

VILNIAUS UNIVERSITETAS

Evelina  
SIAVRIENĖ

Koduojančios ir nekoduojančios  
genomo sekos variantų bei kopijų  
skaičiaus pokyčių molekulinis ir  
funkcinis vertinimas

**DAKTARO DISERTACIJA**

Medicinos ir sveikatos mokslai,  
Medicina (M 001)

---

VILNIUS 2020

Disertacija rengta 2016–2020 metais Vilniaus universiteto Medicinos fakulteto Biomedicinos mokslų instituto Žmogaus ir medicininės genetikos katedroje. Mokslinius tyrimus rėmė Lietuvos mokslo taryba (Nr. S-MIP-17-19/LSS-150000-1179; INGENES projektas). 2019 metais gauta stipendija už akademinį pasiekimą. 2020 metais gauta stipendija už studijų rezultatus.

**Mokslinis vadovas:**

**akad. prof. habil. dr. Vaidutis Kučinskas** (Vilniaus universitetas, medicinos ir sveikatos mokslai, medicina – M 001).

**Mokslinė konsultantė:**

**doc. dr. Eglė Preikšaitienė** (Vilniaus universitetas, medicinos ir sveikatos mokslai, medicina – M 001).

## TURINYS

SANTRUMPŲ SĄRAŠAS .....	6
1. ĮVADAS.....	17
1.1. Tiriamoji problema ir jos aktualumas .....	17
1.2. Tyrimo tikslas .....	18
1.3. Tyrimo uždaviniai.....	18
1.4. Tyrimo naujumas ir reikšmė .....	19
1.5. Ginamieji teiginiai .....	20
1.6. Mokslinio darbo aprobacija .....	20
2. MOKSLINĖS LITERATŪROS APŽVALGA .....	24
2.1. Žmogaus koduojančios ir nekoduojančios genomo sekos variantai ir kopijų skaičiaus pokyčiai .....	24
2.1.1. Patogeniniai DNR sekos variantai ir kopijų skaičiaus pokyčiai .....	27
2.1.1.1. Sukirpimo vietos variantai ir jų pasekmės informacinės RNR lygmeniu.....	29
2.1.2. Lietuvos populiacijos genomo įvairovės tyrimai.....	35
2.2. DNR sekos variantų ir kopijų skaičiaus pokyčių tyrimų strategijos .....	36
2.2.1. <i>In silico</i> analizė.....	39
2.2.2. Funkciniai tyrimo metodai .....	42
2.2.2.1. Informacinės RNR tyrimai .....	44
2.2.2.2. Ląstelių kultūrų ir modelinių organizmų panaudojimas funkciniais tyrimams.....	46
2.2.2.3. CRISPR-Cas9 genomo redagavimo technologija .....	49
2.3. Intelektinės negalios ir įgimtų anomalijų genetinė architektūra ir genomika.....	53
2.3.1. Intelektinės negalios samprata ir klasifikacija.....	53
2.3.2. Intelektinės negalios etiologija .....	55
2.3.3. Sindrominė intelektinė negalia .....	56
2.3.4. Intelektinės negalios ir įgimtų anomalijų tyrimo gairės	58
3. MOKSLINIO DARBO METODIKA .....	61
3.1. Tyrimo strategija ir tiriamieji.....	61
3.2. Tyrimo eiga.....	65
3.3. Tyrimo metodai.....	66

3.3.1.	Genominės DNR gryninimas, švarumo nustatymas bei mėginių ilgalaikis saugojimas .....	66
3.3.2.	Fibroblastų kultūros kultivavimas bei ilgalaikis saugojimas .....	66
3.3.3.	Visuminės RNR gryninimas .....	67
3.3.3.1.	Visuminės RNR gryninimas iš periferinio kraujo ėminių.....	67
3.3.3.2.	Visuminės RNR gryninimas iš fibroblastų kultūros ėminių.....	67
3.3.3.3.	Visuminės RNR koncentracijos ir švarumo nustatymas bei mėginių ilgalaikis saugojimas ..	68
3.3.4.	Kopijinės DNR sintezė .....	68
3.3.5.	Kiekybinė tikro laiko polimerazės grandininė reakcija .	69
3.3.6.	Sekoskaita Sangerio metodu.....	72
3.3.7.	Naujos kartos sekoskaita .....	75
3.3.8.	Fibroblastų kultūros genomo redagavimas CRISPR-Cas9 technologija .....	76
3.3.8.1.	Fibroblastų kultūrų funkciniai tyrimai .....	77
3.3.9.	Statistinė analizė.....	78
3.3.10.	Bioinformacinė duomenų analizė .....	79
4.	TYRIMŲ REZULTATAI IR JŲ APTARIMAS .....	80
4.1.	Nekoduojančios DNR sekos sukirpimo variantų molekulinį ir funkcinių tyrimų rezultatai ir jų aptarimas.....	82
4.1.1.	<i>ARID1B</i> c.5025+2T>C varianto tyrimas .....	84
4.1.2.	<i>GLI3</i> c. 473+3A>T varianto tyrimas .....	87
4.1.3.	<i>SLC9A6</i> c.899+1G>A varianto tyrimas.....	93
4.1.4.	<i>CHD7</i> c.5535-1G>A varianto tyrimas.....	97
4.1.5.	<i>TGFBR2</i> c.1600-2A>G varianto tyrimas.....	101
4.1.6.	<i>BLM</i> c.2308-2A>G varianto tyrimas .....	106
4.1.7.	<i>CAPN3</i> c.1746-20C>G varianto tyrimas .....	110
4.2.	Kopijų skaičiaus pokyčių molekulinį ir funkcinių tyrimų rezultatai ir jų aptarimas .....	114
4.2.1.	Molekulinis, funkcinis ir klinikinis 13q31.3 duplikacijos apibūdinimas.....	115
4.2.2.	Molekulinis, funkcinis ir klinikinis 4q21.22 delecijos apibūdinimas.....	120
4.2.3.	Molekulinis, funkcinis ir klinikinis 12q24.21 delecijos apibūdinimas.....	129
4.3.	Rezultatų apibendrinimas ir praktinės rekomendacijos .....	145

IŠVADOS.....	154
LITERATŪROS SAŖAŠAS.....	156
PRIEDAI .....	180
PADĒKA.....	190

## SANTRUMPŲ SĄRAŠAS

-RT	– neigiama atvirkštinės transkriptazės kontrolė (angl. <i>Negative reverse transcriptase control</i> )
3'ASV	– introno 3' galo akceptorinė sukirpimo vieta
5'DSV	– introno 5' galo donorinė sukirpimo vieta
A	– adeninas
A/PP	– adenino ir polipirimidino turtingos sekos
AR	– aminorūgštis
AT-PGR	– atvirkštinės transkripcijos polimerazės grandininė reakcija
ATP	– adenzino 5'-trifosfatas (angl. <i>Adenosine 5'-Thiotriphosphate</i> )
BAF	– su Brahma susijęs veiksnys (angl. <i>Brahma-Associated Factor</i> )
BAM	– binarinio sekoskaitos duomenų formato failas (angl. <i>Binary file format</i> )
bp	– bazių pora (angl. <i>Base pair</i> )
BWA	– Burrowso–Wheelerio išlyginimo įrankis (angl. <i>Burrows-Wheeler Alignment tool</i> )
C	– citozinas
Cas9	– su CRISPR sistema sąveikaujanti nukleazė 9 (angl. <i>CRISPR associated protein 9 nuclease</i> )
CDK	– nuo ciklino priklausoma kinazė (angl. <i>Cyclin-dependent kinase</i> )
cGMP	– ciklinis guanozino monofosfatas (angl. <i>Cyclic Guanine Monophosphate</i> )
CHARGE sindromas	– akies koloboma, širdies ydos, choanų atrezija, raidos atsilikimas, genitalijų anomalijos, ausų anomalijos ir / ar kurtumas (angl. <i>Coloboma of the eye, heart anomaly, choanal atresia, retardation of mental and somatic developmen, genital abnormalities, ear abnormalities and/or deafness</i> )

CHERISH	– tarptautinio projekto „Protinio atsilikimo diagnostikos gerinimas vaikams iš Rytų Europos ir Centrinės Azijos taikant genetinį apibūdinimą ir bioinformatiką / statistiką“ akronimas (angl. <i>Central Asia through Genetic Characterisation and Bioinformatics/Statistics</i> )
ChIP	– chromatino imunoprecipitacija (angl. <i>Chromatin immunoprecipitation</i> )
ClinVar	– duomenų bazė, kaupianti duomenis apie variantų patogeniškumą (angl. <i>Reports of Clinical Significance of Variants</i> )
CRISPR-Cas9	– genomo redagavimo technologija (angl. <i>CRISPR – Clustered regularly interspaced short palindromic repeats; CRISPR-associated Cas protein 9</i> )
crRNR	– CRISPR sistemos ribonukleorūgštis (angl. <i>CRISPR-derived RNA</i> )
CSS	– Coffino–Siris sindromas (angl. <i>Coffin-Siris syndrome</i> )
Ct	– slenksčio ciklas (angl. <i>Threshold cycle</i> )
CTCF	– CCCTC-surišimo veiksnys (angl. <i>CCCTC-Binding Factor or II Zinc-finger Protein</i> )
dbSNP	– vieno nukleotido polimorfizmų duomenų bazė (angl. <i>Database of single nucleotide polymorphisms</i> )
dbVar	– genominių struktūrinių variantų duomenų bazė (angl. <i>Database of genomic structural variation</i> )
DECIPHER	– žmonių genotipų ir fenotipų duomenų bazė (angl. <i>Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources</i> )
DGV	– genominių variantų duomenų bazė (angl. <i>Database of Genomic Variants</i> )
DMSO	– dimetilsulfoksidas (angl. <i>Dimethyl Sulfoxide</i> )
DNR	– deoksiribonukleorūgštis
DSM V	– Psichikos sutrikimų diagnostikos ir statistikos vadovas, penktasis leidimas (angl. <i>Diagnostic and Statistical Manual of Mental Disorders, 5th Edition</i> )

EKL	– embrioninės kamieninės ląstelės
ELISA	– imunofermentinis baltymų tyrimo metodas (angl. <i>Enzyme-linked immunosorbent assay</i> )
eRF	– egzonų jungties baltymų komplekso subvienetas (angl. <i>Eukaryotic peptide chain release factor</i> )
ESSL	– egzoninė sukirpimo proceso slopiklio vieta (angl. <i>Exonic splicing silencer</i> )
ESST	– egzoninė sukirpimo proceso stipriklio vieta (angl. <i>Exonic splicing enhancer</i> )
F	– tiesioginis oligonukleotidinis pradmuo (angl. <i>Forward primer</i> )
FAM	– fluoroforas 6-karboksifluoresceinas (angl. <i>6-carboxyfluorescein</i> )
FC	– genų raiškos pokytis, išreikštas kartais (angl. <i>Fold change</i> )
FISH	– fluorescencinė <i>in situ</i> hibridizacija (angl. <i>Fluorescence In Situ Hybridization</i> )
G	– guaninas
GATK	– geno analizės įrankis (angl. <i>Genome Analysis Toolkit</i> )
GCPS	– Greigo cefalopolisindaktilija (angl. <i>Greig Cephalopolysyndactyly syndrome</i> )
gDNR	– genomine deoksiribonukleorūgštis
gnomAD	– geno agregacijos duomenų bazė (angl. <i>The Genome Aggregation Database</i> )
gRNR	– gidinė ribonukleorūgštis
GWAS	– plataus masto viso geno sąsajų tyrimai (angl. <i>Genome-Wide Association Studies</i> )
HGMD	– Žmogaus genų variantų duomenų bazė (angl. <i>The Human Gene Mutation Database</i> )
ĮA	– įgimtos anomalijos
IGV	– Integruotas geno vaizdinimo įrankis (angl. <i>Integrative Genomics Viewer</i> )



IN	– intelektinė negalia
INGENES	– nacionalinio projekto „Intelektinės negalios ir įgimtų anomalijų genetinės architektūros identifikavimas“ akronimas (angl. <i>Deciphering the genetic architecture of intellectual disability and congenital anomalies</i> )
iPKL	– indukuotos pluripotentinės kamieninės ląstelės
IQ	– intelekto koeficientas (angl. <i>Intelligence quotient</i> )
iRNR	– informacinė ribonukleorūgštis
ISSL	– introninė sukirpimo proceso slopiklio vieta (angl. <i>Intronic splicing silencer</i> )
ISST	– introninė sukirpimo proceso stipriklio vieta (angl. <i>Intronic splicing enhancer</i> )
ĮŠY	– įgimta širdies yda
Kb	– kilobazė (1 000 bazių porų)
kDNR	– kopijinė deoksiribonukleorūgštis
KMI	– kūno masės indeksas
kPGR	– kiekybinė tikro laiko polimerazės grandininė reakcija
KSP	– DNR sekos kopijų skaičiaus pokyčiai
LDS	– Loeyso–Dietzo sindromas (angl. <i>Loeys-Dietz syndrome</i> )
LGH	– lyginamoji genomo hibridizacija (angl. <i>Comparative Genomic Hybridization</i> )
LGMD	– galūnių ir juosmens raumenų distrofijos (angl. <i>Limb girdle muscular dystrophies</i> )
LITGEN	– nacionalinio projekto „Lietuvos populiacijos genetinė įvairovė ir sandaros kitimai, susiję su evoliucija ir dažniausiai paplitusiomis ligomis“ akronimas (angl. <i>Genetic diversity of the population of Lithuania and changes of its genetic structure associated with evolution and common diseases</i> )
lncRNR	– ilgosios nekoduojančios ribonukleorūgštys (angl. <i>Long non-coding RNAs</i> )
MAF	– retojo alelio dažnis (angl. <i>Minor allele frequency</i> )

Mb	– megabazė (1 000 000 bazių porų)
MFDGA	– Guion-Almeida tipo mandibulofacialinė disostožė (angl. <i>Mandibulofacial dysostosis Guion-Almeida type syndrome</i> )
MGB	– DNR mažąjį griovį atpažįstanti sritis (angl. <i>minor groove binder</i> )
MIM	– pagal Mendelio dėsnius paveldimų ligų katalogas (angl. <i>Mendelian Inheritance in Man</i> )
MIQE	– Gairės tikro laiko PGR tyrimų rezultatų sklaidai (angl. <i>Minimum Information for Publication of Quantitative Real-Time PCR Experiments</i> )
miRNR	– mikroribonukleorūgštis (angl. <i>MicroRNA</i> ),
MPO	– morfolino-fosforodiamidatu modifikuoti priešprasniai oligonukleotidai (angl. <i>Morpholino phosphorodiamidate anti-sense oligonucleotides</i> )
NANOG	– <i>Nanog</i> homeodėžutė (angl. <i>Nanog Homeobox</i> )
NCBI	– Nacionalinio biotechnologijų informacijos centro duomenų bazė (angl. <i>National Center for Biotechnology Information</i> )
NFQ	– nefluorescencinis slopiklis (angl. <i>nonfluorescent quencher</i> )
NKS	– naujos kartos sekoskaita
nt	– nukleotidas
NTC	– neigiama DNR ar RNR matricos kontrolė (angl. <i>No template control</i> )
NVI	– Nacionalinis vėžio institutas
OCT4	– transkripcijos veiksnys 4, sąveikaujantis su oktameru (angl. <i>Octamer-Binding Transcription Factor 4</i> )
OMIM	– pagal Mendelio dėsnius paveldimų ligų internetinis katalogas (angl. <i>Online Mendelian Inheritance in Man</i> )
P300	– histono acetiltransferazė p300 (angl. <i>Histone Acetyltransferase p300</i> )
PAM	– greta protopacerio esantis motyvas (angl. <i>Protospacer-adjacent motif</i> )

PBS	– fosfatinis buferinis tirpalas (angl. <i>Phosphate Buffer Solution</i> )
PGR	– polimerazės grandininė reakcija (angl. <i>Polymerase chain reaction</i> )
PHS	– Pallisterio–Hall sindromas (angl. <i>Pallister-Hall syndrome</i> )
piRNR	– su PIWI baltymais sąveikaujančios ribonukleorūgštys (angl. <i>Piwi-interacting RNAs</i> )
PMID	– unikalus identifikacinis numeris, naudojamas <i>PubMed</i> duomenų bazėje
PolyPhen-2	– bioinformacinės analizės įrankis (angl. <i>Polymorphism Phenotyping v2</i> )
PPD4	– IV tipo preaksialinė polidaktilija (angl. <i>Preaxial polydactyly type IV</i> )
PRA	– psichomotorinės raidos atsilikimas
pre-iRNR	– informacinės ribonukleorūgšties pirmtakas (angl. <i>Precursor messenger RNA</i> )
PROGENET	– nacionalinio projekto „Patogeninių kartotinių sekų reikšmė protinio atsilikimo etiopatogenezeje“ akronimas (angl. <i>Evaluation of pathogenic copy number variation in etiopathology of mental retardation</i> )
PROVEAN	– baltymų variabilumo analizatorius (angl. <i>Protein Variation Effect Analyzer</i> )
PSO	– Pasaulio sveikatos organizacija
PTB	– polipirimidino trakto sukirpimo veiksnys (angl. <i>Polypyrimidine tract-binding protein</i> )
PTV	– baltymą trumpinantys variantai (angl. <i>Protein-truncating variant</i> )
PUBMED	– elektroninė medicinos mokslo publikacijų, gyvybės mokslų žurnalų, elektroninių knygų duomenų bazė (angl. <i>MEDLINE database of publications on life sciences and biomedical topics</i> )

