėtinga „išgryninti“ fenotipą. Remiantis tyrimų duomenimis ir literatūra, yra žinoma, kad genetinę KŠL architektūrą reprezentuoja genetiniai veiksniai, kurie yra arba tiesiogiai asocijuoti su KŠL, arba netiesiogiai – per mechanizmus, lemiančius širdies ir kraujagyslių ligas, lipidų lipoproteinogramos (DTL-cholesterolio, MTL-cholesterolio, triacilglicerolių) pakitimus, MI. Tuo remiantis šiam tyrimui ir atrinktos genetinės sritys KŠL genetinei architektūrai Lietuvos populiacijoje aprašyti. **Lentelėje 3.2** pateikiami trumpai anotuoti (remiantis duomenų bazių informacija) pasirinkti su KŠL susieti genai, su kuriais siejami arba kuriuose yra pasirinkti tyrimui VNP.



### 3.2 lentelė. Su liga susieti genai, genetinės sritys ir jų anotacija

Genas (-ai)	Sritis	Anotacija*	Funkcija**
<i>ABCC6</i>	16p13.1	<i>ATP-binding cassette, sub-family C (CFTR/MRP), member 6</i>	Įvairių molekulių pernaša per ląstelės membranas
<i>ABO</i>	9q34.2	<i>ABO blood group</i>	Glikoziltransferazė, paverčia H antigeną A arba B antigenu
<i>ADAMTS7</i>	15q24.2	<i>ADAM metalloproteinase with thrombospondin type 1 motif, 7</i>	Metaloproteinazė, suardo kremzlių oligomerinio matrikso baltymus
<i>APOB</i>	2p24-p23	<i>apolipoprotein B</i>	Pagrindinė sudedamoji chilomikronų (apo B-48), MTL (apo B-100) ir LMTL (apo B-100) dalis. Prie ląstelių apoB/E receptorių jungiantis MTL dalelėms, Apo B-100 ląstelėse veikia kaip MTL dalelių internalizavimo ląstelėse signalo detektorius***
<i>APOC4</i>	19q13.2	<i>apolipoprotein C-IV</i>	Manoma, kad dalyvauja lipidų metabolizme
<i>APOE</i>	19q13.2	<i>apolipoprotein E</i>	Turtingų triacilgliceroliais lipidų junginių katabolizmas
<i>C6orf105 (ADTRP)</i>	6p24.1	<i>androgen-dependent TFPI-regulating protein</i>	Reguliuoja <i>TFPI</i> raišką ir aktyvumą, didina endotelio apsaugą nuo krešėjimo
<i>CDKN2A/CDKN2B</i>	9p21	<i>cyclin-dependent kinase inhibitor 2A/ cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)</i>	Ląstelės ciklo G1 stadijos kontrolė, auglių slopinantis poveikis
<i>CDKN2B-ASI</i>	9p21.3	<i>CDKN2B antisense RNA 1</i>	Reguliuoja epigenetinį nurašymo proceso slopinimą
<i>CELSR2</i>	1p21	<i>cadherin, EGF LAG seven-pass G-type receptor 2</i>	Manoma, kad veikia kaip receptoriai kontakto sąlygotoje komunikacijoje, ląstelių adhezijoje ir receptorių-ligandų sąveikoje
<i>SORT1</i>	1p13.3; 1p21.3- p13.1	<i>sortilin 1</i>	Prisijungia ligandus, dalyvaujančius įvairiuose ląsteliniuose procesuose
<i>CETP</i>	16q21	<i>cholesteryl ester transfer protein, plasma</i>	Cholesterolio esterių pernaša nuo DTL į kitus lipoproteinus
<i>COL4A1</i>	13q34	<i>collagen, type IV, alpha 1</i>	Užtikrina membranų stabilumą
<i>COL4A2</i>	13q34	<i>collagen, type IV, alpha 2</i>	Užtikrina membranų stabilumą, slopina auglio angiogenezę
<i>CUBN</i>	10p12.31	<i>cubilin (intrinsic factor-cobalamin receptor)</i>	Faktoriaus-vitamino B12 komplekso receptoriaus
<i>CXCL12</i>	10q11.1	<i>chemokine (C-X-C motif) ligand 12</i>	G baltymo porinio receptoriaus ligandas, dalyvauja embriogenezėje, imuniteto palaikyme, uždegiminiame atsake, audinių homeostazėje, auglio augime ir metastazėje
<i>FADS1</i>	1q21.3	<i>flavin adenine dinucleotide synthetase 1</i>	Katalizuoja kofermento FAD susidarymą
<i>FURIN</i>	15q26.1	<i>furin (paired basic amino acid cleaving enzyme)</i>	Baltymų pirmtakų pernaša sekrecijos būdu
<i>FES</i>	15q26.1	<i>feline sarcoma oncogene</i>	Ląstelių transformacija, hematopoezės procesas, augimo veiksnį ir citokinių receptorių signalo pernaša
<i>GCKR</i>	2p23	<i>glucokinase (hexokinase 4) regulator</i>	Slopina gliukokinazę kepenų ir kasos salelių ląstelėse
<i>ITGB3</i>	17q21.32	<i>integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)</i>	Ląstelių adhezija, ląstelių paviršiaus nulemta signalo pernaša
<i>HDAC9</i>	7p21.1	<i>histone deacetylase 9</i>	Galbūt dalyvauja hemopoezėje
<i>HNF1A</i>	12q24.2	<i>HNF1 homeobox A</i>	Transkripcijos veiksnys, dalyvauja kelių specifinių kepenų genų raiškoje
<i>C12orf43</i>	12q24.31	<i>chromosome 12 open reading frame 43</i>	Nėra duomenų
<i>IL6R</i>	1q21	<i>interleukin 6 receptor</i>	Reguliuoja ląstelės augimą ir diferenciaciją, svarbus imuniniam atsakui
<i>LCAT</i>	16q22.1	<i>lecithin-cholesterol acyltransferase</i>	Ląstelės išorės fermentas, suteikiantis cholesteroliui esterio grupę, reikalingą jo



			pernašai
<b>3.2 lentelės tęsinys</b>			
<b>Genas (-ai)</b>	<b>Sritis</b>	<b>Anotacija*</b>	<b>Funkcija**</b>
<i>LDLR</i>	19p13.2	<i>low density lipoprotein receptor</i>	Receptorių nulemta specifinių ligandų endocitozė, svarbus cholesterolio kiekio sintezei reguliuoti
<i>LIPA</i>	10q23.2-q23.3	<i>lipase A, lysosomal acid, cholesterol esterase</i>	Lizosomose katalizuoja cholesterolio esterų ir triacilglicerolių hidrolizę
<i>LIPC</i>	15q21-q23	<i>lipase, hepatic</i>	Triacilglicerolių hidrolazė ir ligandas/jungiantis veiksnys receptoriniame lipoproteinų paėmimo mechanizme
<i>LPL</i>	8p22	<i>lipoprotein lipase</i>	Triacilglicerolių hidrolazė ir ligandas/jungiantis veiksnys receptoriniame lipoproteinų paėmimo mechanizme
<i>MIA3</i>	1q41	<i>melanoma inhibitory activity family, member 3</i>	COL7A1 šalinimas iš endoplazminio tinklo ir auglio slopinimo funkcija
<i>MLXIPL</i>	7q11.23	<i>MLX interacting protein-like</i>	Triacilglicerolių sintezės genų promotorių aktyviklis
<i>MRAS</i>	3q22.3	<i>muscle RAS oncogene homolog</i>	Ķlaidelės augimo ir diferenciacijos signalo nešiklis
<i>MTHFR</i>	1p36.3	<i>methylenetetrahydrofolate reductase (NAD(P)H)</i>	Homocisteino metilinimas iki metionino, 5,10-metilentetrahydrofolato katalizinė reakcija į 5-metiltetrahydrofolatą
<i>PCSK9</i>	1p32.3	<i>proprotein convertase subtilisin/kexin type 9</i>	Specifinių receptorių pristatymas lizosominei degradacijai, cholesterolio ir riebalų rūgščių metabolizmas
<i>PDGFD</i>	11q22.3	<i>platelet derived growth factor D</i>	Ķlaidelių migracija, proliferacija ir transformacija
<i>PHACTR1</i>	6p24.1	<i>phosphatase and actin regulator 1</i>	Galbūt yra pagrindinis endotelio Ķlaidelių funkcijų reguliatorius
<i>PLG</i>	6q26	<i>plasminogen</i>	Kraujo žmogėnas, svarbus krešėjimo ir angiogenezės procese
<i>PON1</i>	7q21.3	<i>paraoxonase 1</i>	Hidrolizuoja paroksoną iki p-nitrofenolio, dalyvauja oksidaciniuose Ķlaidelės procesuose
<i>RXRA</i>	9q34.3	<i>retinoid X receptor, alpha</i>	Branduolio receptoriai, dalyvauja retinoidinės rūgšties nulemtose genų aktyvacijoje
<i>SELP</i>	1q22-q25	<i>selectin P (granule membrane protein 140kDa, antigen CD62)</i>	Palengvina aktyvintų endotelio Ķlaidelių ar trombocitų sąveiką su leukocitais
<i>SH2B3</i>	12q24	<i>SH2B adaptor protein 3</i>	Pagrindinis neigiamas citokinų signalinio kelio reguliatorius, svarbus hemopoezei
<i>SLC22A3</i>	6q25.3	<i>solute carrier family 22 (organic cation transporter), member 3</i>	Svarbus endogeninių mažų katijonų, vaistų ir aplinkos toksinų šalinimui
<i>LPAL</i>	6q26-q27	<i>lipoprotein, Lp(a)-like 2, pseudogene</i>	Panašus į lipoproteiną, pseudogenas
<i>LPA</i>	6q26	<i>lipoprotein, Lp(a)</i>	Svarbus aterosklerozinių trūkių ir trombogenezės reguliavimui
<i>SMAD3</i>	15q22.33	<i>SMAD family member 3</i>	Transkripcijos tarpininkas, dalyvauja karcinogenezės reguliacijoje
<i>TOMM40</i>	19q13	<i>translocase of outer mitochondrial membrane 40 homolog (yeast)</i>	Formuoja kanalus baltymų pernašai į mitochondrijas
<i>UBE2Z</i>	17q21.32	<i>ubiquitin-conjugating enzyme E2Z</i>	Baltymų, dalyvaujančių signalų perdavimo keliuose ir apoptozėje, ubikvitilimas
<i>USF1</i>	1q22-q23	<i>upstream transcription factor 1</i>	Ķlaidelinis transkripcijos veiksnys, susietas su šeimine kombinuota hiperlipidemija
<i>ZC3HC1</i>	7q32.2	<i>zinc finger, C3HC-type containing 1</i>	SCF tipo E3 ubikvitinligazės komplekso narys, reguliuoja Ķlaidelių dalijimosi pradžia
<i>ZEB2</i>	2q22.3	<i>zinc finger E-box binding homeobox 2</i>	Jungimosi prie DNR slopiklis transkripcijos procese, sąveikauja su aktyvintu SMADs

\* pagal NCBI duomenų bazės 2014 m. balandžio mėnesio informaciją; \*\* pagal NCBI duomenų bazės RefSeq 2008–2014 m. informaciją; \*\*\* pagal UniProtKB/Swiss-Prot duomenų bazės 2014 m. informaciją.

### **3.2. Genetinės struktūros ir įvairovės nustatymas pagal pasirinktus VNP**

Siekiant įvertinti lietuvių KŠL riziką, 253 asmenims iš bendros Lietuvos lietuvių populiacijos (asmenys iš šešių etnolingvistinių grupių – 140 aukštaičių ir 113 žemaičių, arba 127 moterys, 126 vyrai) buvo atliktas genetinės struktūros ir įvairovės tyrimas pagal atrinktus 60 VNP. Analizei naudoti duomenys, gauti pagal aprašytą metodiką (**2.2.2 skyrelis ir II priedas**) atlikus plataus masto genotipavimą naujos kartos didelio našumo sistema ir taikant statistinius metodus.

Pagal genotipavimo duomenų kokybės vertinimo aprašą (**2.3.4.1 skyrelis, IV priedas**), analizuotiems duomenims pirmiausia atlikta kokybės kontrolė. Kokybės vertinimo rezultatai pateikti **IV priedo 4.2 lentelėje**.

Vadovaujantis **IV priede** nurodytomis gamintojo rekomendacijomis, po genotipavimo eksperimento, gavus fluorescencinio intensyvumo žaliavinius duomenis, buvo atlikta pirminė duomenų analizė ir kokybės vertinimas. Pagal atitinkamus kriterijus ir rekomenduojamas jų reikšmes vertinta genotipuotų mėginių, VNP ir paties genotipavimo proceso kokybė. Atsižvelgiant į vertinimo kriterijus, genotipuotų mėginių kokybė yra gera (išskyrus mėginio LTG-1075, kuris pašalintas iš tolesnės analizės), genotipavimo proceso rodikliai tenkina kriterijus. Atlikus būtinas duomenų kokybės gryninimo procedūras, tolesnei analizei palikti tik tinkamos genotipavimo kokybės VNP.

Visi toliau pateikiami VNP alelių ir genotipų dažniai yra Hardžio ir Vainbergo pusiausvyroje,  $p \geq 0,001$ .

#### **3.2.1. Vidupopuliacinė genetinė struktūra ir įvairovė pagal tiriamus VNP**

Vidupopuliacinė genetinė struktūra ir įvairovė nustatyta įvertinant atrinktų su liga siejamų VNP genotipų bei alelių dažnius bendrai tiriamai

populiacijai, alelių dažnius tiriamos populiacijos viduje pagal lytį (vyrų ir moterų grupėse) ir pagal etnolingvistines grupes (aukštaičių ir žemaičių grupėse). Tiriamiems VNP tokiose pačiose vidupopuliacinėse grupėse įvertinti haplotipai, haplotipų blokai ir jų dažniai, taip pat rizikos alelių skaičius, tenkantis asmeniui.

VNP genotipai ir jų dažniai tirtoje bendroje Lietuvos lietuvių populiacijoje pateikiami **3.3 lentelėje**. Taip pat pateikiami genotipų dažniai pagal etnolingvistines grupes (aukštaičių ir žemaičių) ir pagal lytį (vyrų ir moterų).

**3.3 lentelė.** Nustatyti VNP genotipų dažniai tirtoje bendroje Lietuvos lietuvių populiacijoje

		GENOTIPŲ DAŽNIAI							GENOTIPŲ DAŽNIAI				
VNP	GT	B	M	V	A	Ž	VNP	GT	B	M	V	A	Ž
rs10420 31	AA	0,03	0,03	0,02	0,02	0,04	rs3184 504	AA	0,17	0,17	0,17	0,18	0,16
	AG	0,32	0,33	0,30	0,31	0,33		AG	0,50	0,50	0,49	0,46	0,55
	GG	0,66	0,64	0,67	0,67	0,64		GG	0,33	0,33	0,33	0,36	0,29
rs10757 274	GG	0,22	0,20	0,23	0,23	0,20	rs3217 992	AA	0,16	0,17	0,16	0,18	0,14
	GA	0,47	0,54	0,40	0,46	0,50		AG	0,51	0,55	0,47	0,49	0,54
	AA	0,31	0,25	0,37	0,31	0,30		GG	0,33	0,28	0,37	0,34	0,32
rs11185 660	GG	0,05	0,05	0,05	0,05	0,04	rs3764 261	AA	0,08	0,09	0,07	0,10	0,05
	GA	0,37	0,38	0,35	0,39	0,33		AC	0,42	0,43	0,40	0,41	0,43
	AA	0,59	0,57	0,60	0,56	0,63		CC	0,50	0,48	0,52	0,49	0,51
rs11203 042	AA	0,13	0,13	0,12	0,12	0,13	rs4252 120	GG	0,07	0,09	0,05	0,06	0,07
	AG	0,47	0,50	0,44	0,46	0,49		GA	0,33	0,27	0,39	0,33	0,33
	GG	0,40	0,37	0,44	0,42	0,38		AA	0,60	0,65	0,56	0,61	0,60
rs11206 510	GG	0,02	0,02	0,03	0,04	0,01	rs4420 638	GG	0,02	0,03	0,02	0,03	0,02
	GA	0,26	0,26	0,27	0,30	0,22		GA	0,23	0,27	0,19	0,20	0,27
	AA	0,71	0,72	0,70	0,66	0,77		AA	0,75	0,70	0,79	0,77	0,72
rs11556 924	AA	0,19	0,19	0,20	0,21	0,17	rs4459 25	AA	0,01	0,02	0,01	0,00	0,03
	AG	0,51	0,50	0,51	0,49	0,52		AG	0,17	0,21	0,13	0,15	0,19
	GG	0,30	0,31	0,29	0,29	0,31		GG	0,82	0,77	0,87	0,85	0,78
rs11591 147	AA	0,00	0,00	0,00	0,00	0,00	rs4773 144	GG	0,17	0,14	0,20	0,21	0,12
	AC	0,01	0,02	0,00	0,01	0,00		GA	0,52	0,54	0,50	0,47	0,57
	CC	0,99	0,98	1,00	0,99	1,00		AA	0,31	0,32	0,30	0,32	0,31
rs12603 26	AA	0,12	0,10	0,13	0,09	0,15	rs4775 041	GG	0,08	0,07	0,09	0,04	0,12
	AG	0,53	0,50	0,56	0,54	0,52		GC	0,42	0,45	0,39	0,42	0,42
	GG	0,35	0,39	0,30	0,36	0,33		CC	0,50	0,48	0,52	0,54	0,46
rs14124 44	AA	0,09	0,10	0,07	0,09	0,09	rs4775 065	AA	0,04	0,05	0,02	0,04	0,03
	AG	0,47	0,50	0,44	0,49	0,45		AG	0,32	0,28	0,35	0,28	0,37
	GG	0,44	0,40	0,48	0,43	0,46		GG	0,65	0,67	0,62	0,68	0,61

3.3 lentelės tęsinys													
		GENOTIPŲ DAŽNIAI							GENOTIPŲ DAŽNIAI				
VNP	GT	B	M	V	A	Ž	VNP	GT	B	M	V	A	Ž
rs1504 68	CC	0,02	0,01	0,03	0,02	0,02	rs4845 625	GG	0,23	0,25	0,21	0,21	0,25
	CA	0,24	0,24	0,24	0,21	0,28		GA	0,50	0,50	0,50	0,52	0,47
	AA	0,74	0,75	0,73	0,77	0,70		AA	0,27	0,25	0,29	0,26	0,28
rs1556 3	GG	0,24	0,28	0,19	0,29	0,18	rs4977 574	GG	0,22	0,21	0,23	0,23	0,21
	GA	0,48	0,44	0,51	0,47	0,49		GA	0,47	0,54	0,40	0,46	0,49
	AA	0,29	0,28	0,30	0,24	0,34		AA	0,31	0,25	0,37	0,31	0,30
rs1714 5738	AA	0,01	0,02	0,01	0,01	0,02	rs5011 20	GG	0,01	0,00	0,02	0,01	0,01
	AG	0,23	0,20	0,25	0,21	0,25		GA	0,22	0,25	0,20	0,24	0,21
	GG	0,76	0,79	0,74	0,79	0,73		AA	0,77	0,75	0,79	0,76	0,79
rs1722 8212	GG	0,06	0,05	0,08	0,06	0,07	rs5623 38	AA	0,02	0,03	0,02	0,01	0,04
	GA	0,39	0,37	0,40	0,36	0,42		AG	0,30	0,28	0,32	0,26	0,34
	AA	0,55	0,58	0,52	0,58	0,51		GG	0,68	0,69	0,67	0,72	0,63
rs1745 47	GG	0,08	0,06	0,10	0,07	0,08	rs5794 59	GG	0,07	0,06	0,07	0,06	0,07
	GA	0,38	0,36	0,40	0,39	0,36		GA	0,40	0,39	0,41	0,38	0,42
	AA	0,55	0,58	0,51	0,54	0,56		AA	0,53	0,55	0,52	0,56	0,50
rs1746 4857	CC	0,04	0,03	0,04	0,04	0,03	rs5882	GG	0,10	0,09	0,12	0,09	0,12
	CA	0,31	0,33	0,29	0,30	0,33		GA	0,40	0,40	0,39	0,39	0,41
	AA	0,65	0,64	0,67	0,66	0,65		AA	0,50	0,51	0,49	0,52	0,48
rs1746 5637	AA	0,12	0,10	0,14	0,11	0,13	rs5918	GG	0,01	0,02	0,01	0,01	0,02
	AC	0,42	0,43	0,42	0,45	0,39		GA	0,28	0,30	0,26	0,27	0,28
	CC	0,46	0,47	0,45	0,45	0,48		AA	0,71	0,69	0,74	0,72	0,70
rs1751 4846	CC	0,26	0,25	0,26	0,27	0,25	rs5998 39	GG	0,06	0,05	0,06	0,09	0,02
	CA	0,48	0,51	0,45	0,45	0,51		GA	0,37	0,40	0,34	0,37	0,38
	AA	0,26	0,24	0,29	0,28	0,24		AA	0,57	0,55	0,60	0,54	0,61
rs1801 133	AA	0,07	0,07	0,07	0,06	0,08	rs6467 76	GG	0,06	0,06	0,06	0,09	0,03
	AG	0,45	0,46	0,44	0,41	0,49		GA	0,37	0,40	0,34	0,36	0,38
	GG	0,48	0,47	0,49	0,52	0,43		AA	0,57	0,54	0,60	0,56	0,59
rs2023 938	GG	0,01	0,02	0,00	0,01	0,01	rs662	GG	0,08	0,09	0,06	0,07	0,08
	GA	0,12	0,09	0,15	0,12	0,12		GA	0,34	0,35	0,33	0,37	0,31
	AA	0,87	0,89	0,85	0,87	0,87		AA	0,58	0,55	0,61	0,56	0,61
rs2048 327	GG	0,11	0,11	0,10	0,12	0,09	rs688	AA	0,16	0,16	0,16	0,18	0,13
	GA	0,48	0,48	0,48	0,51	0,45		AG	0,50	0,52	0,48	0,51	0,49
	AA	0,41	0,41	0,41	0,37	0,46		GG	0,34	0,32	0,36	0,31	0,38
rs2075 650	GG	0,02	0,02	0,02	0,02	0,03	rs6903 956	AA	0,09	0,07	0,10	0,06	0,12
	GA	0,29	0,32	0,25	0,26	0,32		AG	0,44	0,40	0,47	0,43	0,45
	AA	0,69	0,65	0,72	0,71	0,65		GG	0,48	0,52	0,43	0,51	0,44
rs2197 089	GG	0,21	0,21	0,21	0,20	0,22	rs693	GG	0,21	0,24	0,17	0,23	0,19
	GA	0,50	0,55	0,44	0,48	0,51		GA	0,57	0,57	0,56	0,58	0,55
	AA	0,29	0,24	0,35	0,32	0,27		AA	0,23	0,19	0,26	0,19	0,27
rs2252 641	GG	0,15	0,17	0,12	0,14	0,16	rs7173 743	GG	0,20	0,24	0,17	0,21	0,19
	GA	0,39	0,35	0,43	0,39	0,39		GA	0,52	0,51	0,54	0,50	0,55
	AA	0,47	0,48	0,45	0,48	0,45		AA	0,27	0,25	0,29	0,29	0,26
rs2259 816	AA	0,14	0,19	0,10	0,11	0,19	rs7323 14	GG	0,18	0,16	0,20	0,18	0,18
	AC	0,43	0,40	0,47	0,44	0,43		GA	0,45	0,43	0,47	0,43	0,48
	CC	0,42	0,41	0,44	0,46	0,38		AA	0,37	0,41	0,33	0,39	0,35
rs2271 293	AA	0,01	0,02	0,01	0,01	0,01	rs7800 94	AA	0,12	0,10	0,13	0,09	0,15
	AG	0,21	0,23	0,20	0,18	0,26		AG	0,53	0,50	0,56	0,54	0,52
	GG	0,77	0,76	0,79	0,81	0,73		GG	0,35	0,40	0,30	0,37	0,33

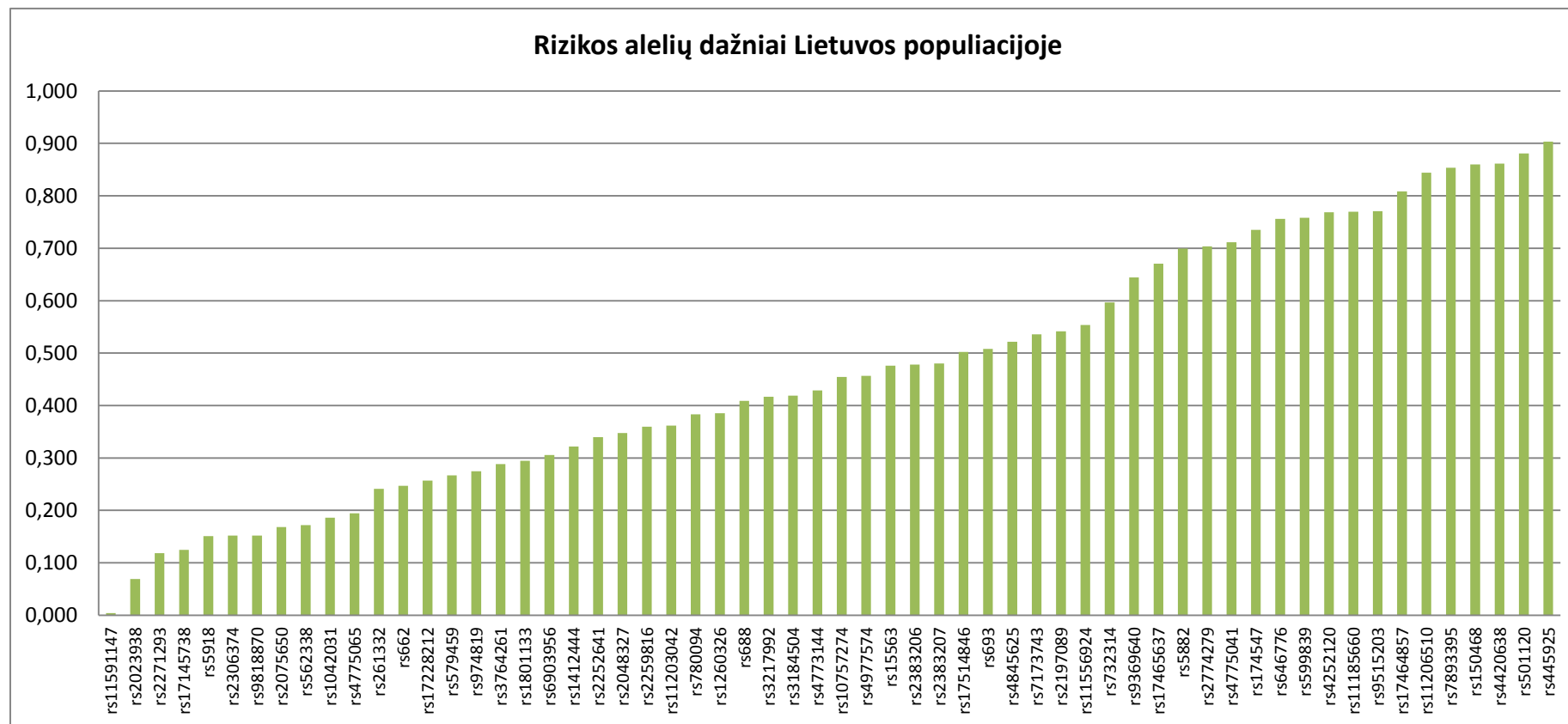
3.3 lentelės tęsinys													
		GENOTIPŲ DAŽNIAI							GENOTIPŲ DAŽNIAI				
VNP	GT	B	M	V	A	Ž	VNP	GT	B	M	V	A	Ž
rs2306 374	GG	0,02	0,02	0,02	0,01	0,03	rs7893 395	AA	0,01	0,02	0,01	0,02	0,00
	GA	0,27	0,30	0,25	0,31	0,23		AG	0,27	0,24	0,29	0,26	0,27
	AA	0,71	0,69	0,74	0,69	0,74		GG	0,72	0,74	0,70	0,71	0,73
rs2383 206	GG	0,23	0,24	0,23	0,24	0,23	rs9369 640	CC	0,10	0,10	0,10	0,11	0,09
	GA	0,49	0,54	0,44	0,50	0,48		CA	0,51	0,49	0,52	0,52	0,49
	AA	0,28	0,22	0,33	0,26	0,29		AA	0,39	0,41	0,37	0,36	0,42
rs2383 207	GG	0,23	0,24	0,23	0,24	0,23	rs9515 203	GG	0,04	0,06	0,03	0,04	0,05
	GA	0,49	0,55	0,44	0,50	0,49		GA	0,37	0,33	0,41	0,39	0,35
	AA	0,27	0,21	0,33	0,26	0,28		AA	0,58	0,61	0,56	0,58	0,59
rs2613 32	AA	0,06	0,06	0,06	0,05	0,06	rs9748 19	AA	0,09	0,09	0,08	0,08	0,10
	AG	0,37	0,39	0,36	0,33	0,42		AG	0,37	0,39	0,36	0,34	0,42
	GG	0,57	0,56	0,59	0,62	0,51		GG	0,54	0,52	0,56	0,58	0,48
rs2774 279	AA	0,09	0,09	0,08	0,08	0,10	rs9818 870	AA	0,02	0,02	0,02	0,01	0,03
	AG	0,42	0,43	0,41	0,41	0,42		AG	0,27	0,30	0,25	0,31	0,23
	GG	0,49	0,48	0,51	0,51	0,48		GG	0,71	0,69	0,74	0,69	0,74

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, GT – genotipas, B – genotipų dažniai tirtoje bendroje Lietuvos lietuvių populiacijos grupėje, M – moterų grupė, V – vyrų grupė, A – aukštaičių grupė, Ž – žemaičių grupė.

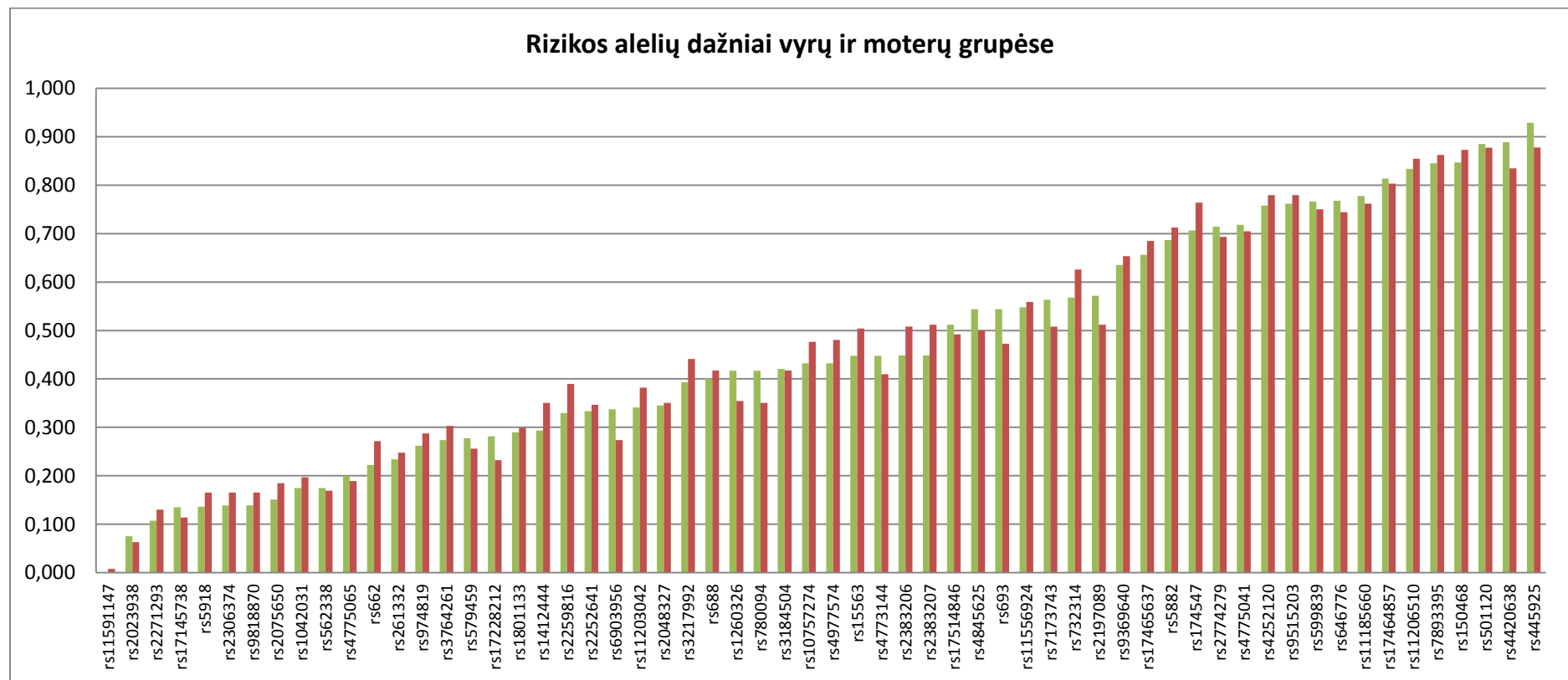
Nustatyti ne tik genotipų, bet ir alelių dažniai pagal tiriamus VNP tiriamoje bendroje Lietuvos lietuvių populiacijoje, tiriamos populiacijos vyrų ir moterų bei aukštaičių ir žemaičių grupėse.

VNP rizikos alelių dažnio pasiskirstymas tirtoje bendroje Lietuvos lietuvių populiacijoje pavaizduotas **3.1 paveiksle** (konkrečias nustatytų rizikos alelių dažnių vertes bendrai Lietuvos lietuvių populiacijai galima rasti **3.10–3.13 lentelėse**). Iš histogramos matyti, kad rečiausiai pasitaikantis rizikos alelis yra VNP rs11591147, esantis 1-oje chromosomoje *PCSK9* geno koduojančioje sekoje. Dažniausiai pasitaikantis rizikos alelis yra VNP rs445925, esantis 19-oje chromosomoje ir siejamas su *APOE* geno sritimi. Dvidešimt penki iš 60 tiriamų VNP (t. y. beveik 42 %) rizikos alelių populiacijoje pasitaikė dažniau negu 50 %.

Tiriamų VNP rizikos alelių dažnio pasiskirstymas vyrų ir moterų grupėse pavaizduotas **3.2 paveiksle**. Iš histogramos matyti, kad rečiausiai pasitaikantis rizikos alelis tiek vyrų, tiek moterų grupėje yra toks pat, kaip ir bendroje populiacijoje – rs11591147. Dažniausiai pasitaikantis rizikos alelis abiem grupėms yra toks pats kaip ir bendrai tiriamai populiacijai – rs445925.



**3.1 paveikslas.** Tiriamų VNP rizikos alelių dažniai tiriamoje populiacijoje. X ašyje – VNP identifikaciniai numeriai, Y ašyje – rizikos alelių dažniai.



**3.2 paveikslas.** Tiriamų VNP rizikos alelių dažniai tiriamos populiacijos vyrų ir moterų grupėse. X ašyje – VNP identifikaciniai numeriai, Y ašyje – rizikos alelių dažniai. Žalia spalva – alelių dažniai vyrų grupėje, raudona – moterų.

Atliktas VNP alelių dažnių palyginimas tirtos populiacijos vyrų ir moterų grupėje (**3.4 lentelė**), pateiktos  $\chi^2$  reikšmės ir statistinis reikšmingumas. Pasirinktas reikšmingumo lygmuo 0,05. Vyrų ir moterų grupėje statistiškai reikšmingų skirtumų tarp alelių dažnių nerasta.

**3.4 lentelė.** Nustatytų VNP alelių dažnių palyginimas vyrų ir moterų grupėse

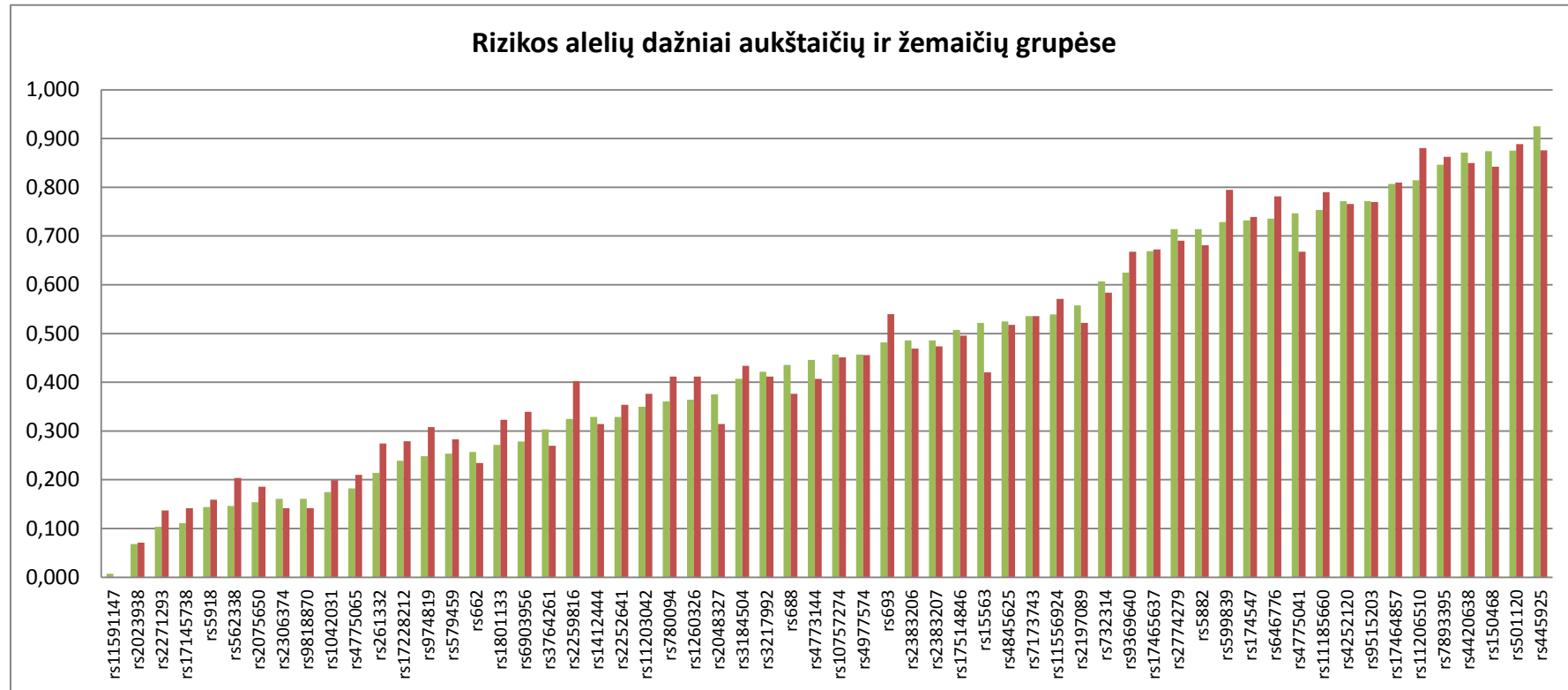
VNP	Rizikos alelis	RAD		$\chi^2$	p
		V	M		
rs1042031	A	0,175	0,197	0,414	0,520
rs10757274	G	0,433	0,476	0,981	0,322
rs11185660	A	0,778	0,762	0,179	0,672
rs11203042	A	0,341	0,382	0,904	0,342
rs11206510	A	0,833	0,854	0,423	0,515
rs11556924	G	0,548	0,559	0,067	0,796
rs11591147	A	0,000	0,008	1,992	0,158
rs1260326	A	0,417	0,354	2,075	0,150
rs1412444	A	0,294	0,350	1,865	0,172
rs150468	A	0,847	0,873	0,715	0,398
rs15563	G	0,448	0,504	1,581	0,209
rs17145738	A	0,135	0,114	0,500	0,480
rs17228212	G	0,282	0,232	1,621	0,203
rs174547	A	0,706	0,764	2,143	0,143
rs17464857	A	0,814	0,803	0,087	0,768
rs17465637	C	0,656	0,685	0,481	0,488
rs17514846	A	0,512	0,492	0,200	0,655
rs1801133	A	0,290	0,299	0,055	0,814
rs2023938	G	0,075	0,063	0,302	0,582
rs2048327	G	0,345	0,350	0,015	0,903
rs2075650	G	0,151	0,185	1,061	0,303
rs2197089	A	0,571	0,512	1,798	0,180
rs2252641	G	0,333	0,347	0,097	0,755
rs2259816	A	0,329	0,390	2,004	0,157
rs2271293	A	0,107	0,130	0,628	0,428
rs2306374	G	0,139	0,165	0,687	0,407
rs2383206	G	0,448	0,508	1,792	0,181
rs2383207	G	0,448	0,512	2,037	0,154
rs261332	A	0,234	0,248	0,134	0,715
rs2774279	G	0,714	0,693	0,277	0,599
rs3184504	A	0,421	0,417	0,006	0,940
rs3217992	A	0,393	0,441	1,203	0,273
rs3764261	A	0,274	0,303	0,531	0,466
rs4252120	A	0,758	0,780	0,332	0,565
rs4420638	A	0,889	0,835	3,122	0,077
rs445925	G	0,929	0,878	3,706	0,054
rs4773144	G	0,448	0,409	0,765	0,382
rs4775041	C	0,718	0,705	0,113	0,737
rs4775065	A	0,200	0,189	0,098	0,755
rs4845625	A	0,544	0,500	0,966	0,326
rs4977574	G	0,433	0,480	1,164	0,281
rs501120	A	0,885	0,877	0,076	0,783
rs562338	A	0,175	0,169	0,025	0,874
rs579459	G	0,278	0,256	0,309	0,578
rs5882	A	0,687	0,713	0,410	0,522
rs5918	G	0,136	0,165	0,848	0,357
rs599839	A	0,766	0,750	0,173	0,677
rs646776	A	0,768	0,744	0,390	0,532
rs662	G	0,222	0,272	1,662	0,197
rs688	A	0,401	0,417	0,143	0,705



3.4 lentelės tęsinys						
VNP	Rizikos alelis	RAD		$\chi^2$	p	
		V	M			
rs6903956	A	0,337	0,274	2,394	0,122	
rs693	A	0,544	0,472	2,566	0,109	
rs7173743	A	0,564	0,508	1,564	0,211	
rs732314	A	0,568	0,626	1,801	0,180	
rs780094	A	0,417	0,350	2,350	0,125	
rs7893395	G	0,845	0,862	0,292	0,589	
rs9369640	A	0,635	0,654	0,191	0,662	
rs9515203	A	0,762	0,780	0,222	0,637	
rs974819	A	0,262	0,287	0,403	0,525	
rs9818870	A	0,139	0,165	0,687	0,407	

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, RAD – rizikos alelio dažnis, V – vyrų grupė, M – moterų grupė,  $\chi^2$  – Pirsono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu.