R	– grįžtamasis oligonukleotidinis pradmuo (angl. <i>Reverse primer</i> )
RENT1	– egzonų jungties baltymų komplekso subvienetas (angl. <i>Regulator of nonsense transcripts 1</i> )
RNP	– ribonukleoproteinas
RNR	– ribonukleorūgštis
rRNR	– ribosominė ribonukleorūgštis (angl. <i>Ribosomal RNA</i> )
RT	– atvirkštinė transkriptazė (angl. <i>Reverse transcriptase</i> )
SHH	– <i>Sonic Hedgehog</i> signalinė molekulė (angl. <i>Sonic Hedgehog Signaling Molecule</i> )
SIFT	– bioinformacinės analizės įrankis (angl. <i>Sorting intolerant from tolerant</i> )
siRNR	– mažosios interferencinės ribonukleorūgštys (angl. <i>Small interfering RNAs</i> )
SMG1	– egzonų jungties baltymų komplekso subvienetas (angl. <i>Serine/threonine-protein kinase SMG1</i> )
SN	– standartinis nuokrypis (angl. <i>Standard deviation</i> )
snoRNR	– mažosios branduolėlio ribonukleorūgštys (angl. <i>Small nucleolar RNAs</i> )
snRNR	– mažosios branduolio ribonukleorūgštys (angl. <i>Small nuclear RNAs</i> )
SOX2	– Y sritis, nulemianti lytį (angl. <i>Sex Determining Region Y, Box 2</i> )
SRE	– sukirpimą reguliuojantys elementai (angl. <i>Splicing-regulatory elements</i> )
SRSF1	– serino / arginino turtingų sekų sukirpimo veiksnys (angl. <i>Serine/arginine-rich splicing factor 1</i> )
SVEIDRA	– Higienos instituto Privalomojo sveikatos draudimo informacinė sistema
SWI/SNF	– baltymų kompleksas, dalyvaujantis nuo ATP priklausomo chromatino struktūros pertvarkymo procese (angl. <i>SWItch/Sucrose Non-Fermentable</i> )

ŠKL	– širdies ir kraujagyslių sistemos ligos
T	– timinas
TALEN	– transkripcijos aktyvintojus primenančios efektorių nukleazės (angl. <i>transcription activator-like effector-based nucleases</i> )
TLK-10-AM	– Tarptautinė statistinė ligų ir sveikatos sutrikimų klasifikacija, dešimtas leidimas, Australijos modifikacija
TPM	– 1 000 000 RNR transkriptų viename mėginyje (angl. <i>Transcripts per million</i> )
tracrRNR	– transaktyvinanti ribonukleorūgštis (angl. <i>trans-activating crRNA</i> )
tRNR	– transportinė ribonukleorūgštis (angl. <i>Transfer RNA</i> )
U	– uracilas
U2AF	– U2 sukirpimo veiksnys (angl. <i>U2 auxiliary factor</i> )
UCSC	– Kalifornijos Santa Kruzo universitetas, JAV (angl. <i>University of California Santa Cruz</i> )
UNIGENE	– tarptautinio projekto „Unikalūs genomo persitvarkymai esant įgimtiems nervų sistemos raidos sutrikimams: kilmė, genominiai mechanizmai, funkcinės ir klinikinės pasekmės“ akronimas (angl. <i>Unique genome variants in congenital neurodevelopmental disorders: origin, genomic mechanisms, functional and clinical consequences</i> )
UTR	– netransliuojama DNR seka (angl. <i>Untranslated region</i> )
VES	– viso egzomo sekoskaita
VGS	– viso genomo sekoskaita
vLGH	– vektorinė lyginamoji genomo hibridizacija (angl. <i>Microarray-based Comparative Genomic Hybridization</i> )
VNP	– vieno nukleotido polimorfizmas (angl. <i>Single nucleotide polymorphism</i> )
VNP-LGH	– vieno nukleotido polimorfizmų lyginamoji genomo hibridizacija (angl. <i>Single Nucleotide Polymorphism–Comparative Genome Hybridization</i> )

- VšĮ VUL SK – Viešosios įstaigos Vilniaus universitetinės ligoninės  
MGC Santaros klinikų Medicininės genetikos centras
- VU GMC – Vilniaus universiteto Gyvybės mokslų centras
- VU MF BMI – Vilniaus universiteto Medicinos fakulteto Biomedicinos  
ŽMGK mokslų instituto Žmogaus ir medicininės genetikos katedra
- WGA – viso genomo pagausinimas (angl. *Whole genome amplification*)
- ZFN – cinko pirštelių nukleazės (angl. *Zinc-finger nucleases*)

## GENŲ SANTRUMPŲ SĄRAŠAS

<i>ACTB</i>	– angl. <i>Actin Beta</i>
<i>ARID1B</i>	– angl. <i>AT-Rich Interaction Domain-Containing Protein 1B</i>
<i>ATM</i>	– angl. <i>Ataxia-Telangiectasia Mutated</i>
<i>BCL2L11</i>	– angl. <i>B-Cell Lymphoma 2 Like Protein 11</i>
<i>BLM</i>	– angl. <i>Bloom Syndrome</i>
<i>C1NH</i>	– angl. <i>Complement Component 1 Inhibitor</i>
<i>CAPN3</i>	– angl. <i>Calpain 3</i>
<i>CARD9</i>	– angl. <i>Caspase Recruitment Domain-Containing Protein 9</i>
<i>CCNC</i>	– angl. <i>Cyclin C</i>
<i>CDKN1A</i>	– angl. <i>Cyclin Dependent Kinase Inhibitor 1A</i>
<i>CDKN1C</i>	– angl. <i>Cyclin Dependent Kinase Inhibitor 1C</i>
<i>CEP290</i>	– angl. <i>Centrosomal Protein, 290-KD</i>
<i>CFTR</i>	– angl. <i>Cystic Fibrosis Transmembrane Conductance Regulator</i>
<i>CHD7</i>	– angl. <i>Chromodomain Helicase DNA-Binding Protein 7</i>
<i>CLPB</i>	– angl. <i>Caseinolytic Peptidase B</i>
<i>COQ4</i>	– angl. <i>Coenzyme Q4</i>
<i>E2F1</i>	– angl. <i>E2F Transcription Factor 1</i>
<i>EFTUD2</i>	– angl. <i>Elongation Factor Tu GTP-binding Domain-containing 2</i>
<i>ENOPH1</i>	– angl. <i>Enolase-Phosphatase 1</i>
<i>FBN1</i>	– angl. <i>Fibrillin 1</i>
<i>FMR1</i>	– angl. <i>Fragile X Mental Retardation 1</i>
<i>GAPDH</i>	– angl. <i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>
<i>GLI3</i>	– angl. <i>Gli-kruppel Family Member 3</i>
<i>GPC5</i>	– angl. <i>Glypican 5</i>

<i>HNRNPD</i>	– angl. <i>Heterogeneous Nuclear Ribonucleoprotein D</i>
<i>HNRNPDL</i>	– angl. <i>Heterogeneous Nuclear Ribonucleoprotein D-Like Protein</i>
<i>HRPT2</i>	– angl. <i>Hyperparathyroidism 2</i>
<i>IL23R</i>	– angl. <i>Interleukin 23 Receptor</i>
<i>KITLG</i>	– angl. <i>Kit Ligand</i>
<i>KMT2D</i>	– angl. <i>Lysine-specific Methyltransferase 2D</i>
<i>LIN54</i>	– angl. <i>Lin-54 Homolog</i>
<i>MED13L</i>	– angl. <i>Mediator Complex Subunit 13-Like</i>
<i>MIR17HG</i>	– angl. <i>Micro RNA 17 Host Gene</i>
<i>NF1</i>	– angl. <i>Neurofibromin 1</i>
<i>NOTCH2</i>	– angl. <i>NOTCH Receptor 2</i>
<i>PRKG2</i>	– angl. <i>Protein Kinase, cGMP-Dependent, Type II</i>
<i>RASGEF1B</i>	– angl. <i>RASGEF Domain Family, Member 1B</i>
<i>RB1</i>	– angl. <i>RB Transcriptional Corepressor 1</i>
<i>RNF186</i>	– angl. <i>Ring Finger Protein 186</i>
<i>SCD5</i>	– angl. <i>Acyl-CoA Desaturase 4</i>
<i>SEC31A</i>	– angl. <i>SEC31 Homolog A, COPII Coat Complex Component</i>
<i>SEMA3E</i>	– angl. <i>Semaphorin-3E</i>
<i>SLC9A6</i>	– angl. <i>Solute Carrier Family 9, Member 6</i>
<i>TAF9B</i>	– angl. <i>TATA-Box Binding Protein Associated Factor 9b</i>
<i>TENM1</i>	– angl. <i>Teneurin Transmembrane Protein 1</i>
<i>TGFBR2</i>	– angl. <i>Transforming Growth Factor-Beta Receptor, Type II</i>
<i>THAP9</i>	– angl. <i>THAP Domain Containing 9</i>
<i>TMEM150C</i>	– angl. <i>Transmembrane Protein 150C</i>
<i>TTN</i>	– angl. <i>Titin</i>



# 1. ĮVADAS

## 1.1. Tiriamoji problema ir jos aktualumas

Žmogaus genomo projektas (angl. *Human Genome Project*) yra vienas žymiausių ir reikšmingiausių tarptautinių mokslinių tyrimų projektų istorijoje. 2003 m., pasibaigus šiam projektui, tarptautinis žmogaus genomo sekoskaitos konsorciumas paskelbė nuskaitytą žmogaus genomo seką ir nustatė, kad DNR seka žmonių populiacijoje yra apie 99,9 proc. vienoda, o likusią dalį sudaro skirtumus tarp žmonių lemiantys DNR sekos pokyčiai, kurių spektras varijuoja nuo taškinių variantų specifiniuose genuose iki didelės apimties chromosomų skaičiaus bei struktūros pokyčių [1].

Sparčiai tobulėjant žmogaus genomo tyrimo technologijoms, pavyzdžiui, lyginamoji genomo hibridizacija (LGH), viso egzomo sekoskaita (VES) ar viso genomo sekoskaita (VGS), genetinių pokyčių nustatymo efektyvumas gerokai išaugo. Naudojant šias pažangias tyrimo technologijas, DNR sekos variantai ir kopijų skaičiaus pokyčiai (KSP) yra nustatomi tiek koduojančioje, tiek nekoduojančioje žmogaus genomo srityje [2]. Nors anksčiau buvo manoma, kad genetiniai pokyčiai, esantys tik baltymus koduojančiose DNR srityse, yra svarbūs funkciniu požiūriu, vykdant *ENCODE* (angl. *Encyclopedia of DNA Elements*) projektą buvo nustatyta, kad daugiau nei 80 proc. netransliuojamos DNR sekos yra transkribuojama ir yra svarbi genų raiškos reguliavimui [3–5].

1 000 genomų projekto duomenimis, žmogaus genome iš viso nustatoma apie 20 000–22 000 variantų sinoniminėse ir nesinoniminėse sekose. Tačiau kliniškai reikšmingų yra tik dalis DNR sekos variantų: 190–210 insercijų bei delecijų, nelemiančių skaitymo rėmelio poslinkio, 220–250 delecijų, lemiančių skaitymo rėmelio poslinkį, 80–100 variantų, sudarančių priešlaikinį baigmės kodoną, bei 40–50 sukirpimo vietos variantų [6, 7]. Tačiau nustatomų variantų ir KSP pokyčių skaičius nuolat didėja. *ClinVar* [8] duomenimis, šiuo metu yra registruota 670 416 ligą lemiančių genetinių pokyčių, o iš jų daugiau nei 40 000 yra galimai susiję su intelektine negalia (IN) [9].

IN nustatoma 1–3 proc. bendros populiacijos ir yra viena svarbiausių socialinių ir ekonominių asmens sveikatos priežiūros problemų visame pasaulyje. IN, kuri dažnai pasireiškia kartu su įgimtomis anomalijomis (IA), būdingi įvairaus spektro ir sunkumo klinikiniai požymiai. IN / IA dažnai pasireiškia kliniškai neatskiriamais ar sunkiai atskiriamais vienas nuo kito fenotipais. Be to, IN / IA lemia įvairūs tiek genetiniai, tiek aplinkos veiksniai.

Dėl heterogeninės IN / ĮA prigimties šią patologiją lemiančių priežasčių nustatymas yra itin sudėtingas, o diagnostinis efektyvumas yra tik apie 50 proc. [10–13].

Lietuvoje, kaip ir visame pasaulyje, IN / ĮA yra aktuali asmens sveikatos priežiūros problema. Todėl nuo 2009 metų VU MF BMI ŽMGK vykdyti nacionaliniai ir tarptautiniai moksliniai projektai – LITGEN, CHERISH, PROGENET, UNIGENE, kurių metu buvo analizuojamas lietuvių, įskaitant asmenis, turinčius IN / ĮA, genomas. Šių mokslinių tyrimų metu, naudojant pažangiausias tyrimo technologijas, ištirta apie 400 asmenų, turinčių sindrominę ir nesindrominę IN, identifikuoti kandidatiniai genai ir nustatyti koduojančios ir nekoduojančios DNR sekos variantai ir KSP.

Nustatytų genetinių pokyčių patogeniškumas dažniausiai yra vertinamas atlikus išsamią *in silico* analizę. Visgi, remiantis *ClinVar* [8] statistiniais duomenimis, tik 12 proc. DNR sekos variantų ir KSP yra klasifikuojama patogeniniais ir 17 proc. nepatogeniniais, o tai reiškia, kad likusią dalį sudaro nežinomos ar iki šiol nepatvirtintos klinikinės reikšmės genetiniai pokyčiai [9]. Todėl dažnai bioinformacinės analizės neužtenka patogeniškumui pagrįsti. Be to, literatūros duomenimis, mažiausiai 10 proc. patvirtintų ligą lemiančių patogeninių variantų iš tiesų neturi jokie poveikio fenotipui [14]. Todėl siekiant įvertinti DNR sekos variantų ir KSP patogeniškumą reikalingi išsamesni genomo, transkriptomo, proteomo, metabolomo ar net interaktomo molekuliniai ir funkciniai tyrimai. Šiame moksliniame darbe IN / ĮA pasirinkta kaip paveldimos ligos modelis etiopatogenezei įvertinti, o šią patologiją lemiantys nežinomos ar iki šiol nepatvirtintos klinikinės reikšmės DNR sekos variantai ir KSP ištirti iRNR lygmeniu.

## 1.2. Tyrimo tikslas

Molekuliniiais ir funkciniais informacinės RNR tyrimo metodais įvertinti DNR sekos sukirpimo vietos variantų ir kopijų skaičiaus pokyčių patogeniškumą bei nustatyti galimą reikšmę intelektinės negalios ir / ar įgimtų anomalijų etiopatogenezei.

## 1.3. Tyrimo uždaviniai

1. Taikant bioinformacinius žmogaus genomo sekos analizės įrankius, išanalizuoti naujos kartos sekoskaitos tyrimų duomenis ir atrinkti nežinomos klinikinės reikšmės ar galimai patogeninius koduojančios

- bei nekoduojančios genomo sekos variantus ir kopijų skaičiaus pokyčius, kurie gali būti siejami su tiriamų asmenų intelektine negalia ir / ar įgimtomis anomalijomis.
2. Ištirti atrinktus DNR sekos variantus ir kopijų skaičiaus pokyčius informacinės RNR lygmeniu, taikant kiekybinės tikro laiko PGR ir Sangerio sekoskaitos metodą tiriamų asmenų kopijinės DNR mėginiuose.
  3. Atlikus mokslinės literatūros ir duomenų bazių analizę, įvertinti nustatytų DNR sekos variantų ir kopijų skaičiaus pokyčių patogeniškumą.
  4. Taikant bioinformacinius žmogaus genomo sekos analizės įrankius, numatyti DNR sekos variantų ir kopijų skaičiaus pokyčių padarinius baltymo lygmeniu.
  5. Patvirtinti DNR sekos pokyčio patogeniškumą, taikant CRISPR-Cas9 technologiją žmogaus fibroblastų kultūroje.
  6. Remiantis klinikiniais, molekuliniais ir funkciniais duomenimis, apibūdinti ryšį tarp genotipo ir fenotipo.

#### 1.4. Tyrimo naujumas ir reikšmė

Šiame moksliniame darbe, kuris yra vienas iš nedaugelio tokio pobūdžio darbų Lietuvoje, užsibrėžtam tikslui ir uždaviniams įgyvendinti, atrinkti nežinomos klinikinės reikšmės DNR sekos sukirpimo variantai ir KSP, kurie yra nauji ir iki šiol literatūroje neaprašyti. Tiriant kopijinę DNR įvertintas nustatytų neaiškios klinikinės reikšmės sukirpimo vietos variantų patogeniškumas ir apibūdintos jų *in silico* prognozuojamos pasekmės baltymui. Tiriant KSP, gauta informacija apie pokyčio pobūdį, dydį ir apimančius genus, kuri yra svarbi nustatant žinomus bei naujus delecijos ar duplikacijos sindromus. Tokie tyrimai taip pat padeda susiaurinti su žinomais sindromais susijusias genomines sritis ar net identifikuoti pagrindinius sindromą lemiančius genus. Šiame darbe nustatytas naujas *MIR17HG* duplikacijos sindromas ir susiaurinta genomine sritis, lemianti 4q21 delecijos sindromą. Taip pat *MED13L* genetinio pokyčio patogeniškumui patvirtinti pirmą kartą fibroblastų kultūroje pritaikyta CRISPR-Cas9 genomo redagavimo technologija.

Funkciniai genomo tyrimai, tokie kaip genų raiškos analizė bei genomo redagavimas, suteikia galimybę geriau suprasti įgimtų sutrikimų etiopatogenezės mechanizmus bei nustatyti ar patvirtinti diagnozę. Šios fundamentinės žinios kartu su molekulinės diagnostikos ir funkcinį tyrimų galimybių plėtra yra svarbios kuriant naujas diagnostikos ir gydymo

strategijas ateityje. Atliekami tyrimai turi tiesioginę naudą tiek tiriamiesiems ir jų šeimoms, tiek gydytojams, kadangi didžiąjai daliai tiriamųjų nustatyta molekulinė IN / IA diagnozė, o tai sudaro efektyvaus genetinio konsultavimo pagrindą.

### 1.5. Ginamieji teiginiai

1. Informacinės RNR tyrimai kiekybinės PGR ir Sangerio sekoskaitos metodu bei ląstelių kultūros genomo redagavimas CRISPR-Cas9 technologija yra efektyvūs nežinomos klinikinės reikšmės genetinių pokyčių patogeniškumui tirti.
2. 5' donorinės sukirpimo vietos sekos variantai dažniausiai lemia egzono, esančio prieš patogeninį variantą, iškritą, o sukirpimo sekos variantai 3' akceptorinėje vietoje lemia egzono, esančio už patogeninio varianto, iškritą.
3. Patogeniniai *MED13L* variantai heterozigotinėje būklėje turi įtakos fibroblastų kultūros gyvybingumui, senėjimui ir ląstelės ciklą valdančių genų raiškai.
4. Klinikinis, molekulinis ir funkcinis kopijų skaičiaus pokyčių charakterizavimas yra kompleksinis būdas tirti ligą lemiančias genomines sritis bei naujus sindromus.

### 1.6. Mokslinio darbo aprobacija

Darbo rezultatai paskelbti Lietuvos ir užsienio leidiniuose, pristatyti nacionalinėse ir tarptautinėse mokslinėse konferencijose.

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

1. **Siavrienė E.**, Preikšaitienė E., Maldžienė Ž., Mikštienė V., Rančelis T., Ambrozaitytė L., Gueneau L., Reymond A., Kučinskas V. *A de novo 13q31.3 microduplication encompassing the miR-17 ~ 92 cluster results in features mirroring those associated with Feingold syndrome* 2. *Gene*. 2020 Aug 30;753:144816. IF=2.319.
2. Pranckėnienė L., **Siavrienė E.**, Gueneau L., Preikšaitienė E., Mikštienė V., Reymond A., Kučinskas V. *De novo splice site variant of ARID1B associated with pathogenesis of Coffin-Siris syndrome*.

- Mol Genet Genomic Med.* 2019 Oct 19:e1006. doi: 10.1002/mgg3.1006. IF=2.448.
3. **Siavrienė E.**, Mikštienė V., Radzevičius D., Maldžienė Ž., Rančelis T., Petraitytė G., Tamulytė G., Kavaliauskienė I., Šarkinas L., Utkus A., Kučinskas V., Preikšaitienė E. *Novel GLI3 variant causes Greig cephalopolysyndactyly syndrome in three generations of a Lithuanian family. Mol Genet Genomic Med.* 2019 Jul 20:e878. doi: 10.1002/mgg3.878. IF=2.448.
  4. **Siavrienė E.**, Petraitytė G., Mikštienė V., Rančelis T., Maldžienė Ž., Morkūnienė A., Byčkova J., Utkus A., Kučinskas V., Preikšaitienė E. *A novel CHD7 variant disrupting acceptor splice site in a patient with mild features of CHARGE syndrome: a case report. BMC Med Genet.* 2019 Jul 17;20(1):127. doi: 10.1186/s12881-019-0859-y. IF=2.051.
  5. **Gasperskaja (Siavrienė) E.**, Kučinskas V. *The most common technologies and tools for functional genome analysis. Acta Med Litu.* 2017; 24 (1): 1-11. doi: 10.6001/actamedica.v24i1.3457.

### **Žodiniai pranešimai:**

1. **Siavrienė E.** (pranešėja), Preikšaitienė E., Kučinskas V. Su intelektine negalia ir įgimtomis anomalijomis siejamų genų variantų funkcinė analizė. Lietuvos mokslų akademijos Biologijos, medicinos ir geomokslų skyriaus 12-ta jaunųjų mokslininkų konferencija „Bioateitis: gamtos ir gyvybės mokslų perspektyvos“, 2019 m. gruodžio 11 d. Kaunas, Lietuva. Gautas I vietos apdovanojimas.
2. **Siavrienė E.** (pranešėja). iRNR splaisingo variantų, susijusių su intelektine negalia ir įgimtomis anomalijomis, funkcinė analizė. Lietuvos mokslų akademijos Biologijos, medicinos ir geomokslų skyriaus 11-ta jaunųjų mokslininkų konferencija „Bioateitis: gamtos ir gyvybės mokslų perspektyvos“, 2018 m. gruodžio 14 d., Vilnius.
3. **Siavrienė E.** (pranešėja), Preikšaitienė E., Kučinskas V. *Functional genome analysis aiming to define the genetic cause of intellectual disability. 56th Polish and 14th International Conference Juvenes Pro Medicina*, 2018 m. gegužės 25–26 d., Lodzė, Lenkija. ISBN 978-83-947627-1-1. Gautas II vietos apdovanojimas.
4. **Siavrienė E.** (pranešėja), Preikšaitienė E., Kučinskas V. *Functional Genome Analysis in Patients with Intellectual Disability. XIV Baltic Congress of Laboratory Medicine*, 2018 m. gegužės 10–12 d. Vilnius. T. 20, Spec. suppl. p. s2. ISSN: 1392-6470.

## Stendiniai pranešimai:

1. **Siavrienė E.**, Mikštienė V., Radzevičius D., Maldžienė Ž., Rančelis T., Petraitytė G., Kavaliauskienė I., Šarkinas L., Utkus A., Kučinskas V., Preikšaitienė E. *Functional analysis of a novel splice site variant in GLI3 gene causing Greig cephalopolysyndactyly syndrome*. Amerikos žmogaus genetikos draugijos konferencija, 2019 m. spalio 15–19 d., Hjustonas, Jungtinės Amerikos Valstijos. Nr. 2484/W
2. **Siavrienė E.**, Mikštienė V., Maldžienė Ž., Petraitytė G., Rančelis T., Utkus A., Preikšaitienė E., Kučinskas V. *Functional delineation of de novo heterozygous intragenic deletion in MED13L*. Europos žmogaus genetikos draugijos konferencija, 2019 m. birželio 15–18 d., Geteborgas, Švedija. Nr. P08.40A.
3. Petraitytė G., **Siavrienė E.**, Mikštienė V., Maldžienė Ž., Rančelis T., Utkus A., Preikšaitienė E., Kučinskas V. *Functional analysis of a novel c.899+1G>A variant in SLC9A6 gene*. Europos žmogaus genetikos draugijos konferencija, 2019 m. birželio 15–18 d., Geteborgas, Švedija. Nr. P08.56A.
4. **Siavrienė E.**, Petraitytė G., Mikštienė V., Maldžienė Ž., Rančelis T., Preikšaitienė E., Kučinskas V. *Novel c.5535-1G>A variant in a patient with a mild features of CHARGE syndrome*. Amerikos žmogaus genetikos draugijos konferencija, 2018 m. spalio 16–20 d., San Diegas, Jungtinės Amerikos Valstijos. Nr. 1188/W.
5. **Siavrienė E.**, Preikšaitienė E., Maldžienė Ž., Ambrozaitytė L., Gueneau L., Reymond A., Kučinskas V. *Microduplication of the 13q31.3 miR17~92 cluster results in a syndrome with features opposite to those associated with Feingold syndrome 2*. Europos žmogaus genetikos draugijos konferencija, 2018 m. birželio 16–19 d., Milanai, Italija. Nr. P08.02B/B.

## Elektroninis pranešimas:

1. Petraitytė G., **Siavrienė E.**, Mikštienė V., Norvilas R., Burnytė B., Utkus A., Preikšaitienė E. *Analysis of the effect of a novel splice site variant in TGFBR2 on mRNA level*. Virtuali Europos žmogaus genetikos draugijos konferencija, 2020 m. birželio 6–9 d.

**Stažuotės:**

1. Lozanos universitete Integratyvosios genomikos centre (angl. *Center for Integrative Genomics hosted at the University of Lausanne*), Lozana, Šveicarija, nuo 2018 m. sausio 8 d. iki 2018 m. vasario 6 d.
2. *The Cell Factory* mokslinių tyrimų ir eksperimentinės plėtros laboratorijoje (angl. *Research and Development*), Nilis, Belgija, nuo 2016 m. vasario 1 d. iki 2016 m. balandžio 31 d.

## 2. MOKSLINĖS LITERATŪROS APŽVALGA

### 2.1. Žmogaus koduojančios ir nekoduojančios genomo sekos variantai ir kopijų skaičiaus pokyčiai

Tarptautinis žmogaus genomo sekoskaitos konsorciumas, pasibaigus Žmogaus genomo projektui (2003 m.), paskelbė nuskaitytą žmogaus genomo seką ir nustatė, kad DNR seka žmonių populiacijoje yra apie 99,9 proc. vienoda, o likusią dalį sudaro skirtumus tarp žmonių lemiantys DNR sekos pokyčiai, kurių spektras varijuoja nuo pavienių nukleotidų iki didelės apimties pokyčių (pvz., delecijos ar duplikacijos) [1].

Variantas – tai branduolio ar mitochondrijos DNR sekos vieno, rečiau kelių, nukleotidų pokytis. Vieno nukleotido pakaita, delecija (angl. *Deletion*; kitaip iškrita) ar insercija (angl. *Insertion*; kitaip intarpas) yra vadinama taškiniu (angl. *Point*) variantu. Pagal poveikį aminorūgščių sekai išskiriami keli taškinių variantų tipai: tylintieji (angl. *Silent*), *missense*, *nonsense* ir skaitymo rėmelio poslinkio (angl. *Frameshift*) variantai. Variantai, kurie nekeičia aminorūgščių sekos polipeptidinėje grandinėje, vadinami sinoniminiais variantais, o keičiantys aminorūgščių seką – nesinoniminiais variantais (2.1 lentelė) [15].

**2.1 lentelė.** Taškinių variantų tipai ir poveikis aminorūgščių sekai [15]

Pavadinimas	Tipas	Poveikis aminorūgščių sekai	Apibrėžimas
Tylintis variantas	Vieno nukleotido pakaita	Sinoniminis	Aminorūgščių seka polipeptidinėje grandinėje nepasikeičia.
<i>Missense</i> variantas	Vieno nukleotido pakaita		Keičiama viena aminorūgštis.
<i>Nonsense</i> variantas	Vieno nukleotido pakaita	Nesinoniminis	Aminorūgštį koduojantis kodonas keičiamas į baigmės (STOP) kodoną.
Skaitymo rėmelio poslinkio variantas	Insercija / delecija		Koduojama nauja aminorūgščių seka.

Didesnės apimties genetiniai pokyčiai yra vadinami kopijų skaičiaus pokyčiais (KSP; angl. *Copy number variation*). Redonas su bendradarbiais (2006) apibrėžė KSP kaip didesnę nei 1 Kb genomine sritį, kuriai, palyginti su referentiniu žmogaus genomu, būdingas įvairus kopijų skaičius [16]. Patobulėjus technologijoms, atsirado galimybė nustatyti mažesnius



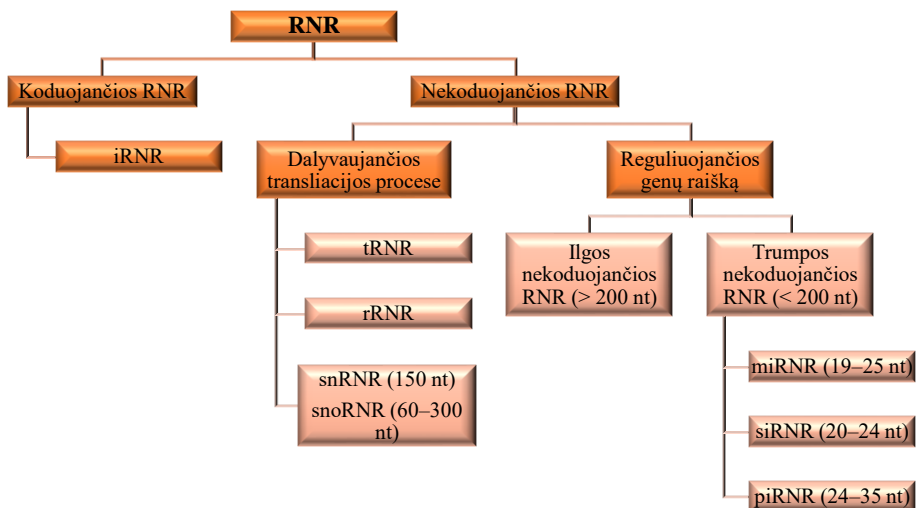
genominius pokyčius, todėl šiuo metu KSP vadinami pokyčiai, kurie yra didesni nei 50 bp [17]. Šiai genolinių pokyčių grupei priklauso delecijos, duplikacijos, insercijos, translokacijos, trinukleotidinių pasikartojimų išsiplėtimas (angl. *Trinucleotide repeat expansion*) bei tandemiškai pasikartojančios sekos [16, 18].

Variantas / KSP, kurio dažnis bendroje populiacijoje yra >1 proc., vadinamas vieno nukleotido polimorfizmu (VNP; angl. *Single nucleotide polymorphism*) ar kopijų skaičiaus polimorfizmu (angl. *Copy number polymorphism*), o retas variantas / KSP, kurio dažnis < 1 proc. – mutacija. Šios plačiai paplitusios sąvokos dažnai yra klaidingai interpretuojamos vertinant varianto / KSP patogeniškumą. Todėl, siekiant suvienodinti klinikinės varianto reikšmės analizės principus visame pasaulyje, Amerikos medicininės genetikos ir genomikos kolegija (angl. *American College of Medical Genetics and Genomics*; ACMG) sukūrė standartus, pagal kuriuos DNR sekos variantai ir KSP yra klasifikuojami į penkias grupes: patogeniniai (angl. *Pathogenic*), galimai patogeniniai (angl. *Likely pathogenic*), neaiškios klinikinės reikšmės (angl. *Uncertain significance*), galimai nepatogeniniai (angl. *Likely benign*) bei nepatogeniniai (angl. *Benign*) variantai. Vadovaujantis šiomis ACMG rekomendacijomis, polimorfizmas reiškia nepatogeninį arba galimai nepatogeninį variantą / KSP, o mutacija – patogeninį arba galimai patogeninį variantą / KSP [19–22].

Žinomos įvairios DNR sekos variantų ir KSP atsiradimo priežastys. DNR sekos variantai gali atsirasti spontaniškai dėl įvairių ląstelinių ar biologinių procesų sutrikimų (pvz., depurinimas, deaminimas, tautomeriniai virsmai) arba gali būti indukuoti aplinkos veiksnių – mutagenų (pvz., cheminės medžiagos, jonizuojančioji spinduliuotė) [15]. O KSP dažniausiai susidaro dėl nealelinės homologinės rekombinacijos proceso mejozės metu [16, 18]. Retesniais atvejais KSP atsiranda dėl nehomologinio galų sujungimo ar trūkio taško replikuojantis DNR grandinei [23].

Šiuo metu naudojant įvairius pažangius molekulinės genetikos metodus – dažniausiai naujos kartos sekoskaitą (NKS; angl. *Next generation sequencing*) – yra nustatomi variantai ir KSP tiek baltymus koduojančiose, tiek nekoduojančiose DNR sekose. Baltymus koduoja apie 2 proc. žmogaus geno. Nepaisant to, anksčiau buvo manoma, kad genetiniai pokyčiai, esantys tik šiose baltymus koduojančiose DNR srityse, yra svarbūs funkciniu požiūriu, o likusi geno dalis buvo laikoma nereikšminga (angl. *Junk DNA*). Tačiau ENCODE (angl. *Encyclopedia of DNA Elements*) projekto metu buvo nustatyta, kad daugiau nei 80 proc. tokios netransliuojamos DNR sekos yra transkribuojama ir yra labai svarbi genų raiškos reguliavimui [3–5]. Nekoduojamoje DNR sekoje yra reguliuojančios, transkripcijos veiksnių

jungimosi, sukirpimo (angl. *Splicing*) bei 5' ir 3' netransliuojamos sritys. Reguluojančioje DNR sekoje yra transkribuojami ne tik genų promotoriai, stiprikliai (angl. *Enhancers*), slopikliai (angl. *Silencers*) bei tarpikliai (angl. *Insulators*), bet taip pat nekoduojančios RNR molekulės (2.1 pav.), pavyzdžiui, tRNR, miRNR, siRNR, snRNR, lncRNR, kurios dalyvauja translacijos, chromatino pertvarkos, genų raiškos reguliacijos ir kituose gyvybiškai svarbiuose procesuose [24–26]. Farhas su bendradarbiais (2015), atlikdami plataus masto viso genomo sąsajų (angl. *Genome-wide association studies*; GWAS) tyrimą, nustatė, kad net 90 proc. autoimuninių ligų atvejų yra nulemta variantų, esančių šiose nekoduojančiose sekose, o dažniausios ligų priežastys – patogeniniai variantai, esantys geno stipriklio srityje [27]. Nepatogeniniai nekoduojančios sekos variantai taip pat gali turėti įtakos fenotipui. Guenther su kolegomis (2014) aprašė Europos populiacijai būdingą dažną ( $MAF^1 = 0,12$ ) VNP (rs12821256), esantį *KITLG* (MIM #184745) geno stipriklio sekoje, kuris nulemia šviesių plaukų spalvą [28].