Atlikus vidupopuliacinį alelių dažnių palyginimą pagal lytis, buvo apskaičiuoti tiriamų VNP rizikos alelių dažniai dviejose etnolingvistinėse Lietuvos grupėse – aukštaičių ir žemaičių. Nustatytų VNP rizikos alelių dažnio pasiskirstymas aukštaičių ir žemaičių grupėse vaizduojamas **3.3 paveiksle**. Iš histogramos matyti, kad rečiausiai pasitaikantis rizikos alelis abiejose etnolingvistinėse grupėse yra toks pat, kaip ir bendroje populiacijoje – rs11591147. Dažniausiai pasitaikantis rizikos alelis aukštaičiams yra toks pats, kaip ir bendrai tiriamai populiacijai – rs445925, tačiau žemaičių grupėje dažniausias alelis buvo kitas, VNP rs501120 esantis 10-oje chromosomoje ir siejamas su *CXCL12* genu.



**3.3 paveikslas.** Tiriamų VNP rizikos alelių dažniai bendros Lietuvos lietuvių populiacijos etnolingvistinėse grupėse. X ašyje – VNP identifikaciniai numeriai, Y ašyje – rizikos alelių dažniai. Žalia spalva – alelių dažniai aukštaičių grupėje, raudona – žemaičių.

Atliktas tiriamų VNP alelių dažnių palyginimas tirtos populiacijos aukštaičių ir žemaičių grupėje (**3.5 lentelė**), pateiktos  $\chi^2$  reikšmės ir statistinis reikšmingumas. Pasirinktas reikšmingumo lygmuo 0,05. Nustatyti statistiškai reikšmingi skirtumai tarp dviejų VNP alelių dažnių: rs11206510, esančio 1-oje chromosomoje bei siejamo su jau minėtu genu *PCSK9*, ir rs15563, esančio 17-oje chromosomoje *UBE2Z* geno 3' NTS.

**3.5 lentelė.** Nustatytų VNP alelių dažnių palyginimas aukštaičių ir žemaičių grupėse

VNP	Rizikos alelis	RAD		$\chi^2$	p
		AUK	ŽEM		
rs1042031	A	0,175	0,199	0,481	0,488
rs10757274	G	0,457	0,451	0,017	0,896
rs11185660	A	0,754	0,790	0,941	0,332
rs11203042	A	0,350	0,376	0,369	0,543
rs11206510	A	<b>0,814</b>	<b>0,881</b>	<b>4,166</b>	<b>0,041</b>
rs11556924	G	0,539	0,571	0,502	0,478
rs11591147	A	0,007	0,000	1,621	0,203
rs1260326	A	0,364	0,412	1,177	0,278
rs1412444	A	0,329	0,314	0,119	0,730
rs150468	A	0,874	0,842	1,034	0,309
rs15563	G	<b>0,522</b>	<b>0,420</b>	<b>5,121</b>	<b>0,024</b>
rs17145738	A	0,111	0,142	1,094	0,296
rs17228212	G	0,239	0,279	1,021	0,312
rs174547	A	0,732	0,739	0,030	0,863
rs17464857	A	0,807	0,810	0,005	0,941
rs17465637	C	0,669	0,673	0,007	0,934
rs17514846	A	0,507	0,496	0,068	0,795
rs1801133	A	0,271	0,323	1,602	0,206
rs2023938	G	0,068	0,071	0,017	0,897
rs2048327	G	0,375	0,314	2,041	0,153
rs2075650	G	0,154	0,186	0,932	0,334
rs2197089	A	0,558	0,522	0,630	0,427
rs2252641	G	0,329	0,354	0,360	0,549
rs2259816	A	0,325	0,403	3,274	0,070
rs2271293	A	0,104	0,137	1,351	0,245
rs2306374	G	0,161	0,142	0,354	0,552
rs2383206	G	0,486	0,469	0,140	0,709
rs2383207	G	0,486	0,474	0,075	0,784
rs261332	A	0,214	0,274	2,465	0,116
rs2774279	G	0,714	0,690	0,346	0,556
rs3184504	A	0,407	0,434	0,360	0,548
rs3217992	A	0,421	0,412	0,051	0,822
rs3764261	A	0,304	0,270	0,690	0,406
rs4252120	A	0,771	0,766	0,025	0,875
rs4420638	A	0,871	0,850	0,502	0,479
rs445925	G	0,925	0,876	3,418	0,064
rs4773144	G	0,446	0,407	0,773	0,379
rs4775041	C	0,746	0,668	3,734	0,053
rs4775065	A	0,182	0,210	0,609	0,435
rs4845625	A	0,525	0,518	0,027	0,870
rs4977574	G	0,457	0,456	0,001	0,975
rs501120	A	0,875	0,888	0,213	0,645
rs562338	A	0,146	0,204	2,865	0,091
rs579459	G	0,254	0,283	0,561	0,454
rs5882	A	0,714	0,681	0,643	0,423
rs5918	G	0,144	0,159	0,231	0,631

**3.5 lentelės tęsinys**

VNP	Rizikos alelis	RAD		$\chi^2$	p
		AUK	ŽEM		
rs599839	A	0,729	0,795	2,961	0,085
rs646776	A	0,736	0,781	1,399	0,237
rs662	G	0,257	0,235	0,344	0,557
rs688	A	0,436	0,376	1,838	0,175
rs6903956	A	0,279	0,339	2,162	0,142
rs693	A	0,482	0,540	1,665	0,197
rs7173743	A	0,536	0,536	0,000	1,000
rs732314	A	0,607	0,584	0,277	0,599
rs780094	A	0,361	0,412	1,365	0,243
rs7893395	G	0,846	0,863	0,270	0,604
rs9369640	A	0,625	0,668	1,016	0,314
rs9515203	A	0,771	0,770	0,002	0,968
rs974819	A	0,248	0,308	2,228	0,136
rs9818870	A	0,161	0,142	0,354	0,552

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, RAD – rizikos alelio dažnis, AUK – aukštaičių grupė, ŽEM – žemaičių grupė,  $\chi^2$  – Pirsono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, **paryškintos pasviros reikšmės** – nustatyti statistiškai reikšmingi skirtumai.

Siekiant nustatyti **haplotipų** įvairovę ir sandarą, tyrimui atrinkta 60 VNP, klasikiniu būdu suskirstytų į grupes pagal priklausomybę chromosomoms. Iš viso sudaryta 14 grupių, o informacija pateikiama **3.6 lentelėje**. Iš haplotipų analizės pašalintas vienas VNP rs2197089, nes jo nebuvo įmanoma priskirti grupei, jis 8-oje chromosomoje buvo vienas.

**3.6 lentelė.** Haplotipų analizei VNP sugrupuoti pagal priklausomybę chromosomai

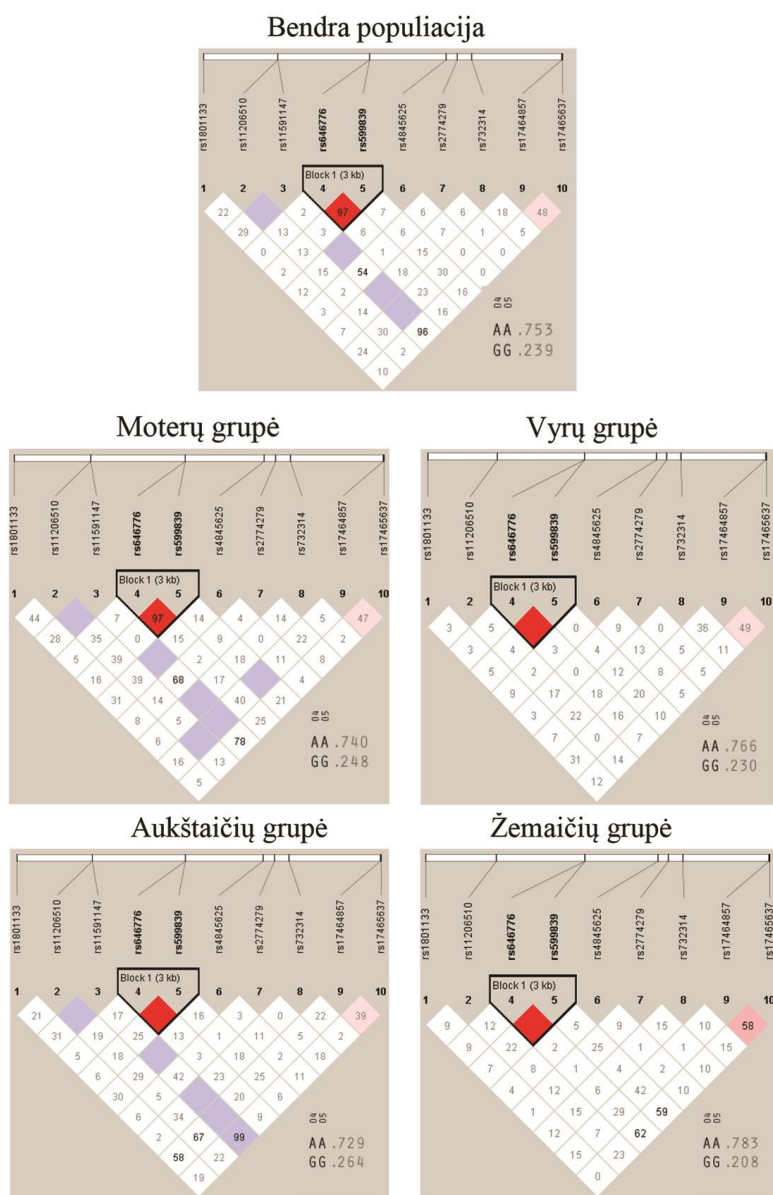
Chr	VNP sk.	VNP
1	10	rs646776, rs599839, rs17465637, rs732314, rs4845625, rs2774279, rs1801133, rs17464857, rs11591147, rs11206510
2	6	rs780094, rs693, rs562338, rs2252641, rs1260326, rs1042031
3	2	rs9818870, rs2306374
6	4	rs6903956, rs9369640, rs4252120, rs2048327
7	4	rs662, rs2023938, rs17145738, rs11556924
9	7	rs2383207, rs2383206, rs11185660, rs579459, rs4977574, rs3217992, rs10757274
10	4	rs501120, rs7893395, rs1412444, rs11203042
11	2	rs974819, rs174547
12	2	rs3184504, rs2259816
13	2	rs4773144, rs9515203
15	6	rs7173743, rs17514846, rs4775065, rs4775041, rs261332, rs17228212
16	4	rs150468, rs5882, rs3764261, rs2271293
17	2	rs15563, rs5918
19	4	rs688, rs445925, rs4420638, rs2075650

Chr – chromosoma, VNP sk. – VNP skaičius priklausantis tai pačiai chromosomai, VNP – vieno nukleotido polimorfizmo identifikaciniai numeriai.

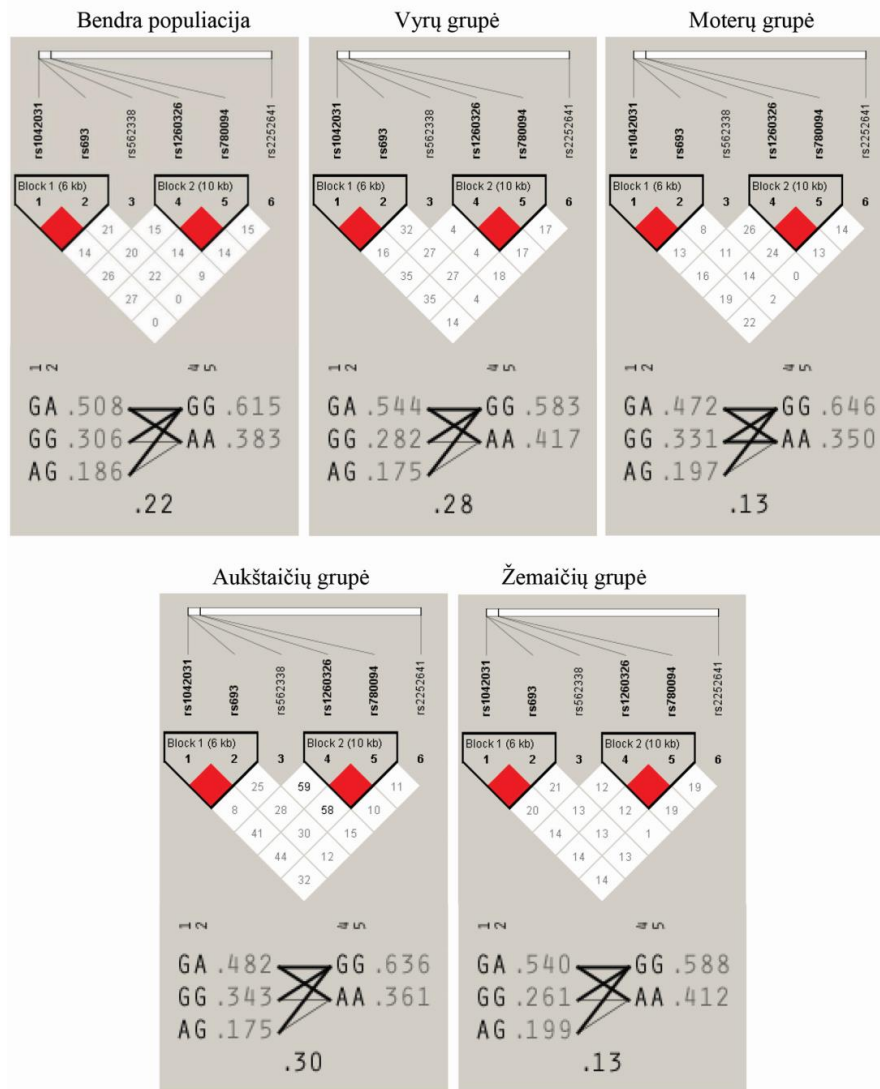
Programa *Haploview v4.2* (Barrett JC et al., 2005) pagal tiriamos populiacijos genotipus ir pateiktas VNP grupes su VNP koordinatėmis genome tiriamai populiacijai ir jos grupėms pagal lytį ir etnolingvistinę grupę sukonstruoti haplotipai, apskaičiuoti jų dažniai, įvertinta sankibos pusiausvyra, nustatyti haplotipų blokai.

Haplotipų blokai nustatyti penkioms VNP grupėms, suskirstytoms pagal chromosomas: 1–3, 9 ir 19. Pirmos chromosomos tirtiems 10 VNP visose tirtose grupėse sukonstruotas vienas ir tas pats haplotipų blokas tarp VNP rs646776 ir rs599839, asocijuotų su sritimi, kurioje yra *CELSR2* ir *SORT1* genai (**3.4 paveikslas**). Visose tirtose grupėse nustatyti du haplotipai: AA ir GG. Haplotipo AA dažnis yra 0,72–0,78 intervalo ribose visose tirtose grupėse, alternatyvių alelių haplotipo (GG) dažnis yra 0,21–0,26 intervalo ribose, vadinasi, dažnesnis yra rizikos alelių haplotipas AA. Antros chromosomos tirtiems 6 VNP visose tirtose grupėse sukonstruoti du tokie patys haplotipų blokai (**3.5 paveikslas**). Pirmasis blokas tarp VNP rs1042031 ir rs693 *APOB* gene, antrasis – rs1260326 ir rs780094 *GCKR* gene. Pirmojo bloko atveju visose tirtose grupėse nustatyti trys haplotipai: GA, AG ir GG, abiejų rizikos alelių haplotipas AA nenustatytas. Visose tirtose grupėse nustatyti tik vieną rizikos alelį turintys haplotipai GA (dažniai grupėse 0,47–0,54) ir AG (dažniai grupėse 0,18–0,2) bei RA neturintis haplotipas GG (dažniai grupėse 0,26–0,34). Antrojo bloko atveju visose tirtose grupėse nustatyti du haplotipai: GG ir AA. Visose tirtose grupėse abu RA turintis haplotipas AA (dažniai grupėse 0,35–0,42) yra retesnis negu alternatyvių alelių haplotipas GG (dažniai grupėse 0,58–0,65). Tarp dviejų haplotipų blokų yra galima nepusiausvyra sankiba, taip pat galimi ir mainai (rekombinacija). Dažniausiai šių dviejų haplotipų blokų atveju visose tiriamose grupėse vienodu dažnumu rekombinuoja visi blokų haplotipai, išskyrus antro bloko RA (AA) haplotipą, kuris rečiau rekombinuoja su pirmojo bloko AG ir GG haplotipais. Išimtis yra moterų grupė, kurioje rečiau rekombinuoja tik su AG haplotipu.

Apskritai vertinant haplotipų blokų sankibą, mažiausia ji yra moterų ir žemaičių grupėse ( $D' = 0,13$ ), didžiausia – aukštaičių ( $D' = 0,30$ ).



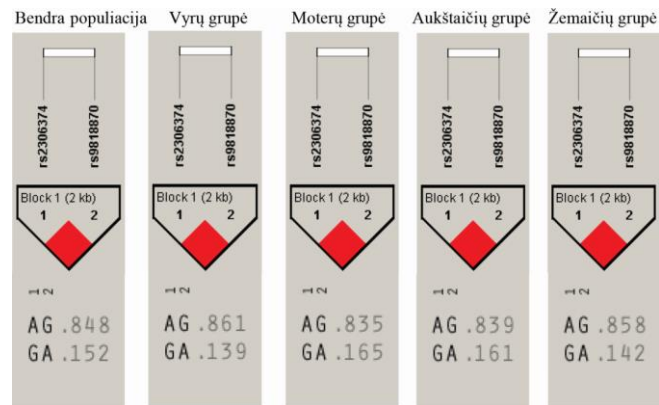
**3.4 paveikslas.** Nepusiausviros sankibos genolapis ir haplotipų blokas 1-ai chromosomai pagal tiriamus 10 VNP tiriamajai populiacijai ir vidupopuliacinėms grupėms pagal lytį ir etnolingvistinę grupę. Viršuje pavaizduoti VNP ir jų išsidėstymas vienas kito atžvilgiu. Ryški juoda linija rėmina haplotipų bloką. Skaičiai kvadratėliuose nurodo  $D'$  statistikos reikšmę dviem atitinkamiems VNP. Apatiniame dešiniajame kampe yra nurodyti haplotipo bloko haplotipai ir jų dažniai tiriamoje grupėje. Kvadratėlių spalvos: balta –  $D' < 1$ ,  $LOD < 2$ ; melsva –  $D' = 1$ ,  $LOD < 2$ ; rožiniai atspalviai –  $D' < 1$ ,  $LOD \geq 2$ ; raudona –  $D' = 1$ ,  $LOD \geq 2$  ( $D'$  – santykinis nepusiausviros sankibos koeficientas,  $LOD$  – nepusiausviros sankibos tikimybės įvertis).



**3.5 paveikslas.** Nepusiausviros sankibos genolapis ir haplotipų blokai 2-ai chromosomai pagal tiriamus 6 VNP tiriamajai populiacijai ir vidupopuliacinėms grupėms pagal lytį ir etnolingvistinę grupę. Viršuje pavaizduoti VNP ir jų išsidėstymas vienas kito atžvilgiu. Ryški juoda linija rėmina haplotipų blokus. Skaičiai kvadratėliuose nurodo  $D'$  statistikos reikšmę dviem atitinkamiems VNP. Apačioje yra nurodyti haplotipų blokų haplotipai ir jų dažniai tiriamoje grupėje. Linijos tarp dviejų haplotipų blokų žymi mainų tarp haplotipų blokų dažnumą (storesnės linijos atitinka dažnesnius mainus, plonesnės – retesnius). Po linijomis pateikiama daugiasritė  $D'$  vertė, kuri šiuo atveju yra haplotipų blokų nepusiausviros sankibos matas. Balta kvadratėlių spalva nerodo statistiškai patikimos nepusiausviros sankibos –  $D' < 1$ ,  $LOD < 2$ , raudona rodo statistiškai patikimą nepusiausvirą sankibą –  $D' = 1$ ,  $LOD \geq 2$  ( $D'$  – santykinis nepusiausviros sankibos koeficientas,  $LOD$  – nepusiausviros sankibos tikimybės įvertis).

Trečios chromosomos dviem tirtiems VNP rs9818870 ir rs2306374, priklausantiems *MRAS* genui, visose tirtose grupėse nustatytas vienas haplotipų blokas (3.6 paveikslas). Nustatyti du haplotipai: AG ir GA. Rizikos alelių

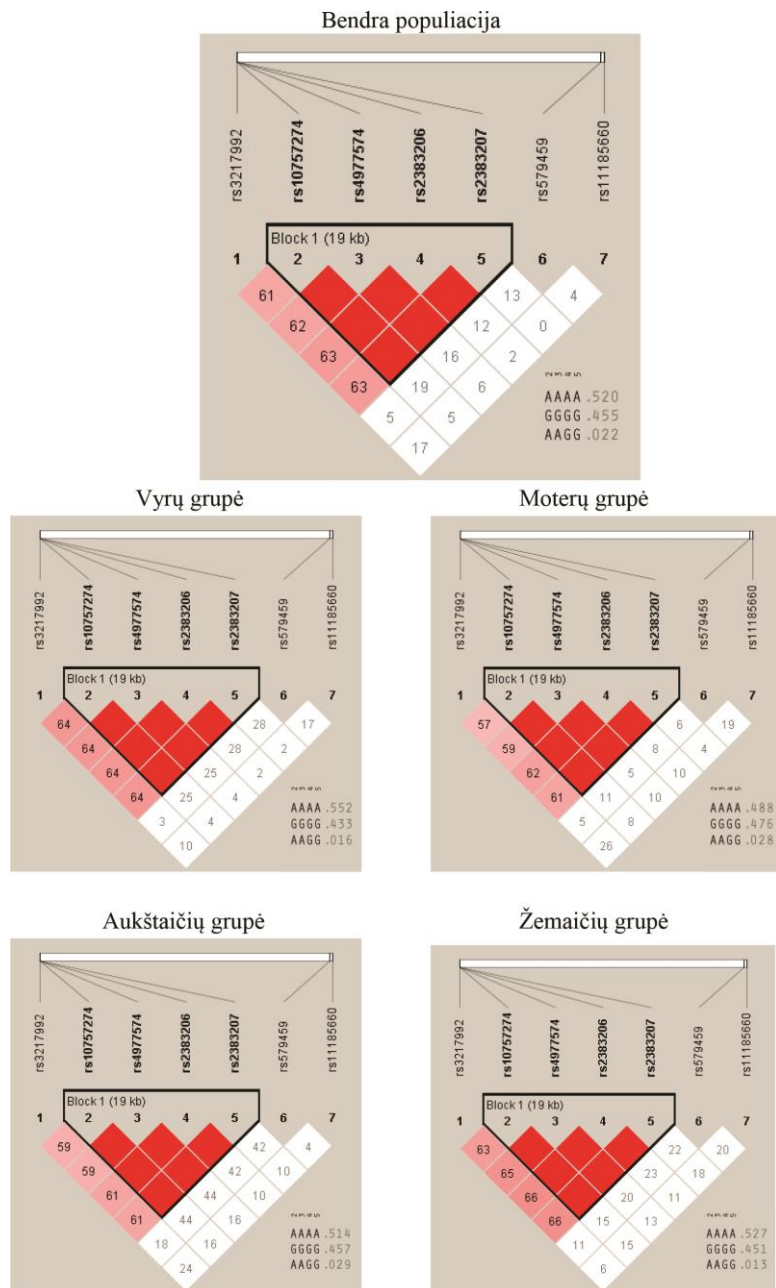
haplotipas GA visose tirtose grupėse yra aptinkamas rečiau (dažniai grupėse 0,14–0,17) negu alternatyvių alelių AG (dažniai grupėse 0,84–0,86).



**3.6 paveikslas.** Nepusiausviros sankibos genolapis ir haplotipų blokai 3-iai chromosomai pagal tiriamus 2 VNP tiriamajai populiacijai ir vidupopuliacinėms grupėms pagal lytį ir etnolingvistinę grupę. Viršuje pavaizduoti VNP ir jų išsidėstymas vienas kito atžvilgiu. Ryški juoda linija rėmina haplotipų bloką. Apačioje yra nurodyti haplotipų bloko haplotipai ir jų dažniai tiriamoje grupėje. Raudona kvadratėlio spalva rodo statistiškai patikimą nepusiausvyrą snakibą –  $D' = 1$ ,  $LOD \geq 2$  ( $D'$  – santykinis nepusiausviros sankibos koeficientas,  $LOD$  – nepusiausviros sankibos tikimybės įvertis).

Devintos chromosomos septyniems tirtiems VNP nustatytas didžiausias haplotipų blokas, kuris susideda iš keturių VNP rs10757274, rs4977574, rs2383206 ir rs2383207 esančių *CDKN2B-AS1* gene (**3.7 paveikslas**). Visose tirtose grupėse nustatyti trys haplotipai: GGGG, AAAA ir AAGG. Rizikos alelių haplotipas GGGG visose tiriamose grupėse yra nustatomas beveik vienodu dažnumu (dažnis grupėse 0,43–0,48), kaip ir alternatyvių alelių haplotipas AAAA (dažnis grupėse 0,49–0,55), rečiau nustatomas haplotipas AAGG (dažnis grupėse 0,01–0,03). Pažymėtina, kad VNP rs3217992, esantis šalia haplotipo bloko, visose tirtose grupėse demonstruoja  $LOD \geq 2$  ir  $D' > 0,5$  su visais haplotipų bloko VNP, kas leistų įtarti nepusiausviros sankibos galimybę.

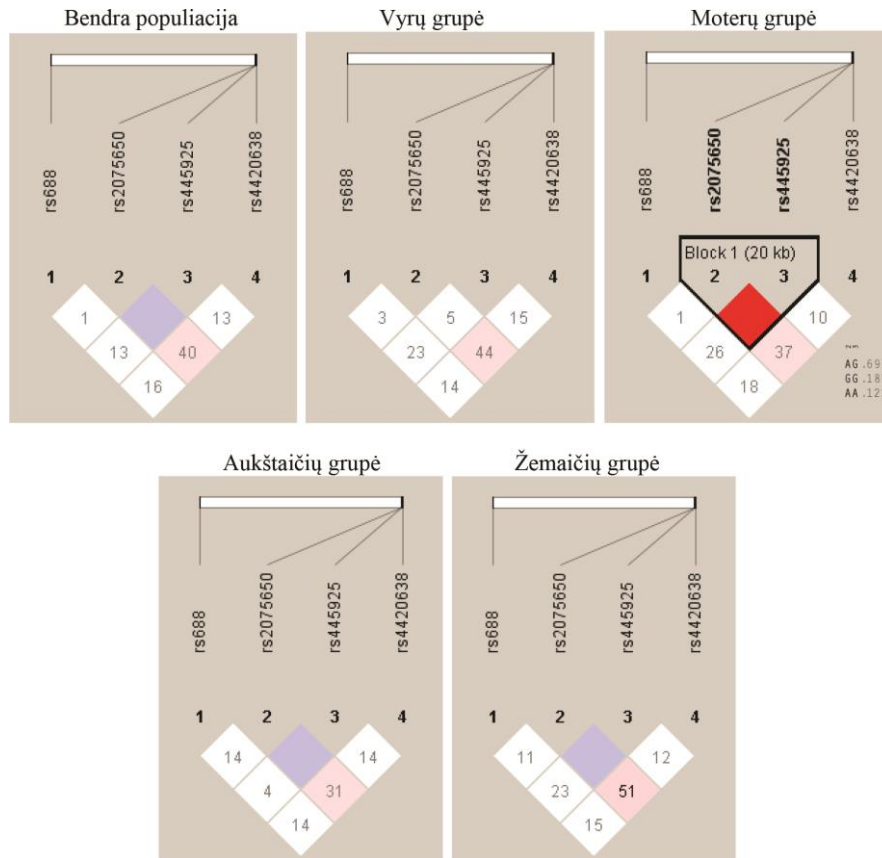




**3.7 paveikslas.** Nepusiausviros sankibos genolapis ir haplotipų blokas 9-ai chromosomai pagal tiriamus 7 VNP tiriamajai populiacijai ir vidupopuliacinėms grupėms pagal lytį ir etnolingvistinę grupę. Viršuje pavaizduoti VNP ir jų išsidėstymas vienas kito atžvilgiu. Ryški juoda linija rėmina haplotipų bloką. Skaičiai kvadratėliuose nurodo  $D'$  statistikos reikšmę dviem atitinkamiems VNP. Apatiniame dešiniajame kampe yra nurodyti haplotipo bloko haplotipai ir jų dažniai tiriamoje grupėje. Kvadratėlių spalvos: balta –  $D' < 1$ ,  $LOD < 2$ ; rožiniai atspalviai –  $D' < 1$ ,  $LOD \geq 2$ ; raudona –  $D' = 1$ ,  $LOD \geq 2$  ( $D'$  – santykinis nepusiausviros sankibos koeficientas,  $LOD$  – nepusiausviros sankibos tikimybės įvertis).

Devynioliktos chromosomos tirtiems keturiems VNP haplotipų blokas nustatytas tik moterų grupėje tarp dviejų VNP rs2075650 (*TOMM40*) ir rs445925 (*APOC1*), esančių srityje, kurioje yra vadinamasis *APOE* genų klasteris (**3.8 paveikslas**). Haplotipų bloke nustatyti trys haplotipai: AG, GG ir

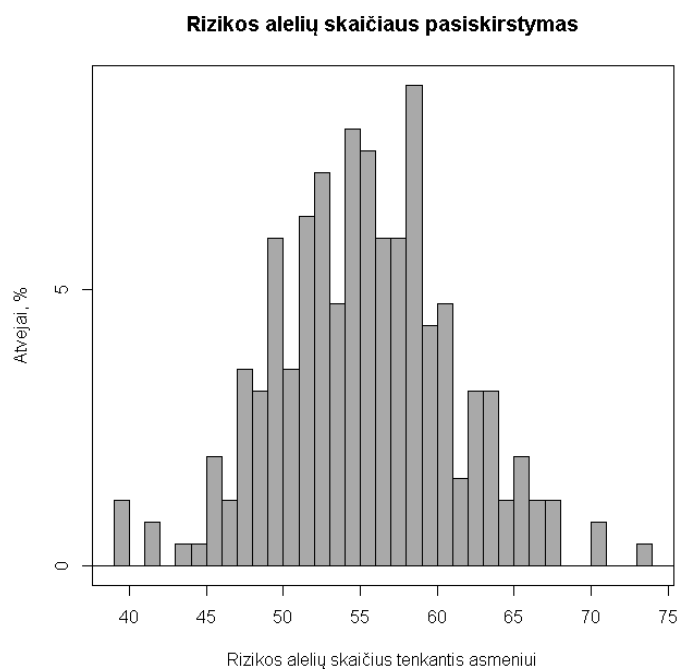
AA. Rizikos alelių haplotipas GG (dažnis grupėje 0,19), kaip ir haplotipas AA (dažnis grupėje 0,12), yra retesnis, palyginti su AG haplotipu, kurio dažnis moterų grupėje yra 0,69.



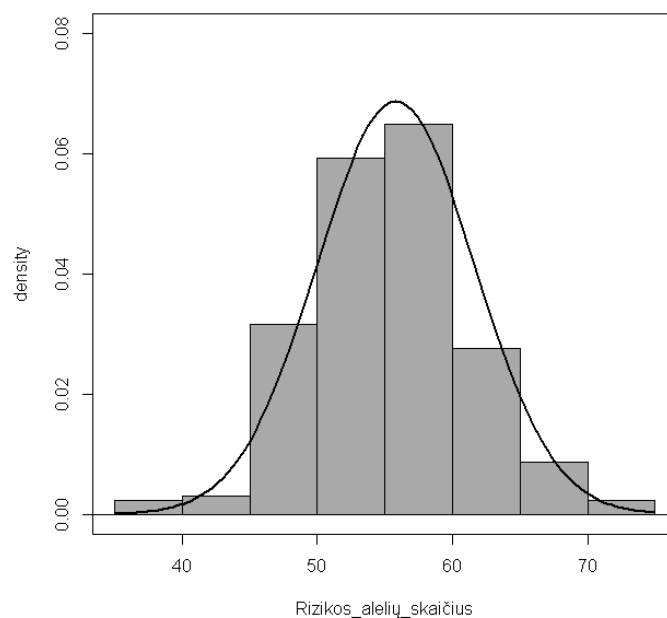
**3.8 paveikslas.** Nepusiausviro sankibos genolapis ir haplotipų blokas 19-ai chromosomai pagal tiriamus 4 VNP moterų grupei. Bendros populiacijos ir kitų vidupopuliacinių grupių haplotipai nesudarė haplotipų bloko. Viršuje pavaizduoti VNP ir jų išsidėstymas vienas kito atžvilgiu. Ryški juoda linija rėmina haplotipų bloką. Skaičiai kvadratėliuose nurodo  $D'$  statistikos reikšmę dviem atitinkamiems VNP. Apatiniame dešiniajame kampe moterų grupėje yra nurodyti haplotipo bloko haplotipai ir jų dažniai tiriamoje grupėje. Kvadratėlių spalvos: balta –  $D' < 1$ ,  $LOD < 2$ ; mėlva –  $D' = 1$ ,  $LOD < 2$ ; rožiniai atspalviai –  $D' < 1$ ,  $LOD \geq 2$ ; raudona –  $D' = 1$ ,  $LOD \geq 2$  ( $D'$  – santykinis nepusiausviro sankibos koeficientas,  $LOD$  – nepusiausviro sankibos tikimybės įvertis).

Tiriamai populiacijai ir vidupopuliacinėse grupėse apskaičiuotas tiriamų 60-ies VNP **rizikos alelių skaičius, tenkantis kiekvienam asmeniui**. Skaičiuotas kiekvienas riziką lemiantis alelis, vadinasi, 60-iai pasirinktų VNP maksimalus rizikos alelių skaičius būtų 120, jei asmuo būtų homozigotinis pagal visus rizikos alelius. Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas tiriamoje bendroje Lietuvos lietuvių populiacijoje pateikiamas **3.9 (a) paveiksle**. Duomenys aprašomi normaliuoju skirstiniu (**3.9 (b)**)

**paveikslas).** Rizikos alelių skaičiaus, tenkančio asmeniui, aprašomoji statistika pateikiama **3.7 lentelėje**.



**(a)**



**(b)**

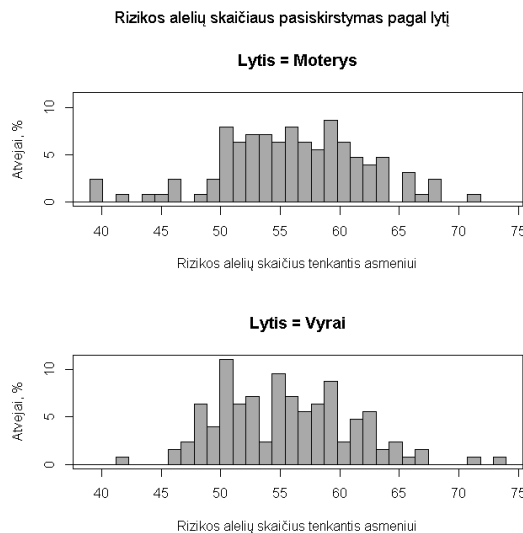
**3.9 paveikslas. (a)** Rizikos alelių, tenkančių asmeniui, pasiskirstymas tiriamoje bendroje Lietuvos lietuvių populiacijoje (%); **(b)** grafinis duomenų palyginimas su normaliojo skirstinio tankiu, *density* – tankis.

**3.7 lentelė.** Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas tiriamoje bendroje Lietuvos lietuvių populiacijoje

Statistika	Vertė
Rizikos alelių skaičiaus, tenkančio asmeniui, vidurkis	55,81
Mediana	56
Moda	59
Mažiausia reikšmė	39
Kvartilis 1, 25 %	52
Kvartilis 2, 50 %	56
Kvartilis 3, 75 %	59
Didžiausia reikšmė	74

Rizikos alelių skaičiaus, tenkančio asmeniui, aritmetinis vidurkis yra 55,81, mediana – 56, o dažniausiai pasitaikantis rizikos alelių skaičius – 59. Mažiausią (39) ir didžiausią (74) rizikos alelių skaičių turėjo po 0,4 % tirtų asmenų. Atliktas pasirinktų VNP rizikos alelių skaičiaus, tenkančio asmeniui, palyginimas tiriamoje bendroje Lietuvos lietuvių populiacijoje pagal grupes, t. y. pagal lytį (vyrai ir moterys) ir pagal etnolingvistinę grupę (aukštaičiai ir žemaičiai).

Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas vyrų ir moterų grupėse vaizduojamas **3.10 paveiksle**.



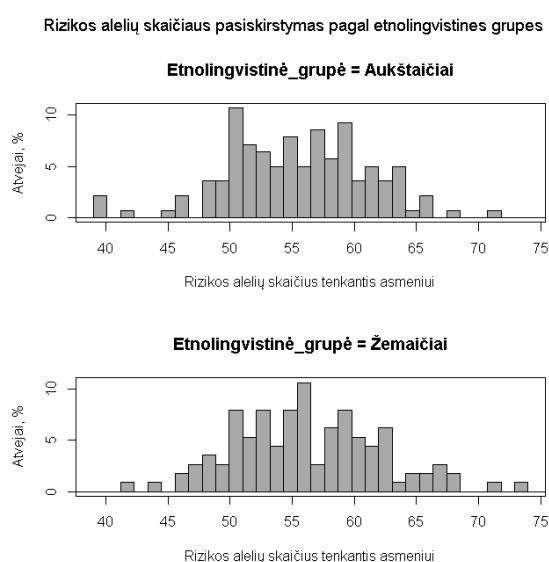
**3.10 paveikslas.** Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas tiriamos populiacijos vyrų ir moterų grupėse (%)

**Lentelėje 3.8** nurodyta aprašomoji statistika vyrų ir moterų grupėms. Atlikus rizikos alelių skaičiaus, tenkančio asmeniui, vidurkio ir dispersijos palyginimus, statistiškai reikšmingo skirtumo tarp vyrų ir moterų grupių nenustatyta (dviejų nepriklausomų imčių  $t$  testas,  $t = 0,81$ ,  $df = 251$ ,  $p = 0,42$ ;  $F$  testas dviejų dispersijų palyginimui  $F = 1,14$ , num  $df = 126$ , denom  $df = 125$ ,  $p = 0,46$ ).

**3.8 lentelė.** Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas vyrų ir moterų grupėse

Statistika	Vertė moterų gr.	Vertė vyrų gr.
Rizikos alelių skaičiaus, tenkančio asmeniui, vidurkis	56,10	55,51
Mediana	56	55
Moda	59	55
Mažiausia reikšmė	39	42
Kvartilis 1, 25 %	53	51
Kvartilis 2, 50 %	56	55
Kvartilis 3, 75 %	60	59
Didžiausia reikšmė	71	74

Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas aukštaičių ir žemaičių grupėse vaizduojamas **3.11 paveiksle**.



**3.11 paveikslas.** Rizikos alelių, tenkančių asmeniui, pasiskirstymas tiriamos populiacijos aukštaičių ir žemaičių grupėse (%)

**Lentelėje 3.9** nurodyta aprašomoji statistika aukštaičių ir žemaičių grupėms. Atlikus rizikos alelių skaičiaus, tenkančio asmeniui, vidurkio ir dispersijos palyginimus, statistiškai reikšmingo skirtumo tarp aukštaičių ir žemaičių grupių nerasta (dviejų nepriklausomų imčių  $t$  testas,  $t = -1,08$ ,  $df = 251$ ,  $p = 0,28$ ;  $F$  testas dviejų dispersijų palyginimui,  $F = 0,93$ , num  $df = 139$ , denom  $df = 112$ ,  $p = 0,68$ ).

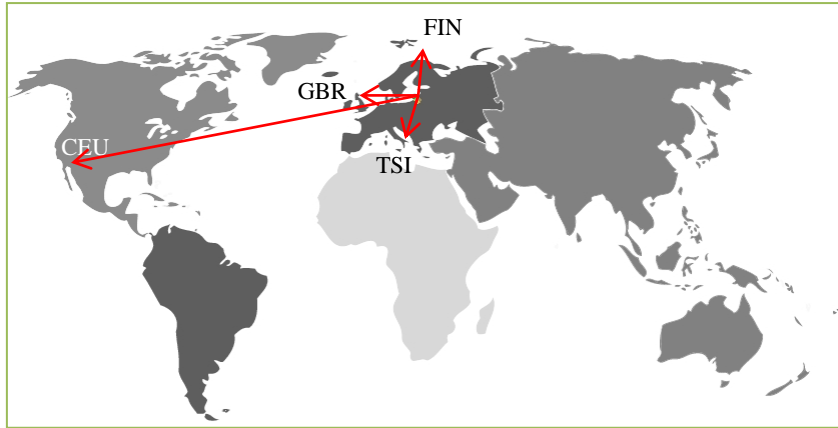
**3.9 lentelė.** Rizikos alelių skaičiaus, tenkančio asmeniui, pasiskirstymas aukštaičių ir žemaičių grupėse

Statistika	Vertė aukštaičių gr.	Vertė žemaičių gr.
Rizikos alelių skaičiaus, tenkančio asmeniui, vidurkis	55,45	56,25
Mediana	55,5	56
Moda	59	56
Mažiausia reikšmė	39	42
Kvartilis 1, 25 %	52	52
Kvartilis 2, 50 %	55,5	56
Kvartilis 3, 75 %	59	60
Didžiausia reikšmė	71	74

### 3.2.2. Tarpopuliacinė genetinė struktūra ir įvairovė pagal tiriamus VNP

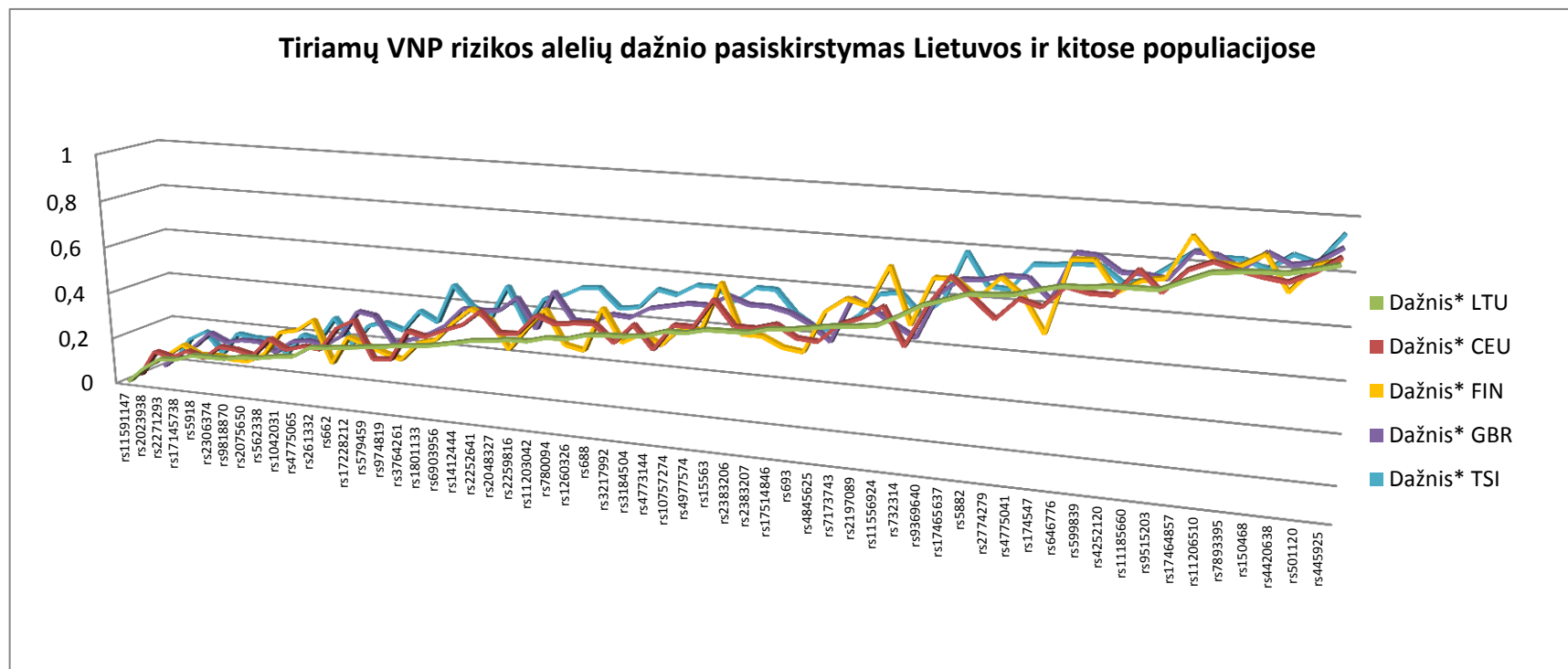
Tarppopuliacinis genetinis profilis nustatytas įvertinant atrinktų su liga siejamų VNP alelių dažnius bendroje Lietuvos lietuvių populiacijoje ir palyginus su 1000 genomų projekto atitinkamais duomenimis. Pagal duomenų prieinamumą ir imtį, populiacijų protėvius ir geografinę padėtį pasirinktos keturios Europos protėvių (angl. *European Ancestry*) populiacijos palyginimui pagal tiriamus VNP. Vakarų Europą reprezentuoja Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai (angl. *Utah residents with Northern and Western European ancestry* – CEU) bei britai Anglijoje ir Škotijoje (angl. *British in England and Scotland* – GBR), Šiaurės Europą – suomiai Suomijoje (angl. *Finnish in Finland* – FIN), Pietų Europą – toskaniečiai Italijoje (angl. *Toscani in Italy* – TSI)), t. y. trys geografiniai poliai: šiaurė, vakarai, pietūs (**3.12 paveikslas**). Alelių dažniams nustatyti ir palyginti naudojami

chromosomų skaičiai: tiriamos bendros Lietuvos lietuvių populiacijos (LTU) – 506, CEU – 170, GBR – 178, FIN – 186, TSI – 196. Deja, Rytų Europos prieinamų duomenų analizei gauti nepavyko.



**3.12 paveikslas.** Pasaulio žemėlapyje pavaizduotos populiacijos, su kuriomis palyginti tiriamos bendros Lietuvos lietuvių populiacijos pasirinktų VNP rizikos alelių dažniai. CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai, FIN – suomia Suomijoje, GBR – britai Anglijoje ir Škotijoje, TSI – toskaniečiai Italijoje. Žemėlapio šabloną parengė A. Pranculis.

Nustatytų VNP rizikos alelių dažnio pasiskirstymas tiriamoje bendroje Lietuvos lietuvių populiacijoje ir palyginimui CEU, FIN, GBR bei TSI grupėse vaizduojamas **3.13 paveiksle**. Iš kreivių matyti, kad rizikos alelių dažniai yra įvairūs ir skiriasi populiacijose. Rizikos alelių dažnio reikšmių pasiskirstymo poliuose matoma tam tikra tendencija. Rečiausi ir dažniausi aleliai populiacijose yra pasiskirstę tolygiau. Rečiausiai pasitaikantis alelis visose populiacijose buvo tas pats VNP rs11591147 *PCSK9* gene. Palyginus bendros Lietuvos lietuvių populiacijos šio VNP (rs11591147) alelių dažnius su kitų populiacijų, statistiškai reikšmingai skyrėsi tik lietuvių ir suomių populiacijų rizikos alelio dažniai, lietuvių populiacijoje alelis buvo retesnis (žr. **3.11 lentelę**). Dažniausiai pasitaikęs rizikos alelis visose populiacijose buvo tas pats VNP rs445925, siejamas su *APOE* geno sritimi, išskyrus suomių populiaciją – joje dažniausias iš tiriamų VNP yra rs17464857 rizikos alelis *MIA3* gene 1-oje chromosomoje.



**3.13 paveikslas.** Tiriamos bendros Lietuvos lietuvių populiacijos rizikos alelių dažnių pasiskirstymas, palyginti su kitomis populiacijomis. LTU – tiriama bendra Lietuvos lietuvių populiacija, CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai, FIN – suomia Suomijoje, GBR – britai Anglijoje ir Škotijoje, TSI – toskaniečiai Italijoje. \* – rizikos alelio dažnis



Atliktas tiriamų VNP alelių dažnių palyginimas tirtos bendros Lietuvos lietuvių populiacijos su kitomis populiacijomis: CEU (**3.10 lentelė**), FIN (**3.11 lentelė**), GBR (**3.12 lentelė**) ir TSI (**3.13 lentelė**). Pateiktos  $\chi^2$  kriterijaus reikšmės ir statistinis reikšmingumas. Pasirinktas reikšmingumo lygmuo 0,05.

**3.10 lentelė.** Bendrai Lietuvos lietuvių populiacijai nustatytų atrinktų VNP alelių dažnių palyginimas su CEU

VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	CEU		
rs1042031	A	0,186	0,194	0,058	0,810
rs10757274	G	0,455	0,459	0,009	0,923
rs11185660	A	0,770	0,818	1,704	0,192
rs11203042	A	<b>0,362</b>	<b>0,447</b>	<b>3,927</b>	<b>0,048</b>
rs11206510	A	0,844	0,865	0,431	0,512
rs11556924	G	0,553	0,606	1,430	0,232
rs11591147	A	0,004	0,012	1,320	0,251
rs1260326	A	0,385	0,429	1,031	0,310
rs1412444	A	0,322	0,376	1,684	0,194
rs150468	A	0,860	0,824	1,329	0,249
rs15563	G	<b>0,476</b>	<b>0,576</b>	<b>5,113</b>	<b>0,024</b>
rs17145738	A	0,125	0,141	0,315	0,574
rs17228212	G	<b>0,257</b>	<b>0,353</b>	<b>5,807</b>	<b>0,016</b>
rs174547	A	<b>0,735</b>	<b>0,653</b>	<b>4,219</b>	<b>0,040</b>
rs17464857	A	0,808	0,829	0,374	0,541
rs17465637	C	0,671	0,741	2,949	0,086
rs17514846	A	0,502	0,500	0,002	0,964
rs1801133	A	0,295	0,312	0,182	0,670
rs2023938	G	<b>0,069</b>	<b>0,124</b>	<b>4,949</b>	<b>0,026</b>
rs2048327	G	0,348	0,359	0,068	0,795
rs2075650	G	0,168	0,153	0,210	0,647
rs2197089	A	0,542	0,553	0,065	0,799
rs2252641	G	<b>0,340</b>	<b>0,441</b>	<b>5,626</b>	<b>0,018</b>
rs2259816	A	0,360	0,365	0,014	0,906
rs2271293	A	0,119	0,106	0,201	0,654
rs2306374	G	0,152	0,176	0,564	0,453
rs2383206	G	0,478	0,476	0,002	0,968
rs2383207	G	0,480	0,476	0,007	0,932
rs261332	A	0,241	0,212	0,612	0,434
rs2774279	G	<b>0,704</b>	<b>0,594</b>	<b>6,959</b>	<b>0,008</b>
rs3184504	A	0,419	0,447	0,411	0,522
rs3217992	A	0,417	0,365	1,445	0,229
rs3764261	A	0,289	0,329	1,015	0,314
rs4252120	A	0,769	0,718	1,803	0,179
rs4420638	A	0,862	0,812	2,473	0,116
rs445925	G	0,903	0,912	0,110	0,740
rs4773144	G	0,429	0,353	3,007	0,083
rs4775041	C	0,712	0,676	0,746	0,388
rs4775065	A	0,194	0,218	0,427	0,513
rs4845625	A	0,522	0,447	2,839	0,092
rs4977574	G	0,457	0,459	0,003	0,958
rs501120	A	0,881	0,853	0,905	0,341
rs562338	A	0,172	0,235	3,348	0,067
rs579459	G	0,267	0,194	3,599	0,058
rs5882	A	0,700	0,665	0,726	0,394
rs5918	G	0,151	0,118	1,144	0,285
rs599839	A	0,758	0,718	1,094	0,296
rs646776	A	0,756	0,735	0,290	0,590
rs662	G	0,247	0,312	2,749	0,097
rs688	A	0,409	0,435	0,360	0,549

3.10 lentelės tęsinys					
VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	CEU		
rs6903956	A	0,306	0,347	1,013	0,314
rs693	A	0,508	0,453	1,538	0,215
rs7173743	A	0,536	0,524	0,076	0,783
rs732314	A	<b>0,597</b>	<b>0,459</b>	<b>9,847</b>	<b>0,002</b>
rs780094	A	0,383	0,418	0,626	0,429
rs7893395	G	0,854	0,841	0,158	0,691
rs9369640	A	0,644	0,624	0,237	0,626
rs9515203	A	0,771	0,741	0,616	0,433
rs974819	A	0,275	0,200	3,741	0,053
rs9818870	A	0,152	0,171	0,326	0,568

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, LTU – tiriama bendra Lietuvos lietuvių populiacija, CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai, RAD – rizikos alelio dažnis,  $\chi^2$  – Pirsono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, **paryškintos pasiviro reikšmės** – nustatyti statistiškai reikšmingi skirtumai.

### 3.11 lentelė. Bendrai Lietuvos lietuvių populiacijai nustatytų atrinktų VNP alelių dažnių palyginimas su FIN

VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	FIN		
rs1042031	A	<b>0,186</b>	<b>0,258</b>	<b>4,358</b>	<b>0,037</b>
rs10757274	G	0,455	0,43	0,329	0,567
rs11185660	A	0,770	0,758	0,105	0,746
rs11203042	A	<b>0,362</b>	<b>0,452</b>	<b>4,644</b>	<b>0,031</b>
rs11206510	A	0,844	0,844	0,000	0,995
rs11556924	G	<b>0,553</b>	<b>0,742</b>	<b>20,222</b>	<b><math>6,897 \times 10^{-6}</math></b>
rs11591147	A	<b>0,004</b>	<b>0,038</b>	<b>12,020</b>	<b><math>5,264 \times 10^{-4}</math></b>
rs1260326	A	<b>0,385</b>	<b>0,296</b>	<b>4,740</b>	<b>0,029</b>
rs1412444	A	<b>0,322</b>	<b>0,425</b>	<b>6,295</b>	<b>0,012</b>
rs150468	A	0,860	0,887	0,865	0,352
rs15563	G	<b>0,476</b>	<b>0,624</b>	<b>11,830</b>	<b><math>5,829 \times 10^{-4}</math></b>
rs17145738	A	0,125	0,102	0,651	0,420
rs17228212	G	0,257	0,226	0,705	0,401
rs174547	A	<b>0,735</b>	<b>0,532</b>	<b>25,764</b>	<b><math>3,859 \times 10^{-7}</math></b>
rs17464857	A	<b>0,808</b>	<b>0,935</b>	<b>16,578</b>	<b><math>4,669 \times 10^{-5}</math></b>
rs17465637	C	0,671	0,71	0,954	0,329
rs17514846	A	<b>0,502</b>	<b>0,387</b>	<b>7,191</b>	<b>0,007</b>
rs1801133	A	0,295	0,28	0,146	0,702
rs2023938	G	0,069	0,086	0,566	0,452
rs2048327	G	<b>0,348</b>	<b>0,269</b>	<b>3,860</b>	<b>0,049</b>
rs2075650	G	0,168	0,14	0,803	0,370
rs2197089	A	0,542	0,581	0,835	0,361
rs2252641	G	0,340	0,387	1,326	0,250
rs2259816	A	0,360	0,366	0,021	0,886
rs2271293	A	0,119	0,14	0,562	0,453
rs2306374	G	0,152	0,097	3,524	0,060
rs2383206	G	0,478	0,425	1,567	0,211
rs2383207	G	0,480	0,425	1,684	0,194
rs261332	A	<b>0,241</b>	<b>0,129</b>	<b>10,262</b>	<b>0,001</b>
rs2774279	G	0,704	0,731	0,505	0,477
rs3184504	A	0,419	0,387	0,571	0,450
rs3217992	A	0,417	0,344	3,019	0,082
rs3764261	A	0,289	0,258	0,626	0,429
rs4252120	A	0,769	0,71	2,553	0,110
rs4420638	A	<b>0,862</b>	<b>0,758</b>	<b>10,534</b>	<b>0,001</b>
rs445925	G	0,903	0,887	0,386	0,534
rs4773144	G	0,429	0,349	3,522	0,061
rs4775041	C	0,712	0,667	1,298	0,255
rs4775065	A	<b>0,194</b>	<b>0,312</b>	<b>10,699</b>	<b>0,001</b>
rs4845625	A	0,522	0,543	0,247	0,619
rs4977574	G	0,457	0,425	0,556	0,456

3.11 lentelės tęsinys					
VNP	Rizikos alelis	RAD		$\chi^2$	<i>p</i>
		LTU	FIN		
rs501120	A	0,881	0,86	0,536	0,464
rs562338	A	<b>0,172</b>	<b>0,242</b>	<b>4,317</b>	<b>0,038</b>
rs579459	G	0,267	0,199	3,355	0,067
rs5882	A	0,700	0,645	1,868	0,172
rs5918	G	0,151	0,108	2,124	0,145
rs599839	A	0,758	0,823	3,260	0,071
rs646776	A	0,756	0,817	2,900	0,089
rs662	G	0,247	0,253	0,023	0,879
rs688	A	0,409	0,478	2,676	0,102
rs6903956	A	0,306	0,36	1,865	0,172
rs693	A	<b>0,508</b>	<b>0,376</b>	<b>9,445</b>	<b>0,002</b>
rs7173743	A	0,536	0,602	2,427	0,119
rs732314	A	0,597	0,522	3,162	0,075
rs780094	A	0,383	0,312	3,009	0,083
rs7893395	G	0,854	0,839	0,241	0,623
rs9369640	A	0,644	0,71	2,600	0,107
rs9515203	A	0,771	0,774	0,009	0,924
rs974819	A	<b>0,275</b>	<b>0,177</b>	<b>6,905</b>	<b>0,009</b>
rs9818870	A	0,152	0,097	3,524	0,060

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, LTU – tiriama bendra Lietuvos lietuvių populiacija, FIN – suomia Suomijoje, RAD – rizikos alelio dažnis,  $\chi^2$  – Pirsono chi kvadrato kriterijus, *p* – *p* reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, **paryškintos pasviros reikšmės** – nustatyti statistiškai reikšmingi skirtumai.

**3.12 lentelė.** Bendrai Lietuvos lietuvių populiacijai nustatytų atrinktų VNP alelių dažnių palyginimas su GBR

VNP	Rizikos alelis	RAD		$\chi^2$	<i>p</i>
		LTU	GBR		
rs1042031	A	0,186	0,191	0,024	0,878
rs10757274	G	0,455	0,506	1,380	0,240
rs11185660	A	0,770	0,764	0,025	0,875
rs11203042	A	<b>0,362</b>	<b>0,506</b>	<b>11,379</b>	<b>7,429×10<sup>-4</sup></b>
rs11206510	A	0,844	0,854	0,103	0,749
rs11556924	G	0,553	0,494	1,843	0,175
rs11591147	A	0,004	0,006	0,084	0,772
rs1260326	A	0,385	0,393	0,035	0,853
rs1412444	A	0,322	0,399	3,446	0,063
rs150468	A	0,860	0,876	0,301	0,583
rs15563	G	0,476	0,551	2,911	0,088
rs17145738	A	0,125	0,180	3,363	0,067
rs17228212	G	0,257	0,331	3,659	0,056
rs174547	A	<b>0,735</b>	<b>0,640</b>	<b>5,745</b>	<b>0,017</b>
rs17464857	A	0,808	0,860	2,360	0,125
rs17465637	C	0,671	0,691	0,249	0,618
rs17514846	A	0,502	0,506	0,007	0,934
rs1801133	A	0,295	0,320	0,415	0,519
rs2023938	G	0,069	0,067	0,006	0,937
rs2048327	G	<b>0,348</b>	<b>0,466</b>	<b>7,855</b>	<b>0,005</b>
rs2075650	G	0,168	0,124	1,966	0,161
rs2197089	A	0,542	0,545	0,006	0,940
rs2252641	G	0,340	0,404	2,393	0,122
rs2259816	A	0,360	0,343	0,166	0,684
rs2271293	A	0,119	0,112	0,049	0,824
rs2306374	G	0,152	0,163	0,116	0,733
rs2383206	G	0,478	0,522	1,030	0,310
rs2383207	G	0,480	0,522	0,940	0,332
rs261332	A	0,241	0,230	0,084	0,772
rs2774279	G	0,704	0,719	0,154	0,695
rs3184504	A	0,419	0,472	1,503	0,220
rs3217992	A	0,417	0,427	0,054	0,817
rs3764261	A	0,289	0,275	0,114	0,736

**3.12 lentelės tęsinys**

VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	GBR		
rs4252120	A	0,769	0,764	0,017	0,898
rs4420638	A	0,862	0,837	0,643	0,423
rs445925	G	0,903	0,910	0,074	0,786
rs4773144	G	0,429	0,489	1,930	0,165
rs4775041	C	0,712	0,719	0,038	0,846
rs4775065	A	0,194	0,180	0,184	0,668
rs4845625	A	<b>0,522</b>	<b>0,404</b>	<b>7,243</b>	<b>0,007</b>
rs4977574	G	0,457	0,506	1,274	0,259
rs501120	A	0,881	0,854	0,872	0,351
rs562338	A	0,172	0,180	0,056	0,812
rs579459	G	0,267	0,219	1,579	0,209
rs5882	A	0,700	0,697	0,006	0,941
rs5918	G	0,151	0,146	0,023	0,879
rs599839	A	0,758	0,820	2,919	0,088
rs646776	A	0,756	0,826	3,676	0,055
rs662	G	<b>0,247</b>	<b>0,343</b>	<b>6,086</b>	<b>0,014</b>
rs688	A	0,409	0,433	0,299	0,584
rs6903956	A	0,306	0,399	5,182	0,023
rs693	A	0,508	0,466	0,912	0,340
rs7173743	A	0,536	0,584	1,252	0,263
rs732314	A	<b>0,597</b>	<b>0,449</b>	<b>11,603</b>	<b>6,586×10<sup>-4</sup></b>
rs780094	A	0,383	0,393	0,054	0,816
rs7893395	G	0,854	0,809	1,985	0,159
rs9369640	A	0,644	0,629	0,130	0,719
rs9515203	A	0,771	0,753	0,237	0,627
rs974819	A	0,275	0,242	0,747	0,387
rs9818870	A	0,152	0,163	0,116	0,733

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, LTU – tiriama bendra Lietuvos lietuvių populiacija, GBR – britai Anglijoje ir Škotijoje, RAD – rizikos alelio dažnis,  $\chi^2$  – Pirsono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, **paryškintos pasviros reikšmės** – nustatyti statistiškai reikšmingi skirtumai.

**3.13 lentelė.** Bendrai Lietuvos lietuvių populiacijai nustatytų atrinktų VNP alelių dažnių palyginimas su TSI

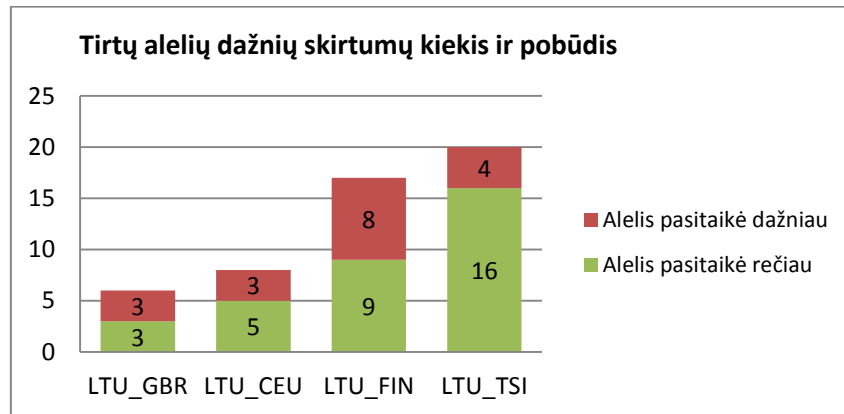
VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	TSI		
rs1042031	A	0,186	0,173	0,143	0,705
rs10757274	G	<b>0,455</b>	<b>0,561</b>	<b>6,438</b>	<b>0,011</b>
rs11185660	A	0,770	0,724	1,579	0,209
rs11203042	A	<b>0,362</b>	<b>0,469</b>	<b>6,881</b>	<b>0,009</b>
rs11206510	A	0,844	0,827	0,315	0,575
rs11556924	G	0,553	0,602	1,364	0,243
rs11591147	A	0,004	0,015	2,575	0,109
rs1260326	A	<b>0,385</b>	<b>0,515</b>	<b>9,780</b>	<b>0,002</b>
rs1412444	A	0,322	0,347	0,394	0,530
rs150468	A	<b>0,860</b>	<b>0,796</b>	<b>4,345</b>	<b>0,037</b>
rs15563	G	0,476	0,510	0,654	0,419
rs17145738	A	<b>0,125</b>	<b>0,066</b>	<b>4,953</b>	<b>0,026</b>
rs17228212	G	0,257	0,281	0,409	0,523
rs174547	A	0,735	0,755	0,292	0,589
rs17464857	A	0,808	0,847	1,423	0,233
rs17465637	C	<b>0,671</b>	<b>0,781</b>	<b>8,150</b>	<b>0,004</b>
rs17514846	A	0,502	0,480	0,283	0,595
rs1801133	A	<b>0,295</b>	<b>0,474</b>	<b>20,269</b>	<b>6,729×10<sup>-6</sup></b>
rs2023938	G	<b>0,069</b>	<b>0,117</b>	<b>4,326</b>	<b>0,038</b>
rs2048327	G	0,348	0,327	0,285	0,594
rs2075650	G	<b>0,168</b>	<b>0,097</b>	<b>5,650</b>	<b>0,017</b>
rs2197089	A	0,542	0,587	1,160	0,281
rs2252641	G	<b>0,340</b>	<b>0,490</b>	<b>13,446</b>	<b>2,456×10<sup>-4</sup></b>
rs2259816	A	<b>0,360</b>	<b>0,449</b>	<b>4,760</b>	<b>0,029</b>

3.13 lentelės tęsinys

VNP	Rizikos alelis	RAD		$\chi^2$	p
		LTU	TSI		
rs2271293	A	0,119	0,158	1,962	0,161
rs2306374	G	0,152	0,153	0,001	0,977
rs2383206	G	<b>0,478</b>	<b>0,571</b>	<b>4,906</b>	<b>0,027</b>
rs2383207	G	<b>0,480</b>	<b>0,571</b>	<b>4,700</b>	<b>0,030</b>
rs261332	A	0,241	0,173	3,739	0,053
rs2774279	G	0,704	0,648	2,035	0,154
rs3184504	A	<b>0,419</b>	<b>0,531</b>	<b>7,114</b>	<b>0,008</b>
rs3217992	A	0,417	0,449	0,591	0,442
rs3764261	A	0,289	0,306	0,211	0,646
rs4252120	A	0,769	0,704	3,159	0,075
rs4420638	A	0,862	0,857	0,024	0,877
rs445925	G	0,903	0,944	2,996	0,083
rs4773144	G	<b>0,429</b>	<b>0,515</b>	<b>4,285</b>	<b>0,038</b>
rs4775041	C	0,712	0,750	1,046	0,307
rs4775065	A	<b>0,194</b>	<b>0,281</b>	<b>6,135</b>	<b>0,013</b>
rs4845625	A	<b>0,522</b>	<b>0,434</b>	<b>4,383</b>	<b>0,036</b>
rs4977574	G	<b>0,457</b>	<b>0,561</b>	<b>6,200</b>	<b>0,013</b>
rs501120	A	0,881	0,832	2,980	0,084
rs562338	A	0,172	0,189	0,275	0,600
rs579459	G	0,267	0,255	0,100	0,752
rs5882	A	0,700	0,653	1,422	0,233
rs5918	G	0,151	0,163	0,168	0,682
rs599839	A	0,758	0,765	0,042	0,838
rs646776	A	0,756	0,765	0,067	0,795
rs662	G	0,247	0,260	0,130	0,718
rs688	A	0,409	0,439	0,512	0,474
rs6903956	A	<b>0,306</b>	<b>0,388</b>	<b>4,322</b>	<b>0,038</b>
rs693	A	0,508	0,429	3,560	0,059
rs7173743	A	0,536	0,495	0,943	0,332
rs732314	A	0,597	0,515	3,841	0,050
rs780094	A	<b>0,383</b>	<b>0,510</b>	<b>9,333</b>	<b>0,002</b>
rs7893395	G	0,854	0,827	0,804	0,370
rs9369640	A	0,644	0,587	2,002	0,157
rs9515203	A	0,771	0,786	0,181	0,670
rs974819	A	0,275	0,347	3,517	0,061
rs9818870	A	0,152	0,153	0,001	0,977

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, LTU – tiriama bendra Lietuvos lietuvių populiacija, TSI – toskaniečiai Italijoje, RAD – rizikos alelio dažnis,  $\chi^2$  – Pirono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, **paryškintos pasviros reikšmės** – nustatyti statistiškai reikšmingi skirtumai.

Nustatyti statistiškai reikšmingi skirtumai tarp tirtų VNP alelių dažnių tiriamoje bendroje Lietuvos lietuvių populiacijoje, palyginti su kitomis populiacijomis (CEU, FIN, GBR, TSI). Daugiausiai statistiškai reikšmingų skirtumų pagal tiriamų VNP rizikos alelių dažnius nustatyta tarp lietuvių ir toskaniečių (20 skirtumų), mažiausiai (6 skirtumai) – lietuvių ir britų (žr. **3.14 paveikslą**).



**3.14 paveikslas.** Statistiškai reikšmingai besiskiriančių tirtų VNP alelių dažnių atvejai bendroje Lietuvos lietuvių populiacijoje, palyginti su kitomis populiacijomis. LTU – tiriamos bendros Lietuvos lietuvių populiacija, CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai, FIN – suomiaai Suomijoje, GBR – britai Anglijoje ir Škotijoje, TSI – toskaniečiai Italijoje.

Pagal 60 tiriamų VNP alelių dažnį tarp tiriamos bendros Lietuvos lietuvių populiacijos ir kitų populiacijų iš viso nustatytas 51 statistiškai reikšmingas skirtumas. Iš jų 19,61 % buvo aleliai, kurių dažnių skirtumas buvo nustatytas keliose populiacijose, su kuriomis buvo palyginta lietuvių populiacija (**3.14 lentelė**). Alelių dažnių skirtumai yra įvairaus pobūdžio, t. y. kai kuriais atvejais rizikos alelis lietuvių populiacijoje buvo retesnis, kai kuriais – dažnesnis. Palyginus su CEU populiacija, iš aštuonių nustatytų reikšmingų alelių dažnių skirtumų trimis atvejais rizikos alelis buvo dažnesnis lietuvių populiacijoje (37,5 %), penkiais atvejais – retesnis (62,5 %). Palyginus su FIN populiacija, iš 17 nustatytų reikšmingų alelių dažnių skirtumų aštuoniais atvejais rizikos alelis buvo dažnesnis lietuvių populiacijoje (47,1 %), devyniais atvejais – retesnis (52,9 %). Palyginus su GBR populiacija, iš šešių nustatytų reikšmingų alelių dažnių skirtumų trimis atvejais rizikos alelis buvo dažnesnis lietuvių populiacijoje (50 %), trimis atvejais – retesnis (50 %). Palyginus su TSI populiacija, su kuria nustatyta daugiausia skirtumų, iš 20-ies nustatytų reikšmingų alelių dažnių skirtumų keturiais atvejais rizikos alelis buvo dažnesnis lietuvių populiacijoje (20 %), 16 atvejų – retesnis (80 %). Taigi

apskritai iš nustatyto 51 alelių dažnio skirtumų atvejo 18 atvejų (31,29 %) rizikos alelių dažniai lietuvių populiacijoje buvo dažnesni nei tose populiacijose, su kuriomis palyginta, o 33 atvejais (64,71 %) – retesni. Tai dar vienas populiacijos genetinės įvairovės ir unikalumo įrodymas. Vieno nukleotido polimorfizmo *LIPA* gene (rs11203042) alelių dažniai lietuvių populiacijoje statistiškai reikšmingai skyrėsi nuo visų populiacijų, su kuriomis buvo atliktas alelių dažnių palyginimas. Visais atvejais rizikos alelis buvo retesnis lietuvių populiacijoje.

**3.14 lentelė.** Dešimt VNP rizikos alelių dažnių skirtumų, kurie nustatyti daugiau nei vienoje lyginamojoje populiacijoje

Genas (-ai)	VNP	RA	LTU *	CEU*	<i>p</i>	FIN*	<i>p</i>	GBR*	<i>p</i>	TSI*	<i>p</i>	Gr
<i>LIPA</i>	rs11203042	A	<b>0,36</b>	0,45	<b>0,05</b>	0,45	<b>0,03</b>	0,51	<b><math>7,43 \times 10^{-4}</math></b>	0,47	<b>0,01</b>	4+
<i>GCKR</i>	rs1260326	A	<b>0,39</b>	0,43	0,31	0,30	<b>0,03</b>	0,39	0,85	0,52	<b>0,00</b>	2+ /-
<i>UBE2Z</i>	rs15563	G	<b>0,48</b>	0,58	<b>0,02</b>	0,62	<b><math>5,83 \times 10^{-4}</math></b>	0,55	0,09	0,51	0,42	2+
<i>FADS1</i>	rs174547	A	<b>0,74</b>	0,65	<b>0,04</b>	0,53	<b><math>3,86 \times 10^{-7}</math></b>	0,64	<b>0,02</b>	0,76	0,59	3-
<i>HDAC9</i>	rs2023938	G	<b>0,07</b>	0,12	<b>0,03</b>	0,09	0,45	0,07	0,94	0,12	<b>0,04</b>	2+
<i>SLC22A3/LPAL2/LPA</i>	rs2048327	G	<b>0,35</b>	0,36	0,79	0,27	<b>0,05</b>	0,47	<b>0,01</b>	0,33	0,59	2+ /-
<i>ZEB2</i>	rs2252641	G	<b>0,34</b>	0,44	<b>0,02</b>	0,39	0,25	0,40	0,12	0,49	<b><math>2,46 \times 10^{-4}</math></b>	2+
<i>LIPC</i>	rs4775065	A	<b>0,19</b>	0,22	0,51	0,31	<b><math>1 \times 10^{-3}</math></b>	0,18	0,67	0,28	<b>0,01</b>	2+
<i>IL6R</i>	rs4845625	A	<b>0,52</b>	0,45	0,09	0,54	0,62	0,40	<b>0,01</b>	0,43	<b>0,04</b>	2-
<i>SELP</i>	rs732314	A	<b>0,60</b>	0,46	<b><math>2 \times 10^{-3}</math></b>	0,52	0,08	0,45	<b><math>6,59 \times 10^{-4}</math></b>	0,52	0,05	2-

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, RA – rizikos alelis, LTU – tiriama bendra Lietuvos lietuvių populiacija, CEU – Šiaurės ir Vakarų Europos protėvių palikuonys Jutos gyventojai, FIN – suomiai Suomijoje, GBR – britai Anglijoje ir Škotijoje, TSI – toskaniečiai Italijoje, \* – rizikos alelio dažnis, *p* – *p* reikšmė  $\chi^2$  statistikai su 1 laisvės laipsniu, Gr – sudarytos grupės: skaičius nurodo, keliose populiacijose yra nustatyti skirtumai, ženklas „+“ arba „-“ nurodo atitinkamai alelis LTU yra retesnis nei lyginamosiose populiacijose arba dažnesnis, ženklas „+/-“ nurodo nevienareikšmį skirtumą.