**2.1 pav.** RNR klasifikacija. Schema adaptuota pagal mokslinį straipsnį [26]

<sup>1</sup> MAF – retojo alelio dažnis (angl. *Minor allele frequency*).

### 2.1.1. Patogeniniai DNR sekos variantai ir kopijų skaičiaus pokyčiai

Žmogaus genome pasireiškiantys VNP ir KSP, kurių dažnis bendroje populiacijoje >1 proc., yra evoliucijos varomoji jėga, kadangi lemia fenotipų įvairovę bei turi įtakos adaptacijos procesams. Nepaisant to, daugelio pasaulio mokslininkų tyrimo objektas yra reti DNR sekos variantai ir KSP, kurių dažnis populiacijoje < 1 proc., kadangi šie genetiniai pokyčiai gali turėti įtakos genų raiškai ar dozei, pažeisti genų struktūrą arba genus reguliuojančias sekas. Visa tai lemia įvairius genetinius sutrikimus, įskaitant IN ir IA [16, 29, 30].

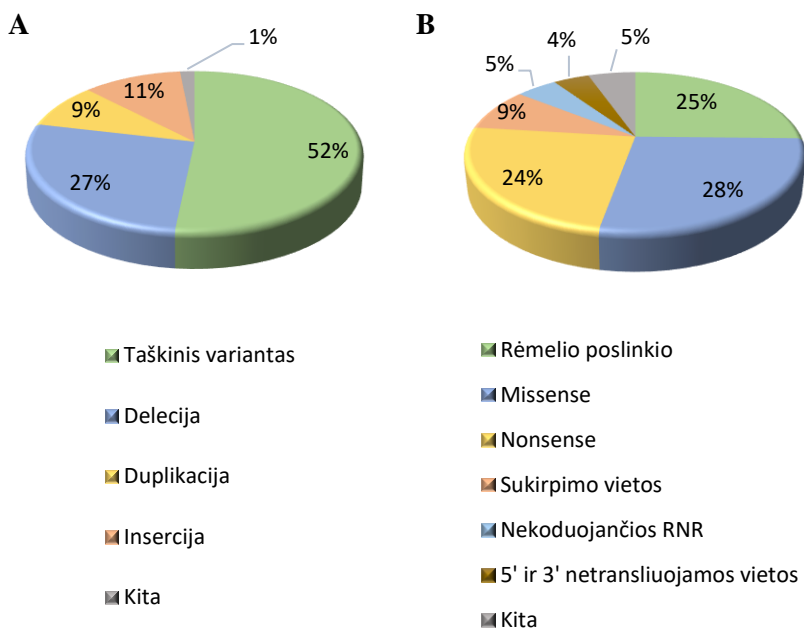
2013 metais baigėsi 1 000 genomų projektas, kurio pagrindinis tikslas buvo ištirti retus žmogaus genomo sekos variantus įvairiose pasaulio populiacijose. Remiantis tyrimų rezultatais, žmogaus genome buvo nustatyta apie 10 000–11 000 variantų nesinoniminėse sekose. Panašus skaičius DNR sekos variantų buvo nustatytas taip pat sinoniminėse sekose. Projekto metu nustatyta, kad kliniškai reikšmingų yra tik dalis variantų: 190–210 insercijų bei delecijų, nelemiančių rėmelio poslinkio, 220–250 delecijų, lemiančių skaitymo rėmelio poslinkį, 80–100 variantų, sudarančių priešlaikinį baigmės kodoną, 40–50 DNR sekos sukirpimo vietos variantų [6, 7].

Redonas su bendradarbiais (2006) *HapMap* projekto metu, tirdami skirtingas pasaulio populiacijas, sukūrė žmogaus genomo KSP katalogą. Vykdydami mokslinį projektą, mokslininkai net 1 447 skirtingose žmogaus genomo srityse nustatė KSP. Šios genomines sritys apėmė 360 Mb (12 proc.) viso žmogaus genomo. Taip pat nustatyta, kad kiekvienas žmogus vidutiniškai turi 70 KSP, todėl šie struktūriniai genomo pokyčiai apibūdinti kaip labiausiai paplitusi žmogaus genomo įvairovės forma [16]. Nepaisant KSP paplitimo žmogaus genome, didžioji dalis KSP yra nepatogeniniai. Zarrei su kolegomis (2015) nustatė, kad maždaug 100 genų homozigotinės delecijos žmogaus genome neturi jokio poveikio fenotipui [17].

Patogeninių KSP įtaka fenotipui pirmą kartą buvo nustatyta 1936 metais tiriant *Drosophila melanogaster*. Tyrimo metu dėl duplikacijos *Bar* gene buvo stebimos neįprastos juostos formos akys (angl. *Bar eye*) [31]. Visgi su didesniu patogeniškumo potencialu yra siejamos delecijos, kadangi įrodyta, jog organizmas lengviau toleruoja genų dozės perteklių nei trūkumą [32, 33].

Nustatomų variantų ir KSP skaičius nuolat didėja. Remiantis Nacionalinio biotechnologijų informacijos centro duomenų bazės (angl. *National Center for Biotechnology Information*; NCBI) *ClinVar* [8] duomenimis, šiuo metu yra įregistruota: 78 726 (12 proc.) patogeniniai, 36 641 (5 proc.) galimai patogeninis, 115 273 (17 proc.) nepatogeniniai, 177 894 (27 proc.) galimai nepatogeniniai bei 261 882 (39 proc.) nežinomos klinikinės reikšmės variantai ir KSP [9]. Dažniausi patogeniniai genetiniai

pokyčiai yra taškiniai variantai (52 proc.) arba KSP: delecijos (27 proc.), insercijos (11 proc.) ir duplikacijos (9 proc.) (2.2 pav. A). Atsižvelgiant į genetinių pokyčių poveikį aminorūgščių sekai, dažniausiai nustatomi rėmelio poslinkio, *missense* bei *nonsense* tipo variantai (2.2 pav. B). Šie variantai taip pat kaip dažniausi nurodomi žmogaus genų variantų duomenų bazėje *HGMD* (angl. *Human Gene Mutation Database*) [34], kurioje iš viso yra pateikta 275 716 variantų, siejamų su paveldimomis žmogaus ligomis [35].



**2.2 pav. A)** Patogeninių variantų pasiskirstymas pagal varianto tipą; **B)** Patogeninių variantų pasiskirstymas pagal poveikį aminorūgščių sekai (skaitymo rėmelio poslinkio, *missense*, *nonsense* variantai) ir vietą DNR sekoje (sukirpimo vietos, nekoduojančios RNR, 5' ir 3' netransliuojamų vietų variantai). Skritulinės diagramos sudarytos remiantis *ClinVar* duomenimis [8, 9]

Procesas, kuomet pirimidinas (C/T) yra keičiamas kitu pirimidinu, o purinas (A/G) kitu purinu, vadinamas tranzicija. Pirimidino keitimas purinu (ir atvirkščiai) vadinamas transversija [15]. Tranzicijos dažnis yra žymiai didesnis, palyginti su transversija: 1,4:1 sinoniminėse sekose bei 2:1 nesinoniminėse sekose [36]. *ClinVar* [8] duomenimis, apie 47 proc. sveikatos sutrikimų ar ligos atvejų lemia tranzicija (A>G). Kiti patogeniniai taškiniai variantai įvyksta genome daug rečiau, pavyzdžiui, transversija (A>C) ir tranzicija (C>T) siejama su ligomis, kurios atitinkamai sudaro 16 proc. ir 14 proc. atvejų [37].

*Nonsense*, rėmelio poslinkio bei sukirpimo vietos variantai dažnai lemia baltymo sutrumpėjimą, tokiais atvejais šie DNR sekos variantai yra vadinami baltymą trumpinančiais variantais (angl. *Protein-truncating variants*; PTV). Šie variantai turi tam tikrą poveikį fenotipui dėl dviejų priežasčių. Visų pirma, PTV gali lemti baltymo funkcijos praradimą (angl. *Loss-of-function*). Tokiais atvejais geno produktas tinkamai neatlieka savo funkcijos arba jo funkcionalumas yra sumažėjęs. Retesniais atvejais PTV gali lemti baltymo funkcijos įgijimą (angl. *Gain-of-function*) – geno produktas atlieka jam nebūdingą funkciją. Visa tai gali turėti neigiamą poveikį fenotipui: gali sutrikti metabolizmas, pigmentacija, elgsena ar net būti ligos atsiradimo priežastis [38, 39].

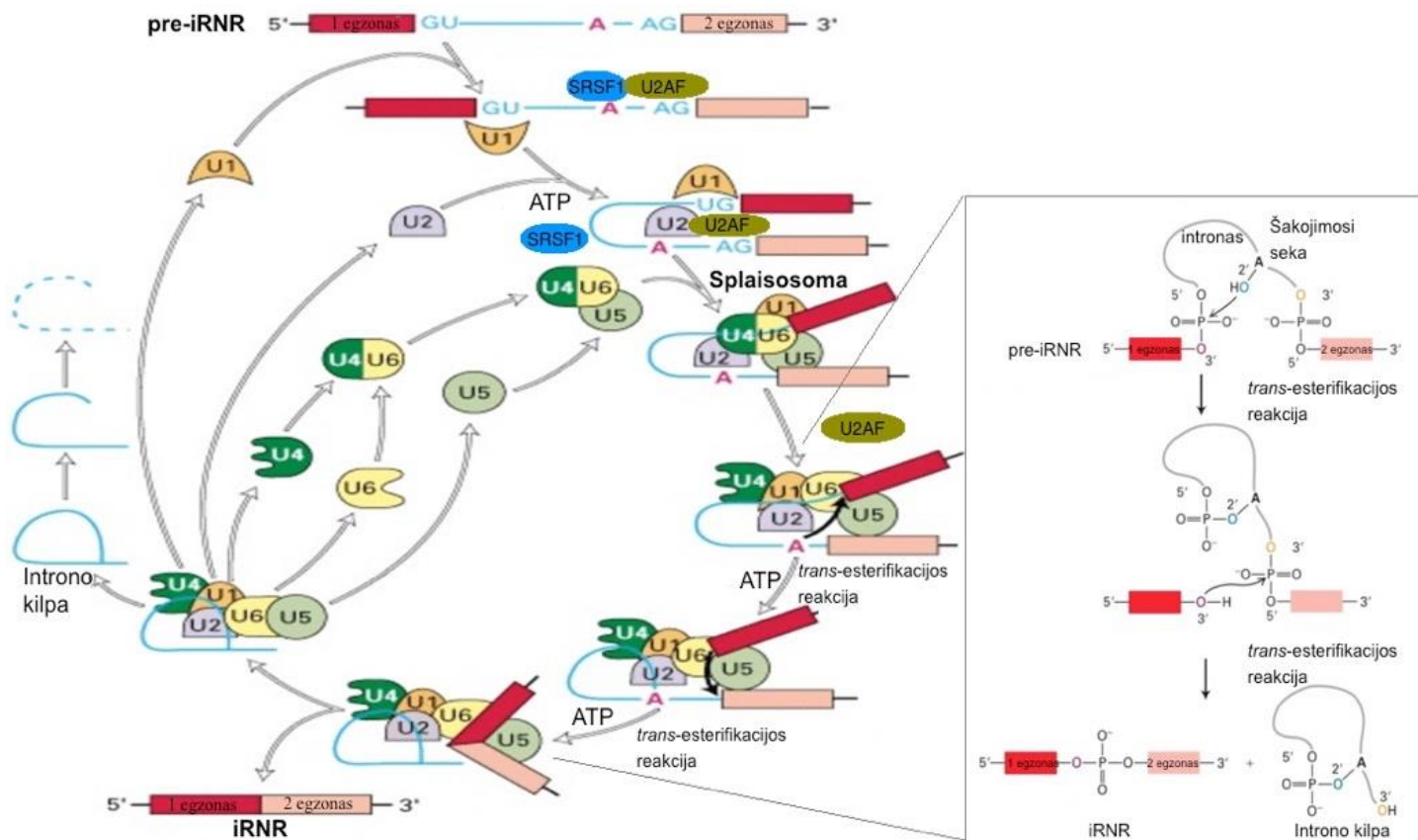
Nors kiekvieno žmogaus DNR sekoje yra apie 50–100 heterozigotinių variantų, kurie charakterizuojami kaip PTV, o *HGMD* [34] duomenų bazėje klasifikuojami kaip sukeliantys paveldimus sutrikimus, ne visi žmonės suserga šiomis ligomis [6]. Tai gali būti aiškinama tuo, kad egzistuoja variantų taisymo mechanizmai, tokie kaip tiesioginis (angl. *Direct repair*), bazių ar nukleotidų ekscizijos (angl. *Base or nucleotide excision repair*), neatitinkančių nukleotidų (angl. *Mismatch repair*), homologinės rekombinacijos (angl. *Homologous recombination*) ir nehomologinio galų sujungimo (angl. *Nonhomologous end joining*) [15]. Žinomas taip pat mechanizmas, veikiantis *posttranskripciniu* lygmeniu, kuris lemia iRNR, turinčios prieš laikiną baigmės kodoną, skaidymą (angl. *Nonsense-mediated decay*; NMD). Veikiant šiam mechanizmui, specifinis egzono jungties baltymų (SMG1, RENT1, eRF1, eRF3) kompleksas atpažįsta prieš laikiną baigmės kodoną, kuris 50–55 nt atstumu yra nutolęs nuo paskutinės egzono poros jungties vietos ir, sąveikaujant su kitais papildomais baltymais, inicijuoja iRNR skaidymo procesą [40]. Šiuo metu taip pat daugėja mokslinių publikacijų apie apsaugančius (angl. *Protective*) variantus, kurie apsaugo žmogaus genomą nuo ligos pasireiškimo [41]. Pavyzdžiui, moksliniais tyrimais patvirtinta, kad *CARD9* (MIM #607212), *RNF186* (MIM #617163) ir *IL23R* (MIM #607562) genų tam tikri variantai apsaugo nuo susirgimo Krono liga [39].

#### 2.1.1.1. Sukirpimo vietos variantai ir jų pasekmės informacinės RNR lygmeniu

iRNR pirmtako (pre-iRNR; angl. *Precursor messenger RNA*) sukirpimo arba kitaip vadinamas brendimo procesas yra svarbus genų raiškai reguliuoti bei baltymų įvairovei užtikrinti [42]. Identifikuota 20 000–25 000 baltymą koduojančių genų, tačiau žmogaus proteomą sudaro keliasdešimt

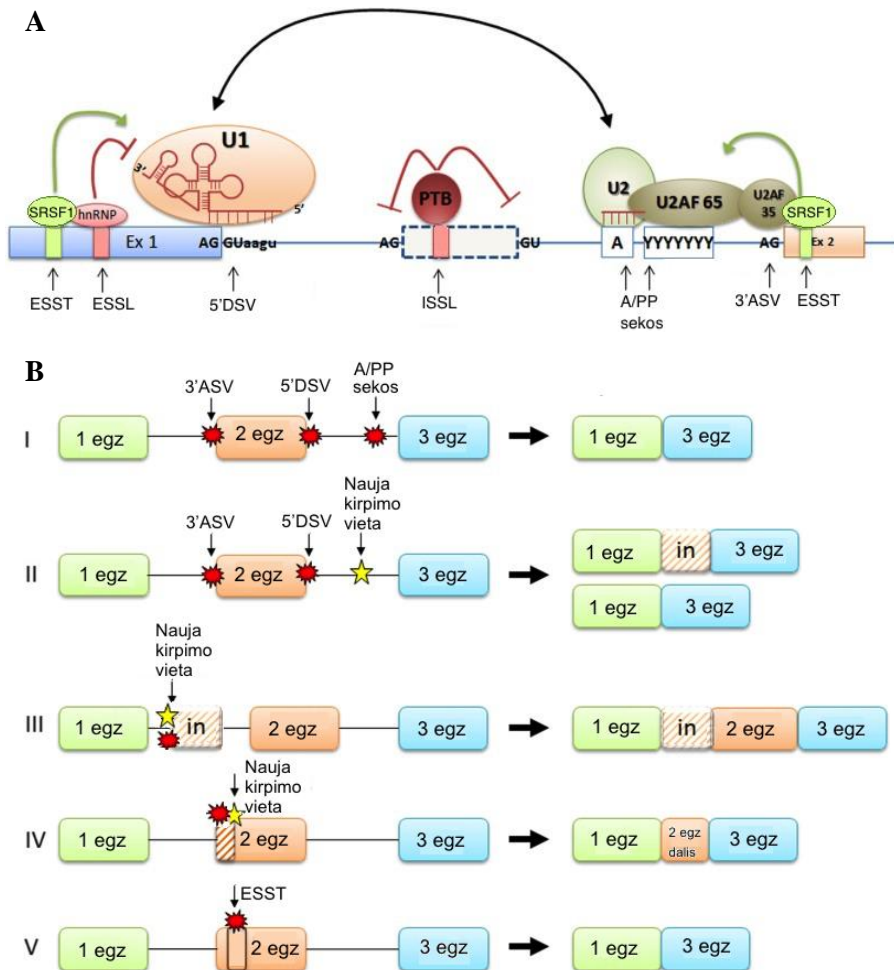
kartų daugiau baltymų, kadangi alternatyvaus pre-iRNR sukirpimo proceso metu skirtingais deriniais pašalinami intronai, atskiriantys gretimus egzonus. Tokiu būdu sukuriamos įvairios baltymų izoformos [43]. Pavyzdžiui, žinoma, kad dėl alternatyvaus sukirpimo proceso susidaro net 46 *NF1* (MIM # 613113) geno koduojami transkriptai [44]. Nors *TTN* (MIM #188840) žinomas kaip daugiausiai egzonų (363) turintis genas, žmogaus genai vidutiniškai turi 8,8 egzono, o intronai sudaro apie 90 proc. pre-iRNR sekos [45, 46].