Vieno nukleotido polimorfizmo *FADS1* gene (rs174547) alelių dažniai bendroje Lietuvos lietuvių populiacijoje statistiškai reikšmingai skyrėsi nuo visų, išskyrus Italijos, populiacijų. Visais skirtumo atvejais rizikos alelis buvo dažnesnis bendroje Lietuvos lietuvių populiacijoje. Aštuoniais VNP alelių dažnio skirtumų atvejais bendra Lietuvos lietuvių populiacija skyrėsi nuo dviejų populiacijų, o skirtumo pobūdis, kaip matyti iš **3.14 lentelės**, yra įvairus, kai kuriais atvejais rizikos alelis bendroje Lietuvos lietuvių populiacijoje buvo retesnis, kai kuriais – dažnesnis.

### **3.3. Naujų genetinių sričių, susijusių su ligos genetinė architektūra, paieška**

#### **3.3.1. Šeimų viso genomo žymenų asociacijos tyrimas**

Plataus masto viso genomo genotipavimui ir asociacijų analizei buvo atrinkta 32 triados, t. y. vaikas ir tėvai (96 asmenys), iš bendros Lietuvos populiacijos, kuriose probandas – priešlaikine KŠL sergantis asmuo. Pagal aprašytą metodiką (**2.2.2 skyrelis**) atlikus plataus masto genotipavimą naujos kartos didelio našumo sistema ir įvertinus duomenų kokybę (**IV priedas, 4.1 lentelė**), tolesnei analizei iš 32 asmenų liko 31 vyriškosios lyties pacientas ir 63 tėvai. Atlikus duomenų filtravimą iš genotipuotų 731412 VNP tolesnei analizei palikta 646445 VNP. Dėl standartinio nepusiausviro perdavimo testo algoritmo nesuderinamumo su genotipų duomenimis lyties chromosomose analizuotos tik autosomos.

Gauti genotipavimo duomenys perkelti į duomenų valdymo platformą *BC/Gene (Biocomputing Platforms Ltd., FI)*, kurioje toliau vykdyta duomenų kokybės kontrolė ir plataus masto genomo asociacijų analizė integruotu programiniu paketu *PLINK v1.07 (Purcell S et al., 2007)*.

Nustatyta 12 VNP statistiškai reikšmingai asocijuotų su KŠL fenotipu ( $p < 0,0001$ ). Trumpa VNP anotacija (remiantis *NCBI* duomenų bazės duomenimis, 2013 liepa),  $\chi^2$  statistika, statistinis reikšmingumas  $p$ , šansų

santykis ir empirinė galia (skaičiavimas remiasi imties dydžiu) pateikiami **3.15 lentelėje**. Priimtinos testo galios reikšmės buvo daugiau negu 0,65 ir neretai pakliuvo į pageidaujama intervalą tarp 0,8 ir 0,95 (*Whitley E, Ball J, 2002*). Statistiškai reikšmingai asocijuoto VNP rs1321936 aptikimo galios reikšmė buvo netenkinanti, todėl žymuo buvo pašalintas iš tolesnės analizės. Šansų santykis nurodo poveikio dydį. Kuo šansų santykio reikšmė didesnė už vieneta, tuo reikšmingesnis yra rezultatas. Labiausiai išsiskiriančios šansų santykio reikšmės buvo dviejų VNP atveju: rs17046570 *RTN4* gene ir rs11743737 *FBXL17* gene.

**3.15 lentelė.** Statistiškai reikšmingi VNP, asocijuoti su koronarine širdies liga

Chr	Genas	Funkcija	VNP	AI	$\chi^2$	<i>p</i>	ŠS (95 % PI)	G*
2p16.1	<i>RTN4</i>	<i>Neuroendokrini nė sekrecija; apoptozės procesas</i>	rs17046570	A	15,21	$9,62 \times 10^{-5}$	18 (2,4-134,8)	0,74
3p23	–	–	rs294314	A	15,38	$8,77 \times 10^{-5}$	0,13 (0,04-0,43)	0,65
5q31.1	<i>LOC340073; LOC100996485</i>	neapibūdintas	rs1346440	G	18,24	$1,946 \times 10^{-5}$	0,12 (0,03-0,38)	0,82
5q31.1	<i>LOC340073; LOC100996485</i>	neapibūdintas	rs2019973	G	16,33	$5,312 \times 10^{-5}$	0,13 (0,04-0,42)	0,68
5q21.3	<i>FBXL17</i>	<i>SCF kompleksas; baltymų ubikvitilinimas</i>	rs11743737	A	15,38	$8,77 \times 10^{-5}$	7,67 (2,30-25,53)	0,65
9q22.23	–	–	rs10819695	G	15,70	$7,439 \times 10^{-5}$	0,10 (0,02-0,41)	0,69
12p11.23	<i>ITPR2</i>	Viduląstelinio Ca <sup>2+</sup> atpalaidavimas	rs11048567	A	17,19	$3,38 \times 10^{-5}$	0,05 (0,01-0,37)	0,81
20q11.21	–	–	rs6141273	A	16,20	$5,70 \times 10^{-5}$	0,05 (0,01-0,39)	0,77
20p12	–	–	rs1321936	G	16,03	$6,23 \times 10^{-5}$	0,18 (0,07-0,46)	0,64

3.15 lentelės tęsinys								
Galimi artefaktai								
Chr	Genas	Funkcija	VNP	Al	$\chi^2$	<i>p</i>	ŠS (95 % PI)	G*
1q32.1	<i>PPP1R12B</i>	Miozino fosfatazės aktyvumo reguliavimas; raumenų sutraukiamojo mechanizmo jautrumo Ca <sup>2+</sup> didinimas	rs12734338	G	26,13	3,19×10 <sup>-7</sup>	29,0 (3,95-212,9)	1,00
4q25	–	–	rs3853444	G	20,83	5,02×10 <sup>-5</sup>	0,13 (0,05-0,37)	0,88
15q25.2	<i>UBE2Q2P1</i>	pseudogenas	rs3883013	G	26,13	3,19×10 <sup>-5</sup>	29,0 (3,95-212,9)	1,00

Chr – chromosomos regionas, kuriam priklauso VNP, VNP – asocijuoto vieno nukleotido polimorfizmo identifikacinis numeris, Al – perduotas alelis, ŠS – šansų santykis, PI – pasikliauties intervalas,  $\chi^2$  – Pirsono chi kvadrato kriterijus, *p* – *p* reikšmė  $\chi^2$  statistikai, G\* – kriterijaus empirinė galia, paryškintos eilutės ir pasviras šriftas nurodo potencialius genus kandidatus, pasirinktus tolesnei patvirtinimo analizei.

Grafinis asocijuotų VNP atvaizdavimas pateikiamas klasikiniu *Manhattan* grafiku **3.15 paveiksle**. Grafike matyti keletas reikšmingai asocijuotų, tačiau labai nutolusių nuo kitų VNP žymenų – rs12734338, rs3883013, rs3853444. Įprastai asociacijos atveju matoma koreliacija, tačiau šių VNP atveju greta nėra kitų VNP, todėl šie atsiskyrę VNP laikomi artefaktais ir yra neanalizuojami.

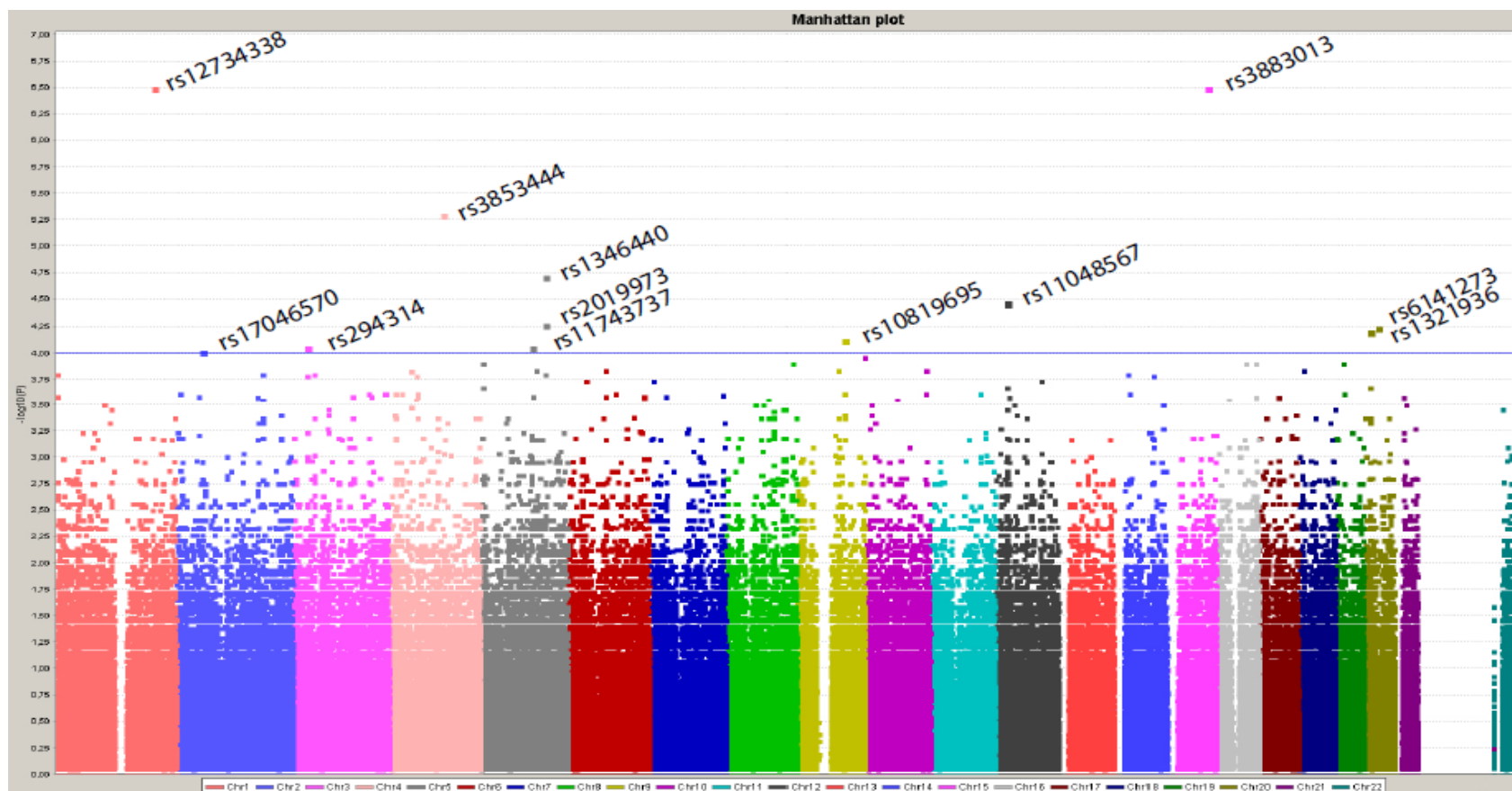
Papildomai statistiškai reikšmingai asocijuotų VNP alelių dažniai pacientų grupėje (29 asmenys) palyginti su tiriamoje bendroje Lietuvos lietuvių populiacijoje (253 asmenys) nustatytais šių VNP alelių dažniais. Rezultatai pateikiami **3.16 lentelėje**. Nustatyti keturi reikšmingi skirtumai (VNP rs1346440, rs2019973, rs11048567, rs6141273), tuo patvirtinant prieš tai nustatytas asociacijas (**3.15 lentelė**) ir atkartojant rezultatus. Kitų keturių VNP atveju asociacijos nepavyko atkartoti: tai dvi sritys (3p23 VNP rs294314 ir 20p12 VNP rs1321936), kuriose nėra genų, bei dvi sritys (2p16.1 VNP rs17046570 ir 5q21.3 VNP rs11743737), kuriose yra genai. Atlikus literatūros (*Chen Y et al., 2010; Jin J et al., 2004; Depre C et al., 2010*) ir duomenų bazių

(NCBI, ENSEMBL) analizę, paaikėjo, kad *RTN4* ir *FBXL17* savo funkcija potencialiai galėtų atlikti vaidmenį aterosklerozės ir kituose su KŠL netiesiogiai susijusiuose procesuose. Todėl iš keturių (3p23, 20p12, 2p16.1 ir 5q21.3) pakartotine asociacijos analize nepatvirtintų sričių tolesnei analizei pasirinktos dvi sritys 2p16.1 ir 5q21.3, kuriose atitinkamai yra genai *RTN4* ir *FBXL17*, galintys turėti funkcinę reikšmę tiriamajam fenotipui. Tai leidžia formuluoti išankstinę genų kandidatų tyrimo hipotezę. Likusių dviejų sričių nepasirinkta analizuoti dėl nepakankamo pagrindo išankstinei hipotezei formuluoti, priešingai nei 2p16.1 ir 5q21.3 sričių atveju. Manant, kad šių sričių genai turi funkcinę pagrindą, asociacija galėjo neatsikartoti, nes galbūt kiti šiuose genuose esantys VNP yra tinkamesni asociacijai nustatyti. Hipotezei patikrinti pasirinkta tolesnė tyrimo strategija.

**3.16 lentelė.** Asociuotų VNP alelių dažnių palyginimas pacientų grupėje ir bendroje Lietuvos lietuvių populiacijoje

Chr	Genas	VNP	Al.	Dažn LTU	Dažn Pac	$\chi^2$	df1, p	ŠS (95 % PI)
2p16.1	<i>RTN4</i>	rs17046570	A	0,39	0,43	0,46	0,50	0,83 (0,46-1,50)
3p23	-	rs294314	A	0,47	0,41	0,77	0,38	1,28 (0,71-2,32)
5q21.3	<i>FBXL17</i>	rs11743737	A	0,45	0,54	1,41	0,24	0,72 (0,40-1,29)
5q31.1	<i>LOC340073;</i> <i>LOC100996485</i>	rs1346440	G	0,54	0,29	13,06	3,01×10 <sup>-4</sup>	2,87 (1,54-5,53)
		rs2019973	G	0,51	0,26	13,57	2,30×10 <sup>-4</sup>	3,02 (1,60-6,02)
9q22.23	-	rs10819695	G	<i>Neįtrauktas į palyginimą, kadangi nėra Illumina Infinium® HD HumanOmniExpress-12 v1.1 VNP luste</i>				
12p11.23	<i>ITPR2</i>	rs11048567	A	0,23	0,05	10,01	1,56×10 <sup>-3</sup>	5,50 (1,73-28,02)
20p12	-	rs1321936	G	0,39	0,29	2,17	0,14	1,56 (0,84-3,01)
20q11.21	-	rs6141273	A	0,12	0,02	5,81	0,02	7,94 (1,32-324,51)

Chr – chromosomos regionas, kuriam priklauso VNP, VNP – asocijuoto vieno nukleotido polimorfizmo identifikacinis numeris, Al. – perduotas alelis, ŠS – šansų santykis, PI – pasikliauties intervalas,  $\chi^2$  – Pirono chi kvadrato kriterijus, p – p reikšmė  $\chi^2$  statistikai, Dažn LTU – bendroje Lietuvos lietuvių populiacijoje nustatytas alelio dažnis, Dažn Pac – pacientų grupėje nustatytas alelio dažnis. Raudonai paryškintos statistiškai reikšmingos reikšmės.



**3.15 paveikslas.** Vieno nukleotido polimorfizmų asociacijos su KŠL fenotipu analizės rezultatai (*Manhattan* grafikas). Y ašyje logaritminė  $p$  reikšmių transformacija. X ašyje – autosomų reprezentacija. Atidėtos VNP pozicijos pagal vietą chromosomoje ir  $-\log_{10}(p)$  vertes. Horizontali linija nurodo pasirinktą reikšmingumo lygmenį ( $\alpha = 0,0001$ ).

### 3.3.2. Potencialių genų kandidatų patvirtinimo tyrimas

Siekiant nustatyti kitus pokyčius *RTN4* ir *FBXL17* genuose ir galimą tų pokyčių asociaciją su KŠL fenotipu bei patvirtinti arba paneigti šių genų kaip potencialių genų kandidatų, susijusių su KŠL, statusą, numatyta tolesnė atvejo ir kontrolės asociacijos analizė. Atlikta *RTN4* ir *FBXL17* genų koduojančios ir gretutinių sekų sekoskaita (*Sanger* metodu) genotipuotiems probandams iš bandomosios asociacijų analizės (31 vyras, 1 moteris), taip pat 61 papildomam KŠL pacientui (55 vyrai ir 6 moterys) ir 48 kontroliniams asmenims (atsitiktiniai, ligos nenurodę 24 vyrai, 24 moterys) iš tiriamos bendros Lietuvos lietuvių populiacijos.

*RTN4* ir *FBXL17* geną sudaro po 9 egzonus, kurie PGR pradmenų kūrimui *RTN4* geno atveju sudalinti į 14 amplikonų, *FBXL17* – 10. Programomis *Primer3 v4.0.0* (Koressaar T et Remm M, 2007; Untergrasser A et al., 2012) ir *Primer-BLAST* (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) sukurti PGR pradmenys ir eksperimentiškai nustatytos tyrimo sąlygos aprašytos **2.2.3 skyrelyje** ir **III priede**. Taikant įvairius eksperimentinius metodus ir skirtingus reagentų derinius, abiejų genų pirmo egzono PGR produkto gavimo sąlygų optimizuoti nepavyko ir specifinis produktas negautas. Todėl pirmas abiejų genų egzonas netirtas.

Nustatyti pokyčiai *RTN4* ir *FBXL17* genuose pateikiami **3.17 lentelėje**. *FBXL17* geno tiriamoji seka nebuvo variabili, nustatyti du pokyčiai: devintame egzone (rs34990078, sinoniminis) ir 3' NTS (rs61749621). *RTN4* geno tirtoje sekoje nustatyti 9 pokyčiai, iš kurių du buvo egzonuose, visi kiti – intronuose. Pagal nustatytus konkrečius VNP (išskyrus rs201547975, esantį *RTN4* introne) kryptingai atlikta papildomos pacientų ir kontrolinės grupių sekoskaita VNP asociacijai su fenotipu įvertinti.

**3.17 lentelė.** Potencialiuose genuose kandidatuose nustatyti pokyčiai bei atvejo ir kontrolės asociacijų analizės rezultatai

Genas	VNP	Vieta gene	MAF_Pacientai	MAF_Kontrolės	$\chi^2$	<i>p</i>	ŠS (95 % PI)	<i>p</i>	RS	<i>p</i>	
<b>RTN4*</b>	rs13002841	2 intronas	0,13	0,22	3,93	0,14	0,47 PI (0,22-1,00)	0,05	0,6 PI (0,36-0,99)	0,05	
	rs11677099	3 egzonas	0,04	0,01	2,25	0,13	4,42 PI (0,54-36,46)	0,17	4,13 PI (0,53-32,06)	0,18	
	rs80121116	3 egzonas	0,04	0	3,80	0,05	8,41 PI (0,47-150,47)	0,15	7,82 PI (0,46-134,08)	0,16	
	rs13401937	3 intronas	0,08	0,03	2,78	0,10	2,89 PI (0,79-10,51)	0,11	2,58 PI (0,79-8,48)	0,12	
	rs2580770	3 intronas	0,23	0,28	1,11	0,58	0,69 PI (0,34-1,39)	0,29	0,81 PI (0,55-1,19)	0,28	
	rs2580769	4 intronas	0,26	0,31	1,82	0,40	0,64 PI (0,32-1,23)	0,21	0,79 PI (0,56-1,13)	0,2	
	rs2255026	6 intronas	0,23	0,29	1,73	0,42	0,63 PI (0,31-1,26)	0,2	0,77 PI (0,53-1,13)	0,19	
	rs201547975	6 intronas	<i>nustatytas vienam asmeniui, netirtas praplėstoj pacientų ir kontrolinių asmenų grupėse</i>								
	rs2255112	7 intronas	0,35	0,36	1,14	0,57	1,13 PI (0,56-2,3)	0,73	1,05 PI (0,79-1,4)	0,74	
<b>FBXL17**</b>	rs34990078	9 egzonas	0,2	0,11	4,38	0,04	2,31 PI (1,04-5,09)	0,04	1,77 PI (1,0-3,15)	0,05	
	rs61749621	3' NTS	0,2	0,11	4,37	0,04	2,31 PI (1,04-5,09)	0,04	1,77 PI (1,0-3,15)	0,05	

VNP – vieno nukleotido polimorfizmo identifikacinis numeris, MAF – retesniojo alelio dažnis, ŠS – šansų santykis, RS – rizikos santykis, PI – pasikliauties intervalas,  $\chi^2$  – Pirsono chi kvadrato kriterijus, *p* – *p* reikšmė atitinkamai  $\chi^2$  statistikai, šansų santykiui ir rizikos santykiui. \* – ištirta 93 asmenys, sergantys KŠL, ir 48 kontroliniai asmenys, \*\* – ištirta 91 asmuo, sergantis KŠL, ir 48 kontroliniai asmenys.

Pagal tiriamus variantus nustatčius genotipus, tiek pacientų, tiek kontrolinėje grupėse atliktas Hardžio ir Vainbergo pusiausvyros įvertinimas. Visi žymenys buvo pusiausviri ( $p > 0,001$ ).

Nustatytiems VNP apskaičiuoti genotipų ir alelių dažniai. Retesniojo alelio dažniai pateikiami **3.17 lentelėje**. Pagal tiriamų VNP genotipų dažnius atliktas KŠL pacientų ir kontrolinių asmenų grupių palyginimas. Nustatyta, kad statistiškai reikšmingai šios grupės skiriasi pagal *FBXL17* geno pokyčius ( $p < 0,05$ ), todėl šie pokyčiai ir genas laikomi susijusiais su KŠL fenotipu. Apskaičiavus šansų ir rizikos santykius (nustatyti ir pasikliauties intervalai) visiems tirtiems VNP, *RTN4* geno variantai reikšmingos įtakos neturi, tačiau rs34990078 ir rs61749621 *FBXL17* gene įverčiai (ŠS 2,31 PI (1,04–5,09),  $p < 0,05$ ) yra reikšmingi ir neatsitiktiniai.



## 4. TYRIMO REZULTATŲ APTARIMAS

Kaip jau minėta įvade, šio tyrimo tikslas yra įvertinti bendrą vieno nukleotido polimorfizmų rizikos alelių profilį ir atlikti daugiaveiksnių ligos genetinės architektūros analizę bendroje Lietuvos populiacijoje. Šiam tikslui pasiekti parinkta strategija, darbo modelis (aprašytas **2.1 skyrelyje**), kuris gali būti taikomas daugumai kitų daugiaveiksnių ligų tirti.

Vykdamas tyrimą pagal modelį, pasirinkta daugiaveiksni liga – KŠL, atrinktos žinomos genetinės sritys – 60 VNP, lemiančių riziką susirgti KŠL. Pagal šiuos genetinius žymenis ištirta bendros Lietuvos lietuvių populiacijos 253 asmenys. Siekiant nustatyti naujas genetines sritis, siejamas su KŠL, atlikta plataus masto genomo asociacijų analizė 31 šeimoje, kurioje probandas yra sergantis KŠL. Vėliau atliktas asocijuotų sričių patvirtinimas taikant genų kandidatų asociacijų analizę didesnėje atvejo ir kontrolės imtyje.

Tokio pobūdžio tyrimų mokslinėje literatūroje yra sudėtinga aptikti, juo labiau Lietuvoje. Tai bene pirmasis tokio pobūdžio tyrimas Lietuvoje ir kartu unikalus pasaulyje.

### **4.1. Genetinė įvairovė, alelių dažnių, haplotipų skirtumai populiacijos viduje ir tarp populiacijų**

Genotipavus 253 asmenis iš bendros Lietuvos lietuvių populiacijos plataus masto genominiams VNP lustams, atlikus duomenų kokybės vertinimo procedūras, tolesniam tyrimui atrinkta 60 VNP, lemiančių padidėjusią riziką susirgti KŠL. Pagal šiuos 60 VNP tiriamiems asmenims nustatyti genotipų, alelių ir haplotipų dažniai, sukonstruoti haplotipų blokai, atliktas alelių dažnių ir haplotipų blokų palyginimas populiacijos viduje (palygintos vyrų ir moterų, aukštaičių ir žemaičių grupės), alelių dažniai palyginti su kitomis populiacijomis (bendra Lietuvos lietuvių populiacija palyginta su CEU, FIN, GBR ir TSI populiacijomis), nustatytas rizikos alelių, tenkančių vienam asmeniui, pasiskirstymas populiacijoje ir tarp vidupopuliacinių grupių.

Aptariant šio tyrimo etapo metu gautus rezultatus, reikėtų turėti mintyje tai, kad bendra Lietuvos lietuvių populiacija yra reprezentuojama kaip etninė lietuvių populiacija. Yra daug plataus masto genomo įvairovės tyrimų, kurie nagrinėja bendras baltaodžių populiacijas arba tiria jas ne pagal konkrečius ligai genomo žymenis. Tačiau pastaruoju metu daugėja atskirų populiacijų tyrimų. Šių tyrimų svarba ir poreikis yra paminėti įvade ir literatūros apžvalgoje. Toks yra ir šis tyrimas. Genetinė įvairovė yra tiriama ne tik atskiroje lietuvių populiacijoje, bet ir pagal konkrečius ligą lemiančius VNP rizikos alelius, ir tai lemia geresnį daugiaveiksnės ligos genetinės architektūros suvokimą. Tik atlikus šį tyrimą bendroje populiacijoje, galima imtis diagnostikos ir prognozės vertinimo tyrimų.

### **Alelių dažniai tirtoje bendroje Lietuvos lietuvių populiacijoje**

Šio tyrimo rezultatai rodo, kad rečiausiai populiacijoje pasitaikantis rizikos alelis, kurio dažnis 0,004, yra VNP rs11591147 (A), esantis 1-oje chromosomoje *PCSK9* geno koduojančioje sekoje (**3.1 paveikslas**). Palyginus su 1000 genomų projekto pateikiama bendros žmonių populiacijos statistika *NCBI dbSNP* duomenų bazėje (prieiga internete <http://www.ncbi.nlm.nih.gov/projects/SNP>), šis alelis taip pat yra retas (MAF = 0,009). Dažniausiai pasitaikantis rizikos alelis, kurio dažnis 0,903, yra VNP rs445925 (G), esantis 19-oje chromosomoje ir siejamas su *APOE* geno sritimi. Palyginus su 1000 genomų projekto pateikiama bendros žmonių populiacijos statistika *NCBI dbSNP* duomenų bazėje, jis taip pat yra neretas (dažnis 0,869). Vidupopuliacinėse grupėse, vyrų ir moterų (**3.2 paveikslas**), aukštaičių ir žemaičių (**3.3 paveikslas**) rečiausias ir dažniausias rizikos alelis buvo tas pats, išskyrus žemaičių grupę, kurioje dažniausias alelis buvo kitas – jo dažnis 0,881. Tai VNP rs501120 (A), esantis 10-oje chromosomoje ir siejamas su *CXCL12* genu. Pagal 60 VNP rizikos alelių, lemiančių polinkį KŠL, dažnių palyginimą vyrų ir moterų grupėje statistiškai reikšmingų skirtumų nerasta (**3.4 lentelė**), kas leidžia teigti, jog pagal tiriamus žymenis lytis įtakos nustatytai genetinei įvairovei neturi. Pagal tuos pačius žymenis palyginus aukštaičius ir

žemaičius, nustatyti du statistiškai reikšmingi skirtumai (**3.5 lentelė**). Tarp šių rizikos alelių – rs11206510 (A), esančio 1-oje chromosomoje ir siejamo su jau minėtu genu *PCSK9* ( $p = 0,041$ ), ir rs15563 (G), esančio 17-oje chromosomoje *UBE2Z* geno 3' NTS ( $p = 0,024$ ), – nustatyti dažnių skirtumai leidžia manyti, kad priklausomai nuo etnolingvistinės grupės genetinė įvairovė pagal polinkį sirgti KŠL lemiančius VNP alelius gali skirtis. Šį teiginį dar labiau sustiprina haplotipų analizė.

### **Haplotipai tirtoje bendroje Lietuvos lietuvių populiacijoje**

Kadangi yra tiriama 60 genetinių žymenų, daugiaveiksnių ligų analizei yra reikalinga **haplotipų analizė**, apimanti daugelio žymenų informaciją, kitais žodžiais, susisteminanti pavienių išsibarsčiusių po genomą žymenų informaciją. Haplotipų blokai yra genomo dalys, pasižyminčios minimalia įvairove. Haplotipų blokų sudarymas – tai atskiro genetinio profilio nustatymas, pagal kurį galima grupuoti asmenis ir tirti koreliaciją su vienokiu ar kitokiu fenotipu (*Zhang K et al., 2003*). Tiriamus VNP suskirsčius į grupes pagal chromosomas (**3.6 lentelė**), šiame tyrime buvo sukonstruoti haplotipų blokai. Iš 14-os chromosomų (1–3, 6, 7, 9, 10–13, 15–17, 19) tiriamų VNP genotipavimo duomenų sukonstruoti haplotipų blokai 1–3, 9 ir 19 chromosomose bendroje Lietuvos lietuvių populiacijoje, vyrų ir moterų bei aukštaičių ir žemaičių grupėse. Daugiausia žymenų (4 VNP) apimantis haplotipų blokas nustatytas devintoje chromosomoje (**3.7 paveikslas**), kiti haplotipų blokai apėmė po du VNP (**3.4, 3.5, 3.6, 3.8 paveikslai**). Devintos chromosomos haplotipų blokas nesiskiria nuo kitose populiacijose nustatomų (*Correia M et al., 2013*), tuo patvirtinama hipotezė apie bendrus žmonių populiacijoms ligą lemiančius genetinius veiksnius (*Visscher PM et al., 2012 (1)*), tačiau nepamirštant skirtingos tų genetinių veiksnių kombinavimosi tikimybės priklausomai nuo populiacijos. Vertinant haplotipų įvairovę sukonstruotuose haplotipų blokuose, visuose yra po du haplotipus, išskyrus devintos, devynioliktos chromosomų ir antros chromosomos pirmąjį haplotipų bloką, kuriuose nustatyta po tris haplotipus. Tik vieno haplotipų bloko atveju

(VNP rs599839 ir rs646776) 1-oje chromosomoje (**3.4 paveikslas**) haplotipas, turintis abu rizikos alelius (AA), buvo dažnesnis (dažnis grupėse 0,72–0,78) negu alternatyvių alelių haplotipas visose tiriamose grupėse. Šis haplotipas panašiu dažniu (0,72) taip pat nustatytas Azijos indų tiriamųjų kohortoje ir atlikus raiškos tyrimus susietas su KŠL bei plazmos cholesterolio kiekio reguliavimu (*Arvind P et al., 2014*). Daugiausia haplotipų blokų, t. y. du (rs1042031 ir rs693; rs1260326 ir rs780094), nustatyta antroje chromosomoje (**3.5 paveikslas**), visose likusiose chromosomose – po vieną haplotipų bloką. Antros chromosomos antras haplotipų blokas (rs1260326 ir rs780094) yra nustatytas ir danų bei prancūzų populiacijose, kurių tyrimais taip pat patvirtinta retesnių, šiuo atveju – rizikos alelių (AA) haplotipų asociacija su KŠL rizikos veiksniais (*Tam CH et al., 2009*). Tarp dviejų haplotipų blokų yra galima nepusiausvira sankiba, taip pat galimi ir mainai (rekombinacija). Toks reiškinys vertintinas antros chromosomos abiejų nustatytų haplotipų blokų atveju. Remiantis rezultatais, visose tiriamose grupėse tarp haplotipų yra vienoda rekombinacijos tendencija, išskyrus moterų grupę, kurioje yra tik viena rečiau rekombinuojanti haplotipų pora. Pagal  $D'$  statistiką vertinant haplotipų blokų sankibos pusiausvirumą, didesnis pusiausvirumas yra moterų ir žemaičių grupėse ( $D' = 0,13$ ), mažesnis – vyrų ( $D' = 0,28$ ) ir aukštaičių ( $D' = 0,30$ ), t. y. vyrų ir aukštaičių grupių haplotipų blokams yra didesnė nepusiausviro sankibos tikimybė. Toks rezultatas gali rodyti ankstesnę žemaičių gyvenimą šiose teritorijose, palyginti su aukštaičiais, nes yra žinoma, kad kuo ilgiau populiacija gyvuoja, tuo daugiau yra laiko įvykti rekombinacijai ir tuo labiau didėja genetinė įvairovė ir mažinama nepusiausvira sankiba (*The International HapMap Consortium, 2005*). Moterų ir vyrų atveju nustatytasis skirtumas gali atspindėti biologinę dažnesnės rekombinacijos moterų chromosomų mainuose savybę (*Coop G, Przeworski M, 2007*). Devynioliktoje chromosomoje iš tiriamų VNP yra sudarytas vienas haplotipų blokas, tik moterų grupėje, kitose grupėse toks blokas nenustatytas (**3.8 paveikslas**). Haplotipų bloką sudaro du VNP, rs2075650 (genas *TOMM40*) ir rs445925 (genas *APOC1*) ( $D' = 1$ ), esantys vadinamojoje *APOE* srityje. Jau minėta, kad

šiam blokui nustatyti trys haplotipai: AG, GG ir AA, kur GG yra abiejų rizikos alelių haplotipas. Nepusiausvira sankiba rodo neatsitiktinį ryšį tarp šių dviejų VNP moterų grupėje. Šių dviejų VNP sankibą nurodo ir kitų mokslininkų tyrimai CEU imtyje, kur apskaičiuota D' irgi yra lygi vienetai (arba 100 %) (*Bekris LM et al., 2012*). Vadinasi, yra su lytimi susijusi priežastis, kodėl šie du VNP išliko neatskirti rekombinacijos tiriamoje bendroje Lietuvos lietuvių populiacijoje. Yra nustatyta, kad rs2075650, esantis *TOMM40* gene, kartu ir šis genas, yra susiję su Alzheimerio ligos kiekybiniais požymiais. Nustatyta nepusiausvyra sankiba tarp geno *TOMM40* VNP ir VNP esančių *APOE* srityje (sykiu yra ir mūsų analizuojamo haplotipo bloko VNP rs445925) bei atlikti funkciniai tyrimai, autorių teigimu, leidžia manyti, kad *TOMM40* ir *APOE* raiška yra veikiama regioninių reguliacinių haplotipų, kurie daro biologinį poveikį Alzheimerio ligos patogenezėi (*Bekris LM et al., 2012*). Pastebėtas ryšys tarp Alzheimerio ligos ir lyties. Pelių tyrimais parodyta, kad estrogenai gali reguliuoti apoE raišką priklausomai nuo audinio (*Srivastava RA et al., 1996*). Neseniai įrodyta, kad lytis modifikuoja su *APOE* siejamą polinkį Alzheimerio ligai, todėl moterys serga dažniau (*Altmann A et al., 2014*). Tokiu atveju pagal nepusiausviro sankibos rezultatus reikšmingas gali būti ne tik *APOE* genas ir jo variantai, bet ir *TOMM40*. Aptariant likusį haplotipų bloką trečioje chromosomoje (**3.6 paveikslas**), kurį sudaro VNP rs9818870 ir rs2306374, reikėtų paminėti tai, ką nustatė mokslininkai plataus masto genomo asociacijų tyrimais. Pirmasis variantas, susietas su KŠL, buvo rs9818870, kurio rizikos alelis (A) *cis* padėtyje slopina *MRAS* raišką kraujagyslių audiniuose (*Folkersen L et al., 2010*), tačiau *CARDIoGRAM* tyrimo metu stipriausią asociaciją su KŠL rodė kitas tos pačios srities VNP – rs2306374, kurio rizikos alelis (G) didina *FAIM* (angl. *Fas apoptotic inhibitory molecule*, apoptozę stabdanti molekulė Fas) raišką. Tokie rezultatai iškėlė klausimą, kuris variantas, o gal abu, tarpininkauja nulemiant ligos riziką (*Maouche S, Schunkert H, 2012*). Mūsų tyrimo duomenimis, nustatyta nepusiausvira sankiba tarp minėtų abiejų VNP rodo, kad abu variantai gali būti susiję su KŠL rizika.

Apibendrinant yra manytina, kad lietuvių tautą suformavo dvi gentys – aukštaičiai ir žemaičiai, tačiau pagal antropologinius duomenis skirtumai tarp šių grupių išnyko viduramžiais, ir nuo to laiko Lietuvos populiacija ir Rytų Europos bei visos Europos genetinės įvairovės požiūriu yra labai homogeniška (*Česnys G, 1991*). Tačiau, genetinių tyrimų duomenimis, lietuvių genų fonde vis dar jaučiama išnykusių baltų genčių – prūsų, jotvingių ir kuršių – įtaka. <...> Lyginant žemaičius ir aukštaičius akivaizdu, kad šioms grupėms būdingas genetinis heterogeniškumas. Tai paaiškinama tuo, kad žemaičiai yra autochtonai, o aukštaičiai – vėliau atsikraustę rytiniai baltai (*Kučinskas V, 2004*). Mūsų tyrimo rezultatai pagal genetinius KŠL žymenis taip pat įrodo žemaičių kaip Lietuvos senbuvių istoriją.

### **Bendroje Lietuvos lietuvių populiacijoje tirtų alelių dažnių palyginimas su kitomis populiacijomis**

Dvidešimt penki (t. y. beveik 42 %) iš 60 tiriamų VNP rizikos alelių tiriamoje bendroje Lietuvos lietuvių populiacijoje pasitaikė didesniu negu 0,5 dažniu. Palyginus su 1000 genomų projekto pateikiamais tokių alelių dažniais bendroje žmonių populiacijoje, dažnio tendencija pasitvirtino, tokių alelių dažniai buvo > 0,5, išskyrus dviejų VNP atveju (rs693 (A) ir rs4845625 (A)). Taip pat pastebėtas dėsniumas, kad tie rizikos aleliai, kurie yra nustatomi rečiausiai arba dažniausiai bendroje Lietuvos lietuvių populiacijoje, t. y. esantys brėžiamos įstrižainės galuose (**3.13 paveikslas**), pasižymi dažnio tolygumu tarp lyginamų populiacijų CEU, FIN, GBR, TSI ir LTU, o rizikos aleliai, esantys tarp įstrižainės galų, pagal dažnį yra linkę labiau barstyti. Tai gali būti bendra savybė, paaiškinama tuo, kad reti ir dažni riziką lemiantys aleliai yra būdingi žmonių populiacijai bendrai, o genetinė įvairovė, esanti viduryje tarp retų ir dažnų rizikos alelių – unikalus derinys, priklausantis nuo kiekvienos individualios populiacijos.