Mažosios branduolio RNR (snRNR) dėl sekų, turtingų uracilo, vadinamos kaip U-RNR (U1, U2, U4, U5, U6), atlieka pagrindinį vaidmenį šalinant intronus. Šios snRNR jungdamosi su specifiniais baltymais (pvz., SRSF1, U2AF65, U2AF35, PTB) sudaro ribonukleoproteinų (RNP) kompleksą, vadinamą splaisosoma, kuri yra pagrindinis iRNR kirpimo įrankis (2.3 pav.). Splaisosomos U1 ir U2 funkciniai subvienetai atpažįsta intronų specifines sekas (angl. *Consensus sequence*). U1 snRNR atpažįsta ir jungiasi introno 5' gale, kuriame yra konservatyvūs GU dinukleotidai. U2 snRNR atpažįsta ir jungiasi adenino (A) turtingoje sekoje – šakojimosi srityje (angl. *Branch site*), kuri yra nutolusi apie 20 nt atstumu nuo konservatyvių AG dinukleotidų introno 3' gale. Šalia šios sekos yra kita atpažinimui svarbi vieta – polipirimidino seka (angl. *Polypyrimidine tract*), kurią atpažįsta U2 RNP. U1 ir U2 RNP sąveikaudami tarpusavyje suartina 5' ir 3' kirpimo sekas, o U4, U5 ir U6 RNP užtikrina teisingą konfigūraciją ir formuoja katalizinę splaisosomos kompleksą. U2 ir U6 RNP dalyvauja transesterifikacijos reakcijose. Pirmoje transesterifikacijos reakcijoje adenino, esančio šakojimosi sekoje, 2' hidroksi grupė sąveikauja su fosfatu, esančiu introno 5' gale. Susidarius kilpai (angl. *Lariat*), nukirpto egzono 3' hidroksi grupė sąveikauja su fosfatu, kuris yra introno 3' gale. Kilpa pašalinama, o egzonai sujungiami ligavimo reakcijoje panaudojant adenzino 5'-trifosfatą (angl. *Adenosine 5'-thiotriphosphate*; ATP) kaip energijos šaltinį. Splaisosomos kompleksas disocijuoja, todėl mechanizmo veikimo procesas gali būti kartojamas su kita pre-iRNR (2.3 pav.) [47, 48].



2.3 pav. Alternatyvaus pre-iRNR sukirpimo procesas. Bendroji sukirpimo mechanizmo schema adaptuota pagal mokslinę literatūrą [47, 48]

Sukirpimo procesas yra reguliuojamas baltymų, vadinamų *trans*-veiksniais (angl. *Trans-acting factors*), ir *cis*-reguliuojančių sekų (angl. *Cis-regulatory sequences*). Pagrindinės *cis*-reguliuojančios sekos yra introno 5' galo donorinė ir 3' galo akceptorinė sukirpimo seka bei introno 3' galo A turtinga šakojimosi ir polipirimidino seka. Šioms sekoms taip pat priskiriami vadinamieji sukirpimą reguliuojantys elementai (SRE; angl. *Splicing-regulatory elements*), t. y. sukirpimo proceso stiprikliai ir slopikliai, kurie yra tiek egzonišose (ESST, ESSL), tiek intronišose (ISST, ISSL) sekose. SRE sąveikaujant su specifiniais *trans*-veiksniais, sukirpimo mechanizmas yra atitinkamai aktyvuojamas arba nuslopinamas (2.4 pav. A). *Trans*-veiksnių grupei priklauso serino ir arginino turtingi baltymai (pvz., SRSF1 ar U2AF), taip pat anksčiau išvardytos snRNR molekulės, sudarančios RNP kompleksą [49].





**2.4 pav. A)** *Cis* ir *trans* sukirpimo procesą reguliuojantys elementai: 5' DSV – introno 5' galo donorinė ir 3' ASV – 3' galo akceptorinė sukirpimo vieta, ESST – egzoniinė sukirpimo proceso stipriklio vieta, ESSL – egzoniinė sukirpimo proceso slopiklio vieta, ISSL – introniinė sukirpimo proceso slopiklio vieta, A/PP – adenino ir polipirimidino turtingos sekos; **B)** Sukirpimo proceso patogeninių variantų pasekmės. Žymėjimas: egz – egzonas, in – intronas, raudona žvaigždutė – patogeninio sukirpimo varianto vieta, geltona žvaigždutė – nauja kirpimo vieta; schemos adaptuotos pagal mokslinę literatūrą [48]

Remiantis moksline literatūra galima teigti, kad apie 10–15 proc. sukirpimo proceso patogeninių variantų *cis*-reguliuojančiose sekose ar *trans*-veiksnius koduojančiose sekose yra genetinių ligų priežastis [50, 51]. Žmogaus genų patogeninių variantų duomenų bazės *HGMD* [34] duomenimis, su genetinėmis ligomis siejami 9 proc. (23 354/269 419) sukirpimo vietos variantų [35]. *ClinVar* [8] duomenų bazėje šiuo metu yra registruojama mažiau, t. y. 6 865 (9 proc.) patogeniniai sukirpimo vietos variantai, iš kurių 754 variantai siejami su IN [9]. Šie patogeniniai variantai sutrikdo normalią sukirpimo mechanizmo veiklą.

Dažniausiai dėl patogeninio varianto, ypač introno 5' donorinėje arba 3' akceptorinėje bei A/PP sekose, yra pašalinamas egzonas (2.4 pav. B I) [36, 48]. Krawczakas su bendradarbiais (2007), atlikę 38 genuose 478 patogeninių variantų metaanalizę, nustatė, kad donorinės sukirpimo vietos patogeniniai variantai yra maždaug pusantro karto dažnesni nei akceptorinės vietos variantai [51]. Taip pat šie patogeniniai variantai dažnai lemia kriptinį (angl. *Cryptic*) alternatyvų 5' ir 3' galo sukirpimą gretimame egzone arba introne (2.4 pav. B II) [50, 52]. Tokiais atvejais vietoj įprastų GU dinukleotidų introno 5' gale ir AG – 3' gale pasirenkama kita alternatyvi kirpimo vieta, kuri dažniausiai lemia dalies introno įterpimą arba viso egzono praradimą. Pavyzdžiui, Eng su bendradarbiais (2004) nustatė patogeninį variantą c.4437-1G>A *ATM* (MIM #607585) gene, kuris lėmė kriptinės alternatyvios sukirpimo vietos susidarymą vienu nukleotidu toliau nuo pirminės vietos, kadangi pirmasis gretimo egzono nukleotidas yra G [53]. Remiantis šiuo tyrimu, galima būtų teigti, jog, įvykus pokyčiui 3' akceptorinėje sekoje, U2 snRNR atpažįsta kitą AG dinukleotidų porą, kuri yra arčiausiai nutolusi nuo pirminės 3' akceptorinės sukirpimo vietos. Tačiau Moonas su kolegomis (2005) moksliniame straipsnyje aprašė du *HRPT2* (MIM #145001) geno pakitusius transkriptus, kurie susidarė dėl patogeninio varianto c.238-1G>A antrajame introne. Pirmajame transkripte buvo pašalintas trečiasis egzonas, o kitame transkripte buvo pašalinta tik dalis (23 nt) trečiojo egzono [54]. Dėl šios priežasties iki šiol nėra žinomas tikslus mechanizmas, koku būdu įvykus

pokyčiui sukirpimo vietose pasirenkama alternatyvi kirpimo vieta [50, 52, 55]. Manoma, kad tam gali turėti įtakos genomo vieta, ESST ir ESSL tankis genome ar net pre-iRNR antrinės struktūros [48].

Retesniais atvejais patogeniniai variantai introninėse sekose, įskaitant ISST ir ISSL, lemia dalies introno intarpą (2.4 pav. B III). Paprastai introninės sekos išlaikymas iRNR trukdo iRNR transportavimui iš branduolio į citozolį, todėl tokios iRNR nėra transliuojamos [36]. Pavyzdžiui, c.3718-2477C>T variantas, kuris yra dažniausiai nustatomas patogeninis variantas *CFTR* (MIM # 602421) gene Lenkijos populiacijoje. Įvykus šiam pokyčiui 19 introne, į gretimą egzoną yra įterpiama 84 nt ilgio introno seka, dėl kurios susidaro skaitymo rėmelio poslinkis ir priešlaikinis baigmės (*STOP*) kodonas [56, 57]. Panašiai patogeniniai variantai egzoninėse sekose, įskaitant ESST ir ESSL, lemia egzono intarpą, kuris gali būti pilnas arba dalinis (2.4 pav. B IV). Nustatyta, kad patogeniniai variantai ESST, taip pat kaip donorinėje bei akceptorinėje sukirpimo vietoje, dažniausiai lemia viso egzono iškritą (2.4 pav. B V) [48].

*Trans*-reguliuojančių veiksmų patogeniniai variantai, kurie turi įtakos sukirpimo procesui, nustatomi rečiau, palyginti su *cis*-reguliuojančių sekų variantais. Anczuków su kolegomis (2012) aprašė pelių epitelinių ląstelių kultūros funkcinį tyrimą, kuriame buvo dirbtinai sutrikdyta SRSF1 baltymą koduojančio geno raiška. Nustatyta, kad šio geno padidėjusi raiška turi neigiamą poveikį *BCL2L11* (MIM #603827) geno sukirpimo mechanizmui: vietoj 4 egzono įterpiamas 3 egzonas. Susidaręs alternatyvus transkriptas yra atsparus apoptozei, todėl skatinamas krūties vėžio vystymosi procesas [49, 58].

Patogeniniai sukirpimo variantai *cis* ir *trans* reguliuojančiuose elementuose, priklausomai nuo tipo ir vietos genome, turi įvairių pasekmių žmogaus proteomui. Dažniausiai dėl šių variantų susidaro skaitymo rėmelio poslinkis ir priešlaikinis (*STOP*) kodonas. *Post*transkripciniu lygmeniu veikiantis apsaugantis mechanizmas – NMD dažnai atpažįsta iRNR transkriptus, turinčius priešlaikinį baigmės kodoną, ir skatina jų skaidymą. Todėl dėl DNR sekos varianto polipeptidas gali nesusidaryti. Tokiais atvejais fenotipiniai požymiai yra aiškinami haplonepakankamumu. Kita vertus, kai NMD mechanizmas neatpažįsta priešlaikinio baigmės kodono arba patogeninis variantas, nesukeliantis rėmelio poslinkio, lemia dalies aminorūgščių sekos praradimą, yra gaminamas trumpesnis baltymas, kuris gali būti pakitusių funkcijų arba iš viso nefunkcionalus [36, 59].

Nors sukirpimo vietos variantai nėra tokia dažna genetinių ligų ir sutrikimų, įskaitant IN, priežastis kaip kiti taškiniai variantai ar KSP [60], sukirpimo variantų biologinės pasekmės bei molekuliniai etiopatogeninės

mechanizmai yra svarbūs ir vis dar reikalauja daugiau ir išsamesnių mokslinių tyrimų [52].

### 2.1.2. Lietuvos populiacijos genomo įvairovės tyrimai

Kiekviena pasaulio populiacija yra unikali dėl savo kultūros geografinės padėties ir evoliuciškai vykusių mutacijų, genų dreifo, gamtinės atrankos ar migracijos. Dėl šios priežasties populiacijų tyrimai yra labai svarbūs, siekiant nustatyti populiacijai būdingus genetinius variantus [61]. Lietuvos populiacija yra išskirtinė, kadangi jos dydis išliko gana stabilus nepaisant demografinių pokyčių, nulemtų gamtinių ir socialinių veiksnių. Taip pat Lietuva yra viena iš dviejų likusių šalių, kurių kalbos priklauso baltų indoeuropiečių kalbų grupei [62].

Siekiant identifikuoti Lietuvos populiacijai būdingas genomines sritis, 2011–2015 metais buvo vykdomas LITGEN projektas. Atlikus plataus masto genominius tyrimus asmenims iš šešių etnolingvistinių Lietuvos regionų (Rytų, Pietų, Vakarų Aukštaitija ir Šiaurės, Pietų, Vakarų Žemaitija), buvo gauti reikšmingi rezultatai. Pavyzdžiui, nustatyta, kad Lietuvos populiacija pasižymi homogeniškumu, kadangi genetiniai skirtumai tarp šešių tirtų grupių sudarė < 2 proc. Nustatyta, kad Lietuvos populiacija yra artimiausia Latvijos ir Estijos populiacijoms, o padėtis kitų analizuotų populiacijų atžvilgiu priklauso nuo geografinių atstumų [62, 63]. Lietuvos populiacijos genetinė struktūra taip pat buvo tirta dr. I. Kavaliauskienės, dr. T. Rančelio, dr. L. Pranckėnienės, doc. dr. Ž. Maldžienės ir kitų Lietuvos mokslininkų. Dr. I. Kavaliauskienė savo moksliniame darbe nustatė, kad genetinės įvairovės skirtumai tarp tirtų Lietuvos etnolingvistinių grupių matomi pagal VNP alelių, mitochondrijos DNR, Y chromosomos haplogrupių ir haplotipų pasiskirstymą, o didžiausi skirtumai būdingi Y chromosomos genetinei įvairovei [64]. Kituose moksliniuose darbuose buvo tirtas išskirtinai patogeninių genomo variantų (dr. T. Rančelis [65]), *de novo* variantų (dr. L. Pranckėnienė [66]) ir KSP (doc. dr. Ž. Maldžienė [67]) pasiskirstymas Lietuvos populiacijoje.

Žmonių populiacijos genotipų ir genominių sričių tyrimai yra svarbūs ne tik siekiant charakterizuoti kilmę ir populiacijai būdingus ypatumus dėl besikeičiančių aplinkos sąlygų, bet taip pat sveikatai reikšmingoms genominiams sritims ir rizikos veiksniams identifikuoti bei tinkamai gydymo strategijai parinkti. Moksliniais tyrimais patvirtinta, kad skirtingos populiacijos gali skirtingai toleruoti naudojamo vaisto veikliąją medžiagą [68]. Taip pat patvirtintas skirtingas ligų paplitimo dažnis tarp skirtingų populiacijų. Pavyzdžiui, pjautuvine anemija (angl. *Sickle-cell*; MIM

#603903) žymiai dažniau serga Afrikos populiacija nei Europos, o cistinė fibrozė (MIM #219700) ir hemochromatozė (MIM #231100), atvirkščiai, yra dažniau būdinga europiečiams [68].

Lietuvoje dažniausia mirties priežastis yra širdies ir kraujagyslių sistemos ligos (ŠKL), o, lyginant skirtingas Europos populiacijas, Lietuvoje mirtingumas nuo šių ligų yra beveik du kartus didesnis. LITGEN projekto metu nustatyta, kad pagrindiniai ŠKL rizikos veiksniai yra padidėjęs KMI ir dislipidemija [69]. Koronarinės širdies ligos genetinę architektūrą bendroje Lietuvos populiacijoje savo moksliniame darbe, panaudojusi dažnus VNP genomo žymenis, ištyrė dr. I. Domarkienė [70]. Kituose moksliniuose darbuose tirti genetiniai veiksniai, dalyvaujantys kitų ligų patogenezėje. Pavyzdžiui, doc. dr. L. Ambrozaitytė nustatė stipriausius genus kandidatus, lemiančius polinkį lūpos ir / ar gomurio nesuaugimams [71], o dr. V. Mikštienė įvertino genetinių veiksnių reikšmę klausos sutrikimo patogenezei ir fenotipinei išraiškai [72].

VU MF BMI ŽMGK yra vykdomi moksliniai darbai bei nacionaliniai ir tarptautiniai projektai, skirti įvertinti IN ir / ar IA etiopatogenezei. Pavyzdžiui, doc. dr. Ž. Maldžienė nustatė genetinių sričių kiekybinius pokyčius bei jose esančius stiprius genus kandidatus, kurių dozės pokytis lemia IN [67], o doc. dr. E. Preikšaitienė įvertino patogeninių DNR sekos KSP įtaką asmenų, turinčių IN, fenotipui ir centrinės nervų sistemos raidos bei funkcijos sutrikimams [32]. Nuo 2009 metų vykdomų projektų (pvz., UNIGENE, PROGENET, CHERISH, INGENES) metu, naudojant pažangiausias tyrimo technologijas, buvo ištirta daugiau nei 400 asmenų, turinčių sindrominę ir nesindrominę IN, identifikuoti IN kandidatiniai genai, patikslinta žinomų sindromų fenotipinių požymių įvairovė bei išanalizuoti patogeniniai variantai ir KSP, kurie yra siejami su IN ir / ar IA [73]. UNIGENE projektas suteikė pagrindą funkciniam tyrimams Lietuvoje, kadangi Šveicarijos mokslininkai variantų, siejamų su IN, ištyrimui taikė funkcinės genomikos metodus, o dalis jų buvo taikomi pastaruoju metu vykdant INGENES projektą, kurio pagrindinis tikslas – nustatyti ir patvirtinti IN ir / ar IA lemiančius naujus DNR sekos variantus ir KSP, atliekant klinikinį, molekulinį ir funkcinį jų charakterizavimą [74].

## 2.2. DNR sekos variantų ir kopijų skaičiaus pokyčių tyrimų strategijos

Lyginamosios genomo hibridizacijos (LGH) ir sekoskaita vadinamų DNR nukleotidų sekos nustatymo metodų atsiradimas ir vystymasis lėmė proveržį medicinos srityje [75].

Nuo 2004 metų pradėjus taikyti pažangias molekulinio kariotipavimo technologijas, tokias kaip vektorinė lyginamoji genomo hibridizacija (vLGH, angl. *Array-based comparative genomic hybridization*) ir vieno nukleotido polimorfizmo lyginamoji genomo hibridizacija (VNP-LGH; angl. *Single nucleotide polymorphism genotyping array*), suteikta galimybė ne tik efektyviai nustatyti nesubalansuotus struktūrinius genominius pokyčius – delecijas ir duplikacijas visame žmogaus genome, bet ir patikslinti pokyčių dydį ir vietą genome. LGH metodų principas pagrįstas skirtingai pažymėta DNR hibridizacija su specifiniais žymenimis: žinomos sekos oligonukleotidiniais žymenimis (vLGH) arba vieno nukleotido polimorfizmą žymenimis (VNP-LGH). Šių metodų skiriamoji geba, priklausomai nuo naudojamo žymens, gali būti nuo 1 nukleotido (VNP-LGH) iki kelių šimtų nukleotidų ar net didesnių KSP (vLGH) [36, 76, 77].

Nežinomai genominei sričiai ištirti visame pasaulyje yra taip pat plačiai naudojamos tiek pirmos (Sangerio), tiek antros (naujos) kartos sekoskaitos technologijos [38]. 1977 metais Frederickas Sangeris su kolegomis sukūrė grandinės terminacijos principu paremtą sekoskaitos metodą, kuris daugelį metų buvo laikomas populiariausiu ir efektyviausiu DNR sekos variantų tyrimo metodu. Šis metodas pagrįstas fluorescuojančiais žymenimis žymėtų terminalinių dideksinukleotidtrifosfatų (ddNTP) naudojimu, kuriems prisijungus prie polimerazės grandininėje reakcijoje (PGR; angl. *Polymerase chain reaction*) amplifikuojamos grandinės nutrūksta tolesnis grandinės ilginimo procesas. Šio tyrimo metu naudojamas tik vienas pradmuo, todėl, atlikus sekoskaitos PGR, gaunamas statistinis viengrandžių fragmentų mišinys, kurių ilgis skiriasi tik vienu nukleotidu. Toks fragmentų mišinys toliau tiriamas automatinės kapiliarinės elektroforezės metodu, kuriuo pagal fluorescuojantį žymenį nustatoma tiesioginė nukleotidų seka [36, 78, 79]. Sangerio metodas yra tinkamas tirti specifines genetines sritis, kadangi vieno tyrimo metu tiriamas DNR fragmentas, kurio ilgis yra apie 700 nt [80]. Taip pat Sangerio metodu įprasta atlikti triados (probando ir jo tėvų) segregacijos tyrimą, siekiant įvertinti analizuojamo varianto kilmę (*de novo* ar paveldėtas) [81].

2003 m. pasibaigus Žmogaus genomo projektui pradėjo sparčiai vystytis didelio našumo NKS technologijos, kurios yra skirtos atlikti plataus masto genomo tyrimams. Šiomis technologijomis yra nustatomi ne tik smulkūs KSP, kurie negali būti nustatomi LGH metodu, bet taip pat koduojančios ir nekoduojančios DNR sekos variantai [75, 82].

Priklausomai nuo tyrimo apimties, žinomos kelios NKS DNR sekos tyrimo strategijos: viso genomo (VGS; angl. *Whole genome*), viso egzomo (VES; angl. *Whole exome*) ar su tam tikru požymiu (pvz., IN) siejamų genų

grupės (angl. *Gene panel*) sekoskaita. Nors egzomas sudaro nedidelę (2 proc.) viso žmogaus genomo dalį, apie 85 proc. patogeninių variantų, lemiančių ligas, paveldimas pagal Mendelio dėsnius, nustatomi koduojančioje genomo dalyje. Dėl šios priežasties žymiai dažniau yra atliekama VES ar specifinių genų grupės sekoskaita nei VGS [83]. Siekdami palyginti VES bei su IN siejamų 500 genų grupės sekoskaitos technologijas, Sunas su kolegomis (2015), pritaikę abu metodus, ištyrė devynių tiriamųjų, turinčių IN, mėginius. Išanalizavus ir palyginus abiem metodais gautus duomenis, buvo nustatyti tie patys, galimai su patologija siejami, variantai [84]. Palyginti su VGS, VES yra pranašesnė dėl mažesnės kainos, didesnio DNR fragmentų padengimo bei analizuojamų duomenų apimties [83, 84]. Svarbu pažymėti, kad NKS taikymo galimybės nėra apribotos vien tik DNR sekos tyrimais. NKS metodais galima taip pat tirti transkriptomą (iRNR sekoskaita) ar net proteomą (*ChIP* sekoskaita) [36].

Nepriklausomai nuo taikomos NKS strategijos, išskiriami keli pagrindiniai sekoskaitos etapai: bibliotekos paruošimas, sekoskaita, vizualizavimas ir duomenų analizė [36]. Sekoskaitos principai skiriasi, priklausomai nuo komercinių NKS sistemų (2.2 lentelė). Šiuo metu plačiausiai naudojamas yra *Illumina* genetinis analizatorius, kuriuo atliekama sinteze pagrįsta sekoskaita (angl. *Sequencing by synthesis*) [36, 85].

## 2.2 lentelė. Populiariausios NKS sistemos [85]

Pavadinimas	Sekoskaitos principas	Gamintojas
<i>454/Roche</i>	Sekoskaita pagrįsta sinteze pagal pirofosfatų susidarymą, kitaip vadinama pirosekoskaita.	<i>454 Life Sciences</i> (JAV)
<i>Illumina/Solexa</i>	Sekoskaita pagrįsta sinteze.	<i>Illumina</i> (JAV)
<i>Ion Torrent</i>	Sekoskaita pagrįsta vandenilio jonų (H <sup>+</sup> ), kurie yra atpalaiduojami kaip pašalinis produktas DNR sintezės metu, nustatymu.	<i>Thermo Fisher Scientific</i> (JAV)
<i>HeliScope</i>	Vienos molekulės sekoskaita, kuri atliekama be amplifikacijos etapo.	<i>Helicos BioSciences</i> (JAV)
<i>MinION</i>	Vienos molekulės sekoskaita, kuri pagrįsta	<i>Oxford Nanopore</i>
<i>Nanopore sequencing</i>	DNR sekoskaita per nanoporą, įterptą į laidžią elektrai polimero membraną.	<i>Technologies</i> (Jungtinė Karalystė)
<i>10x Genomics</i>	Vienos ląstelės sekoskaita.	<i>NASDAQ</i> (JAV)

NKS, palyginti su pirmos kartos sekoskaitos metodu, yra pranašesnė atliekant didelio masto tyrimą, tačiau galimi klaidingai teigiami (angl. *False-positive*) rezultatai, todėl iki šiol Sangerio sekoskaitą rekomenduojama taikyti siekiant patvirtinti NKS rezultatus [86].

Genetiniai variantai, kurie yra siejami su tam tikrais klinikiniais požymiais ar ligomis, efektyviai nustatomi taikant plataus masto genomo sąsajos (angl. *Genome-wide association study*; GWAS) tyrimus [87]. Naudojant specifinius lustus su VNP, kurie yra plačiai paplitę visame genome, yra tiriamos sergančiųjų ir sveikųjų asmenų grupių genetinės sritys, susijusios su konkrečia tiriama patologija ar specifiniu fenotipiniu požymiu. Jei VNP yra statistiškai reikšmingai dažnesnis sergantiesiems, o tyrimo rezultatai atkartojami kituose tyrimuose, toks VNP vertinamas kaip susijęs su liga ar požymiu [36, 88].

Per pastaruosius dešimtmečius buvo gerokai patobulinti žmogaus genomo tyrimo metodai, todėl DNR sekos variantų ir KSP nustatymas tapo įprasta procedūra tiek mokslinėje, tiek medicininėje praktikoje. Pritaikius šiuos metodus labai išaugo tiriamų asmenų, įskaitant tuos, kurie turi IN ir IA, diagnostinis efektyvumas [75, 89].

### 2.2.1. *In silico* analizė

Nustačius genetinį variantą ar KSP, svarbu atlikti išsamią *in silico* analizę, apimančią įvairius bioinformacinius įrankius, duomenų bazes bei mokslinę literatūrą, siekiant įvertinti genetinio varianto ar KSP patogeniškumą ir įtaką fenotipui [20, 22].

Genetinių pokyčių įvertinimas yra ypač svarbus klinikinėje diagnostikoje, kadangi, literatūros duomenimis, mažiausiai 10 proc. patvirtintų ligų lemiančių patogeninių variantų iš tiesų neturi jokio poveikio fenotipui [14]. Siekiant išvengti tokių klaidų ir suvienodinti klinikinės varianto reikšmės analizės principus, ACMG sukūrė koduojančios ir nekoduojančios genomo sekos variantų bei KSP (žr. 2.1 skyrių) interpretavimo gaires ir kriterijus, pagal kuriuos genomo variantai turėtų būti priskirti vienai ar kitai kategorijai [20, 22].

DNR sekos variantai klasifikuojami į patogeninių, galimai patogeninių, nepatogeninių, galimai nepatogeninių ir neaiškios klinikinės reikšmės variantų kategorijas, atsižvelgiant į varianto tipą (pvz., rėmelio poslinkio, iRNR sukirpimo vietos variantas), poveikį aminorūgščių sekai (sinoniminis ar nesinoniminis variantas), kilmę (*de novo* ar paveldėtas), dažnį populiacijoje, vietą genome (pvz., karštajame mutacijų taške ar funkciškai svarbų baltymo domeną koduojančioje sekoje). Reti *de novo* DNR sekos variantai, kurie sukuria priešlaikinį baigmės kodoną, sudaro skaitymo rėmelio poslinkį ar keičia iRNR sukirpimo vietą, dažnai yra vertinami kaip patogeniniai [20]. KSP patogeniškumas vertinamas panašiai, papildomai atsižvelgiama į pokyčio persidengimą su polimorfinėmis sritimis ir patvirtintais ligų

lemiančiais variantais, genų skaičių analizuojamoje srityje, pokyčio dydį bei pobūdį (delecija ar duplikacija) [22, 33]. Dažniausiai patogeniniai KSP yra *de novo*, nepersidengiantys su polimorfinėmis sritimis, tačiau persidengiantys su žinomais ligą lemiančiais variantais bei genais [32].

Vertinant genetinių pokyčių patogeniškumą, rekomenduojama išnagrinėti mokslinę literatūrą ir duomenų bazes, kadangi analizuojamas variantas gali būti anksčiau nustatytas ir net aprašytas kitų tyrėjų, o didesnis skaičius aprašytų pacientų yra vienas svarbiausių kriterijų patvirtinant naujus kandidatinius genus.

Didžiausia ir populiariausia medicinos ir gyvybės mokslų publikacijų duomenų bazė yra *PubMed* [90]. Kliniškai reikšmingų variantų duomenų bazėse (pvz., *ClinVar* [8], *OMIM* [91], *HGMD* [34]) pateikiama informacija apie patogeninius variantus, kurie yra siejami su žmogaus paveldimomis ligomis. Patogeniškumui pagrįsti vienas svarbiausių kriterijų yra varianto dažnis populiacijoje, kuris gali būti patikrintas *dbSNP* [92], *dbVar* [93], *DGV* [94], 1 000 genomų projekto [6] duomenų bazėse ar *ENSEMBL* [95], *gnomAD* [96] genomo naršyklėse. Vienintelis mokslinių duomenų bazių trūkumas yra tas, kad nėra vykdoma skelbiamų variantų peržiūra pagal anksčiau išvardytus kriterijus, todėl galimi neteisingai klasifikuotų variantų atvejai. Dėl šios priežasties, naudojant duomenų bazes, svarbu įvertinti visą teikiamą informaciją: kokiais tyrimo metodais varianto patogeniškumas buvo pagrįstas, kaip dažnai duomenų bazės informacija yra atnaujinama ir kt. [20].

Siekiant supaprastinti variantų patogeniškumo vertinimą yra sukurti tiek komerciniai, tiek viešai prieinami internetiniai įrankiai: *SIFT* [97], *PolyPhen-2* [98], *MutationTaster2* [99], *PROVEAN* [100] ir kt. Taip pat yra sukurti bioinformaciniai įrankiai, skirti sukirpimo variantų padariniams prognozuoti. Plačiausiai naudojamas yra *Human Splicing Finder v.3.1* [101]. Naudojant, pavyzdžiui, *ExPASy* [102] įrankį yra galimybė ne tik palyginti tiriamąją genominę seką su referentine seka, bet ir paversti nukleotidų seką baltymų seka. Tokią baltymų seką, nuskaitytą pagal genominę seką su tiriamu patogeniniu variantu, patogu analizuoti vertinant patogeninio varianto poveikį baltymo struktūrai. Populiariausios duomenų bazės, kaupiančios informaciją apie baltymų struktūrą ir jų funkcijas, yra *UniProt* [103] bei *Pfam v 32.0* [104].

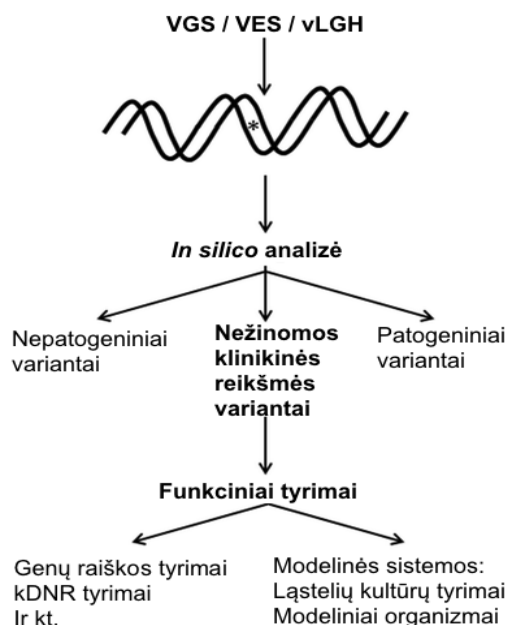
Vis daugėja mokslinių publikacijų apie *in silico* analizės klaidingai teigiamus rezultatus [105]. Pavyzdžiui, Miosge su kolegomis (2015) tirdami *de novo* taškinius variantus, lemiančius imuninės sistemos sutrikimus, nustatė, kad net 73 proc. analizuojamų variantų, anksčiau bioinformaciniais įrankiais įvertintų kaip patogeniniai, iš tiesų neturi jokio poveikio baltymo struktūrai ir limfocitų patogenezėi [106]. Be to, visi bioinformaciniai įrankiai skiriasi tarpusavyje priklausomai nuo sukurto algoritmo ar analizei pasirinkto



statistinio metodo, todėl vienu įrankiu variantas gali būti prognozuojamas kaip patogeninis, o kitu – nepatogeninis. Dėl šių priežasčių, atliekant *in silico* analizę, svarbu išsamiai įvertinti teikiamą informaciją ir vadovautis daugiau nei vienu įrankiu [107].

Paprastai, įvertinus genetinio pokyčio patogeniškumą, jo poveikį iRNR ir / ar baltymo struktūrai, toliau atliekama genotipo ir fenotipo ryšio analizė. Analizuojant KSP, yra identifikuojami asmenys su persidengiančiais struktūriniais pokyčiais bei yra lyginami jų klinikiniai požymiai. Viena populiariausių ir didžiausių tarptautinių duomenų bazių, kurioje yra pildoma informacija apie KSP ir fenotipinius požymius, – *DECIPHER* (angl. *Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources*) [108]. Tikslus klinikinių požymių, įskaitant IN ir / ar IA, apibūdinimas daugeliu atvejų padeda efektyviai identifikuoti genetinį sindromą. Šiam tikslui pasiekti yra sukurtos pažangios programinės įrangos (pvz., *Face2Gene* [109]) bei internetiniai įrankiai (pvz., *Phenolyzer* [110], *The Human Phenotype Ontology* [111]), kurie pagal klinikinius požymius ir / ar kandidatinius genus pateikia galimai susijusių genetinių sutrikimų ir sindromų sąrašą [112].

Bioinformaciniai įrankiai, duomenų bazės ir mokslinės literatūros analizė yra labai naudinga vertinant genominių pokyčių patogeniškumą ar poveikį sukirpimo mechanizmui bei atliekant genotipo ir fenotipo ryšio analizę. Nepaisant to, tai yra sudėtingas procesas, kadangi dažniausiai ši analizė atliekama skirtingais būdais priklausomai nuo konkretaus klinikinio atvejo [22]. Kai kuriais atvejais, atlikus išsamią *in silico* analizę, surenkama pakankamai įrodymų, pagrindžiančių varianto patogeniškumą ir jo įtaką fenotipui. Visgi dažnai tai yra tik tarpinis etapas, po kurio būtina atlikti funkcinius eksperimentus, siekiant įvertinti nežinomos klinikinės reikšmės varianto patogeniškumą ar patvirtinti tyrimo rezultatus (2.5 pav.) [82, 113].