Atlikus bendros Lietuvos lietuvių populiacijos 60 VNP alelių dažnių palyginimą su keturiomis Europos protėvių populiacijomis (CEU, FIN,

GBR ir TSI; **3.10–3.13 lentelės**), daugiausiai skirtumų nustatyta tarp LTU ir TSI (20 skirtumų), mažiausiai – tarp LTU ir GBR (6 skirtumai) populiacijų. Tarp LTU ir FIN nustatyta 17, o tarp LTU ir CEU – 8 skirtumai. Tad pagal 60 VNP, siejamų su KŠL, genetinę struktūrą ir įvairovę galima išvelgti tendenciją, kad daugiausia skirtumų yra šiaurės–pietų (vertikalios) geografinės ašies atžvilgiu, mažiausia – vakarų (horizontalios). Tai būtų galima paaiškinti geografinių platumų atitikimu. Be to, remiantis *L. L. Cavalli-Sforza* ir bendraautorių pagrindinių genetinių komponentių analize, tiriant Europos populiacijų genetinę istoriją pagal antrąją komponentę (kuri paaiškina 22 % visos Europos genetinės įvairovės), susidaro gradientas iš šiaurės į pietus. Šios genetinės komponentės ir geografinės platumos sąsaja yra akivaizdi, todėl geriausiai gali būti paaiškinta atskiroms geografinėms platumoms būdingomis klimatinėmis ir ekologinėmis sąlygomis (*Cavalli-Sforza LL et al., 1994*). Buvo tiriama europiečių genetinė struktūra, kartu ir lietuvių (*Nelis M et al., 2009*). Nustatyta, kad Europos populiacijų genetinė struktūra koreliuoja su šalių geografija, genetinė įvairovė išsidėsto gradientiškai atitinkamai pradėdant Europos šiaurės vakarais, o baigiant pietryčiais. Tokiu būdu susidaro trikampis genetinis žemėlapis (**1.3 paveikslas**), kurio centre yra Centrinė ir Vakarų Europa, o viršūnėse (1) Suomija, (2) Pabaltijys, Lenkija ir Vakarų Rusija ir (3) Italija (*Nelis M et al., 2009*). Šis suskirstymas taip pat atitinka gautus mūsų tyrimo duomenis. Tačiau remiantis mūsų tyrimo rezultatais, galima teigti, kad bendra Lietuvos lietuvių populiacija nuo kitų populiacijų skyrėsi skirtingais variantais, vienu atveju rizikos aleliai buvo dažnesni LTU populiacijoje, kitu atveju – retesni, palyginti su kitomis populiacijomis; buvo tokių VNP, kuriais LTU skyrėsi nuo kelių ar net visų populiacijų, su kuriomis palyginta (**3.14 lentelė**). Vadinasi, bendros Lietuvos lietuvių populiacijos genetinė įvairovė yra ir panaši, ir kartu skirtinga, palyginti su kitomis Europos protėvių populiacijomis. Ir jeigu genetinė įvairovė atitinka geografinę populiacijos padėtį, vadinasi, yra konkretus genetinės įvairovės derinys, dar tiksliau aprašantis būtent tos vietovės, kurioje populiacija gyvena, koordinatas. Greičiausiai yra taip, kad daugiaveiksnes ligas lemiantys rizikos aleliai yra tie

patys ir bendri visoms žmonių populiacijoms. Tačiau tarp populiacijų gali skirtis alelių poveikio dydis ir alelių dažnių deriniai (*Visscher PM et al., 2012(1)*). Tai ir įrodo mūsų tyrimas. Alelinis heterogeniškumas pabrėžia skirtingų populiacijų įtraukimo į tyrimus svarbą, ypač tiriant genetines daugiaveiksnių požymių, tokių kaip lipidų koncentracija, asociacijas (*Coram MA et al., 2013*). Neseniai pradėti genetinės įvairovės tyrimai globaliu, žemyniniu, regioniniu, etninių grupių bei atskirų šalių mastu, tačiau retai kur šalių imtys tinkamai reprezentuojamos (*Esko T et al., 2013*).

#### **4.2. Genetinės struktūros ir įvairovės pagal tirtus VNP reikšmė bendroje Lietuvos lietuvių populiacijoje**

Kalbant apie genetinę struktūrą ir įvairovę, didžioji dalis rezultatų jau aptarta ankstesniame skyrelyje, kadangi genetinę struktūrą ir įvairovę, kaip žinoma, sudaro ir genotipų, ir alelių, ir haplotipų dažniai, jų įvairovė populiacijos viduje ir kitų populiacijų kontekste. Šis skyrelis skirtas aptarti asmeniui tenkančio rizikos alelių skaičiaus statistiką, jos reikšmę ir rezultatų panaudojimo galimybes.

Pagal 60 tiriamų VNP, lemiančių KŠL riziką, didžiausias galimas rizikos alelių skaičius asmeniui gali būti 120. Tačiau tokių asmenų šio tyrimo metu bendroje Lietuvos lietuvių populiacijoje neaptikta. Kraštinės nustatytos reikšmės buvo 39 ir 74 aleliai. Apskaičiuotas aritmetinis vidurkis parodė, kad LTU populiacijoje asmuo turi vidutiniškai ~56 rizikos alelius, t. y. beveik pusę visų galimų, pagal kuriuos tiriama. Apskaičiuota trečiojo kvartilio reikšmė rodo, kad ketvirtadalis (25 %) LTU populiacijos turi didesnę rizikos alelių skaičių negu kiti trys ketvirtadaliai tiriamųjų. Suskirsčius tiriamuosius į grupes pagal alelių skaičių: 39–50 (1 grupė); 51–62 (2 grupė); 63–74 (3 grupė), daugiausia asmenų priklauso antrai grupei (68 %), mažiausiai – trečiai grupei (13 %). Tuo remiantis bendros Lietuvos lietuvių populiacijos asmenis būtų galima suskirstyti taip: santykinai mažą rizikos alelių skaičių, santykinai vidutinį rizikos alelių skaičių ir santykinai didelį rizikos alelių skaičių turinčių asmenų. Daugelio autorių duomenimis, daugiaveiksnių ligų, iš jų ir KŠL,



atveju dažni VNP rizikos aleliai turi mažą poveikį fenotipui, t. y. šansų santykis priklauso intervalui nuo 1 iki 2, tačiau poveikis yra suminis (*Manolio TA et al., 2009*). Remiantis mūsų tyrimo rezultatais, jeigu prielaida apie kiekvieno dažno VNP rizikos alelio suminį poveikį fenotipui yra teisinga, išvada tokia, kad pagal tiriamų VNP rizikos alelių skaičių, tenkantį asmeniui, daugiausia asmenų turi vidutinę riziką susirgti KŠL (**3.9 paveikslas**). Šie duomenys sutampa ir su kitų mokslininkų atliktų prospektyviųjų tyrimų rezultatais: pagal asmeniui tenkantį alelių skaičių daugiausia tiriamųjų priklauso vidutinės rizikos susirgti KŠL grupei (*Humphries SE et al., 2010*).

Atlikus asmeniui tenkančio alelių skaičiaus vidurkių ir dispersijų palyginimą tarp tiriamos populiacijos grupių – vyrų ir moterų bei aukštaičių ir žemaičių, statistiškai reikšmingų skirtumų nerasta. Vadinasi, galima teigti, kad pagal tiriamų KŠL rizikos alelių skaičių, tenkantį asmeniui, populiacija yra homogeniška.

Atsižvelgus į nustatytus 60 KŠL rizikos alelių dažnius populiacijoje ir asmeniui tenkančio rizikos alelių skaičiaus statistiką, papildomai įtraukiant į tyrimą pacientus, būtų galima sudaryti rizikos vertinimo (prognozės) algoritmą pagal genetinius veiksnius. Žinant esamų KŠL rizikos prognozės algoritmų neuniversalumą ir ribotumą, palyginus genetinį rizikos vertinimo algoritmo rezultatą su klasikinių veiksnių prognozės algoritmų rezultatais, būtų galima sukurti bendrą algoritmą, padedantį tiksliau pateikti KŠL rizikos prognozę Lietuvos populiacijai.

### **4.3. Asociacijų rezultatai ir jų reikšmė**

Daugiaveiksnės ligos genetinė architektūra neturi galutinės apibrėžties. Ne viskas ištirta bendrai žmonių populiacijai, juo labiau ne viskas ištirta ir nustatyta specifinėse unikaliose populiacijose. Galbūt egzistuoja panašus, tačiau visgi skirtingas genetinis profilis, lemiantis ligą. Tad siekiant papildyti žinias apie ligos genetinę architektūrą, reikia nustatyti ne tik jau žinomų riziką lemiančių VNP genetinį profilį populiacijoje, bet ir ieškoti naujų, galbūt populiacijai specifinių genetinių sričių, sietinų su liga.

Įgyvendinant šį tikslą buvo atlikta bandomoji plataus masto genomo žymenų asociacijos su KŠL fenotipu analizė šeimų grupėje. Nustatytos devynios genetinės sritys, sietinos su KŠL (**3.15 lentelė**). Kaip minėta literatūros apžvalgoje, plataus masto genomo asociacijų tyrimų rezultatus būtina patvirtinti, todėl atliktas asociacijos patvirtinimo tyrimas (**3.16 lentelė**). Iš devynių VNP keturiems (rs1346440, rs2019973, rs11048567, rs6141273) buvo patvirtinta reikšminga asociacija, vienas VNP neįtrauktas į analizę dėl techninių galimybių, keturiems VNP (rs294314, rs1321936, rs17046570, rs11743737) reikšminga asociacija nepatvirtinta. Tačiau iš keturių sričių, kurių atveju nepavyko atkartoti asociacijos, dvejose srityse, 2p16.1 (VNP rs17046570) ir 5q21.3 (VNP rs11743737), yra genai, atitinkamai *RTN4* ir *FBXL17*, kurie savo funkcija galėtų atlikti vaidmenį aterosklerozės ir kituose su KŠL netiesiogiai susijusiuose procesuose ir tai sudaro pagrindą formuluoti išankstinę hipotezę, priešingai nei rs294314 ir rs1321936 atveju. Galbūt kiti šiuose genuose nustatyti pokyčiai galėtų būti asocijuoti su KŠL fenotipu. Hipotezei patikrinti atlikta *RTN4* ir *FBXL17* genų sekoskaita ir nustatytų pokyčių asociacijos su KŠL fenotipu analizė.

Iš trijų sričių, kurioms patvirtinta asociacija, tik vienoje srityje (20q11.21) nėra genų. Kitose srityse yra genai (12p11.23 – *ITPR2*; 5q31.1 – *LOC340073*; *LOC100996485*), apie kurių VNP asociaciją su KŠL informacijos duomenų bazėse nėra. *LOC340073* ir *LOC100996485* yra nekoduojančios RNR, kurios gali turėti įtakos su KŠL susijusiuose biologinių kelių reguliacijos procesuose, t. y. veikti ne tiesiogiai per baltymus, o reguliuoti genų raišką. Tačiau konkretniems teiginiams apie šių RNR genų dalyvavimą KŠL patogenezėje duomenų nepakanka. Genas *ITPR2* (angl. *inositol 1,4,5-trisphosphate receptor, type 2*) koduoja inozitolio 1,4,5-trifosfato antro tipo receptorių. Ląstelėse *ITPR2* funkcionuoja kaip  $Ca^{2+}$  permembraninis nešiklis, nuo inozitolio 1,4,5-trifosfato priklausantis kalcį atpalaiduojantis kanalas ir sąveikauja su fosfatidilo inozoliu (Genų ontologijos duomenų bazė (angl. *The Gene Ontology Database*); prieiga internete [www.geneontology.org](http://www.geneontology.org)).

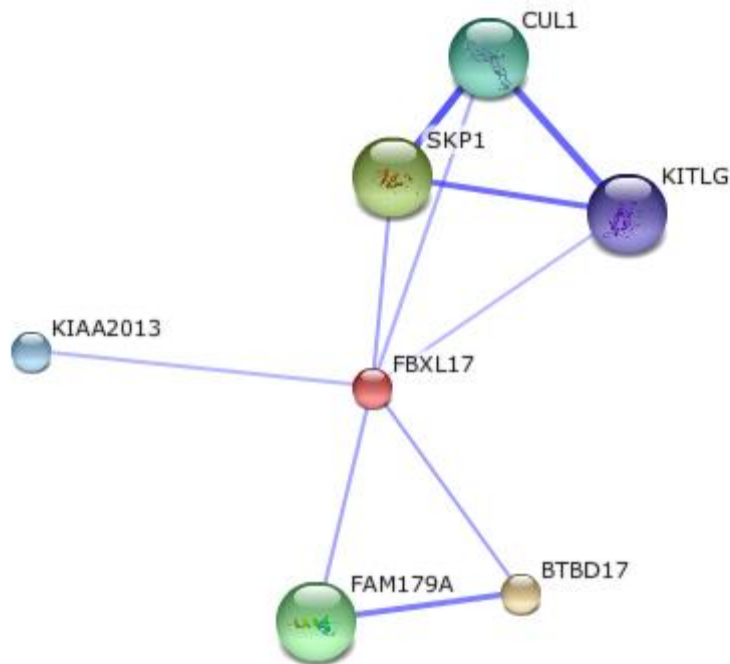
Mokslininkai yra nustatę asociaciją tarp *ITPR2* geno VNP ir uždegiminių procesų bei endotelio funkcijos. Manoma, kad *ITPR2* genas gali turėti įtakos širdies ir kraujagyslių ligų vystymuisi (*Wilker EH et al., 2009*). Taip pat yra parodyta, kad *ITPR2* dalyvauja atsake į oksidacinį stresą ir ląstelės ciklo reguliavime (*Warnatz HJ et al., 2011*), šie procesai taip pat yra susiję su ateroskleroze ir KŠL.

Antroje chromosomoje (2p16.1) esantis *RTN 4* (angl. *reticulon 4*) genas koduoja retikuloną 4. Retikulonai yra siejami su endoplazminio tinklo darniu funkcionavimu, dalyvauja neuroendokrininėje sekrecijoje, su neuroendokrininių ląstelių membrana susijusioje pernašoje, apoptozės procesuose. *RTN 4* baltymas yra potencialus neuronų augimo inhibitorius, taip pat galintis slopinti centrinės nervų sistemos regeneraciją aukštesniuose stuburiniuose (*NCBI duomenų bazė, Chen Y et al., 2010*). *RTN4* taip pat laikomas kandidatiniu kraujagyslių ląstelių apoptozės ir aterosklerozės procesų tarpininku (*Chen Y et al., 2010*).

Penktoje chromosomoje (5q21.3) esantis *FBXL17* (angl. *F-box and leucine-rich repeat protein 17*) genas koduoja baltymą *FBXL17*, kuris turi F sritį, būdingą F sritį turinčiai baltymų šeimai. F sritis – tai apie 40-ies aminorūgščių motyvas. F sritį turintys baltymai kartu su baltymais *SKP1* (angl. *S-phase kinase-associated protein 1*, su ląstelės S-fazės kinaze susijęs baltymas 1) bei *CUL1* (angl. *culin1*, kulinas 1) sudaro kompleksą SCF, kuris yra baltymų ubikvitino ligazė (*Jin J et al., 2004*). Šis kompleksas yra vienas svarbiausių ubikvitino-proteosomų sistemoje (UPS), kuri dalyvauja 70–90 % ląstelės baltymų degradacijos procesų. Tarp degraduojamų baltymų yra širdies ir kraujagyslių sistemai svarbių baltymų. UPS taip pat kontroliuoja endotelinių ląstelių ciklą. Oksidacinio streso poveikis SCF kompleksui sutrikdo UPS ir kartu jos reguliuojamą endotelio funkciją, kuri svarbi kraujagyslių būklei (*Depre C et al., 2010*).

Atlikus *RTN4* ir *FBXL17*, potencialių genų kandidatų, asociacijos analizę atvejo ir kontrolės grupėje pagal šiuose genuose nustatytus pokyčius, statistiškai reikšminga asociacija su KŠL pasitvirtino tik *FBXL17* geno tirtų VNP atveju (rs34990078 ir rs61749621).

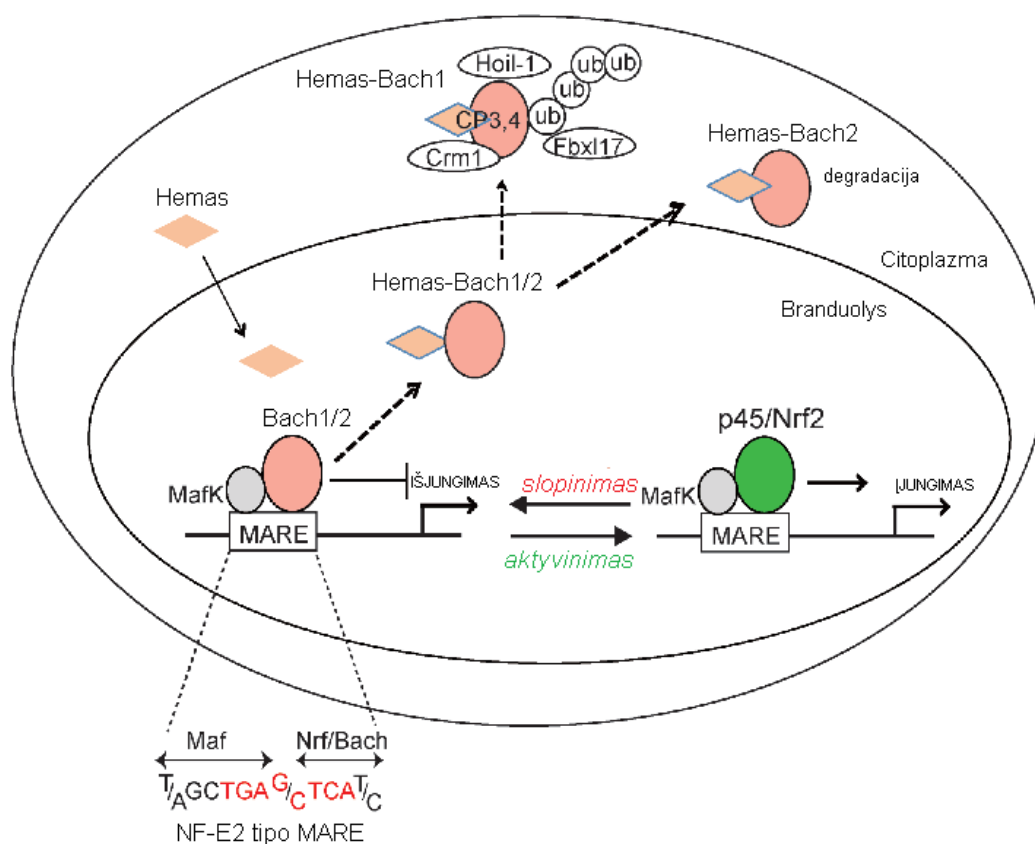
Nagrinėjant, kaip *FBXL17* genas, o tiksliau *FBXL17* baltymas, gali būti susijęs su KŠL ar aterosklerozės procesais, atlikus bioinformacinę analizę programiniu paketu *STRING v9.1* (Franceschini A et al., 2013) nustatyti funkciniai *FBXL17* partneriai (**4.1 paveikslas**). Genų ontologijos duomenų bazės duomenimis, *FBXL17* funkcinis molekulinės sąveikos pobūdis yra jungimasis prie kitų baltymų. Kaip jau minėta, *FBXL17* dalyvauja SCF komplekse, kurio funkciniai partneriai yra *CUL1* ir *SKP1*. Šis kompleksas svarbus ubikvitilinant kitus baltymus. Kiti funkciniai *FBXL17* partneriai: *BTBD17* (angl. *BTB (POZ) domain containing 17*, BTB domeną turintis baltymas 17), *FAM179A* (angl. *family with sequence similarity 179, member A*, 179 baltymų šeimos su sekos panašumu baltymas narys A), *KIAA2013* (angl. *KIAA2013*), *KITLG* (angl. *KIT ligand*, KIT ligandas). Iš baltymų pavadinimų matyti, kad yra mažai žinių apie juos ir jų funkcijas, išskyrus *KITLG*, kuris yra žinomas kaip putliųjų ląstelių proliferacijos aktyviklis, kuris taip pat gali sustiprinti mieloidinių ir limfoidinių ląstelių pirmtakių proliferaciją kaulų čiulpų ląstelių kultūroje. Taip pat šis baltymas dalyvauja adhezinėje ląstelės–ląstelės sąveikoje ir sinergiškai veikia su kitais citokiniais, manoma, kad su interleukiniais.



**4.1 paveikslas.** Funkciniai *FBXL17* geno koduojamo baltymo partneriai. Skrituliai vaizduoja baltymus, jungiančios linijos rodo asociacijos tarp baltymų stiprumą (kuo storesnė linija, tuo asociacija stipresnė)

Duomenų apie *FBXL17* nėra daug, genas kartografuotas ir jo koduojamas baltymas apibūdintas 2004 metais (*Jin J et al., 2004*). Naujausiais duomenimis yra manoma, kad *FBXL17* per *BACH1* (angl. *BTB and CNC homology 1, basic leucine zipper transcription factor 1*, „leucino užtrauktuko“ sekų motyvo transkripcijos veiksnys 1), transkripcijos slopiklį, kontroliuoja nuo *NRF2* (angl. *nuclear factor, erythroid 2-like 2*, eritroidinis antro tipo branduolio veiksnys 2) priklausomą genų aktyvinimą (*Tan MK et al., 2013*). *NRF2* yra transkripcijos veiksnys, reguliuojantis genus, kurie turi oksidacinio atsako elementus promotoriuose. Dauguma šių genų dalyvauja biologiniame atsako kelyje, kuris reaguoja į žaizdas ar uždegimus ir su jais siejamus laisvuosius radikalus (*NCBI duomenų bazės duomenimis, 2014 balandis*). *BACH1* taip pat reguliuoja ir jau aukščiau minėto *ITPR2* geno raišką, veikdamas atsaką į oksidacinį stresą, ląstelės ciklo bei apoptozės procesus (*Warnatz HJ et al., 2011*). Pelių tyrimais yra parodyta, kad *Bach 1* dalyvauja širdies ir kraujagyslių

sistemos ligų procesuose. Pelėms, kurios buvo šeriamos riebalų gausiu daviniu ir kurių *ApoE* genas buvo išveiklintas, nesant Bach1 ateriosklerozė nesivystė. Nėgana to, gydant peles alavo protoporfirinu (hemo oksigenazės 1 inhibitorius), kurioms trūksta tiek ApoE, tiek Bach1, išnyko ateriosklerozinės plokštelės ir kiti su liga susiję požymiai (Watarai Y et al., 2008). Šie rezultatai rodo FBXL17-BACH1-NRF2 (4.2 paveikslas) ir oksidacinio streso baltymų svarbą ateriosklerozės, kartu ir KŠL, procesui ir reikalauja tolesnių tyrimų siekiant nustatyti konkretų biologinį kelią.



**4.2 paveikslas.** Hipotetinis Bach1 reguliacinis veikimo mechanizmas pelių ląstelėse (Igarashi K et Watanabe-Matsui M, 2014)

Mūsų tyrimo metu tiriant *FBXL17* geno asociaciją, vienodu dažniu nustatyti du šio geno pokyčiai: rs34990078 ir rs61749621. Abu šie variantai yra nepusiausviroje sankiboje  $D' = 0,96$  (1000 genomų pirmos fazės CEU populiacijos tyrimo duomenimis). Abiejų pokyčių atveju mūsų atliktame tyrimo retesniojo alelio homozigotų nenustatyta. Pakaita rs34990078, c.2432

G>A yra devintame geno egzone ir priskiriama sinoniminių pokyčių tipui, nes nekeičia aminorūgšties baltyme (p.Thr675=). Kita nustatyta pakaita, rs61749621, c.1\*G>A, yra *FBXL17* geno 3' netransliuojamoje srityje, ir priskiriama 3' NTS pokyčių tipui. Abu nustatyti pokyčiai, *Ensembl* duomenų bazės duomenimis, yra reguliacinėje baltymo dalyje, todėl neatmestinas šių pokyčių poveikis baltymo *FBXL17* funkcionalumui.

Galima teigti, kad *FBXL17* genas yra galimas genas kandidatas, dalyvaujantis KŠL ir aterosklerozės patogenezėje. Tiriama *RTN4* ir *FBXL17* genų kaip potencialių kandidatų asociaciją su KŠL, tiriamoje imtyje dėl techninių aplinkybių nebuvo tiriami pirmi šių genų egzonai, kas neleidžia atmesti *RTN4* geno kaip potencialaus kandidato. Be to, tyrimo tikslumą riboja imtis, kurią padidinus būtų gauti dar tikslesni rezultatai.

Atliekant naujų KŠL genetinių veiksnių paiešką, nustatytos potencialios kandidatines genetines sritys, kurios iki šiol dar nebuvo asocijuotos su KŠL. Tokiu būdu pagilintos žinios apie KŠL architektūrą. Taip pat buvo nustatytas genetinis žinomų KŠL rizikos VNP profilis, genetinė struktūra ir įvairovė Lietuvos populiacijoje. Šis tyrimas yra vienas iš pavyzdžių, kaip galima tirti daugiaveiksnių ligų genetinę architektūrą. Akivaizdu, kad bendra Lietuvos lietuvių populiacija pagal tiriamus KŠL riziką lemiančius alelius, palyginti su kitomis populiacijomis, yra unikali, kaip savo ruožtu yra unikalios ir kitos populiacijos. Taigi, nors daug kas jau žinoma ir kas bendra žmonių populiacijai, kiekviena tauta, norėdama pakeisti esamą sergamumą ir mirštamumą, turės pati sau atsakyti į priežastinius klausimus, kas, kaip ir kodėl lemia ligą ir kaip tai galima paveikti. Siekiant tikslesnių rezultatų, Lietuvoje būtų tikslinga tęsti šio tipo tyrimus.

## IŠVADOS

1. Sumodeliuota daugiaveiksnės ligos genetinės architektūros ištyrimo strategija, remiantis pasirinktu koronarinės širdies ligos pavyzdžiu. Taikant panašų modelį galima tirti daugelį daugiaveiksnių ligų.
2. Atrinkta 60 polinkį sirgti koronarine širdies liga lemiančių genomo VNP žymenų.
3. Pagal atrinktus 60 VNP nustatyta tirtos bendros Lietuvos lietuvių populiacijos genetinė struktūra ir įvairovė:
  - a. Pagal tirtų alelių dažnius lietuvių populiacijoje statistiškai reikšmingų skirtumų vyrų ir moterų grupėje nenustatyta, aukštaičių ir žemaičių grupėje nustatyti du alelių dažnio skirtumai.
  - b. Pagal haplotipų blokų analizę vyrų ir moterų grupės bei aukštaičių ir žemaičių grupės skiriasi, nustatyti skirtingi haplotipų dažniai ir nepusiausvira sankiba, o haplotipų blokas 19-oje chromosomoje rastas tik moterų grupėje.
  - c. Pagal rizikos alelių skaičių, tenkantį asmeniui, populiaciją būtų galima suskirstyti į tris dalis: 39–50 alelių (1 grupė), 51–62 aleliai (2 grupė), 63–74 aleliai (3 grupė); daugiausia asmenų (68 %) priklauso antrai grupei, mažiausiai (13 %) – trečiai. Taigi pagal nustatytą rizikos alelių skaičių, tenkantį asmeniui, bendros Lietuvos lietuvių populiacijos didžiajai asmenų daliai būdinga santykinai vidutinė koronarinės širdies ligos rizika.
  - d. Pagal lietuvių populiacijoje tirtų VNP genetinio profilio palyginimą su kitomis populiacijomis (CEU, FIN, GBR, TSI), daugiausia alelių dažnių skirtumų yra nustatyta šiaurės–pietų geografinės ašies atžvilgiu (LTU-FIN 17 skirtumų, LTU-TSI 20 skirtumų), mažiausia – vakarų (LTU-CEU 8 skirtumai, LTU-GBR 6 skirtumai). Tai įrodo geografinių platumų įtaką genetinei



įvairovei, kiekvienos populiacijos unikalumą daugiaveiksnių ligų atveju ir jų tyrimo svarbą.

4. Atliekant bandomąjį plataus masto genomo tyrimą nustatyti nauji VNP, asocijuoti su koronarine širdies liga:
  - a. Atlikus plataus masto genomo asociacijų analizę šeimose, siekiant rasti naujų koronarinės širdies ligos genetinių veiksnių, nustatytos aštuonios genetinės sritys, statistiškai reikšmingai asocijuotos su koronarine širdies liga. Iš jų pakartotinai patvirtintos trys sritys: 5q31.1 (rs1346440, rs2019973 RNR genuose *LOC340073* ir *LOC10099648*), 12p11.23 (rs11048567 gene *ITPR2*) ir 20q11.21 (rs6141273).
  - b. Atlikus genų kandidatų patikrą atvejo ir kontrolės asociacijos analize, *RTN4* genas nepatvirtintas, bet neatmestinas kaip kandidatas, o *FBXL17* geno atveju nustatyta silpna statistiškai reikšminga asociacija, patvirtinanti jį kaip geną kandidatą, galintį dalyvauti KŠL ir aterosklerozės patogenezėje.

## LITERATŪROS SĄRAŠAS

1. Altman DG. Practical statistics for medical research. Chapman and Hall, 1991, London, UK.
2. Altmann A, Tian L, Henderson VW, Greicius MD. Alzheimer's Disease Neuroimaging Initiative Investigators. Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann Neurol*. 2014 Mar 13. doi: 10.1002/ana.24135. (Epub ahead of print)
3. Arnett DK. The need for genetics/genomics literacy in cardiovascular/stroke care. 2012, pranešimas <http://www.vdebooks.com/cardiovascular-genetics-and-genomics-PDF-4419897/>
4. Arvind P, Nair J, Jambunathan S, Kakkar VV, Shanker J. CELSR2-PSRC1-SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort. *J Cardiol*. 2014 Mar 24. pii: S0914-5087(14)00067-7. doi:10.1016/j.jjcc.2014.02.012. (Epub ahead of print)
5. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005 Jan 15;21(2):263-5.
6. Bekris LM, Lutz F, Yu CE. Functional analysis of APOE locus genetic variation implicates regional enhancers in the regulation of both TOMM40 and APOE. *J Hum Genet*. 2012 Jan; 57(1): 18–25.
7. Bhopal R. Glossary of terms relating to ethnicity and race: for reflection and debate. *Journal of Epidemiology and Community Health* 2004; 58: 441-5.
8. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet*. 2008; 40: 695–701.
9. Bolton JL, Stewart MCW, Wilson JF, Anderson N, Price JF. Improvement in Prediction of Coronary Heart Disease Risk over Conventional Risk Factors Using SNPs Identified in Genome-Wide Association Studies. *PLoS ONE*. 2013; 8(2): e57310.

10. Boudi B. Risk Factors for Coronary Artery Disease. Medscape, 2013.
11. Cavalli-Sforza LL, Menozzi P, Piazza A. The History and Geography of Human Genes. Princeton University Press, 1994, Princeton, US.
12. Chan L, Boerwinkle E. Gene-environment interactions and gene therapy in atherosclerosis. *Cardiology in Review* 1994; 2: 130-7.
13. Chen Y, Zhao S, Xiang R. RTN3 and RTN4: Candidate modulators in vascular cell apoptosis and atherosclerosis. *J Cell Biochem.* 2010 Nov 1;111(4):797–800.
14. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* 2004 Aug 6; 305(5685): 869-72.
15. Coop G, Przeworski M. An evolutionary view of human recombination. *Nat Rev Genet.* 2007 Jan; 8(1): 23-34.
16. Cooper RS, Kaufman JS, Ward R. Race and genomics. *N Engl J Med.* 2003; 348: 1166-70.
17. Coram MA, Duan Q, Hoffmann TJ, Thornton T, Knowles JW, Johnson NA, Ochs-Balcom HM, Donlon TA, Martin LW, Eaton CB, Robinson JG, Risch NJ, Zhu X, Kooperberg C, Li Y, Reiner AP, Tang H. Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. *Am J Hum Genet.* 2013 Jun 6; 92(6): 904-16.
18. Correia M, Branco CC, Bruffaerts N, Balagué A, Mota-Vieira L. Genetic risk assessment for cardiovascular disease in Azoreans (Portugal): a general population-based study. *Gene.* 2013 Dec 10; 532(1): 132-9.
19. Česnys G. Antropologinės mūšų šaknys. *Mokslas ir Lietuva.* 1991; 1:4-10.
20. Česnys G. Rasės samprata istorinėje antropologijoje. Vilnius: Vilniaus universiteto leidykla, 2009.
21. Dahlback B. Advances in understanding pathogenic mechanisms of thrombophilic disorders. *Blood.* 2008 Jul 1; 112(1): 19-27.

22. Daly MJ, Rioux JD, Schaffner SF et al. High-resolution haplotype structure in the human genome. *Nature Genetics* 2001; 29: 229-32.
23. Dandona P, Aljada A, Chaudhuri A, Mohanty P, Garg R. Metabolic syndrome: a comprehensive perspective based on interactions between obesity, diabetes, and inflammation. *Circulation*. 2005; 111: 1448-54.
24. Deloukas P, The CARDIoGRAMplusC4D Consortium et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013 Jan; 45(1): 25-33.
25. Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schäfer H. Gene-environment interactions for complex traits: definitions, methodological requirements and challenges. *Eur J Hum Genet*. 2008 Oct; 16(10): 1164-72.
26. Dent T. Predicting the risk of coronary heart disease with conventional, genetic and novel molecular biomarkers. 2009, PHG Foundation, Cambridge, UK.
27. Depre C, Powell SR, Wang X. The role of the ubiquitin-proteasome pathway in cardiovascular disease. *Cardiovasc Res*. 2010; 85(2): 251-2.
28. Di Rienzo A, Hudson RR. An evolutionary framework for common diseases: the ancestral-susceptibility model. *Trends Genet*. 2005; 21:596–601.
29. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*. 2010; 8: e1000294.
30. Ding K, Kullo IJ. Evolutionary Genetics of Coronary Heart Disease. *Circulation*. 2009; 119: 459-67.
31. DNR išskyrimo rinkinio techninis aprašas: MagneSil® Genomic, Large Volume System, Instructions For Use of Products A4080, A4082, A4085 Technical Bulletin.
32. Domarkienė I, Pranculis A, Germanas Š, Jakaitienė A, Vitkus D, Dženkevičiūtė V, Kučinskienė ZA, Kučinskas V. *RTN4* and *FBXL17* genes are associated with coronary heart disease in genome-wide

- association analysis of Lithuanian families. *Balkan Journal of Medical Genetics*. 2013; 16( 2): 17–22.
33. Drenos F, Whittaker JC, Humphries SE. The use of meta-analysis risk estimates for candidate genes in combination to predict coronary heart disease risk. *Ann Hum Genet*. 2007 Sep; 71(Pt 5): 611-9.
34. Dumitrescu L, Carty CL, Taylor K, Schumacher FR, Hindorff LA, Ambite JL, Anderson G, Best LG, Brown-Gentry K, Bůžková P, Carlson CS, Cochran B, Cole SA, Devereux RB, Duggan D, Eaton CB, Fornage M, Franceschini N, Haessler J, Howard BV, Johnson KC, Laston S, Kolonel LN, Lee ET, MacCluer JW, Manolio TA, Pendergrass SA, Quibrera M, Shohet RV, Wilkens LR, Haiman CA, Le Marchand L, Buyske S, Kooperberg C, North KE, Crawford DC. Genetic determinants of lipid traits in diverse populations from the population architecture using genomics and epidemiology (PAGE) study. *PLoS Genet*. 2011 Jun; 7(6): e1002138.
35. Eichler K, Puhon MA, Steurer J, Bachmann LM. Prediction of first coronary events with the Framingham score: A systematic review. *American Heart Journal*. 2007; 153: 722–31.
36. Esko T, Mezzavilla M, Nelis M, Borel C, Debniak T, Jakkula E, Julia A, Karachanak S, Khrunin A, Kisfali P, Krulisova V, Aušrelė Kučinskienė Z, Rehnström K, Traglia M, Nikitina-Zake L, Zimprich F, Antonarakis SE, Estivill X, Glavač D, Gut I, Klovins J, Krawczak M, Kučinskas V, Lathrop M, Macek M, Marsal S, Meitinger T, Melegh B, Limborska S, Lubinski J, Paolotie A, Schreiber S, Toncheva D, Toniolo D, Wichmann HE, Zimprich A, Metspalu M, Gasparini P, Metspalu A, D'Adamo P. Genetic Structure of Europeans: A View from the North–East Genetic characterization of northeastern Italian population isolates in the context of broader European genetic diversity. *European Journal of Human Genetics*. 2013; 21: 659-65.
37. Folkersen L, van't Hooft F, Chernogubova E, Agardh HE, Hansson GK, Hedin U, Liska J, Syvänen AC, Paulsson-Berne G, Paulsson-Berne G,

- Franco-Cereceda A, Hamsten A, Gabrielsen A, Eriksson P, groups BaAs. Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. *Circ Cardiovasc Genet.* 2010; 3: 365-73.
38. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 2013 Jan;41(Database issue): D808-15.
39. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science.* 2002 Jun 21; 296(5576): 2225-9.
40. Genest JJ, Martin-Munley SS, McNamara JR et al. Familial lipoprotein disorders in patients with premature coronary artery disease. *Circulation* 1992; 85: 2025-33.
41. Germanas Š, Jakaitiene A. Power approximation of the transmission disequilibrium test. *Lithuanian Mathemat J.* 2012; 53(Series B): 56-61.
42. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet.* 2012 Jan 18; 13(2): 135-45.
43. Girdenis A, Zinkevičius Z. Dėl lietuvių kalbos tarmių klasifikacijos. *Kalbotyra.* 1966; 14: 139-47.
44. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee.

- Heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation*. 2013 Jan 1; 127(1): e6-e245.
45. Grigalionienė K, Berezniakovas A, Kučinskas V, Kučinskienė V, Petrulionienė Ž. PON1 ir IL-6 genų genetinių žymenų asociacijos su paveldimu polinkiu aterosklerozei analizė Lietuvos populiacijoje. *Laboratorinė medicina*. 2007; 9(3): 109-14. (1)
46. Grigalionienė K, Pauža P, Kučinskas V, Kučinskienė V, Petrulionienė Ž. Apolipoproteino E, kepenų lipazės genų ir jų alelinių variantų asociacija su ateroskleroze Lietuvos populiacijoje. *Laboratorinė medicina*. 2007; 9(3): 124-9. (2)
47. Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G, Di Rienzo A. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet*. 2008; 4: e32.
48. Hardy J, Singleton A. Genomewide association studies and human disease. *N Engl J Med*. 2009 Apr 23; 360(17): 1759-68.
49. Henn BM, Gravel S, Moreno-Estrada A, Acevedo-Acevedo S, Bustamante CD. Fine-scale population structure and the era of next-generation sequencing. *Hum Mol Genet*. 2010 Oct 15; 19(R2): R221-6.
50. Hern LM, Bidichandani SI. What Mendel did not discover: exceptions in Mendelian genetics and their role in inherited human disease. *J Okla State Med Assoc*. 2004 Jan; 97(1): 12-7.
51. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007 Jul 21; 335(7611): 136.
52. Holdt LM, Sass K, Gabel G, Bergert H, Thiery J, Teupser D. Expression of Chr9p21 genes CDKN2B (p15INK4b), CDKN2A (p16INK4a, p14ARF) and MTAP in human atherosclerotic plaque. *Atherosclerosis* 2011; 214: 264-70.
53. Houle D. Comparing evolvability and variability of quantitative traits. *Genetics*. 1992; 130: 195-204.