**2.5 pav.** Pagrindinė genetinių pokyčių tyrimo strategija. Schema adaptuota pagal mokslinę literatūrą [113]

### 2.2.2. Funkciniai tyrimo metodai

Daugybė molekulinį mechanizmų, tokių kaip transkripcijos ir translacijos procesų reguliavimas, *post*transliacinės modifikacijos, iRNR arba baltymo transportavimas ir suardymas, veikia DNR-baltymo kelyje. Be abejonės, dar daugiau tokių biologinių mechanizmų veikia DNR-organizmo kelyje. Įvykus patogeniniam pokyčiui DNR sekoje, sutrinkdama biologinių tarpusavyje susijusių veiksmų grandinė. Dėl šios priežasties nežinomos klinikinės reikšmės genetinių pokyčių bei genotipo ir fenotipo ryšio charakterizavimui turėtų būti taikomi eksperimentiniai genomo, epigenomo, transkriptomo, proteomo, metabolomo ar net interaktomo tyrimai, apimantys ne tik skirtingus funkcinius tyrimo metodus, bet ir skirtingas gyvūnų ar ląstelių kultūrų modelines sistemas [114, 115].

XXI amžius kitaip vadinamas technologijų kūrimo amžiumi, todėl nenuostabu, kad genetikos mokslo srityje taip pat gausu įvairių technologijų, skirtų atlikti išsamiai molekulinei ir funkcinei genetinių pokyčių analizei. Taikant pirmos / antros kartos sekoskaitos, LGH ar GWAS tyrimų technologijas, efektyviai nustatomi DNR sekos variantai ir / ar KSP [75, 89]. Pavyzdžiui, Huo su bendradarbiais (2019), atlikę GWAS, nustatė 180 genetinių sričių, siejamų su šizofrenija. Taip pat mokslininkai nustatė 132

VNP, dėl kurių sutriko transkripcijos veiksnių jungimosi procesas, bei 97 VNP, kurie sutrikdė genų raišką, vykstančią žmogaus smegenyse [116].

Klasikiniu ir plačiausiai taikomu specifinių genų raiškos tyrimo metodu laikoma kiekybinė tikro laiko polimerazės grandininė reakcija (kPGR; angl. *Quantitative real-time polymerase chain reaction*; plačiau 2.2.2.1 skyriuje) [117]. O visas genų raiškos profilis paprastai yra tiriamas atliekant RNR sekoskaitą taikant NKS technologijas. Šis metodas ne tik suteikia informaciją apie viso transkriptomo raišką, bet ir nustatomi nauji transkriptai ar sukirpimo variantų pasekmės transkriptomo lygmeniu [115]. Šia technologija ištyręs maždaug 1 000 asmenų iš skirtingų audinių išgrynintą RNR, įskaitant iRNR, rRNR, tRNR ir net lncRNR, Genotipo-audinių raiškos (angl. *Genotype-Tissue Expression*; GTEx) konsorciumas gavo duomenis apie audiniui specifinę genų raišką, kuri yra kaupiama ir viešai prieinama GTEx duomenų bazėje [118]. Sėkmingo RNR sekoskaitos pritaikymo pavyzdys yra Cummings ir bendradarbių (2017) atliktas tyrimas. Ištyrus 50 asmenų, turinčių raumenų įgimtas anomalijas, ir išanalizavus gautus duomenis, net 35 proc. tiriamųjų buvo nustatyta ir patvirtinta klinikinė diagnozė. Svarbu pažymėti, kad nežinomų atvejų bendras genetinis diagnostinis efektyvumas yra 25–50 proc., todėl šis tyrimas patvirtina RNR sekoskaitos tinkamumą tiriant nežinomos klinikinės reikšmės variantus [119].

Genų raiškos pokyčiai galimi ne vien tik įvykus patogeniniam variantui ar KSP baltymą koduojančiose ar reguliuojančiose sekose. Cheminės modifikacijos, tokios kaip DNR metilinimas bei histonų kovalentinės modifikacijos, reguliuoja genų raišką be pasikeitimų pirminėje DNR sekoje. Tokie genų raiškos pokyčiai gali būti efektyviai ištirti epigenetiniais tyrimo metodais identifikavus pakitusias sritis. Pavyzdžiui, chromatino imunoprecipitacijos metodu nustatomos DNR sekos, kuriose yra išsidėsčiusios nukleosomos, histonų variantai arba vyksta histonų kovalentinės modifikacijos, o DNR imunoprecipitacijos metodu, kuris pagrįstas antikūnų, atpažįstančių metilintą DNR seką, prisijungimu, nustatomos metilinimui jautrios DNR sekos [120].

Daugeliu atvejų transkriptomo tyrimai yra labai informatyvūs ir pakankami klinicinei diagnozei patvirtinti. Tačiau baltymų struktūrai ir funkcijoms įtakos gali turėti ne tik DNR sekos pakitimai, bet taip pat ir sutrikę *posttransliaciniai* reguliaciniai mechanizmai, iRNR sukirpimo mechanizmas, baltymų modifikacijos (pvz., fosforilinimas, ubikvitilinimas, ribozilinimas) ar netipinės baltymų tarpusavio sąveikos, todėl dažnai būtina įvertinti pokyčius baltymų lygmeniu [121].

Priklausomai nuo tyrimo tikslo, žinomos dvi pagrindinės baltymų tyrimų strategijos. Viena iš tyrimo strategijų yra pagrįsta imuninės reakcijos,

kurioje dalyvauja antigenas ir antikūnas, principu. Pavyzdžiui, imunofermentiniu *ELISA* (angl. *enzyme-linked immunosorbent assay*) metodu, kuris pasižymi aukštu jautrumu ir specifiskumu, kiekybiškai vertinami baltymai, antikūnui atpažinus ir prisijungus prie baltymo ligando [122]. Kitas plačiai naudojamas ir taip pat antikūnų principu pagrįstas metodas yra *Western blotting* (arba trumpai *Western blot*). Tai pusiau kiekybinis tyrimo metodas, kuris susideda iš dviejų pagrindinių etapų. Pirmajame etape vykdoma baltymų elektroforezė denatūruojančiomis sąlygomis. Išsifractionavusius baltymus pernešus ant specifinės membranos (pvz., nitroceliuliozinė), baltymai inkubuojami su pirminiais antikūnais. Antrajame etape antriniai antikūnai inkubuojami su pirminiais, kurių sąveikos yra vizualizuojamos imunofluorescencine arba chemiliuminescencine mikroskopija [123].

Kai kuriais atvejais nėra galimybės sukurti antikūno tiriamam baltymui, todėl yra sukurta kita baltymų tyrimo strategija be antikūnų naudojimo. Tokių metodų pavyzdys yra masės spektrometrija, kurios pagrindinis principas – įkrautos molekulės masės ir krūvio santykio matavimas. Masės spektrometrija, naudojama kartu su, pavyzdžiui, dujų ar skysčių chromatografija, yra tinkama technologija tirti ne tik proteomą, bet ir metabolomą, todėl šis metodas ypač dažnai taikomas paveldimų medžiagų apykaitos ligų diagnostikoje [124, 125]. Taip pat masės spektrometrija neretai yra naudojama kartu su imunoprecipitacijos technologijomis. Šių metodų kombinacija yra efektyvi tiriant baltymų sąveikas – interaktomą [126].

Šiuo metu egzistuoja didelis metodų, tinkančių tiksliai ir išsamiai funkinei genetinei analizei atlikti, pasirinkimas. Kiekvienas tyrimo metodas turi privalumų ir trūkumų, todėl, prieš pasirenkant metodą funkinei analizei, svarbu įvertinti ne tik tyrimo tikslą ir pobūdį, bet ir metodo apribojimus. Tinkamai pasirinktas tyrimo metodas padeda atlikti išsamią funkcinę analizę ir suprasti biologinius tarpusavyje susijusius mechanizmus tiek ląstelių, tiek organizmo lygmeniu [114, 115]. Šiame disertaciniame darbe nežinomos klinikinės reikšmės variantai buvo tiriami iRNR lygmeniu, todėl detaliau apžvelgiami žmogaus ląstelių iRNR funkciniai tyrimo metodai.

#### 2.2.2.1. Informacinės RNR tyrimai

Vis daugėja mokslinių publikacijų, rekomenduojančių nustatytus koduojančios ir nekoduojančios DNR sekos pokyčius patvirtinti iRNR tyrimais. Dėl cheminių savybių iRNR, palyginti su DNR, yra mažiau stabili ir labiau jautri hidrolizei. Taip pat iRNR būdingas didelis koncentracijų variabilumas tarp skirtingų tiriamųjų mėginių, kadangi iRNR koncentracijai

turi įtakos daugelis veiksnių, pavyzdžiui, tiriamo asmens lytis, amžius, sveikatos būklė, fizinis krūvis ar audinys, iš kurio gryninama iRNR. Tačiau, atvirkštinės transkriptazės PGR metodu susintetinus kDNR pagal matricinę RNR, šių problemų išvengiama ir tolesniems tyrimams dažniau naudojama būtent kDNR [15, 38, 48, 127].

Kiekybinė tikro laiko PGR (kPGR) technologija, pasižyminti dideliu jautrumu, yra laikoma „auksiniu standartu“ vertinant specifinių genų raišką ar KSP. kPGR yra klasikinės polimerazės grandininės reakcijos modifikacija, kurios metu ne tik padauginamas tikslinis DNR ar kDNR fragmentas, bet ir matuojamas fluorescencijos signalas, kurio intensyvumas yra tiesiogiai proporcingas susintetinto produkto kiekiui. Fluorescencijos signalą skleidžia naudojami fluorescuojantys žymenys: į dvigrandę DNR interkalijuojantys dažai (pvz., *SYBR Green*) arba taikiniui specifiniai žymenys su fluoroforais (pvz., *TaqMan*), todėl tikru laiku stebimas fluorescencijos signalo kitimas, pagal kurį kiekybiškai nustatomas pradinis DNR ar kDNR kiekis reakcijoje. Kiekybinei analizei taikomos dvi strategijos: santykinė (angl. *Relative quantification*) ir absoliuti (angl. *Absolute quantification*). Taikant absoliutųjį metodą, KSP ar genų raiškos pokytis vertinamas pagal kalibracinę kreivę (angl. *Standard curve*), kuri sudaroma atliekant serijinius skiedimus mėginio, kuriame yra žinomas tiriamo DNR ar kDNR fragmento kiekis. O santykiniu (kitaip vadinamu  $\Delta\Delta Ct$ ) metodu KSP ar geno raiškos pokytis vertinamas lyginant su kontroliniu (referentiniu) mėginiu ir normalizuojant pagal namų ruošos geną (angl. *Housekeeping gene*). Tiriamo asmens geno raiškos pokytis vertinamas pokyčio kartais (angl. *Fold change*; FC), t. y. kiek kartų tiriamajame mėginyje KSP ar geno raiška skiriasi, lyginant su kontroliniu mėginiu. FC apskaičiuojamas pagal formulę:

$$\Delta Ct_1 = Ct(\text{tiriamo geno}) - Ct(\text{normalizuojančio geno})$$

$$\Delta Ct_2 = Ct(\text{tiriamo geno}) - Ct(\text{normalizuojančio geno})$$

$$\Delta\Delta Ct = \Delta Ct_1 - \Delta Ct_2$$

$$FC = 2^{\Delta\Delta Ct}, \text{ kur}$$

Ct (angl. *threshold cycle*) – ciklas, kuriame mėginio fluorescencijos lygis reikšmingai viršija bazinį, pradinių reakcijų ciklą, fluorescencijos lygį.  $\Delta Ct_1$  – tiriamo asmens Ct,  $\Delta Ct_2$  – kontrolinio asmens Ct [117, 128, 129].

Priklausomai nuo išorinių veiksnių (pvz., lytis ar amžius) iRNR būdingas didelis išeišos variabilumas tarp skirtingų tiriamųjų. Todėl atliekant kPGR neretai susiduriama su problema, kaip tinkamai suplanuoti tyrimą ir patikimai interpretuoti gautus rezultatus. Siekdami išvengti šių problemų bei norint užtikrinti kPGR tyrimų efektyvumą, atkartojamumą ir nuoseklumą, Bustinas su kolegomis (2009) paskelbė *MIQE* (angl. *Minimum Information for Publication of Quantitative Real-Time PCR Experiments*) gaires, kuriomis

rekomenduojama vadovautis atliekant kPGR eksperimentus ir skelbiant gautus rezultatus [130, 131].

kPGR – patikimas ir efektyvus metodas KSP ar genų raiškos analizei, tačiau šiuo metodu yra labai ribotos galimybės iširti sukirpimo variantų pasekmes iRNR lygmeniu. Remiantis moksline literatūra, vienas iš efektyvių būdų tirti tokius variantus yra kDNR sekoskaita Sangerio metodu. Dirbant pagal šią metodiką, ypač svarbu užtikrinti, kad tiriamajame mėginyje nebūtų genominės DNR priemaišų, priešingu atveju tai kliudys gauti patikimus rezultatus [38, 48]. Šią tyrimo strategiją sėkmingai pritaikė van Haelst su kolegomis (2015) tirdami *de novo* akceptorinės sukirpimo vietos variantą c.480-1G>T *MED13L* (MIM #608771) gene. Mokslininkai nustatė, kad dėl kriptinės alternatyvios kirpimo vietos penktas egzonas sutrumpėjo 45 nt, o baltymo lygmeniu nebuvo transliuojama 15 aminorūgščių [132]. Panašiai Fusco ir kt. (2019) ištyrė du asmenis, turinčius introninius sukirpimo vietos variantus c.6872-24T>A ir c.7571-12T>A *FBNI* (MIM #134797) gene. Atlikus kDNR sekoskaitą Sangerio metodu, pirmuoju atveju buvo nustatytas 22 nt ilgio introno fragmento intarpas į gretimą 57 egzoną, o antruoju atveju buvo nustatytas kito 10 nt ilgio introninio fragmento intarpas į gretimą 62 egzoną. Abiem atvejais dėl kriptinės alternatyvios sukirpimo vietos susidarymo variantai lėmė rėmelio poslinkį ir priešlaikinio baigmės kodono susidarymą, dėl kurių susidarė trumpesni baltymai [133].

Taigi, molekuliniai ir funkciniai iRNR (ar kDNR) tyrimai yra reikalingi siekiant suprasti tiriamo varianto etiopatogenezės mechanizmus iRNR lygmeniu. O šiuo metu taikomi modernūs genų terapijos metodai ir besivystančios genomo redagavimo technologijos suteikia galimybę šiuos mechanizmus išsamiau išanalizuoti *in vitro* ląstelių kultūrose ar *in vivo* modeliniuose gyvūnuose, tokiu būdu atspindint analizuojamą žmogaus fenotipą [131, 134].

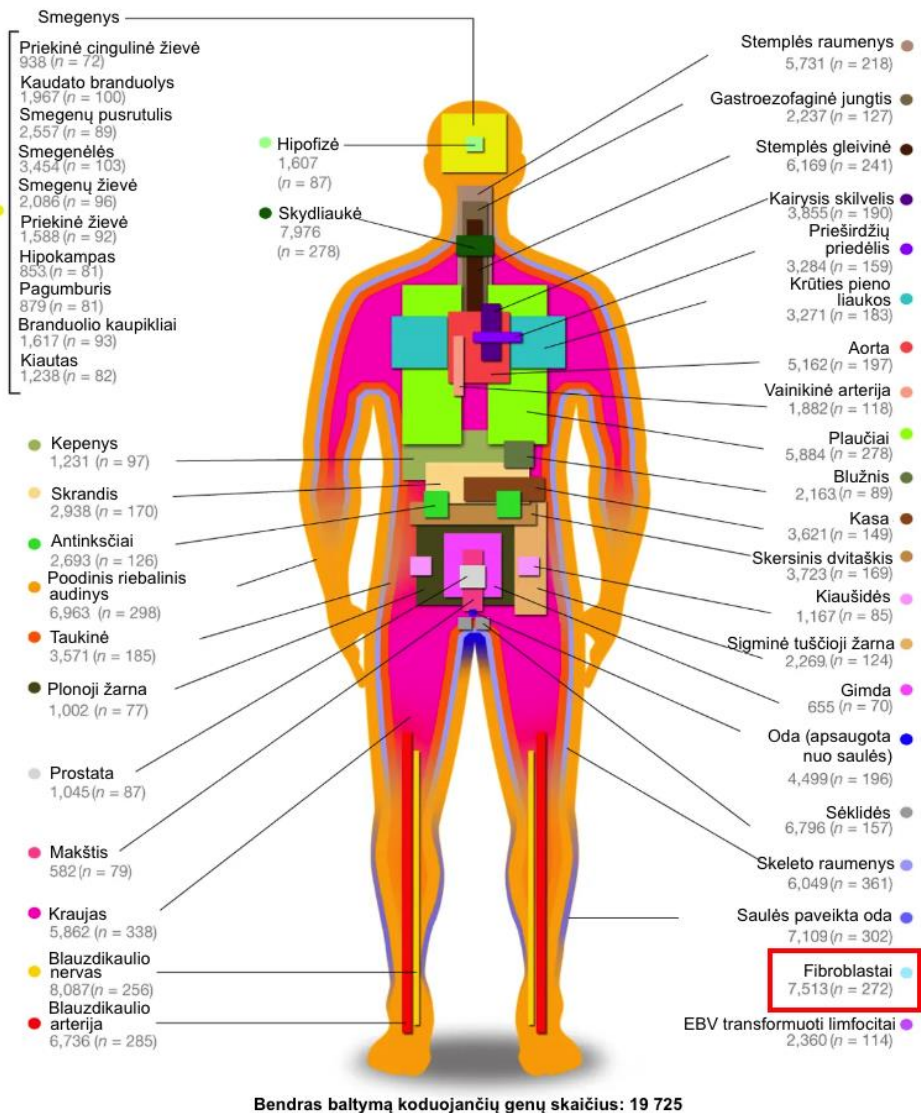
#### 2.2.2.2. Ląstelių kultūrų ir modelinių organizmų panaudojimas funkciniais tyrimams

Ląstelių kultūrų ar modelinių organizmų eksperimentai yra seniai taikomi, siekiant nustatyti biologinius ir patologinius mechanizmus, taip pat sukurti efektyvius gydymo būdus. Pasirenkant modelinį organizmą, atsižvelgiama į kelis kriterijus. Visų pirma, organizmas turi būti nepavojingas, lengvai išlaikomas ir prižiūrimas laboratorijoje, o vienos generacijos laikas turi būti trumpas. Taip pat turi būti lengvai sukeliama genomo modifikacijos bei nustatomos grynavislės homozigotinės organizmo linijos [135].