54. Humphries SE, Drenos F, Ken-Dror G, Talmud PJ. Coronary heart disease risk prediction in the era of genome-wide association studies: current status and what the future holds. *Circulation*. 2010 May 25; 121(20): 2235-48.
55. Hunter DJ, Kraft P. Drinking from the fire hose--statistical issues in genomewide association studies. *N Engl J Med*. 2007 Aug 2; 357(5): 436-9.
56. Hunter DJ. Gene–environment interactions in human diseases. *Nature Reviews Genetics*. 2005; 6: 287-98.
57. Igarashi K, Watanabe-Matsui M. Wearing red for signaling: the heme-bach axis in heme metabolism, oxidative stress response and iron immunology. *Tohoku J Exp Med*. 2014; 232(4): 229-53.
58. Inouye M, Kettunen J, Soininen P, Silander K, Ripatti S, Kumpula LS, Hamalainen E, Jousilahti P, Kangas AJ, Mannisto S, Savolainen MJ, Jula A, Leiviska J, Palotie A, Salomaa V, Perola M, Ala-Korpela M, Peltonen L. Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol Syst Biol*. 2010 Dec 21; 6: 441.
59. Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet*. 2011 Jun; 43(6): 519-25.
60. Jakaitienė A, Kučinskas V. Lietuvių populiacijos struktūros įvertinimas pagal etnolingvistines grupes. *Lietuvos statistikos darbai*. 2013; 52(1): 34-38.
61. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS, McPherson R. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol*. 2009 Oct; 29(10): 1671-7.



62. Jin J, Cardozo T, Lovering RC, Elledge SJ, Pagano M, Harper JW. Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev.* 2004 Nov 1; 18(21): 2573-80.
63. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA. Haplotype tagging for the identification of common disease genes. *Nat Genet.* 2001 Oct; 29(2): 233-7.
64. Kittles RA, Weiss KM. Race, ancestry, and genes: implications for defining disease risk. *Annu Rev Genomics Hum Genet.* 2003; 4: 33–67.
65. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. *Science.* 2005 Apr 15; 308(5720): 385-9.
66. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics.* 2007; 23(10): 1289-91.
67. Krimpenfort P, Ijpenberg A, Song JY, van der Valk M, Nawijn M, Zevenhoven J, Berns A. p15Ink4b is a critical tumour suppressor in the absence of p16Ink4a. *Nature* 2007; 448: 943–6.
68. Kučinskis V. *Genomo įvairovė: Lietuviai Europoje.* Vilnius: Spalvų šalis, 2004.
69. Lewis CM, Knight J. Introduction to genetic association studies. *Cold Spring Harb Protoc.* 2012 Mar 1; 2012(3): 297-306.
70. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest.* 2008 May; 118(5): 1590-605.
71. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G,

- Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature*. 2009 Oct 8; 461(7265): 747-53.
72. Maouche S, Schunkert H. Strategies beyond genome-wide association studies for atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2012 Feb; 32(2): 170-81.
73. Martin C, Taylor P, Potts H. Construction of an odds model of coronary heart disease using published information: the Cardiovascular Health Improvement Model (CHIME). *BMC Medical Informatics and Decision Making*. 2008; 8: 49.
74. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007 Jun 8; 316(5830): 1488-91.
75. Morkūnas K. Lietuvių kalbos atlasas I. Mokslas, 1977, Vilnius, žemėlapis Nr.1, prieiga internete [http://www.tarmes.lt/images/tarmiu\\_skirstymas/zem\\_1\\_nauja%20klasifikacija.gif](http://www.tarmes.lt/images/tarmiu_skirstymas/zem_1_nauja%20klasifikacija.gif)
76. Nakajima T, Wooding S, Sakagami T, Emi M, Tokunaga K, Tamiya G, Ishigami T, Umemura S, Munkhbat B, Jin F, Guan-Jun J, Hayasaka I, Ishida T, Saitou N, Pavelka K, Lalouel JM, Jorde LB, Inoue I. Natural selection and population history in the human angiotensinogen gene (AGT): 736 complete AGT sequences in chromosomes from around the world. *Am J Hum Genet*. 2004; 74: 898-916.
77. Neel JV. Diabetes mellitus: a “thrifty” genotype rendered detrimental by “progress”? *Am J Hum Genet*. 1962; 14: 353-62.
78. Nelis M, Esko T, Mañgi R, Zimprich F, Zimprich A, et al. Genetic Structure of Europeans: A View from the North–East. *PLoS ONE*. 2009; 4(5): e5472.

79. Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG. Recent and ongoing selection in the human genome. *Nat Rev Genet.* 2007; 8: 857-68.
80. Peloso GM, Demissie S, Collins D, Mirel DB, Gabriel SB, Cupples LA, Robins SJ, Schaefer EJ, Brousseau ME. Common genetic variation in multiple metabolic pathways influences susceptibility to low HDL-cholesterol and coronary heart disease. *J Lipid Res.* 2010 Dec; 51(12): 3524-32.
81. Pepalytė I, Kučinskienė ZA, Grigalionienė K, Petrulionienė Ž, Dženkevičiūtė V, Bagdonaitė L, Kučinskas V. Genetic variants that participate in oxidation processes and/or oxidative stress and are associated with atherosclerosis. *European Medical, Health and Pharmaceutical Journal.* 2012; 3: 13-16.
82. Perk J, De Backer G, Gohlke H, Graham I, Reiner Z, Verschuren M, Albus C, Benlian P, Boysen G, Cifkova R, Deaton C, Ebrahim S, Fisher M, Germano G, Hobbs R, Hoes A, Karadeniz S, Mezzani A, Prescott E, Ryden L, Scherer M, Syvanne M, Scholte op Reimer WJ, Vrints C, Wood D, Zamorano JL, Zannad F; European Association for Cardiovascular Prevention & Rehabilitation (EACPR); ESC Committee for Practice Guidelines (CPG). European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Eur Heart J.* 2012 Jul; 33(13): 1635-701.
83. Petkevičienė J, Smalinskienė A, Lukšienė DI, Jurėnienė K, Ramažauskienė V, Klumbienė J, Lesauskaitė V. Associations between apolipoprotein E genotype, diet, body mass index, and serum lipids in Lithuanian adult population. *PLoS One.* 2012; 7(7): e41525.
84. Preuss M, König IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg S, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin

- E, Hengstenberg C, Holm H, Laaksonen R, Li M, März W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AF, Thorleifsson G, Voight BF, Wells GA, Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H; CARDIoGRAM Consortium. Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet*. 2010 Oct; 3(5): 475-83.
85. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet*. 2001; 69: 124-37.
86. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007 Sep; 81(3): 559-75.
87. Qin ZS, Niu T, Liu JS. Partition-ligation-expectation-maximization algorithm for haplotype inference with single-nucleotide polymorphisms. *Am J Hum Genet*. 2002 Nov; 71(5): 1242-7.
88. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.
89. Rakauskas R. Rūšies sąvoka biologijoje. Vilnius: Vilniaus universiteto leidykla, 2001.
90. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet*. 2001;17: 502-10.
91. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki ML, Nieminen MS, Melander O, Salomaa V, Peltonen L, Kathiresan S. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet*. 2010 Oct 23; 376(9750): 1393-400.

92. Roberts R, Stewart AFR. 9p21 and the genetic revolution for coronary artery disease. *Clin Chem*. 2012 Jan; 58(1): 104-12.
93. Roberts R, Stewart AFR. Personalized medicine: a future prerequisite for the prevention of coronary artery disease. *American Heart Journal* 2006; 4: 222-7.
94. Roden D. Cardiovascular genetics and genomics. Wiley-Blackwell, 2009, Dallas, US.
95. Rodriguez S, Gaunt TR, Day INM. Hardy-Weinberg Equilibrium Testing of Biological Ascertainment for Mendelian Randomization Studies. *Am J Epidemiol*. 2009 Feb 15; 169(4): 505-14.
96. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson JF, Campbell H. Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet*. 2011 Nov 11; 89(5): 607-18.
97. Smalinskienė A, Petkevičienė J, Lukšienė D, Jurėnienė K, Klumbienė J, Lesauskaitė V. Association between APOE, SCARB1, PPAR $\alpha$  polymorphisms and serum lipids in a population of Lithuanian adults. *Lipids Health Dis*. 2013 Aug 6; 12: 120.
98. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. *Annals of Human Genetics* 1996; 59: 983-9.
99. Srivastava RA, Bhasin N, Srivastava N. Apolipoprotein E gene expression in various tissues of mouse and regulation by estrogen. *Biochem Mol Biol Int*. 1996 Feb; 38(1): 91-101.
100. Stanislovaitienė D, Lesauskaitė V, Žaliūnienė D, Smalinskienė A, Gustienė O, Žaliaduonytė-Pekšienė D, Tamošiūnas A, Lukšienė D, Petkevičienė J, Žaliūnas R. SCARB1 single nucleotide polymorphism (rs5888) is associated with serum lipid profile and myocardial infarction in an age- and gender-dependent manner. *Lipids Health Dis*. 2013 Mar 5; 12: 24.

101. Strachan T, Read A. Human molecular genetics. 4th edition. Garland Science, 2011, Abingdon, UK.
102. Tam CH, Ma RC, So WY, Wang Y, Lam VK, Germer S, Martin M, Chan JC, Ng MC. Interaction effect of genetic polymorphisms in glucokinase (GCK) and glucokinase regulatory protein (GCKR) on metabolic traits in healthy Chinese adults and adolescents. *Diabetes*. 2009 Mar; 58(3): 765-9.
103. Tan MK, Lim HJ, Bennett EJ, Shi Y, Harper JW. Parallel SCF adaptor capture proteomics reveals a role for SCFFBXL17 in NRF2 activation via BACH1 repressor turnover. *Mol Cell*. 2013 Oct 10; 52(1): 9-24.
104. The International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005 Oct 27; 437(7063): 1299-320.
105. Tikkanen E, Havulinna AS, Palotie A, Salomaa V, Ripatti S. Genetic risk prediction and a 2-stage risk screening strategy for coronary heart disease. *Arterioscler Thromb Vasc Biol*. 2013 Sep; 33(9): 2261-6.
106. Timinskas A, Kučinskienė ZA, Kučinskas V. Atherosclerosis: alterations in cell communication. *Acta Medica Lituanica*. 2007; 14(1): 26-31.
107. Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG Primer3 - new capabilities and interfaces. *Nucleic Acids Research*. 2012; 40(15): e115.
108. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012 Jan 13; 90(1): 7-24. (1)
109. Visscher PM, Goddard ME, Derks EM, Wray NR. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry*. 2012 May; 17(5): 474-85. (2)
110. Wald NJ, Law MR. A strategy to reduce cardiovascular disease by more than 80%. *BMJ* 2003; 326: 1419-23.

111. Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. Interpretation of association signals and identification of causal variants from genome-wide association studies. *Am J Hum Genet.* 2010 May 14; 86(5): 730-42.
112. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet.* 2005 Feb; 6(2): 109-18.
113. Warnatz HJ, Schmidt D, Manke T, Piccini I, Sultan M, Borodina T, Balzereit D, Wruck W, Soldatov A, Vingron M, Lehrach H, Yaspo ML. The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle. *J Biol Chem.* 2011 Jul 1; 286(26): 23521-32.
114. Watari Y, Yamamoto Y, Brydun A, Ishida T, Mito S, Yoshizumi M, Igarashi K, Chayama K, Ohshima T, Ozono R. Ablation of the *bach1* gene leads to the suppression of atherosclerosis in *bach1* and apolipoprotein E double knockout mice. *Hypertens Res.* 2008 Apr; 31(4): 783-92.
115. Whitley E, Ball J. Statistics review 4: sample size calculations. *Crit Care.* 2002; 6(4): 335-341.
116. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet.* 2005 May; 76(5): 887-93.
117. Wilker EH, Alexeeff SE, Poon A, Litonjua AA, Sparrow D, Vokonas PS, Mittleman MA, Schwartz J. Candidate genes for respiratory disease associated with markers of inflammation and endothelial dysfunction in elderly men. *Atherosclerosis.* 2009 Oct; 206(2): 480-5.
118. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton

- P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008 Feb; 40(2): 161-9.
119. Williams GC, Nesse RM. The dawn of Darwinian medicine. *Q Rev Biol.* 1991; 66: 1–22.
120. Williams GC. Pleiotropy, natural selection, and the evolution of senescence. *Evolution.* 1957; 11: 398–411.
121. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998 May 12; 97(18): 1837-47.
122. Woodward M, Brindle P, Tunstall-Pedoe H; SIGN group on risk estimation. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart.* 2007 Feb; 93(2): 172-6.
123. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk of complex disease. *Curr Opin Genet Dev.* 2008; 18: 257–263.
124. Zeller T, Blankenberg S, Diemert P. Genomewide association studies in cardiovascular disease – an update 2011. *Clin Chem.* 2012 Jan; 58(1): 92-103.
125. Zhang K, Sun F, Waterman MS, Chen T. Haplotype block partition with limited resources and applications to human chromosome 21 haplotype data. *Am J Hum Genet.* 2003 Jul; 73(1): 63-73.
126. Zivelin A, Griffin JH, Xu X, Pabinger I, Samama M, Conard J, Brenner B, Eldor A, Seligsohn U. A single genetic origin for a common



Caucasian risk factor for venous thrombosis. *Blood*. 1997 Jan 15; 89(2): 397-402.

127. Žaliaduonytė-Pekšienė D, Šimonytė S, Lesauskaitė V, Vaškelytė J, Gustienė O, Mizarienė V, Jurkevičius R, Jarienė G, Tamošiūnas A, Žaliūnas R. Left ventricular remodelling after acute myocardial infarction: Impact of clinical, echocardiographic parameters and polymorphism of angiotensinogen gene. *J Renin Angiotensin Aldosterone Syst*. 2013 Jan 2. [Epub ahead of print]

## **PADĖKA**

Esu dėkinga savo darbo vadovui prof. habil. dr. akad. Vaidučiui Kučinskui už suteiktą galimybę atlikti šį darbą, pasitikėjimą, vertingą pagalbą ir patarimus.

Už pagalbą, vertingus patarimus ir palaikymą nuoširdžiai ačiū prof. habil. dr. akad. Zitai Aušrelei Kučinskienei, prof. (HP) Algirdui Utkui, dr. Laimai Ambrozaitytei.

Taip pat dėkoju savo brangiems kolegoms, kurie prisidėjo įgyvendinant šį darbą: Ingridai Uktverytei, Aidui Pranculiui, Justinai Petraškaitei, Raimondai Maconytei, Alinai Urnikytei, Daivai Kazlauskaitei, Dalytei Pliaugo, dr. Aušrai Morkūnienei, Kristinai Grigalionienei, Violetai Mikštienei, dr. Almai Molytei, doc. dr. Audronei Jakaitienei, Šarūnui Germanui, dr. Valentinai Ginevičienei, doc. dr. Loretai Bagdonaitei, prof. habil. dr. Žanetai Petrulionienei, dr. Vilmai Dženkevičiūtei ir visiems VU MF ŽMGK bei VUL SK MGC darbuotojams, LITGEN projekto darbuotojų kolektyvui, Estijos Tartu universiteto Molekulinės ir ląstelės biologijos instituto Biotechnologijos skyriaus kolektyvui.

Išskirtinė padėka yra tiems asmenims, kurie sutiko duoti savo kraujo mėginius ir dalyvauti šiame moksliniame tyrime, taip pat mano šeimai už palaikymą, supratimą ir visokeriopą pagalbą.

## **PRIEDAI**

**I PRIEDAS.** DNR išskyrimo ir DNR koncentracijos bei švarumo nustatymo protokoliai

### **DNR IŠSKYRIMO FENOLIO IR CHLOROFORMO IŠSKYRIMO METODU PROTOKOLAS**

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

- ✓ **10 % natrio dodecilsulfatas**, gamintojas *AppliChem GmbH*, DE
- ✓ **2 % proteinazė K**, gamintojas *Thermo Fisher Scientific, Inc.*, US
- ✓ **3 M natrio acetatas**, pH = 5,2, gamintojas *AppliChem GmbH*, DE
- ✓ **70 %, 96 % etanolis**, gamintojas AB Vilniaus degtinė, LT
- ✓ **Branduolių pernešimo buferis** (ruošiamas, 10 mM Tris-HCl, pH = 10,5; 1 mM EDTA; 0,15 mM NaCl)
- ✓ **Chloroformas**, gamintojas *Carl Roth GmbH + Co.KG*, DE
- ✓ **Chloroformo ir izoamilio alkoholio mišinys** (24:1), gamintojas *Carl Roth GmbH + Co.KG*, DE
- ✓ **ddH<sub>2</sub>O**
- ✓ **Fenolis**, gamintojas *Carl Roth GmbH + Co.KG*, DE
- ✓ **Lizuojantis buferis** (ruošiamas, 0,32 M sacharozės; 5 mM MgCl<sub>2</sub>; 1 % Triton X – 100; 10 mM Tris-HCl, pH = 7,5)
- ✓ **TE buferis** (10 mM Tris-HCl, 1 mM EDTA, pH=8), *AppliChem GmbH*, DE
- ✓ **Vakuuminis mėgintuvėlis (3–10 mL) su EDTA** ir paimtu veniniu krauju

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

- ✓ **Centrifuga** (1,5 mL, 15 mL, 50 mL mėgintuvėliams), modelis *EBA 12* ir modelis *Universal32R* (gamintojas *Hettich Zentrifugen*, DE)
- ✓ **Latekso pirštinės**
- ✓ **Liofilizatorius**, angl. *Eppendorf Concentrator 5301*, gamintojas *Eppendorf AG*, DE
- ✓ **Mikrocentrifuga** (0,5 mL, 1,5 mL, 2 mL), gamintojas *Carl Roth GmbH + Co.KG*, TW

- ✓ **Mikropipetės**, vienakanalės (2–20  $\mu\text{L}$ , 10–100  $\mu\text{L}$ , 100–1000  $\mu\text{L}$ )
- ✓ **Sterilios Pastero pipetės** (3 mL)
- ✓ **Sterilūs antgaliai** (20  $\mu\text{L}$ , 100  $\mu\text{L}$ , 1000  $\mu\text{L}$ )
- ✓ **Sterilūs mėgintuvėliai** (0,5 mL, 1,5 mL, 2 mL, 50 mL)
- ✓ **Termostatas**, gamintojas *WTB Binder GmbH*, DE

### Darbo eiga

1. Veninį kraują iš vakuuminio mėgintuvėlio (3–10 mL) su antikoagulantu 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 supipetuoti 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 supipetuoti 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  $\mu\text{L}$  ląstelių branduolių pernešimo buferio.
6. Visą tūrį sterilia Pastero pipete perkelti į sterilų 2 mL tūrio mėgintuvėlį, įpilti 20  $\mu\text{L}$  10 % natrio dodecilsulfato tirpalo ir 10  $\mu\text{L}$  2 % proteinazės K tirpalo.
7. Inkubuoti termostate 37 °C temperatūroje 16 val. arba 55 °C temperatūroje 3 val.
8. Po inkubavimo įpilti 400  $\mu\text{L}$  fenolio, mėgintuvėlį su mišiniu vartyti 10 min, vėliau centrifuguoti 5000 aps./min 2 min.
9. Viršutinį sluoksnį nusiurbti ir įpilti į naują sterilų 2 mL tūrio mėgintuvėlį.
10. Įpilti 200  $\mu\text{L}$  chloroformo ir 200  $\mu\text{L}$  fenolio, mėgintuvėlį su mišiniu vartyti 5–10 min, centrifuguoti 5000 aps./min 2 min.

11. Viršutinį sluoksnį nusiurbti ir įpilti į naują sterilų 2 mL tūrio mėgintuvėlį.
12. Įpilti 500 µL chloroformo ir izoamilo alkoholio mišinio, mėgintuvėlį su mišiniu vartyti 5 min, centrifuguoti 5000 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, centrifuguoti 5000 aps./min 2 min.
15. Viršutinį sluoksnį nusiurbti, visą mišinio tūrį padalinti į dvi lygias dalis ir jas 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 DNR, mėgintuvėlius centrifuguoti 10 000–13 000 aps./min 5 min.
17. Viršnuosėdinį skystį nusiurbti, įpilti po 500 µL 70 % šalto (-20 °C) etanolio, centrifuguoti 10000–13000 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 DNR saugoti -20 °C arba -80 °C temperatūroje.
20. Nusiurbti 1,5 mL tūrio mėgintuvėlio viršnuosėdinį sluoksnį, liofilizuoti 30 °C temperatūroje.
21. Į mėgintuvėlį su liofilizuota DNR įpilti 100 µL TE buferio ir laikyti 4 °C temperatūroje ne trumpiau nei 24 val., kad išskirta DNR pakankamai ištirptų.

## DNR IŠSKYRIMO AUTOMATIZUOTA ROBOTINE SISTEMA *TECAN* *FREEDOM EVO*<sup>®</sup> 200 PROTOKOLAS

### Reagentai ir cheminės medžiagos

- ✓ **0,25 M NaOH**
- ✓ **2-propanolis**, angl. *2-propanol, iso-propyl alcohol*, gamintojas **Sigma-Aldrich Co., US**
- ✓ **96 % etanolis**, gamintojas AB „Vilniaus degtinė“, LT
- ✓ **ddH<sub>2</sub>O**
- ✓ **DNR išskyrimo rinkinys**, angl. *MagneSil<sup>®</sup> Genomic, Large Volume System*, gamintojas *Promega Corp., US*
- ✓ **Išplovimo buferis**, angl. *Elution Buffer*
- ✓ **Laštelių lizės tirpalas**, angl. *Cell Lysis Solution (Genomic Purification)*, gamintojas *Promega Corp., US*
- ✓ **Lizuojantis buferis**, angl. *eLysis Buffer (Large Volume System)*
- ✓ **Magnetinės dalelės**, angl. *MagneSil<sup>®</sup> Paramagnetic Particles*
- ✓ **Plovimo buferis etanolio pagrindu**, angl. *Alcohol wash*
- ✓ **Vakuuminis mėgintuvėlis (3–10 mL) su EDTA** ir paimtu veniniu krauju

### Iranga ir laboratoriniai reikmenys

- ✓ **Kaitinimo blokas su adapteriu**, angl. *VWR<sup>™</sup> Digital Heatblock*, gamintojas *VWR International LLC, US*, angl. *Heat block Adapter*, gamintojas *Promega Corp., US*
- ✓ **Latekso pirštinės**
- ✓ **Magnetinis stovas**, angl. *MagnaBot<sup>®</sup> Large Volume Magnetic Separation Device*, gamintojas *Promega Corp., US*
- ✓ **Mėgintuvėlių (50 mL) stovas**, angl. *Tube Holder, 50 mL Tubes*, gamintojas *Promega Corp., US*
- ✓ **NALGENE<sup>®</sup> indas (1 L talpos)**, angl. *NALGENE<sup>®</sup> 1000 mL PP jar*, gamintojas *Nalge Nunc International Corp., US*

- ✓ **Purtyklė su adapteriu**, angl. *IKA® KS 130 control*, gamintojas *IKA-Werke GmbH & Co. KG, DE*, angl. *Shaker Top Adapter*, gamintojas *Promega Corp., US*
- ✓ **Sterilūs mėgintuvėliai** (1,5 mL)
- ✓ **Sterilūs mėgintuvėliai** (50 mL), angl. *Corning 50 ml Centrifuge Tube*, gamintojas *Corning Inc., US*
- ✓ **TECAN FREEDOM EVO® 200** prietaisas – automatizuota robotinė DNR išskyrimo sistema, angl. *TECAN FREEDOM EVO® 200 robotic platform*, gamintojas *Tecan Schweiz AG, CH*

### **Darbo eiga**

1. Pagal veninio kraujo ėminio tūrį (mL) ir ėminių kiekį – 8, 16 arba 24, paruošti laboratorinius reikmenis ir reagentus (**2.1 lentelė**, **2.2 lentelė**).

**2.1 lentelė.** Reagentų kiekiai, reikalingi išskirti gDNR iš 3 mL veninio kraujo ėminio

Ėminiai (vnt.)	Lizuojantis buferis (mL)	Plovimo buferis (mL)	Išplovimo buferis (mL)	Magnetinės dalelės (mL)	Ląstelių lizės tirpalas (mL)
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)	-

**2.2 lentelė.** Reagentų kiekiai, reikalingi išskirti gDNR iš 10 mL veninio kraujo ėminio

Ėminiai (vnt.)	Lizuojantis buferis (mL)	Plovimo buferis (mL)	Išplovimo buferis (mL)	Magnetinės dalelės (mL)	Ląstelių lizės tirpalas (mL)
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)	-

2. Paruošti plovimo buferį, į buferio tirpalą (500 mL) įpilti 250 mL 2-propanolio ir 250 mL 96 % etanolio.
3. Ląstelių lizės tirpalą praskiesti ddH<sub>2</sub>O santykiu 1:4.
4. Magnetines daleles inkubuoti termostate 37 °C 35 min.
5. Įsitikinti, kad prietaiso *TECAN FREEDOM EVO® 200* atliekų indai (angl. *Waste bottle*) švarūs ir tušti.

6. Prietaiso *TECAN FREEDOM EVO*<sup>®</sup> 200 skysčių sistemos (angl. *Liquid System*) indą pripildyti ddH<sub>2</sub>O.
7. Patikrinti prietaiso švirkštus, jei yra vandens lašų, nuvalyti beplauše popierine servetėle.
8. Atsižvelgiant į kraujo ėminių skaičių, į prietaiso 1 ir (ar) 2 pozicijas įstatyti sterilius 1,5 mL mėgintuvėlius.
9. Atsižvelgiant į kraujo ėminių skaičių, į prietaiso 3 ir (ar) 4 pozicijas be dangtelių įstatyti vakuuminius mėgintuvėlius su ėminiais (10 mm/3 mL arba 16 mm/10 mL).
10. Į prietaiso 10–14 pozicijas įstatyti tris *NALGENE*<sup>®</sup> 1 L talpos indus.
11. Į prietaiso 17–25 pozicijas įstatyti du mėgintuvėlių (50 mL) stovus, kurių kiekviename įstatyta po aštuonis sterilius 50 mL mėgintuvėlius be dangtelių.
12. Į prietaiso 28–29 poziciją įstatyti indą su lizuojančiu buferiu.
13. Į prietaiso 30–31 poziciją įstatyti indą su paruoštu plovimo buferiu.
14. Į prietaiso 34–39 pozicijas įstatyti stovą su aštuoniais 50 mL mėgintuvėliais, pripildytais magnetinių dalelių.
15. Į prietaiso 42–43 poziciją įstatyti du indus: vieną su ląstelių lizės tirpalu, kitą su išplovimo buferiu.
16. Įjungti 50–60 pozicijoje esančią purtyklę su adapteriu.
17. Įjungti 62–69 pozicijoje esantį kaitinimo bloką su adapteriu, nustatyti 95 °C temperatūrą.

### **Prietaiso ir programos aktyvavimas**

1. Uždaryti apsauginį stiklą, įjungti prietaisą (mygtukas *ON*).
2. Įjungti programą *Freedom EVOware*<sup>®</sup> *Standard 2.3* (*EVOware 2.3*).
3. Pasirinkti *Run an existing script*.
4. Pasirinkti *Start your selection* (žalia rodyklė).
5. Kataloge *Favorite* pasirinkti dokumentą  
*NEW\_Tecan\_gDNA\_XL\_16\_mm\_tubes*.
6. Pasirinkti *Run*.



7. Pasirodžius įspėjimui *Fill up system liquid!!!* spausti *Ok*.
8. Nurodyti prašomą tūrį *Volume in waste* ir *Volume in cleaner* įvedant 30.
9. Pasirodžius klausimui *Start recovery mode?* (*0 = no; 1 = yes*) įvesti 0.
10. Pasirodžius klausimui *Read barcode?* (*1 = yes; 0 = no*) įvesti 0.
11. Pasirodžius klausimui *How many sample would you like to process?* įvesti aktualų ėminių skaičių 8, 16 arba 24.
12. Pasirodžius klausimui *Start with blood in 50 mL processing tube* (*0 = no; 1 = yes*) įvesti 0.
13. Pasirodžius klausimui *All required racks, plates and buffers correctly prepared on to worktable?* dar kartą patikrinti, ar reagentai ir visi darbui reikalingi reikmenys tinkamai paruošti, įsitikinus spausti *Ok*.
14. Pasirodžius klausimui *Is the heat block turned on and set up 95 °C?* patikrinti, ar kaitinimo bloko temperatūra pasiekusi 95 °C, įsitikinus spausti *Ok*.
15. Pasirodžius klausimui *Would you like 1 or 2 elution to be performed?* įvesti 1.

### **Vykdomi DNR išskyrimo etapai**

1. Ėminio ląstelių lizavimas ląstelių lizės tirpalu.
2. gDNR prijungimas prie magnetinių dalelių.
3. gDNR plovimas lizuojančio ir plovimo buferio mišiniu tris kartus.
4. gDNR plovimas plovimo buferiu tris kartus.
5. gDNR pašalinimas nuo magnetinių dalelių išplovimo buferiu.

### **Prietaiso priežiūra**

1. Po kiekvieno gDNR išskyrimo atlikti prietaiso priežiūrą: prietaiso išorinius ir vidinius paviršius nuvalyti ir dezinfekuoti.
2. Į prietaiso 42 poziciją įstatyti du indus, vieną su 96 % etanoliu, kitą su 0,25 M NaOH.
3. Įjungti *EVOware 2.3* programą.
4. Pasirinkus *Run maintenance*, toliau pasirinkti *Clean\_Fix\_Tips\_Inside* ir spausti *Start your selection* (žalia rodyklė), galiausiai spausti *Run*.

## DNR KONCENTRACIJOS IR ŠVARUMO MATAVIMO PROTOKOLAS

### Reagentai ir cheminės medžiagos

- ✓ ddH<sub>2</sub>O
- ✓ DNR tirpalas
- ✓ TE buferis (10 mM Tris-HCl, 1 mM EDTA, pH=8), gamintojas *AppliChem GmbH, DE*
- ✓ Įranga ir laboratoriniai reikmenys
- ✓ Beplaušės popierinės servetėlės
- ✓ Latekso pirštinės
- ✓ Mikrocentrifuga (0,5 mL, 1,5 mL), gamintojas *Carl Roth GmbH + Co.KG, TW*
- ✓ Mikropipetės, vienakanalės (2–20 µL, 10–100 µL)
- ✓ *NanoDrop*<sup>®</sup> spektrofotometras, angl. *NanoDrop*<sup>®</sup> *ND-1000 Spectrophotometer*, gamintojas *NanoDrop Technologies Inc., US*
- ✓ Purtyklė, modelis *Vortex-T Genie 2*, gamintojas *Scientific Industries, Inc., US*
- ✓ Sterilūs antgaliai (20 µL, 100 µL)
- ✓ Sterilūs mėgintuvėliai (0,5 mL, 1,5 mL)

### Darbo eiga

1. Įjungti programą *ND-1000 v3.5.2*.
2. Pasirinkti *Nucleic Acid*.
3. Pakelti spektrofotometro *NanoDrop*<sup>®</sup> svirtelę (angl. *Sampling arm*), beplauše servetėle, suvilgyta ddH<sub>2</sub>O, nuvalyti apatinį ir viršutinį matavimo paviršius (angl. *Lower and Upper Measurement Pedestal*).
4. Mikropipete užlašinti 2 µL ddH<sub>2</sub>O mėginio ir ekrane pateiktoje informacinėje lentelėje spausti *OK*. Spektrofotometras aktyvuojamas.
5. Pasibaigus spektrofotometro aktyvacijai, nepašalinus nuo matavimo paviršiaus ddH<sub>2</sub>O, spausti *Blank*.

6. Pakelti svirtelę ir švaria beplauše popierine servetėle nuvalyti apatinį ir viršutinį matavimo paviršius.
7. Mėgintuvėlį su išskirtos DNR tirpalu supurtyti ir trumpai centrifuguoti.
8. Mikropipete užlašinti 2  $\mu\text{L}$  DNR tirpalo mėginio ir spausti *Measure*.
9. Įvertinama išskirtos DNR tirpalo koncentracija (DNR tirpalo sugertis esant 260 nm bangos ilgiui ( $A_{260}$ )) ir švarumas (DNR tirpalo, esant 260 nm ir 280 nm bangos ilgiams, sugerties reikšmių santykis ( $A_{260/280}$ )). DNR tirpalo koncentracija nurodoma ng/ $\mu\text{L}$ . Tinkamai išgrynintos DNR švarumo reikšmės priklauso intervalui (1,7–1,9).
10. Pakartotinai spausti *Measure* ir nustatyti dviejų gautų reikšmių vidurkį vienam mėginiui.
11. Esant poreikiui, išskirtos DNR tirpalą skiesti ddH<sub>2</sub>O arba TE buferiu iki reikiamos darbinės koncentracijos.
12. Baigus matavimą, prietaiso matavimo paviršius nuvalomas beplauše popierine servetėle ir išplaunamas ddH<sub>2</sub>O.

## II PRIEDAS. Genotipavimo *Illumina Infinium*<sup>®</sup> HD VNP lustais protokolas

### *Illumina Infinium*<sup>®</sup> HD Assay Ultra PROTOKOLAS

#### Medžiagos ir reagentai

- ✓ **2-propanolis**, 100 %, gamintojas *Sigma-Aldrich, Co.*, US
- ✓ **ATM reagentas**, angl. *Anti-Stain Two-Color Master Mix*, dvispalvis reagentų mišinys fluoroforo žymeį sustiprinti, *Infinium*<sup>®</sup> HD Assay Kit WG-Post 1 LVI rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **Baliklio tirpalas**, 0,525 %, gaminamas iš 99,6 % natrio hipochlorito tirpalo, Kometa „Pelėsių valiklis“, gamintojas *Pufas Werk KG*, DE
- ✓ **ddH<sub>2</sub>O** – didejonizuotasis vanduo
- ✓ **Detergento tirpalas**, 1%, angl. *Alconox*<sup>®</sup> Powdered Precision Cleaner, gamintojas *Alconox, Inc.*, US
- ✓ **Eliucijos buferis**, angl. *Elution buffer, MagneSil*<sup>®</sup> Genomic, Large Volume System rinkinys, gamintojas *Promega Corp.*, US
- ✓ **Etanolis**, 100 %, gamintojas *Sigma-Aldrich, Co.*, US
- ✓ **FA/ EDTA tirpalas**, 95 % formamido ir 1 mM EDTA tirpalas
- ✓ **FMS reagentas**, angl. *Fragmentation Solution*, skaidymo tirpalas, *Infinium*<sup>®</sup> HD Assay Kit WG-Post 1 LVI rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **MA1 reagentas**, angl. *Multi-Sample Amplification 1 Mix*, daugiau nei vieno mėginio gausinimo mišinys 1, *Infinium*<sup>®</sup> HD Assay Kit WGS-Pre LVI rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **MA2 reagentas**, angl. *Multi-Sample Amplification 2 Mix*, daugiau nei vieno mėginio gausinimo mišinys 2, pradmenys, *Infinium*<sup>®</sup> HD Assay Kit WGS-Pre LVI rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **MSM reagentas**, angl. *Multi-Sample Amplification Master Mix*, pagrindinis daugiau nei vieno mėginio gausinimo mišinys, *Infinium*<sup>®</sup> HD Assay Kit WGS-Pre LVI rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **NaOH**, 0,1 N

- ✓ **PB1 reagentas**, angl. *Reagent used to prepare BeadChips for hybridization*, paruošiamasis VNP lustų hibridizacijos reagentas, silpno surišimo tirpalas plovimui, *Infinium<sup>®</sup> Assay Kit Post 4 LV* rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **PB2 reagentas**, angl. *Humidifying buffer used during hybridization*, drėgmę palaikantis hibridizacijos buferis, *Infinium<sup>®</sup> Assay Kit Single Post 3 LV* rinkinys, gamintojas *Illumina, Inc.*, US
- ✓ **PM1 reagentas**, angl. *Precipitation solution*, nusodinimo tirpalas, *Infinium<sup>®</sup> Assay Kit Single Post 3 LV* rinkinys, gamintojas *Illumina, Inc.*, US
- ✓ **RA1 reagentas**, angl. *Resuspension, hybridization and wash solution*, resuspendacijos, hibridizacijos ir plovimo (stipraus surišimo) tirpalas, *Infinium<sup>®</sup> Assay Kit Post 2 LMV* rinkinys, gamintojas *Illumina, Inc.*, US
- ✓ **STM reagentas**, angl. *Superior Two-Color Master Mix*, pagrindinis dvispalvis dažymo reagentų mišinys (žymėjimui fluoroforais), *Infinium<sup>®</sup> HD Assay Kit WG-Post 1 LVI* rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **TE buferis** (10 mM Tris-HCl, 1 mM EDTA, pH = 8), gamintojas *AppliChem GmbH*, DE
- ✓ **TEM reagentas**, angl. *Two-Color Extension Master Mix*, dvispalvis vieno nukleotido pratęsimo reakcijos reagentų mišinys, *Infinium<sup>®</sup> HD Assay Kit WG-Post 1 LVI* rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **VNP lustai**, angl. *BeadChip Array 12-sample HD DNA Analysis BC HumanOmniExpress-12v1-1* ir *HumanOmniExpress-12v1-0*, žmogaus genomines DNR VNP genotipavimo lustai, gamintojas *Illumina Inc.*, US
- ✓ **XC1 reagentas**, angl. *Xstain BeadChip Solution 1*, VNP lustų dažų tirpalas 1, blokuojantis, *Infinium<sup>®</sup> HD Assay Kit WG-Post 1 LVI* rinkinys, gamintojas *Illumina Inc.*, US

- ✓ **XC2 reagentas**, angl. *Xstain BeadChip Solution 2*, VNP lustų dažų tirpalas 2 buferio keitimui, *Infinium<sup>®</sup> HD Assay Kit WG-Post 1 LVI* rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **XC3 reagentas**, angl. *Xstain BeadChip Solution 3*, VNP lustų dažų tirpalas 3 silpno surišimo plovimui, *Infinium<sup>®</sup> Assay Kit Post 4 LV* rinkinys, gamintojas *Illumina Inc.*, US
- ✓ **XC4 reagentas**, angl. *Xstain BeadChip Solution 4*, VNP lustų dažų tirpalas 4 dažų apsaugojimui, *Infinium<sup>®</sup> Assay Kit Post 4 LV* rinkinys, gamintojas *Illumina Inc.*, US

### Įranga ir laboratoriniai reikmenys

- ✓ **Beplaušės servetėlės**, angl. *lint-free wipes*
- ✓ **Beplaušiai sugertukai**, angl. *lint-free absorbent pad*
- ✓ **Centrifuga**, 15 mL mėgintuvėliams, modelis *Universal 32 R* (gamintojas *Hettich Zentrifugen*, DE), 96 šulinėlių plokštelėms modelis *Heraeus Multifuge 3L* (gamintojas *Thermo Electron Corp.*, DE)
- ✓ **Daugiavietis VNP lustų talpiklis**, angl. *Multi-Sample BeadChip Alignment Fixture*, gamintojas *Illumina Inc.*, US
- ✓ **Folijos užlydytuvas**, angl. *Combi Thermo-Sealer ALPS<sup>™</sup> 25 (Heat Sealer)*, plokštelių dengiamosios folijos užlydytuvas, gamintojas *Thermo Fisher Scientific*, UK
- ✓ **Hibridizacijos kameros guminis tarpiklis**, angl. *Hybridization chamber gasket*, gamintojas *Illumina Inc.*, US
- ✓ **Hibridizacijos kameros įdėklai**, angl. *Hybridization chamber inserts*, gamintojas *Illumina Inc.*, US
- ✓ **Hibridizacijos kameros**, angl. *Hybridization chambers*, gamintojas *Illumina Inc.*, US
- ✓ **Hibridizacijos krosnis**, angl. *Illumina<sup>®</sup> Hybridization Oven, Model 5521 with BeadChip Hyb Wheel*, hibridizacijos krosnis su lustų

- hibridizacijos padėklų, modelis 5521, gamintojas *Boehel Scientific*, US
- ✓ **HiScan™SQ sistema, genetinis analizatorius**, angl. *HiScan™SQ system*, gamintojas *Illumina Inc.*, US
  - ✓ **Laikiklis ant VNP lustų talpiklio**, angl. *Multi-Sample BeadChip Alignment Fixture Alignment Bar*, gamintojas *Illumina Inc.*, US
  - ✓ **Latekso pirštinės**
  - ✓ **Mėgintuvėlių stovai**
  - ✓ **Mentelė** gnybtukams nuimti, angl. *dismantling tool*, gamintojas *Illumina Inc.*, US
  - ✓ **Mikropipetės**, vienkanalės ir daugiakanalės, 2–20 µL, 10–100 µL, 20–200 µL, 30–300 µL, 100–1000 µL
  - ✓ **Mikroplokštelių inkubatorius Hybex®**, angl. *Hybex® Microsample Incubator (Heat block)*, gamintojas *SciGene Corp.*, US
  - ✓ **MSA3 plokštelė**, angl. *MSA3 plate*, 96 šulinėlių polipropileno 0,8 mL mikrotitro plokštelė (MIDI), gamintojas *Thermo Scientific*, UK
  - ✓ **MSA3 plokštelės 96 šulinėlių dengiamoji guma**, angl. *96-well cap mat*, gamintojas *Thermo Scientific*, UK
  - ✓ **MSA3 plokštelės dengiamoji folija**, angl. *Heat Sealing Foil Sheets*, gamintojas *Thermo Scientific*, UK
  - ✓ **Nešiojamas termometras OMEGA®**, angl. *OMEGA® 869C Handheld RTD (resistance temperature detector) Thermometer*, OMEGA® nešiojamas RTD termometras 869C, gamintojas *OMEGA Engineering Inc.*, US
  - ✓ **Pincetas** stikleliams perkelti, angl. *Self locking tweezers 6 1/2"*, gamintojas *Aven Inc.*, US
  - ✓ **Plastikiniai antgaliai su filtrais** 20 µL, 100 µL, 200 µL, 1000 µL
  - ✓ **Plastikiniai loveliai** reagentams išpilstyti, angl. *Tip-tubes*
  - ✓ **Plastikiniai mėgintuvėliai** 0,2 mL, 0,5 mL, 1,5 mL

- ✓ **Plokštelių purtyklė**, angl. *Illumina<sup>®</sup> High Speed Microplate Shaker Digital 230V*, didelio greičio 230 V skaitmeninė plokštelių purtyklė, gamintojas *Henry Troemer, LLC, US*
- ✓ **Spektrofotometras** *NanoDrop<sup>®</sup>*, angl. *NanoDrop<sup>®</sup> ND-1000 Spectrophotometer*, gamintojas *NanoDrop Technologies Inc., US*
- ✓ **Stikliniai ir plastikiniai indai**, plovimui, angl. *Wash dish*, gamintojas *Illumina Inc., US*
- ✓ **Suspaustas oras**, nupūsti dulkėms nuo dengiamųjų stiklų, angl. *Air pressure cleaner (Laboratory air gun)*, gamintojas *UAB ACME<sup>®</sup> Europe, Kaunas, LT*
- ✓ **Te-Flow<sup>™</sup> pratekamosios kameros**, angl. *Te-Flow<sup>™</sup> Flow-Through Chambers (with black frames, spacers, glass backplates, and clamps)*, *Te-Flow<sup>™</sup> pratekamosios kameros (su juodais rėmeliais, tarpinėmis, dengiamaisiais stiklais ir gnybtukais)*, gamintojas *Illumina Inc., US*
- ✓ **Te-Flow<sup>™</sup> pratekamųjų kamerų laikiklis Tecan**, 48 vietų, angl. *Tecan Te-Flow<sup>™</sup> Module Chamber Rack 48 Pos.*, gamintojas *Tecan Schweiz AG, CH*
- ✓ **Termostatas ir vandens cirkuliacijos sistema** prie *Tecan* prietaiso, angl. *Advanced Programmable Controller Refrigerated/ Heated Circulating Baths*, pažangus programuojamas temperatūros reguliatorius su cirkuliacijos sistema, gamintojas *VWR<sup>®</sup> International, LLC, US*
- ✓ **Vakuuminė džiovyklė**, angl. *Vakuuum Desiccator "Space Saver"*, vakuuminė džiovyklė *"Space Saver"*, gamintojas *Scienceware<sup>®</sup> Bel-ART products, US*
- ✓ **VNP lustų laikiklis**, metalinis, angl. *Wash rack*, plovimui prieš dažymą, gamintojas *Illumina Inc., US*
- ✓ **VNP lustų laikiklis**, plastmasinis, angl. *Staining rack*, plovimui po dažymo, gamintojas *Illumina Inc., US*
- ✓ **Žirklys** tarpinėms perkirpti, gamintojas *Illumina Inc., US*



## **Darbo eiga**

### **MĖGINIŲ PARUOŠIMAS**

Atrinktų ir į tyrimą įtrauktų asmenų išskirtos ir tirpintos TE arba elucijos buferyje genominės DNR koncentracija ir švarumas nustatomas spektrofotometru *NanoDrop*<sup>®</sup>. DNR mėginiai, jeigu reikia, naujame mėgintuvėlyje skiedžiami ddH<sub>2</sub>O iki protokole nurodytos darbinės koncentracijos (~50 ng/μL). Vienu kartu genotipuojant 96 mėginius rekomenduojama kartu leisti vieną geros kokybės (kontrolinį) DNR mėginį.

### **GENOMINĖS DNR GAUSINIMAS (MSA3 plokštelės paruošimas) – atliekama PGR pilstymo patalpoje, Pre-PGR**

1. Atsižvelgiant į tiriamų DNR mėginių skaičių, į naują MSA3 plokštelę atitinkamai išpilstyti po 20 μL MA1 reagento (daugiau nei vieno mėginio gausinimo mišinys 1).
2. Į atitinkamus (pagal pildomą protokolą) MSA3 plokštelės šulinėlius įpilti po 4 μL DNR (konc. ~50 ng/μL).
3. Į kiekvieną šulinėlį su MA1 ir DNR įpilti po 4 μL 0,1 N NaOH.
4. Uždengti plokštelę dengiamąja guma, purtyti plokštelių purtyklėje 1600 aps./min 1 min, centrifuguoti 280xg 1 min.
5. Inkubuoti plokštelę kambario (20–25 °C) temperatūroje 10 min.
6. Į kiekvieną šulinėlį su mėginiais išpilstyti:
  - po 34 μL MA2 reagento (daugiau nei vieno mėginio gausinimo mišinys 2, pradmenys);
  - po 38 μL MSM reagento (pagrindinis daugiau nei vieno mėginio gausinimo mišinys).
7. Uždengus plokštelę dengiamąja guma, purtyti plokštelių purtyklėje 1600 aps./min 1 min, centrifuguoti 280xg 1 min.
8. Inkubuoti plokštelę hibridizacijos krosnyje 37 °C 20–24 val.

**GENOMINĖS DNR SKAIDYMAS (fermentinis) – šis ir visi tolimesni protokolo žingsniai atliekami Post-PGR patalpoje**

1. Išimtą iš hibridizacijos krosnies MSA3 plokštelę centrifuguoti 50xg 1 min.
2. Į mėginių šulinėlius po 25 µL išpilstyti FMS reagentą (skaidymo tirpalas), uždengti plokštelę dengiamąja guma.
3. Purtyti plokštelę 1600 aps./min 1 min, centrifuguoti 50xg 1 min.
4. Inkubuoti plokštelę mikroplokštelių inkubatoriuje *Hybex*<sup>®</sup> 37 °C 1 val.

**GENOMINĖS DNR NUSODINIMAS**

1. Į MSA3 plokštelės mėginių šulinėlius po 50 µL išpilstyti PM1 reagentą (nusodinimo tirpalas), uždengti plokštelę dengiamąja guma.
2. Purtyti plokštelę 1600 aps./min 1 min.
3. Inkubuoti plokštelę mikroplokštelių inkubatoriuje *Hybex*<sup>®</sup> 37 °C 5 min.
4. Po inkubavimo plokštelę centrifuguoti 50xg 1 min.
5. Į kiekvieną šulinėlį su mėginiu po 155 µL išpilstyti 100 % 2-propanolį, uždengti plokštelę **nauja** dengiamąja guma, ~10 kartų vartyti plokštelę.
6. Inkubuoti plokštelę 4 °C 30 min.
7. Po inkubavimo MSA3 plokštelę centrifuguoti 4 °C 3000xg 20–30 min.  
**NB!** Po centrifugavimo plokštelės šulinėlių su mėginiais dugne turi būti matomas melsvos nuosėdos.
8. **Nedelsiant** išimti plokštelę iš centrifugos, išmesti dengiamąją gumą ir **staigiu judesiu** apversti plokštelę ant beplaušių sugertukų.
9. Neleidžiant viršnuosėdiniam skysčiui pakliūti į kitus šulinėlius stipriai daužyti plokštelę ant absorbuojančių beplaušių sugertukų ~1 min.
10. Apverstą neuždengtą plokštelę ant mėgintuvėlių stovo džiovinti kambario temperatūroje (20–25 °C) 1 val. Prieš resuspendaciją neapvertus plokštelės įsitikinti, ar nuosėdos tikrai išdžiūvo.

## **GENOMINĖS DNR RESUSPENDACIJA**

1. Į MSA3 plokštelę su nuosėdomis po 23 µL išpilstyti iš anksto atitirpintą RA1 reagentą (resuspendacijos, hibridizacijos ir plovimo (aukšto surišimo) tirpalas). Ant plokštelės užlydyti dengiamosios folijos lapą naudojantis folijos užlydytuvu (prieš tai įsitikinti, ar užlydytuvo kaitinimo paviršius švarus). Nuleistą užlydytuvą palaikyti prispaudus prie folijos ~1,5–5 s.
2. Plokštelę inkubuoti hibridizacijos krosnyje 48 °C 1 val.
3. Po inkubavimo MSA3 plokštelę purtyti 1800 aps./min 1 min, trumpai centrifuguoti 280xg.

## **GENOMINĖS DNR HIBRIDIZACIJA ANT VNP LUSTŲ (tipas 12x1)**

### **Hibridizacijos kameros paruošimas**

1. Surinkti hibridizacijos kamerą (dėl dėvėjimosi vienodumo visuomet rinkinio detales naudoti kartu, nemišyti su naujesnėmis ir pan.): į hibridizacijos kameros dugną tinkamai įdėti guminių tarpiklių.
2. Į hibridizacijos kameros dugne esančius rezervuarus įpilti po 400 µL PB2 reagento (drėgmę palaikantis hibridizacijos buferis).
3. Uždengti hibridizacijos kameros dangčiu.

### **Hibridizacija**

1. Resuspenduotus mėginius MSA3 plokštelėje denatūruoti mikroplokštelių inkubatoriuje *Hybex*<sup>®</sup> 95 °C 20 min.
2. Iš šaldytuvo išimti VNP lustus, bet neatidaryti pakuočių. Atidaryti tik prieš pat DNR mėginių įnešimą.
3. Po denatūravimo MSA3 plokštelę vėsinti kambario temperatūroje (20–25 °C) 30 min, trumpai centrifuguoti 280xg.
4. Išpakuoti VNP lustus ir tinkamai (brūkšniniai kodai turi būti jiems numatytoje vietoje) sudėti į hibridizacijos kameros įdėklus, kurie

- užpylus DNR mėginius ant VNP lustų bus dedami į hibridizacijos kamerą.
5. Laikant mikropipetę statmenai, ant atitinkamo įleidimo į lustą laukelio užpilti po 15 µL resuspenduoto DNR mėginio. Mėginius užpilti pagal pildomą protokolą, pažymint protokole jų vietą VNP luste ir VNP lusto brūkšninį kodą
  6. Nedelsiant VNP lustus, esančius hibridizacijos kameros įdėkluose, tinkamai įdėti į hibridizacijos kameroje jiems skirtas vietas ir atsargiai uždengti hibridizacijos kameros dangčiu. Dviem judesiais užspausti hibridizacijos kameros užraktus (po du priešingose pusėse, įstrižai).
  7. Hibridizacijos kamerą padėti ant hibridizacijos krosnies padėklo taip, kad kameros užraktai būtų nukreipti į kairę ir į dešinę krosnies sienas. Hibridizacijos kameros logotipas turėtų būti nukreiptas į hibridizacijos krosnies viršų.
  8. Inkubuoti esant 5 vienetų hibridizacijos krosnies padėklo siūbavimui 48°C 16–24 val.

#### **XC4 reagento (VNP lustų dažų tirpalas 4 dažams apsaugoti) paruošimas**

1. Į XC4 butelį įpilti 330 mL 100 % etanolio. Stipriai kratyti 15 s. Jei nevyksta resuspendacija, purtyti plokštelių purtyklėje 1625 aps./min, kol pasigamins vienalytė suspensija.
2. Palikti resuspenduotis per naktį kambario temperatūroje (20–25 °C).

**NB!** XC4 yra daugkartinio naudojimo, tinkamas 24 lustams padengti. Galioja dvi savaites po pagaminimo. Pagaminus ant butelio būtina pažymėti pagaminimo datą ir panaudojimo kartų skaičių.

#### **VNP LUSTŲ PLOVIMAS (parengimas dažymui)**

1. Hibridizacijos kamerą išimti iš hibridizacijos krosnies, palaukti ~25min, kol atvės kambario temperatūroje (20–25 °C).
2. Pripildyti (~200 mL) du stiklinius indus PB1 reagento (paruošiamasis lustų hibridizacijos reagentas siplno surišimo plovimui). Paruošti

- metalinį (plovimo prieš dažymą) VNP lustų laikiklį ir pamerkti į pirmąjį stiklinį indą su PB1 reagentu.
3. Paruošti daugiavietį VNP lustų talpiklį. Pripildyti jį ~150 mL PB1 reagento (kad VNP lustai būtų visiškai apsemti).
  4. Paruošti *Te-Flow*<sup>TM</sup> pratekamųjų kamerų dalis: juodus rėmelius, plastikines tarpines, dengiamuosius stiklus ir gnybtukus.
  5. Dengiamuosius stiklus paruošti iš anksto:
    - ✓ pirmą kartą naudojant, taip pat kas septynis naudojimo kartus, plauti stiklus 0,525 % baliklio tirpalu pagal *Illumina Infinium*<sup>®</sup> Laboratorijos paruošimo ir procedūrų (*Illumina Infinium*<sup>®</sup> Assay Lab Set Up and Procedures) vadove nurodytą protokolą;
    - ✓ kiekvieną kartą prieš naudojant plauti 1 % detergento tirpalu (*Alconox*<sup>®</sup>) pagal *Illumina Infinium*<sup>®</sup> Laboratorijos paruošimo ir procedūrų (*Illumina Infinium*<sup>®</sup> Assay Lab Set Up and Procedures) vadove nurodytą protokolą.
  6. Įjungti termostatą ir vandens cirkuliacijos sistemą, prijungtą prie *Te-Flow*<sup>TM</sup> pratekamųjų kamerų laikiklio *Tecan*, kad nusistovėtų 44 °C temperatūra (nurodytai temperatūrai pasiekti reikalinga nustatyti 1,5 °C didesnę temperatūrą).
  7. Hibridizacijos kamerai atvėsus, ją atidaryti ir išimti įdėklus su VNP lustais.
  8. Iš įdėklo išimti po vieną VNP lustą, jį laikant nuožulniai brūkšniu kodu arčiau savęs, vienu lėtu įstrižu tęstiniu judesiu nulupti VNP lusto plėvelę (*IntelliHyb Seal*). Neliesti lusto paviršiaus su hibridizuotais mėginiais.
  9. Nulupus plėvelę VNP lustą palengva panardinti į metalinį VNP lustų laikiklį pirmame inde su PB1 reagentu.
  10. Žingsnius 8 ir 9 atlikti kiekvienam VNP lustui.

11. Visiems lustams esant panardintiems metaliniame VNP lustų laikiklyje pirmame inde su PB1 reagentu, laikiklį aktyviai, bet švelniai kilnoti kertant skysčio paviršių ~1 min.
12. Perkelti laikilį su VNP lustais į antrą indą su PB1 reagentu, įsitikinti, kad lustai visiškai panyra į PB1 reagentą, ir laikiklį aktyviai, bet švelniai kilnoti kertant skysčio paviršių ~1 min.

### ***Te-Flow*<sup>TM</sup> pratekamųjų kamerų surinkimas**

1. Jei dar neparuoštas, paruošti daugiavietį VNP lustų talpiklį. Pripildyti jį ~150 mL PB1 reagento (kad VNP lustai būtų visiškai apsemti).
2. Surinkti *Te-Flow*<sup>TM</sup> pratekamašias kameras:
  - ✓ juodus rėmelius įdėti į jiems skirtas vietas daugiaviečiame VNP lustų talpiklyje;
  - ✓ į rėmelius įdėti VNP lustus, įsitikinti, kad jie yra visiškai panirę PB1 reagente;
  - ✓ atskyrus nuo baltų apsauginių plėvelių, permatomas plastikines tarpines tinkamai (pasinaudojus daugiaviečio VNP lustų talpiklio išsikišimais užkabinus paleisti) uždėti ant panardintų VNP lustų;
  - ✓ atitinkamai uždėti talpiklio laikiklį ant VNP lustų talpiklio;
  - ✓ suspausto sauso oro priemone nupūsti dulkeles nuo tinkamai išplautų ir 70 % etanolio nuvalytų dengiamųjų stiklų ir nedelsiant stiklus guldyti ant VNP lustų su plastikinėmis tarpinėmis (stiklo nuožulnioji vieta, skirta reagentų rezervuarui formuoti, turi būti ties VNP lusto brūkšniu kodu, nukreipta į vidų);
  - ✓ kiekvienos surenkamos *Te-Flow*<sup>TM</sup> pratekamosios kameros dviejose vietose užspausti gnybtukus – 5 mm atstumu nuo viršaus ir brūkšninio kodo pusėje 5 mm atstumu nuo reagentų rezervuaro pradžios;

- ✓ žirkklėmis nukirpti išsikišusias plastikinių tarpinių dalis abiejuose surinktos *Te-Flow*<sup>™</sup> pratekamosios kameros pusėse;
  - ✓ surinktas *Te-Flow*<sup>™</sup> pratekamąsias kameras laikyti horizontaliai ant darbastalio ir dėti į *Te-Flow*<sup>™</sup> pratekamųjų kamerų laikiklį *Tecan* tik tuomet, kai viskas paruošta tolesnei dažymo procedūrai.
3. Tuščią hibridizacijos kamerą nedelsiant plauti ddH<sub>2</sub>O, o PB2 reagento rezervuarus kruopščiai patrinti šepetėliu, kad būtų išvengta PB2 pridžiūvimo. Rekomenduojama plovimui panaudoti 1 % detergento (*Alcanox*<sup>®</sup>) tirpalą ir paskui gausiai praskalauti ddH<sub>2</sub>O.

### VNP LUSTŲ DAŽYMAS

1. Iš anksto paruošus vandens cirkuliacijos sistemą ir termostatą, patikrinti, ar sistemoje nėra burbulų, vartant *Te-Flow*<sup>™</sup> pratekamųjų kamerų laikiklį *Tecan*. Į pratekamųjų kamerų laikiklį *Tecan* įstatyti nešiojamą termometrą *OMEGA*<sup>®</sup> aktualiai temperatūrai (44 °C) nustatyti. Pasiekus reikiamą temperatūrą, įsitikinti, ar ir kitose *Tecan* talpiklio vietose temperatūra yra ta pati.
2. Paruošti reagentus: atitirpdžius centrifuguoti 3000xg 3 min, pagal naudojimo eiliškumą sudėti stovė.
3. Resuspendavimui užtikrinti stipriai supurtyti XC4 reagento butelį.

### Vienos bazės pratęsimo reakcija

1. Įsitikinus, kad temperatūra yra 44 °C, įstatyti surinktas *Te-Flow*<sup>™</sup> pratekamąsias kameras į *Te-Flow*<sup>™</sup> pratekamųjų kamerų laikiklį *Tecan*. Esant nedideliame pratekamųjų kamerų skaičiui, įstatyti jas įvairiose laikiklio *Tecan* vietose, esant didesniame skaičiui, įstatyti iš eilės pradedant pirma eile.
2. Į kiekvieną *Te-Flow*<sup>™</sup> pratekamąją kamerą mikropipete įpilti:
  - ✓ 150 µL RA1 reagento (resuspendacijos, hibridizacijos ir plovimo (aukšto surišimo) tirpalo), inkubuoti 30 s, kartoti 5 kartus (iš viso 6 įpylimai);

- ✓ 450 µL XC1 reagento (VNP lustų dažų tirpalas 1, blokuojantis), inkubuoti 10 min;
- ✓ 450 µL XC2 reagento (VNP lustų dažų tirpalas 2, buferio keitimui), inkubuoti 10 min;
- ✓ 200 µL TEM reagento (dvispalvis vieno nukleotido pratęsimo reakcijos reagentų mišinys), inkubuoti 15 min;
- ✓ 450 µL FA/ EDTA reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min;
- ✓ termostato ir vandens cirkuliacijos sistemoje nustatyti temperatūrą, nurodytą ant STM reagento (pagrindinis dvispalvis dažymo reagentų mišinys (žymėjimui fluoroforais)) buteliuko (nustačius 32,5 °C, bus 32 °C), jei nenurodyta – nustatyti 37 °C. Norint, kad greičiau vėstų, į cirkuliacijos sistemą įdėti talpas su šaldytu skystu ledu.
- ✓ 450 µL XC3 reagento (VNP lustų dažų tirpalas 3, žemo surišimo plovimui), inkubuoti 1 min, pakartoti vieną kartą.

### VNP lustų dažymas

1. Jei įvykdžius protokolą numatoma iš karto skenuoti VNP lustus, **įjungti *iScan* skenerį**.
2. Dažymą pradėti tik tuomet, kai pasiekta ant STM reagento pakuotės nurodyta temperatūra.
3. Į kiekvieną *Te-Flow*<sup>™</sup> pratekamąją kamerą mikropipete įpilti:
  - ✓ 250 µL STM reagento, inkubuoti 10 min;
  - ✓ 450 µL XC3 reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min;
  - ✓ 250 µL ATM reagento (dvispalvis reagentų mišinys, fluoroforo žymės sustiprinimui), inkubuoti 10 min;
  - ✓ 450 µL XC3 reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min;
  - ✓ 250 µL STM reagento, inkubuoti 10 min;



- ✓ 450 µL XC3 reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min;
  - ✓ 250 µL ATM reagento, inkubuoti 10 min;
  - ✓ 450 µL XC3 reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min;
  - ✓ 250 µL STM reagento, inkubuoti 10 min;
  - ✓ 450 µL XC3 reagento, inkubuoti 1 min, vieną kartą pakartoti ir inkubuoti 5 min.
4. Nedelsiant išimti *Te-Flow*<sup>TM</sup> pratekamašias kameras iš laikiklio *Tecan*, padėti horizontaliai ant darbastalio.

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

1. Į plastikinį plovimo indą įpilti ~310 mL PB1 reagento, į jį tinkamai įstatyti plastikinį VNP lustų laikiklį plovimui po dažymo (laikiklio rankenos tvirtinimo mechanizmas turėtų būti nukreiptas į save).
2. Išrinkti *Te-Flow*<sup>TM</sup> pratekamašias kameras:
  - ✓ mentele atkabinti gnybtukus;
  - ✓ atsargiai nuimti dengiamuosius stiklus;
  - ✓ stengiantis nepaliesiti VNP lusto mėginių zonų nuimti plastikines tarpines;
  - ✓ VNP lustą brūkšniu kodu nuo savęs nedelsiant panardinti į PB1 tirpale esantį plastikinį laikiklį. Pusę VNP lustų panardinti prieš laikiklio rankeną, pusę – už. Lustai turi būti visiškai panardinti PB1 reagente.
3. Lėtai kertant PB1 reagento paviršių kilnoti plastikinį laikiklį su VNP lustais aukštyn–žemyn 10 kartų. Palaikyti 5 min.
4. Supurtyti XC4 reagentą, įsitikinti, kad jis visiškai resuspendavęs.
5. Į plastikinį plovimo indą įpilti ~310 mL XC4 reagento, neleisti jam inde stovėti ilgiau nei 10 min.
6. Po VNP lustų mirkymo PB1 reagente, plastikinį VNP lustų laikiklį plovimui po dažymo tinkamai panardinti (laikiklio rankenos tvirtinimo mechanizmas turėtų būti nukreiptas į save) į indą su XC4 reagentu.

7. Lėtai kertant XC4 reagento paviršių kilnoti plastikinį laikiklį su VNP lustais aukštyn–žemyn 10 kartų. Palaikyti 5 min.
8. Lėtai ištraukti ir nuvarvinti plastikinį laikiklį su VNP lustais iš XC4 reagento indo, laikiklį padėti ant mėgintuvėlių stovo taip, kad VNP lustų brūkšniniai kodai būtų nukreipti į viršų.
9. Kadangi VNP lustai slidūs, rekomenduojama specialiu užsirakinančiu pincetu paeiliui iš plastikinio laikiklio ištraukti VNP lustus ir išrikiuoti ant mėgintuvėlių stovo brūkšniniais kodais į viršų.
10. VNP lustus džiovinti vakuuminėje džiovyklėje 50–55 min esant 508 mm Hg (0,68 bar).
11. Jei dar neįjungtas, įjungti *iScan* skenerį.
12. Išdžiovintų VNP lustų reversines puses atsargiai nušluostyti nuo XC4 reagento pertekliaus 70 % etanolio suvilgytomis beplaušėmis servetėlėmis.
13. Neketinant iš karto skenuoti, VNP lustus laikyti specialioje dėžutėje vakuuminėje džiovyklėje. Rekomenduojama skenuoti nepraėjus 72 val.
14. Visus VNP lustų paruošimui naudotus laboratorinius reikmenis užmerkti ir gerai išplauti naudojant ddH<sub>2</sub>O.

## VNP LUSTŲ SKENAVIMAS *Illustrina iScan* SKENERIU *HiScan™SQ* SISTEMA

### *iScan* paruošimas:

- ✓ įjungti *Illustrina* kompiuterį;
- ✓ suvesti prisijungimo duomenis;
- ✓ įjungti prietaisą *HiScan™SQ*;
- ✓ dukart pele spragtelėjus ant *iScan Control Software (ICS)* aplikacijos, esančios pradiniam darbalaukiui lange, ir suvesti prisijungimo duomenis.

**NB!** Iš anksto per *DMAP Client* programą parsisiųsti VNP lustų (pagal kiekvieną unikalų brūkšninį kodą) dekodavimo (angl. *decode*) dokumentus, kiekvienam VNP lustui po vieną *.dmap* dokumentą. Juos įkelti į DMAPS katalogą (kompiuterio D diske, D:/DMAPS).

## VNP lustų skenavimas

1. *iScan ICS* aplikacijoje pasirinkti **Start**.
2. Atsidarius prietaiso durelėms ant išvažiavusios platformos į atitinkamas vietas sudėti paruoštus VNP lustus. Paspaudus platformos šonuose esančius juodus mygtukus atitinkamai padėti VNP lustai sukrenta į taisyklingą skenavimui padėtį.
3. Pasirinkus **Next**, prietaiso durelės užsidaro. Skeneris nuskaity VNP lustų brūkšninius kodus. Pasirodo lentelė (3x4) su kiekvieno VNP lusto duomenimis (brūkšninio kodo skaitmenimis, lusto tipu ir skenavimo protokolu bei gausiamo vaizdo formatu). Naudojantis **Settings** mygtukais lentelėje pagal poreikį galima atlikti nustatymų korekcijas. Lentelės apačioje esančiuose laukeliuose būtina nurodyti *.dmap* dokumentų katalogą (**Input path:** D:/DMAPs) ir po skenavimo gausinamų duomenų kaupimo katalogą (**Output path:** D:/Illumina/LITGEN Scans).
4. Jeigu viskas yra atlikta, pasirinkti **Scan**.

### III PRIEDAS. Polimerazės grandininės reakcijos (PGR) ir sekoskaitos protokolas

#### POLIMERAZĖS GRANDININĖS REAKCIJOS (PGR) IR AUTOMATIZUOTOS SEKOSKAITOS SANGER METODU PROTOKOLAS

##### Reagentai ir cheminės medžiagos

- ✓ **1x TBE buferis** (0,09M Tris-borato, 0,002M EDTA, pH=8,0), gamintojas *Carl Roth GmbH + Co.KG*, DE
- ✓ **3130 Polimeras POP-7™**, gamintojas *Applied Biosystems*, US
- ✓ **3730 Kapiliarinės elektroforezės buferis su EDTA (10x)**, angl. *3730 Buffer (10x) with EDTA*, gamintojas *Applied Biosystems*, US
- ✓ **5x sekoskaitos buferis**, angl. *5x Sequencing Buffer*, gamintojas *Applied Biosystems*, UK
- ✓ **6x įvedimo dažas**, angl. *6x DNA Loading Dye*, gamintojas *Thermo Fisher Scientific*, LT
- ✓ **70 %, 96 % etanolis**, gamintojas AB Vilniaus degtinė, LT
- ✓ **Agarozės milteliai** (SeaKem®, LE Agarose), gamintojas *Lonza Group Ltd*, US
- ✓ **ddH<sub>2</sub>O**
- ✓ **DMSO** – dimetilsulfoksidai, gamintojas *Carl Roth GmbH+ Co.KG*, DE
- ✓ **DNR tirpalas**
- ✓ **dNTP**, deoksinukleotidtrifosfatų mišinys, 2M, gamintojas *Thermo Fisher Scientific*, LT
- ✓ **Egzonukleazė I**, angl. *Exonuclease I, ExoI*, gamintojas *UAB Fermentas*, LT
- ✓ **Etidžio bromidas**, gamintojas *Carl Roth GmbH + Co.KG*, DE
- ✓ **Hi-Di™ Formamidas**, gamintojas *Applied Biosystems*, UK
- ✓ **Krevečių šarminė fosfatazė**, angl. *Shrimp Alkaline Phosphatase*, *UAB Fermentas*, LT

- ✓ **Molekulinės DNR masės ir ilgio standartas**, angl. *GeneRuler™ 100bp DNA Ladder*, gamintojas UAB *Fermentas*, LT
- ✓ **Natrio acetatas, 3 M (pH=5,2)**, gamintojas *AppliChem GmbH*, DE
- ✓ **Oligonukleotidiniai pradmenys**, gamintojas *Metabion International AG*, DE
- ✓ **PGR mišinys, 2x**, angl. *Thermo Scientific Dream Taq PCR Master Mix (2x)*, gamintojas *Thermo Fisher Scientific*, LT
- ✓ **PGR reagentų rinkinys**, *Taq* polimerazė, MgCl<sub>2</sub>, 10x PGR buferis (+ (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), gamintojas *Thermo Fisher Scientific*, LT
- ✓ **TE buferis** (10 mM Tris-HCl, 1 mM EDTA, pH=8), gamintojas *AppliChem GmbH*, DE
- ✓ **Terminacijos mišinys**, angl. *BigDye® Terminator v3.1 Cycle Sequencing Kit*, gamintojas *Applied Biosystems*, US

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

- ✓ **96 šulinėlių plokštėlės**, angl. *MicroAmp™ Optical 96-Well Reaction Plate*, gamintojas *Applied Biosystems*, SG
- ✓ **96 vietų dengiamoji guma sekoskaitos plokštelėms**, angl. *Plate Septa 96-Well*, gamintojas *Applied Biosystems*, JP
- ✓ **Centrifuga** 0,5 ir 1,5 mL mėgintuvėliams, modelis *Centrifuge 5424* (gamintojas *Eppendorf AG*, DE) arba modelis *EBA 12* (gamintojas *Hettich Zentrifugen*, DE) ir 96 šulinėlių plokštelėms modelis *Universal 32* (gamintojas *Hettich Zentrifugen*, DE)
- ✓ **Genetinis analizatorius 3130xl, 16 kapiliarų**, angl. *3130xl Genetic Analyzer*, pagal *Applied Biosystems* užsakymą, gamintojas *Hitachi High-Technologies Corp.*, JP
- ✓ **Horizontali elektroforezės sistema**, angl. *Electrophoretic Gel System*, modelis *MIDICELL® PRIMO EC330*, gamintojas *E-C Apparatus Corp.*, US
- ✓ **Latekso pirštinės be talko, nitrilo pirštinės**
- ✓ **Mėgintuvėlių stovai**

- ✓ **Mikrobangų krosnelė**
- ✓ **Mikrocentrifuga** (0,2 mL, 0,5 mL, 1,5 mL), gamintojas *Carl Roth GmbH + Co.KG*, TW
- ✓ **Mikropipetės**, vienakanalės (0,1–2,5 µL, 0,5–10 µL, 2–20 µL, 10–100 µL, 100–1000 µL) ir daugiakanalė (0,5–10 µL)
- ✓ **Parafino juostelė**, angl. *Parafilm „M“<sup>®</sup> Laboratory Film*, gamintojas *Pechiney Plastic Packaging, Inc.*, US
- ✓ **Popierinės servetėlės**
- ✓ **Purtyklė**, modelis *Vortex-T Genie 2*, gamintojas *Scientific Industries, Inc.*, US
- ✓ **Sterilūs antgaliai** (10 µL, 20 µL, 100 µL)
- ✓ **Sterilūs mėgintuvėliai** (0,2 mL, 0,5 mL, 1,5 mL)
- ✓ **Šaldymo stovas**
- ✓ **Termocikleris**, modelis *Mastercycler, Mastercycler pro vapo.protect* (gamintojas *Eppendorf AG*, DE), modelis *DNA Engine TETRAD2* (gamintojas *MJ Research Inc.*, CA), modelis *UNO II* (pagal *BIOMETRA GmbH*, DE užsakymą gamintojas *Biotron*)
- ✓ **UV transiliuminatorius**, modelis *UVT-28 ME*, su vaizdų analizės sistema ir programine įranga *EasyWin 32*, gamintojas *Herolab GmbH*, DE

## **Darbo eiga**

### **Polimerazės grandininė reakcija (PGR)**

Koduojančiai ir aplinkinėms sekoms nustatyti gausinami *RTN4* ir *FBXL17* amplikonai. Oligonukleotidinių pradmenų sekos ir PGR produkto ilgis nurodyti **3.1 lentelėje**. Sukurtiems pradmenims nustatomos optimalios reakcijos sąlygos tam, kad būtų gautas specifinis kiekvieno iš gausinamų fragmentų produktas. Kiekvienas etapas vykdomas steriliai. Eksperimentiškai nustačius reakcijos sąlygas, PGR vykdoma pagal toliau pateiktą protokolą.

**3.1 lentelė. *RTN4* ir *FBXL17* genų amplikonų oligonukleotidiniai pradmenys PGR ir fragmentų ilgiai**

Amplikonas	Pradmens seka	Fragmento ilgis, bp
<b><i>RTN4</i> geno</b>		
1	RTN4_1F_5'-GGAGCCGTCACCACAGTA-3' RTN4_1R_5'-GGCTGCACACAAAAGAGG-3'	761
2	RTN4_2F_5'-GATCTTGCCCTTTCTCTGT-3' RTN4_2R_5'-TCGGAACCATGCTAATTTTT-3'	300
3.1	RTN4_3.1F_5'-TTGTCTTTGCACAGCATGTAA-3' RTN4_3.1R_5'-TGAACGATGATCCCATTTCT-3'	536
3.2	RTN4_3.2F_5'-GGAACACTTCAAGAAAATGTCAG-3' RTN4_3.2R_5'-ATCCTTTATACCTTCTGGCGTA-3'	582
3.3	RTN4_3.3F_5'-AGCAACTTGGAAAGTAAAGTGG-3' RTN4_3.3R_5'-GTCAGGCAAAACTGGTGAAG-3'	600
3.4	RTN4_3.4F_5'-TGGACTTGGTTCAAACATCA-3' RTN4_3.4R_5'-GTGAGACTTTCTTTCACAAGCA-3'	586
3.5	RTN4_3.5F_5'-AAGATTCTCACCTGATTCTGA-3' RTN4_3.5R_5'-AACTGATTTTCTCTTCAACTTTGG-3'	594
3.6	RTN4_3.6F_5'-CCCACAAAAGTGAAATTGCT-3' RTN4_3.6R_5'-AAATCCAAAAGTGGCATGAA-3'	379
4	RTN4_4F_5'-AGAGTGGGTTTAAAATGTGGGT-3' RTN4_4R_5'-CCAATCCTGTTTACACTATTGCC-3'	432
5	RTN4_5F_5'-GACATTTAGGGGAAAATCACTCC-3' RTN4_5R_5'-AGGGAATAGCACTTCTTCATCCT-3'	378
6	RTN4_6F_5'-ACCCAGTCTTCAAGTTGTATGC-3' RTN4_6R_5'-GCCCACATTTTCATCCCCTAC-3'	308
7-8	RTN4_7-8F_5'-AACTGCAGCCTCACACAGTT-3' RTN4_7-8R_5'-CAATCCAGCACACCTTATCC-3'	556
9	RTN4_9F_5'-TCTAATTCATGAGGCCAAGGTG-3' RTN4_9R_5'-CGTCAAGGTTTCGTTCTTCCC-3'	298
<b><i>FBXL17</i> geno</b>		
1.1	FBXL17_1.1F_5'-GTGTGTGTGCGCAGGAAG-3' FBXL17_1.1R_5'-AAGCTGGCCAGGAAGAGACT-3'	619
1.2	FBXL17_1.2F_5'-CTCCTGCTGCAAAGAGTTGG-3' FBXL17_1.2R_5'-GCTTAGGGGGAGGAGAGAGA-3'	667
2	FBXL17_2F_5'-GGCCTCTTTTATTCAAGAAGGTT-3' FBXL17_2R_5'-GGTCATATGAGTGTTACTTAGAATTGG-3'	281
3	FBXL17_3F_5'-TGTTAAGACGGTTTCTCTTTTGA-3' FBXL17_3R_5'-TGGGAACATTTTTAACGTAACCA-3'	498
4	FBXL17_4F_5'-TGTTTTAAGTTATGGATTGCTTACATT-3' FBXL17_4R_5'-TTCGAAGGTTAGGAACAAAATGA-3'	344
5	FBXL17_5F_5'-AGTGGATGTAGGGTCAAGAGA-3' FBXL17_5R_5'-TGGTCTGACTGTCCCTTATTCA-3'	399
6	FBXL17_6F_5'-TGTTGCCATTTCAGATATTCACCA-3' FBXL17_6R_5'-CAATTTGAAAAGGCACAGCTGT-3'	469
7	FBXL17_7F_5'-CCATTAAGAAGCTGTGGCATT-3' FBXL17_7R_5'-GAAAGGATTAGGAAAATGGTTCA-3'	219
8	FBXL17_8F_5'-CGTTCATTTACACACAAAATTTCA-3' FBXL17_8R_5'-TGTAGAAAGCAAACAACCTTGATAGTC-3'	300
9	FBXL17_9F_5'-AACAAGGCGAGGGCTTT-3' FBXL17_9R_5'-TCTGCTGAATGATCCCAGTG-3'	297

1. Apskaičiuojama, kiek reikės PGR mišinio (**3.2. lentelė**) esamam mėginių skaičiui, t. y. gausinamų DNR mėginių skaičius + neigiama kontrolė (dejonizuotas vanduo).

**3.2 lentelė.** PGR mišinys vienam mėginiui atitinkamiems *RTN4* ir *FBXL17* amplikonams

Reagentas	<i>RTN4</i> , 1 mėginiui, $\mu\text{L}$	<i>FBXL17</i> , 1 mėginiui, $\mu\text{L}$		
<b>Amplikonas</b>	<b>2-9</b>	<b>2</b>	<b>3,4</b>	<b>5-9</b>
ddH <sub>2</sub> O	4,75	15,4	9,5	14,4
PGR reakcijos mišinys	6,25	-	12,5	-
gDNR matrica (50–500 $\mu\text{g}/\text{mL}$ )	1	1	1	1
Pradmuo A+B (10 pmol/ $\mu\text{L}$ )	1	2	2	2
DMSO	-	-	-	1
dNTP (2M)	-	2,5	-	2,5
10x PGR buferis +(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	-	2,5	-	2,5
MgCl <sub>2</sub>	-	1,5	-	1,5
<i>Taq</i> DNR polimerazė	-	0,1	-	0,1
<b>Bendras reakcijos tūris, <math>\mu\text{L}</math></b>	<b>13</b>	<b>25</b>	<b>25</b>	<b>25</b>

2. PGR mišinys ruošiamas steriliai, šaldančiajame stovelyje. PGR mišinys išpilstomas į paruoštus 0,2 mL talpos mėgintuvėlius. Į kiekvieną mėgintuvėlį su mišiniu įpilama 1  $\mu\text{L}$  gDNR tirpalo. Bendras reakcijos tūris 13 (arba 25)  $\mu\text{L}$ .
3. Mėgintuvėliai sudedami į termociklerį, kur pagal atitinkamas sąlygas (**3.3 lentelė**) vykdoma PGR reakcija. Po reakcijos mėginiai laikomi +4 °C temperatūroje. Atlikus PGR, reakcijos produktas tikrinamas vykdant elektroforezę 1,5 % agarozės gelyje, siekiant įvertinti jo specifiškumą, kiekį ir ilgį.



### 3.3 lentelė. *RTN4* ir *FBXL17* amplikonų PGR sąlygos

Etapas	Temperatūra, °C		Laikas		Ciklų skaičius	
	<i>RTN4</i>	<i>FBXL17</i>	<i>RTN4</i>	<i>FBXL17</i>	<i>RTN4</i>	<i>FBXL17</i>
Pradinė deantūracija	95	95	3 min	10 min	1	1
Denatūracija	95	95	30 s	1 min	30	30
Pradmens prijungimas	55	X*	30 s	1 min		
Sintezė	72	72	45 s	1 min		
Papildoma sintezė	72	72	7 min	7 min	1	1
Pabaiga	4	4	→∞	→∞	1	1

\* – temperatūra priklausoma nuo pradmenų poros: 2 amplikonui – 60 °C; 3, 4, 8 amplikonams – 58 °C; 5, 6, 7, 9 amplikonams – 55 °C.

#### Elektroforezė 1,5 % agarozės gelyje

Elektroforezė – tai metodas, pagrįstas skirtingu teigiamai ir neigiamai įkrautų mišinio komponentų judrumu tirpale, kuriuo leidžiama elektros srovė. Taikant šį metodą patikrinama, ar tiriami fragmentai po PGR pasigausino bei koks jų specifiškumas. Elektroforezė vykdoma pagal toliau pateiktą metodiką.

1. Paruošiama forma geliui, įdedamos šukutės platiems ar siauriems takeliams formuoti.
2. 1,5 % agarozės tirpalas paruošiamas agarozės miltelius užpylus 1xTBE buferiu. Agarozė tirpinama mikrobangų krosnelėje. Tirpalas atvėsinamas iki 60 °C temperatūros.
3. Į agarozės tirpalą įpilama 0,9 ar 1,8 μL etidžio bromido tirpalo (10 mg/mL).
4. Paruoštas agarozės tirpalas lėtai pilamas į sukonstruotą formą. Sustingti (polimerizuotis) paliekama kambario temperatūroje ~30 min.
5. Geliui sustingus, atsargiai ištraukiamos šukutės. Gelis perkeliamas į horizontalios elektroforezės sistemą, kuri pripildyta 1xTBE buferio.
6. Į gelio šulinėlius mikrodozatoriumi įpilamas PGR produktas ir 3x įvedimo dažas (santykiu 2:1). Kaip kontrolė leidžiamas DNR molekulinės masės ir dydžio standartas.

7. Elektroforezės sistema uždengiama, prijungiamas srovės šaltinis. Nustatoma tinkama įtampa (pagal gelį ir vonelės dydį). Elektroforezė vyksta ~30 min.
8. Pasibaigus elektroforezei, gelis vizualizuojamas ultravioletinėje šviesoje (302 nm), UV transiliumiatoriuje. Gelis analizuojamas *EasyWin 32* gelių analizavimo programa, fotografuojamas. Gaunama elektroforegrama.

Nustačius specifiskus PGR produktus, toliau atliekama jų sekoskaita.

## Sekoskaita

### *1 etapas. PGR produkto valymas Exo-SAP metodu*

1. Apskaičiuojama, kiek reikės Exo-SAP (Egzonukleazė I–Krevečių šarminė fosfatazė) mišinio esamam mėginių skaičiui. Mišinys ruošiamas pagal **3.4 lentelėje** vienam mėginiui nurodytus reakcijai reikalingų medžiagų tūrius.

**3.4 lentelė.** Exo-SAP mišinys vieno mėginio valymui

<b>Exo-SAP mišinys</b>	<b>Tūris, <math>\mu\text{L}</math></b>
TE buferis	1,80
SAP (1U/ $\mu\text{L}$ )	0,15
ExoI (20U/ $\mu\text{L}$ )	0,05
<b>Bendras mišinio tūris</b>	<b>2,00</b>

2. Exo-SAP mišinys išpilstomas po 2  $\mu\text{L}$  į paruoštus 0,2 mL mėgintuvėlius. Į kiekvieną mėgintuvėlį su mišiniu įdedama po 5  $\mu\text{L}$  PGR produkto. Bendras reakcijos tūris 7  $\mu\text{L}$ .
3. Mėgintuvėliai dedami į termociklerį, kur reakcija vyksta pagal **3.5 lentelėje** nurodytą programą. Mėginiai laikomi +4 °C temperatūroje.

**3.5 lentelė.** PGR produkto valymo Exo-SAP mišiniu reakcijos sąlygos

Valymo Exo-SAP etapai	Temperatūra, °C	Trukmė
1. Inkubacija	37	40 min
2. ExoI-SAP inaktyvacija	80	20 min
3. Pabaiga	4	→ ∞

**2 etapas. Ciklinė sekoskaitos reakcija**

1. Apskaičiuojama, kiek reikės sekoskaitos reakcijos mišinio (**3.6 lentelė**) esamam mėginių skaičiui.

**3.6 lentelė.** Sekoskaitos reakcijos mišinys vienam mėginiui

Sekoskaitos reakcijos mišinys	Tūris, μL
ddH <sub>2</sub> O	3,15
5x Sekoskaitos buferis	1,40
Terminacijos mišinys	0,25
Pradmuo (A arba B) (10 pmol/μL)	0,20
<b>Bendras mišinio tūris</b>	<b>5,00</b>

2. Sekoskaitos reakcijos mišinys išpilstomas po 5 μL į paruoštus 0,5 mL mėgintuvėlius. Į kiekvieną mėgintuvėlį su mišiniu įdedama po 2 μL Exo-SAP produkto. Bendras reakcijos tūris 7 μL.
3. Mėgintuvėliai dedami į termociklerį, kur reakcija vyksta pagal **3.7 lentelėje** nurodytą programą. Mėginiai laikomi -20 °C temperatūroje.

**3.7 lentelė.** Sekoskaitos reakcijos sąlygos

Sekoskaitos reakcijos etapai	Temperatūra, °C	Trukmė	Ciklų skaičius
1. Pradinis denatūravimas	96	10 s	1
2. Denatūravimas	96	10 s	25
3. Pradmenų prijungimas	50	5 s	
4. Sintezė	60	4 min	
7. Pabaiga	4	→∞	1

### **3 etapas. Sekoskaitos reakcijos produkto valymas etanoliu ir natrio acetatu**

1. Apskaičiuojama, kiek reikės mišinio esamam mėginių skaičiui (3.8 lentelėje).

**3.8 lentelė.** Mišinys vieno mėginio valymui etanoliu

<b>Valymo etanoliu ir natrio acetatu mišinys</b>	<b>Tūris, µL</b>
Natrio acetatas 3M	3,0
96 % etanolis	62,5
ddH <sub>2</sub> O	14,5
<b>Bendras mišinio tūris</b>	<b>80,0</b>

2. Mišinys išpilstomas po 80 µL į mėgintuvėlius su sekoskaitos reakcijos produktu.
3. Mėgintuvėliai inkubuojami 10 min tamsoje 25 °C temperatūroje.
4. Po inkubacijos centrifuguojama 15 min 12000 aps/min.
5. Viršnuosėdinis mišinys nupilamas, mėgintuvėliai nusausinami beplaušėmis servetėlėmis. Pilama po 200 µL 70 % šalto etanolio. Mėgintuvėliai kruopščiai purtomi 1min.
6. Centrifuguojama 5 min 12000 aps/min.
7. Viršnuosėdinis mišinys nupilamas, mėgintuvėliai nusausinami beplaušėmis servetėlėmis, 5 min džiovinami 30 °C temperatūroje liofilizatoriuje. Iki frakcionavimo, vykdant automatizuotą kapiliarinę elektroforezę, mėginiai laikomi -20 °C temperatūroje.

### **4 etapas. Tiesioginės nukleotidų sekos nustatymas frakcionuojant automatizuotos kapiliarinės elektroforezės metodu**

1. Į kiekvieną mėgintuvėlį su liofilizuota DNR įdedama po 8 µL Hi-Di™ formamido. Kruopščiai pipetuojama ir perkeliama į sekoskaitos plokštelę. Centrifuguojama 1 min 1500 aps/min.
2. Mėginiai, esantys plokštelėje, denatūruojami 4 min 95 °C temperatūroje, atšaldomi ir dedami į genetinį analizatorių 3130xl Genetic Analyzer, prieš tai paspaudus prietaiso mygtuką TRAY.

3. *3130xl Data Collection v3.0* kompiuterinėje programoje sukuriamas sekoskaitos protokolas.
  - a. Plokštelės analizės protokolo ruošimo vadove (*Plate manager* skiltis), pasirinkus *New*, nurodoma:
    - i. Analizės pavadinimas;
    - ii. Analizės paskirtis (*Application*): *Sequencing Analysis*;
    - iii. Plokštelės tipas (*Plate type*): *96-Well*;
    - iv. Atlikimo vieta (*Owner*): *MGC*;
    - v. Tyrėjas (*Operator*): tyrėjo inicialai;
    - vi. Baigiama *OK* paspaudimu.
  - b. Atsidariusiame lange *Sequencing Analysis Plate Manager Editor*, nurodoma:
    - i. Mėginio pavadinimas (*Sample Name*);
    - ii. Rezultatų grupė (*Results group1*): *Ausra*;
    - iii. Instrumento operacijų protokolas (*Instrument Protocol 1*): *SEQ\_STD*;
    - iv. Analizės protokolas (*Analysis Protocol 1*): *3130POP7\_BDTv3-KB-Denovo\_v5.2*;
    - v. Kūrimas baigiamas *OK* paspaudimu.
  - c. Skiltyje *Run Scheduler* paspaudus *Find all* surandamas sukurtas plokštelės protokolas.
  - d. Sukurtas plokštelės protokolas susiejamas su analizatoriumi naudojant *Link* funkciją.
  - e. Paleidžiama elektroforezė *Run* paspaudimu.
4. Vykdoma automatinė kapiliarinė elektroforezė *POP-7<sup>TM</sup>* polimere (prietaiso protokolo modulis nurodytas **3.9 lentelėje**). Elektroforezės eigai naudojama *3130xl Data Collection v3.0* kompiuterinė programa. Elektroforegramai analizuoti naudojama *Sequencing analysis v5.2* kompiuterinė programa.

**3.9 lentelė.** 3130xl sekoskaitos prietaiso protokolo modulis POP-7<sup>TM</sup> polimerui, RapidSeq36\_POP7\_1

Nr.	Parametro pavadinimas	Reikšmė	Intervalas
1	Bloko temperatūra	60	int 18...65 °C
2	Kapiliarų užpildymo tūris	6500	6500...38000 žingsnių
3	Srovės stabilumas	5	int 0...2000 uA
4	Įtampa iki eigos	15	0...15 kV
5	Laikas iki eigos	180	1...1000 s
6	Injekcijos įtampa	1,2	1...15 kV
7	Injekcijos laikas	23	1...600 s
8	Įtampos žingsnių skaičius	20	1...100 nk
9	Įtampos žingsnio intervalas	15	1...60 s
10	Duomenų užlaikymo laikas	120	1...3600 s
11	Eigos įtampa	8,5	0...15 kV
12	Eigos laikas	2780	300...14 000 s

### **5 etapas. Sekoskaitos rezultatų interpretavimas**

Pirminį sekos nustatymą atlieka programa *Sequencing analysis* v5.2. Gaunamas grafinis spalvotas nukleotidų sekos vaizdas (pikų grafikas), kur kiekvienas nukleotidas žymimas tam tikra spalva: A – žalia, T – raudona, G – juoda, C – mėlyna. Kompiuterinė programa vieną piką vertina kaip vieną atitinkamą nukleotidą. Rezultatai interpretuojami atliekant sekoskaitos metodu nustatyto DNR fragmento nukleotidų sekos palyginimą su referentine tokio DNR fragmento seka, pateikiama *Ensembl* duomenų bazėje (GRCh37; *RTN4* geno seka ENSG00000115310, pradmenys kurti pagal DNR nuorašą RTN4-001 ENST00000337526; *FBXL17* geno seka ENSG00000145743, pradmenys kurti pagal DNR nuorašą FBXL17-201 ENST00000542267).

**Viengubų pikų kreivė** gaunama, kai: analizuojama nukleotidų seka yra tapati referentinei nukleotidų sekai; tiriamasis DNR fragmentas yra homozigotinis ar hemizigotinis pagal nukleotidų sekos pokytį. **Dvigubą piką turinti kreivė** gaunama, kai tiriamas DNR fragmentas yra heterozigotinis pagal pakaitą. **Neįskaitoma kreivė** (visi pikai dvigubi) 3' kryptimi nuo tam tikro piko gaunama, kai tiriamasis DNR fragmentas yra heterozigotinis pagal iškritą ar intarpą.

## IV PRIEDAS. Darbo su *GenomeStudio*<sup>™</sup> aprašas

### Tyrimo projekto sukūrimas

1. Atidaroma programa *GenomeStudio*<sup>™</sup>, *File* direktorijoje pasirenkama naujo projekto sukūrimo funkcija *New Project* ir pasirenkama kurti projektą genotipavimo modulyje *Genotyping*.
2. Atsidariusio lango *Projects Repository* laukelyje nurodoma, kur bus išsaugotas sukurtas projektas; *Project Name* laukelyje – norimas projekto pavadinimas; pasirenkama *Next*.
3. Pažymima *Use sample sheet to load sample intensities* ir pasirenkama *Next*.
4. Atsidariusio lango atitinkamame laukelyje nurodoma prašomos bylos vieta:
  - a. Mėginių sąrašo *Sample Sheet* (.csv plėtinys) direktorija;
  - b. Duomenų intensyvumo bylų direktorija *Data Repository*;
  - c. Lusto VNP identifikatorių bylos (.bpm plėtinys) direktorija *Manifest Repository*.
5. Pasirenkama *Next*.
6. Atsidariusiame lange varnele pažymima *Import cluster positions from a cluster file*.
  - a. Pažymėjus aktyvuojamas laukelis, kuriame nurodoma VNP klasterizavimui skirtos bylos *Cluster File* (.egt plėtinys) direktorija.
  - b. *Project Settings/Project Creation Actions* skiltyje pažymimos reikalingos funkcijos:
    - i. VNP klasterizavimo *Cluster SNPs*;
    - ii. Mėginių ir VNP statistikos skaičiavimo *Calculate Sample and SNP Statistics*;
    - iii. Paveldėjimo skaičiavimo (galimas esant giminingiems tiriamiesiems) *Calculate Heritability*;

- iv. Nustatoma parametro *Gen Call Threshold* slenkstinė vertė 0,15 (*Gen Call* įvertis nustato mėginio duomenų ir klasterio atitikimą).
  - c. Pasirenkama *Finish* ir pradedamas generuoti projektas.
7. Programai *GenomeStudio*<sup>™</sup> baigus darbą, sukuriamas projektas plėtiniu *.bsc*.

### **Mėginių, genotipuotų VNP ir genotipavimo proceso kokybės kontrolė**

Kokybės kontrolė sukurtame *GenomeStudio*<sup>™</sup> projekte atliekama keliais etapais. Visi toliau minimi kokybės vertinimo lygmenys ir naudojami kriterijai pateikiami **4.1 lentelėje**. Netenkinančius kriterijų mėginius reikia pašalinti iš sukurto projekto ir pakartoti eksperimentą, VNP – jei įmanoma, pakoreguoti rankiniu būdu, tačiau jei vis tiek netenkina kriterijų – pašalinti iš sukurto projekto. Kokybės kontrolės pagrindiniai etapai:

1. Vertinama gautų skenavimo vaizdų kokybė. Nustačius netinkamą kokybę, galima atlikti pakartotinį to paties lusto skenavimą arba pakartoti eksperimentą. Tačiau prieš kartojant eksperimentą verta atlikti toliau pateikiamas kokybės kontrolės procedūras.
2. Sukūrus genotipavimo projektą *GenomeStudio*<sup>™</sup> pirmiausia atliekamas mėginių kokybės įvertinimas mėginių lentelėje (*Samples table*).
3. Įvertinus mėginių kokybę, atliekamas genotipuotų VNP kokybės vertinimas VNP lentelėje (*SNP table*).
4. Vertinamas genotipavimo procesas, kurio rezultatai pateikiami grafiškai *GenomeStudio*<sup>™</sup>.

Pagal aptartus kokybės kontrolės etapus ir **4.1 lentelėje** nurodomus kriterijus šiam tyrimui naudotų duomenų kokybės vertinimas pateikiamas **4.2 lentelėje**.



**4.1 lentelė.** Mėginių, genotipuotų VNP ir genotipavimo proceso vertinamieji kokybės kontrolės kriterijai ir jų reikšmės

<b>KOKYBĖS VERTINIMAS</b>			
<b>MĖGINIŲ</b>			
<b>Kriterijus</b>	<b>Angl. terminas</b>	<b>Reikšmės</b>	<b>Reikšmių slenksčiai</b>
<b>Genotipų kokybės klasifikavimo santykis</b>	<i>Call* rate</i>	0–100	> 99 (kai kuriuose <i>Illumina</i> dokumentuose nurodoma ~98)
<b>95 A ir B alelio intensyvumo procentilis</b>	<i>p95Gm, p95Red</i>	Fluorescencijos intensyvumo vienetai	Svarbu, kad būtų daugiau už bazinį fluorescencijos lygį
<b>10 procentilis <i>GenCall</i> reikšmių</b>	<i>p10GC</i>	0–1	> 0,15; < 0,2 laikomi nepavykusiais; > 0,7 laikomi labai geros kokybės
<b>Paveldėjimo klaida</b>	<i>Heritability error</i>		Vertinama esant giminingiems asmenims. Esant klaidų patikrinti informaciją mėginių sąrašo byloje, nenustačius neatitikčių, pašalinti mėginį iš analizės
<b>Mėginių grafikas</b>	<i>Sample graph</i>	Grafinis mėginių genotipų vaizdavimas	Grafike įvertinami labiausiai atsiskyrę mėginiai
<b>GENOTIPUOTŲ VNP</b>			
<b>Klasifikavimo dažnis</b>	<i>Call frequency, Call Freq</i>	0–1	> 0,99, taikoma 1 % taisyklė
<b>Paveldėjimo klaida</b>	<i>Heritability error</i>	0→	Esant klaidų, peržiūrėti ir šalinti tuos VNP, kurie turi daug klaidų
<b>VNP klasterio kokybės įvertis</b>	<i>GenTrain score</i>	0–1	Kuo klasteriai labiau nutolę nuo idealių normalizuotų $\theta$ verčių (0; 0,5; 1), tuo <i>GenTrain</i> reikšmė yra arčiau nulio
<b>Heterozigotų perteklius</b>	<i>Het excess</i>	(-1)–1	0, pertekliaus nėra; vertinti (-0,3)–0,2 ribose, t. y. peržiūrėti tuos įrašus, kurie > 0,2, ir tuos, kurie < (-0,3)
<b>VNP genotipų klasterių atsiskyrimas</b>	<i>ClusterSep</i>	0–1	Vertinti esančius <0,4, visus <0,27 anuluoti
<b>AB klasterio centro <math>\theta</math> vertė**</b>	<i>AB T Mean</i>	0–1	0,5 (heterozigotos), t. y. siekiama vertė; peržiūrėti tuos, kurie yra intervaluose 0–0,2 ir 0,8–1
<b>AB klasterio centro R vertė**</b>	<i>AB R Mean</i>	0→	Netinkami yra žemo intensyvumo
<b>Retesnio alelio dažnis</b>	<i>Minor allele frequency, Minor Freq, MAF</i>	0–1	< 0,1 (peržiūrėti dėl apgaulingų homozygotų)
<b>X chr. vyrų grupėje</b>	-	-	Neturėtų būti heterozigotinių reikšmių
<b>GENOTIPAVIMO PROCESO</b>			
<b>NB! Žemiau, visur vertinamas fluorescencijos intensyvumas</b>			

4.1 lentelės tęsinys

Kriterijus	Angl. terminas	Reikšmės	Reikšmių slenksčiai
<b>Santykinis intensyvumas</b>	<i>Relative intensities evaluation</i>	Žemas < 150000 < aukštas	Daugeliu atvejų reikia žiūrėti konkrečiai, nes ir >5000 intensyvumas gali būti vertinamas kaip aukštas
<b>Dažymas</b>	<i>Staining Control</i>	Dažymo efektyvumas, kiekvieno dažo (raudono – DNP; žalio – Biotin) <b>HIGH</b> kontrolės turi būti aukšto intensyvumo	Intensyvumas: „DNP (High)“ kontrolės aukštas kitų trijų – žemas „Biotin (High)“ kontrolės aukštas; kitų trijų – žemas
<b>Nukleotidinės bazės pratęsimas</b>	<i>Extention Control</i>	Nukleotidinės bazės pratęsimas reakcijos efektyvumas, nukleotidų atitinkamo dažo fluorescencijos intensyvumas	Intensyvumas: A, T – aukštas G, C – žemas G, C – aukštas A, T – žemas
<b>Taikinio pašalinimas</b>	<i>Target Removal Control</i>	DNR matricos atjungimo efektyvumas po nukleotidinės bazės pratęsimas reakcijos, kontrolės tik žalio dažo sistemoje	Intensyvumas: Žemas Žemas
<b>Hibridizacija</b>	<i>Hybridization Control</i>	DNR hibridizacijos efektyvumas pagal trijų koncentracijų sintetines DNR, kontrolės tik žalio dažo sistemoje	Intensyvumas: Žemas Trys lygiai: Juodas – žemas Mėlynas – tarpinis Žalias – aukštas
<b>Atkūrimas</b>	<i>Restoration Control</i>	DNR „atkūrimo“ efektyvumas, kontrolės tik žalio dažo sistemoje	Intensyvumas: Žemas Žemas (aukštas tik tuo atveju, jei naudotas <i>Infinium HD FFPE Restore Kit</i> )
<b>Atitikimas, tikslumas</b>	<i>Stringency Control</i>	Hibridizacijai vertinti: <b>PM (Perfect Match)</b> – visiškas atitikimas; <b>MM (Mis-Match)</b> – neatitikimas; kontrolės tik raudono dažo sistemoje	Intensyvumas: PM – aukštas MM – žemas PM ir MM žemas
<b>Specifiškumas</b>	<i>Non-Specific Binding Control</i>	Mėginių kokybės ir rinkinio specifiškumo įvertinimas; žymenys sukurti bakterijų sekų pagrindu, todėl, jei rezultatas specifiškas, intensyvumas turi būti <b>žemas</b>	Intensyvumas: Žemas Žemas
<b>Polimorfiškumas</b>	<i>Non-Polymorphic Control</i>	Mėginių ir genotipavimo kokybės įvertinimas	Intensyvumas: A, T – aukštas G, C – žemas C, G – aukštas A, T – žemas

\*GenomeStudio™ atlieka genotipų calling arba kokybės klasifikavimo procedūrą, kuri yra pagrįsta Bajeso modelių (pgl. *Illumina GenCall Data Analysis Software Technote*), \*\*Vertinama normalizuotose polinėse koordinatėse, Θ – intensyvumo taško kampas nuo X-ašies, R – intensyvumas, spalvos žalia ir raudona nurodo aktualią fluorescencijos spalvinę sistemą.

**4.2 Ienetelė.** Tyrimui naudotų duomenų *GenomeStudio*<sup>™</sup> programa sukurto projekto kokybės vertinimas

<b>KOKYBĖS VERTINIMAS</b>		
<b>MĖGINIŲ</b>		
<b>Kriterijus</b>	<b>Angl. terminas</b>	<b>Pastabos</b>
<b>Genotipų kokybės santykis</b>	<i>Call rate</i>	Intervale 0,97-100; 0,97 priimtina, kadangi kiti kriterijai atitinka geros kokybės reikalavimus; visi mėginiai tinkami
<b>95 A ir B alelio intensyvumo procentilis</b>	<i>p95Gm, p95Red</i>	Raudonas 3874-9931; Žalias 3002-7751; kokybė tinkama
<b>10 procentilis <i>GenCall</i> reikšmių</b>	<i>p10GC</i>	> 0,7 (visi mėginiai tinkami); LTG-1075 netinkamas (0,4)
<b>Paveldėjimo klaida</b>	<i>P-P-C Error Rate</i>	Intervale 0-0,07; kokybė tinkama
<b>Mėginių grafikas</b>	<i>Sample graph</i>	Visi mėginiai tinkami; LTG-1075 netinkamas
<b>GENOTIPUOTŲ VNP</b>		
<b>Klasifikavimo dažnis</b>	<i>Call frequency, Call Freq</i>	Intervale 0,13-1; pagal 1 % taisyklę, visi, kurie < 0,9845, yra eliminuojami, pašalinta 10262 VNP
<b>Paveldėjimo klaida</b>	<i>Heritability error</i>	Rep = 0 (tinkama); P-C kl. = 0 (tinkama); P-P-C kl. yra, žiūrėti pagal poreikį
<b>VNP klasterio kokybės įvertis</b>	<i>GenTrain score</i>	Intervale 0,35-0,98; kokybė tinkama
<b>Heterozigotų perteklius</b>	<i>Het excess</i>	< (-0,3) yra 31 VNP, peržiūrėta, tinkami; > 0,2 yra 25 VNP, peržiūrėta, tinkami (vertinta atmetus X, XY, Y, 0 chromosomas, jos vertintos atskirai)
<b>VNP genotipų klasterių atsiskyrimas</b>	<i>ClusterSep</i>	Pašalinti < 0,27 arba įvertinus palikti kaip tinkami; 0,27 < x < 0,4 nustatyti 2128 atvejai, visi peržiūrėti, kai kurie koreguoti ir palikti kaip tinkami, kai kurie anuliuoti
<b>AB klasterio centro θ vertė</b>	<i>AB T Mean</i>	Peržiūrėta; < 0,3 yra 2209 VNP, tačiau < 0,2 yra 30, įvertinti kaip tinkami; > 0,7 yra 6722 VNP, tačiau > 0,8 nėra; kokybė tinkama
<b>AB klasterio centro R vertė</b>	<i>AB R Mean</i>	Intervale 0,16-4,85; < 0,145 anuliuoti
<b>Retesnio alelio dažnis</b>	<i>Minor Freq</i>	Autosomose yra 140839 VNP, kurių MAF < 0,1, tačiau mažiau negu 0,001 nėra
<b>X ir Y chr. vyrų grupėje</b>		Y chr. įvertinti 1372 žymenys, pakoreguoti visi VNP, kurie netinkami; X chr. žymenų yra 17814, iš jų 14036 yra tokie, kurių AB dažnis nelygus 0. Rekomenduojama analizei eliminuoti heterozigotines reikšmes. Yra 463 XY (PAR) VNP, kurių kokybė tinkama
<b>GENOTIPAVIMO PROCESO</b>		
<b>Dažymas</b>	<i>Staining Control</i>	Intensyvumas: > 3000 (visi mėginiai tinkami); LTG-1075 netinkamas
		> 2000 (visi mėginiai tinkami); LTG-1075 netinkamas
<b>Nukleotidinės bazės pratęsimas</b>	<i>Extention Control</i>	Intensyvumas: > 3000 (visi mėginiai tinkami); LTG-1075 netinkamas
		> 2000 (visi mėginiai tinkami); LTG-1075 netinkamas
<b>Taikinio pašalinimas</b>	<i>Target Removal Control</i>	Intensyvumas: Vertinama tik <i>green</i> sistema
		< 200 (visi mėginiai tinkami); LTG-1075 netinkamas

4.2 lentelės tęsinys		
Kriterijus	Angl. terminas	Pastabos
Hibridizacija	Hybridization Control	Intensyvumas:
		Vertinama tik <i>green</i> sistema
		J ~2000; M ~3000-7000; Ž ~ 5000-12000; (visi mėginiai tinkami); LTG-1075 netinkamas
Atkūrimas	Restoration Control	Intensyvumas:
		Vertinama tik <i>green</i> sistema
		< 500 (visi mėginiai tinkami); LTG-1075 netinkamas
Atitikimas, tikslumas	Stringency Control	Intensyvumas:
		PM > 6000; MM < 1000; visi mėginiai tinkami; LTG-1075 netinkamas
		Vertinama tik <i>red</i> sistema
Specifiškumas	Non-Specific Binding Control	Intensyvumas:
		< 500 (visi mėginiai tinkami); LTG-1075 netinkamas
		< 400 (visi mėginiai tinkami); LTG-1075 netinkamas
Polimorfiškumas	Non-Polymorphic Control	Intensyvumas:
		> 3000 (visi mėginiai tinkami); LTG-1388?; LTG-1075 netinkamas
		> 3000 (visi mėginiai tinkami); LTG-1075 netinkamas

Θ – intensyvumo taško kampas nuo X ašies, R – intensyvumas, spalvos žalia ir raudona nurodo aktualią fluorescencijos spalvinę sistemą.

## Įvesties bylų kūrimas trečių šalių programoms

Įvertinus tyrimo duomenų kokybę, pagal poreikį tolesnei analizei atlikti trečių šalių sukurtais programiniais įrankiais (angl. *3rd party tools*) yra kuriamos įvesties (angl. *input*) bylos. Genotipavimo rezultatų tolesnei analizei šiame darbe buvo naudotas viso genomo asociacijų analizės programinis atviros prieigos įrankių paketas *PLINK v1.07* (Purcell S et al., 2007; prieiga internete <http://pngu.mgh.harvard.edu/purcell/plink/>).

Įvesties bylos kuriamos *GenomeStudio*<sup>™</sup> tyrimo duomenų projekte.

1. *GenomeStudio*<sup>™</sup> tyrimo duomenų projekte psirenkama *Analysis* → *Report Wizard* → *Custom Report* → *PLINK Input Report 2.1.3 by Illumina, Inc. From Illumina, Inc.*
2. Pasirenkama įvesties bylos parametrai *Report Input Parameters*:
  - a. Analizės tipas *Analysis Type: Association*;

- b. Komandų vykdomasis kelias *ExePath*: *c:\plink\plink-1.04-dos\plink.exe*;
  - c. Genealogijos byla *PedigreeFile*: ”-,,;
  - d. Fenotipinių požymių byla *TraitDataFile*: ”-,,;
  - e. Naudojama Illumina DNR grandinės kodavimo sistema *UseForwardStrand: False*.
3. Pasirenkama *Next*.
  4. Pažymima *Selected samples* ir pasirenkama *Next*.
  5. Pažymima *Remove zeroed SNPs from the report* ir pasirenkama *Next*.
  6. Atsiradusio lango laukelyje *Output Path* nurodoma įvesties bylos saugojimo direktorija, laukelyje *Report Name* – įvesties bylos pavadinimas. Pasirenkama *Finish*.

*GenomeStudio*<sup>™</sup> sukuria 5 bylas, iš kurių pagrindinės yra *.map* ir *.ped* plėtiniai. Aprašytuoju principu yra sukuriamos įvesties bylos visoms pageidaujamų tiriamųjų grupių imtims.

## **APIE AUTORE**

**Asmeniniai duomenys:** Ingrida Domarkienė, gimusi 1985 03 01.

**Kontaktinė informacija:** Vilniaus universiteto Medicinos fakulteto Žmogaus ir medicininės genetikos katedra, Santariškių g. 2, LT-08661 Vilnius.

**El. pašto adresas:** [ingrida.domarkiene@mf.vu.lt](mailto:ingrida.domarkiene@mf.vu.lt)

### **Išsilavinimas:**

1992–2004 m. Vilniaus miesto savivaldybės Grigiškių „Šviesos“ vidurinė mokykla (B Nr. 078252)

2004–2008 m. Vilniaus universitetas, Gamtos mokslų fakultetas, molekulinės biologijos studijų programa, biologijos bakalauro kvalifikacinis laipsnis (BA Nr. 0714685)

2008–2010 m. Vilniaus universitetas, Gamtos mokslų fakultetas, genetikos studijų programa, biologijos magistro kvalifikacinis laipsnis (GM Nr. 0640084)

2010–2014 m. Vilniaus universitetas, Medicinos fakultetas, Žmogaus ir medicininė genetikos katedra, doktorantė

### **Darbinė patirtis**

2008–2010 m. VU MF Žmogaus ir medicininės genetikos katedra, pareigos – laborantė;

Nuo 2010 m. VU MF Žmogaus ir medicininės genetikos katedra, pareigos – jaunesnioji mokslo darbuotoja; VŠĮ VUL SK Medicininės genetikos centras Molekulinės genetikos laboratorija, pareigos – medicinos genetikė.

### **Narystė profesinėse organizacijose:**

Europos Žmogaus genetikos draugijos narė

Lietuvos Žmogaus genetikos draugijos narė

**Tobulinimosi kursai, seminarai, konferencijos, stažuotės:**

2009 05 29 – 2009 06 01 Kursai „ESHG/LSHG courses for laboratory medical geneticists: Translating genomics into the clinics“, Vilnius, Lietuva

2010 08 22 – 2010 08 26 Kursai „The 1st OpenGENE Young Investigator Workshop in Baltic Region: GWAS: from genotyping to sequencing“, Tartu, Estija

2011 03 03 „Thermo Fisher Scientific“ seminaras „Mokslo pasaulis – Lietuva“, Vilnius, Lietuva

2011 04 04 – 2011 04 08 Mokslinė išvyka į Tartu universiteto Molekulinės ir ląstelės biologijos institutą Estijoje

2011 05 28 – 2011 05 31 Konferencija „European Human Genetics Conference 2011“, Amsterdamas, Nyderlandai

2011 09 22 Praktinė konferencija „Naujosios genomikos technologijos klinikinėje genetikoje“, Vilnius, Lietuva

2012 02 29 – 2012 03 01 Kursai „5500 LifeScope Bioinformatics training“, Darmstadt, Vokietija

2012 05 10 – 2012 05 12 11-asis Baltijos šalių laboratorinės medicinos kongresas, Vilnius, Lietuva

2012 06 12 – 2012 06 15 Tarptautinė konferencija „Evoliucinė medicina: nauji senųjų problemų sprendimai“, Vilnius, Lietuva

2012 10 05 Pirmoji tarptautinė Baltijos paveldimo vėžio konferencija, Vilnius, Lietuva

2012 12 03 – 2012 12 05 Simpoziumas „New Frontiers in Personal Genomics“ Neimegene, Nyderlanduose

2013 06 26 – 2013 06 27 Kursai „DHPLC techniques & Transgenomic® Navigator software“, Vilnius, Lietuva

2013 06 08 – 2013 06 11 Konferencija „European Human Genetics Conference 2013“, Paryžius, Prancūzija

2013 09 17 Seminaras „Mokslo tyrimai, žiniasklaida ir savo pozicijos išreiškimas“, Vilnius, Lietuva

2013 09 20 Mokslinė-praktinė konferencija „Intelektinės negalios genomo tyrimai. Misija įmanoma?“

2013 10 07 Thermo Fisher Scientific seminaras „Mokslo dienos 2013“, Vilnius, Lietuva

2013 10 11 – 2013 10 12 Šeštasis Baltijos šalių aterosklerozės kongresas, Ryga, Latvija

2013 11 14 Konferencija „Nacionalinių veiklos, susijusios su retomis ligomis, aktualijos, planų ir strategijų įgyvendinimo aktualijos“, Vilnius, Lietuva

2013 12 16 Sveikatos informatikos ir bioinformatikos asociacijos (SVIBILT) konferencija „Duomenys – XXI amžiaus žaliava. Ar turime kuo kasti?“

2014 02 28 Konferencija „Retųjų ligų diena Lietuvoje 2014“, Vilnius, Lietuva

2014 03 21 Konferencija „Pasaulinė Dauno sindromo diena Lietuvoje“, Vilnius, Lietuva

2014 03 27 – 2014 03 28 Konferencija „Baltic metabolic group meeting: inherited glycogen storage disorders and disorders of purines and pyrimidines“, Vilnius, Lietuva



2014 03 06 – 2014 04 17 Kursai „Biostatistika su R paketu“,  
Vilnius, Lietuva

2014 06 03 – 2014 06 06 „Ion Torrent system and Next  
Generation Sequencing applications“ apmokymai, Vilnius,  
Lietuva

2014 09 08 – 2014 09 09 „Ion Torrent World Tour 2014“ *Ion  
Torrent*<sup>TM</sup> vartotojų konferencija, Kopenhaga, Danija