Žmogaus fenotipinių požymių modeliavimui plačiausiai naudojamos pelės (lot. *Mus Muculus*), žiurkės (lot. *Rattus norvegicus*) ir zebražuvės (lot. *Danio rerio*). Dažniausiai patogeniniai variantai lemia geno funkcijos praradimą, todėl šiuose modeliniuose organizmuose, taikant įvairius mutagenezės mechanizmus, patogų užslopinti (angl. *Knock-out*) tiriamo ligos geno raišką ir stebėti tokios modifikacijos poveikį fenotipo lygmeniu. Pavyzdžiui, N-etil-N-nitrozokarbamidu chemiškai indukuojant atsitiktinius DNR sekos taškinis variantus yra kuriamos transgeninių pelių linijos, atspindinčios tiriamą žmogaus požymį ar ligą. Nors transgeninės pelės turi nemažai privalumų, jų didžiausias trūkumas yra tas, jog indukuoti variantai ne visada nulemia identiškus fenotipinius požymius, kurie būdingi tiriamam asmeniui. Todėl alternatyvus funkcinių tyrimų būdas yra morfolino-fosforodiamidatu modifikuotų priešprasmų oligonukleotidų (angl. *Morpholino phosphorodiamidate anti-sense oligonucleotides*; MPO) panaudojimas zebražuvėje. Šiems modifikuotiems oligonukleotidams hibridizuojantis su tiksline iRNR yra nuslopinama baltymo transliacija modeliniame organizme [136]. Pastaruoju metu Wortmann su kolegomis (2015) sėkmingai pritaikė šią strategiją nežinomos klinikinės reikšmės varianto *CLPB* (MIM # 616254) gene patogeniškumui įvertinti. Nustatyta, kad MPO injekcija zebražuvėje sukėlė akivaizdžias smegenėlių anomalijas, kurios taip pat buvo stebimos tiriamiesiems asmenims [137].

Gyvūnų, taip pat kaip ir žmogaus, funkciniai tyrimai kelia daug bioetinių klausimų. Dėl to nemaža dalis žmonių prieštarauja *in vivo* eksperimentams su modeliniais gyvūnais. Tačiau šiuo metu vis daugiau *in vitro* mokslinių tyrimų atliekama su žmogaus ląstelių kultūromis. Pastaruoju metu ypač išpopuliarėjo trijų dimensijų (3D) ląstelių kultūrų kultivavimo sistemos, kuriose vos per porą dienų ląstelių kultūra gali būti padauginta iki kliniškai reikšmingo ląstelių skaičiaus. Taip pat, palyginti su įprastomis dviejų dimensijų ląstelių kultivavimo sistemomis, 3D sistemos biologiniu ir biocheminiu atžvilgiu labiau atspindi žmogaus *in vivo* mikroaplinką. Dėl tokių privalumų *in vitro* žmogaus ląstelių kultūrų funkciniai tyrimai yra efektyvus bioetinių klausimų dėl modelinių gyvūnų sprendimo būdas [134]

Priklausomai nuo tyrimo tikslo ir genų raiškos specifiniame audinyje, funkciniais tyrimams gali būti naudojamos įvairios ląstelių kultūros. Patraukliausia moksliniams tyrimams yra fibroblastų kultūra, kadangi *GTEX* projekto metu nustatyta, kad šiose ląstelėse transliuojama daugiausiai (7 513) genų (2.6 pav.).



**2.6 pav.** *GTEX* projekto metu tirta baltymus koduojančių genų raiška specifiniuose žmonių audiniuose ir ląstelių linijose. Žymėjimas: n – imties dydis, EBV – Epsteino–Barro virusas. Adaptuota pagal mokslinę literatūrą [138]



Be to, pirminė fibroblastų kultūra gaunama iš odos bioptato, kuris paimamas nesudėtingos ir mažai invazinės procedūros metu, o *in vitro* kultivavimas yra nesudėtingas procesas. Fibroblastai yra taip pat svarbūs reguliuojant įvairius fiziologinius procesus: tarpląstelinės medžiagos (angl. *extracellular matrix*) sintezę, augimo veiksnių sekreciją, uždegimines reakcijas bei žaizdų gijimą [138, 139].

Nepaisant to, tam tikrų ligos genų raiška vyksta tik specifiniuose audiniuose, pavyzdžiui, smegenyse ar širdies raumenyje, todėl funkciniais genetiniais tyrimais ištirti šiuos genus ypač sudėtinga, o dažnai iš viso neįmanoma [113]. Dėl šios priežasties funkciniais tyrimams tiktų embrioninės kamieninės ląstelės (EKL), kurios pasižymi dideliu proliferacijos potencialu ir galimybe diferencijuotis į visų tipų ląsteles. Tačiau tokie tyrimai yra ribojami dėl akivaizdžių moralinių bei etinių aspektų. Siekdami išvengti šių apribojimų Takahashi ir Yamanaka (2006) virusinės transformacijos metodu perprogramavo žmogaus somatinės ląstelės ir tokiu būdu sukūrė indukuotas pluripotentes kamienines ląsteles (iPKL; angl. *induced pluripotent stem cells*), kurių proliferacijos ir diferenciacijos potencialas, panašiai kaip EKL, suteikia galimybę tirti skirtingų tipų ląsteles, įskaitant neuronus ar kardiomiocitus [140, 141]. Pavyzdžiui, Romero-Moya su kolegomis (2017) transformavo tiriamų asmenų fibroblastų kultūrą į iPKL, siekdami ištirti DNR variantą *COQ4* (MIM #612898) gene, kuris galimai lėmė kofermento Q10 trūkumą (MIM #616276). Tiriant transformuotą iPKL kultūrą buvo stebimi mitochondrijų anomalijų požymiai, kuriuos nulemia kofermento Q10 trūkumas. Be to, pritaikę CRISPR-Cas9 genomo redagavimo sistemą (žr. 2.2.2.3 skyrių), autoriai ištaisė genetinį variantą *COQ4* gene ir tokiu būdu sėkmingai atkūrė iPKL biocheminį fenotipą [142].

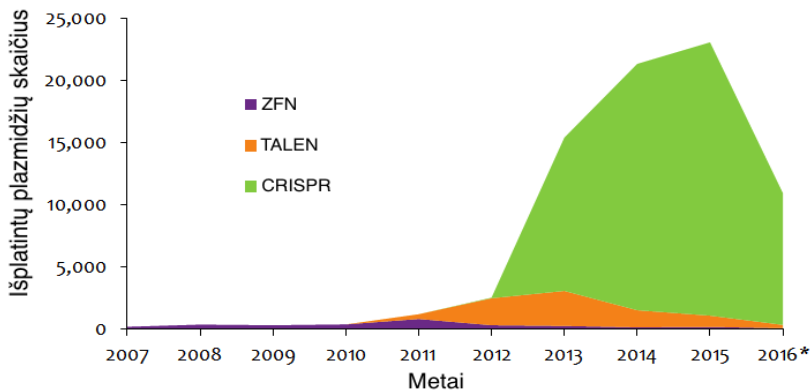
Taigi, funkciniai genetiniai tyrimai modelinėse sistemose yra reikšmingas ir efektyvus būdas charakterizuoti nežinomos klinikinės reikšmės koduojamos ir nekoduojamos genomo sekos variantus. Todėl publikuojama vis daugiau mokslinių straipsnių, teigiančių, jog DNR sekos variantų patogeniškumui įvertinti turi būti taikomi funkciniai tyrimai, atveriantys naujas diagnostikos ir gydymo galimybes medicininėje praktikoje [134].

### 2.2.2.3. CRISPR-Cas9 genomo redagavimo technologija

Genomo redagavimas (angl. *Genome editing*) – tai daug moksliniams tyrimams žadanti technologija, kuri skirta gyvo organizmo DNR sekai įterpti, ištrinti arba pakeisti. Genomo modifikacijos yra atliekamos nuo 1979 metų, kuomet ši technologija pirmą kartą buvo pritaikyta mielėms [143].

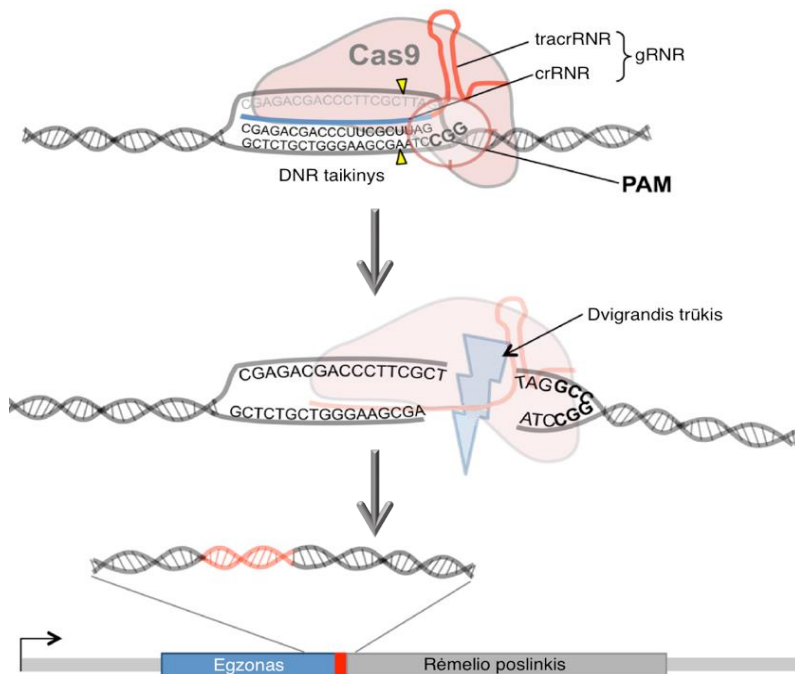
Žinoma įvairių genų redagavimo strategijų. Pavyzdžiui, patogiai ir greitai genai gali būti nutildomi panaudojus siRNR, tačiau tai yra nestabili modifikacija, ribojanti galimybes tirti ilgalaikį poveikį modelinėse sistemose. Šių apribojimų galima išvengti panaudojus kitas technologijas, tokias kaip transkripcijos aktyvintojus primenančias efektorių nukleazes (angl. *Transcription activator-like effector-based nucleases*; TALEN), cinko pirštelių nukleazes (angl. *Zinc-finger nucleases*; ZFN) arba meganukleazes. Šiomis nukleazėmis gana tiksliai ir efektyviai modifikuojami genai, tačiau kiekvienam DNR taikiniui reikia specifinės nukleazės, kurios sukūrimas, sintezė ir patvirtinimas reikalauja daug darbo ir laiko [134, 144].

Revoliucionuojanti medicinos mokslą ir vis plačiau naudojama (2.7 pav.) yra CRISPR (angl. *Clustured Regularly Interspaced Short Polindromic Reapets*) sistema, kuri *Escherichia coli* genome buvo atrasta maždaug prieš 30 metų, tačiau funkcijos ilgą laiką buvo nežinomos. Tik 2011 metais lietuvių mokslininkų grupės buvo nustatyta, kad CRISPR-Cas9 sistemą galima užprogramuoti kirpti DNR seką norimoje vietoje [145], todėl ilgai netrukus ši technologija *Nature* žurnale buvo pripažinta 2011 metų metodu [146].



**2.7 pav.** Addgene ZFN, TALEN ir CRISPR plazmidžių pasiskirstymas pagal metus [147]

CRISPR-Cas (angl. *CRISPR-associated Cas protein*) – tai bakterijų apsaugos nuo virusų sistema, kuri yra sudaryta iš specifinės DNR sekos (CRISPR lokuso) ir Cas baltymų. Moksliniais tyrimais nustatyta, kad modelinėse sistemose panaudojus DNR taikiniui specifinę gidinę RNR (gRNR), sudarytą iš crRNR (angl. *CRISPR-derived RNA*) ir tracrRNR (angl. *trans-activating crRNA*), galima užprogramuoti Cas9 baltymą kirpti DNR seką norimoje vietoje. crRNR atpažįsta tiriamos DNR sekos PAM (angl. *protospacer-adjacent motif*) sritį, o tracrRNR sąveikauja su Cas9 baltymu. Ribonukleoproteino kompleksui kerpant tikslinę DNR seką sudaromas dvigrandis trūkis. Tokiu būdu tikslinis DNR sekos fragmentas gali būti iškirptas (angl. *Knock-out*) arba, įterpiant donorinės DNR fragmentą, pataisytas (angl. *Knock-in*). Susidariusį dvigrandį trūkį užtaiso nehomologinis galų sujungimo arba, rečiau, homologinis rekombinacinis taisymo mechanizmas. Tais atvejais, kai dėl įvykusios vieno nukleotido ar DNR sekos fragmento insercijos ar delecijos susidaro rėmelio poslinkis, geno koduojamo produkto funkcija yra prarandama (angl. *Loss-of-function*) (2.8 pav.). Jei rėmelio poslinkis DNR sekoje visgi neįvyksta, gali pakisti geno koduojamo produkto funkcija (angl. *Gain-of-function*) [148, 149].



**2.8 pav.** CRISPR-Cas9 geno redagavimo sistemos schema. Adaptuota pagal mokslinę literatūrą [148]

CRISPR-Cas9 sistema vis plačiau taikoma modelinėse sistemose, siekiant patvirtinti nežinomos klinikinės reikšmės DNR varianto poveikį fenotipui. Pavyzdžiui, redaguotų ląstelių klonai naudojami geno haplonepakankamumo funkciniam poveikiui įvertinti transkriptomo, proteomo, ląstelės ar net viso organizmo lygmeniu. Tačiau dažnai susiduriama su įvairiais techniniais sunkumais klonuojant redaguotas ląsteles, todėl šis tyrimas turi būti ypač gerai optimizuotas [150]. Norint įvertinti varianto patogeniškumą ir įtaką fenotipui, taip pat sėkmingai yra naudojami gyvūnų modeliai. Pavyzdžiui, zebražuvės buvo pritaikytos lietuvių kilmės sibsų, kuriems pasireiškė IN ir IA, tokios kaip hipotonija, epilepsija, smegenų ir kitos anomalijos, funkciniam tyrimams. VES metodu buvo nustatyti du sudėtiniai (angl. *Compound*) heterozigotiniai variantai *KIAA1109* gene (MIM #611565), kuris anksčiau nebuvo siejamas su jokia liga. Todėl, siekiant patvirtinti geną kandidatą, į tyrimą įtraukta daugiau tiriamųjų, turinčių panašių klinikinių požymių ir genetinių variantų tiriamajame gene. Naudojant MPO ir taikant CRISPR-Cas9 genomo redagavimą buvo sukurtos *kiaa1109* zebražuvių linijos, kurioms buvo stebimas išlenktas kūnas, hidrocefalija ir kiti galvos smegenų defektai. Zebražuvių fenotipas atspindėjo tiriamų asmenų klinikinius požymius, todėl šio tyrimo rezultatai suteikė galimybę apibūdinti naują sindromą, kuris buvo pavadintas Alkurayos–Kučinsko sindromu (MIM #617822) [151].

Kadangi moksliniais tyrimais patvirtinta, kad modelinėse sistemose šis mechanizmas efektyviai veikia, dėl savo unikalių savybių CRISPR-Cas9 sistema yra daug žadanti medicininėje praktikoje. Remiantis *ClinicalTrials.gov* [152] duomenų baze, yra vykdomi klinikiniai tyrimai naudojant šią technologiją. Oficialiai užregistruota 19-ka tokių tyrimų, iš kurių septyni yra pasiekę antrąją klinikinių tyrimų stadiją. Klinikinių tyrimų duomenimis, bandoma gydyti įvairias ligas: žmogaus imunodeficito viruso sukeltą infekciją, pjautuvinę anemiją, hematologinius ir onkologinius susirgimus [153]. 2020 metais pirmą kartą tiesiogiai pritaikyta CRISPR-Cas9 genų terapija asmeniui, sergančiam įgimta Lėberio amauroze (angl. *Leber's congenital amaurosis 10*; MIM #611755). Taikomos genų terapijos metu tikimasi pašalinti *CEP290* (MIM #610142) patogeninį variantą, kuris, manoma, akies tinklainėje lemia rodopsino sintezės sutrikimus [154, 155].

Visuotinai suprantama, kad visos technologijos turi privalumų ir trūkumų, ne išimtis yra ir CRISPR-Cas9 sistema. Svarbiausi šios sistemos trūkumai – redagavimo efektyvumas ir kirpimas šalutinėse vietose (angl. *Off-target*). Kirpimui šalutinėse vietose patikrinti rekomenduojama taikyti NKS arba specifinę neatitikčių tyrimo nukleazėmis metodą (angl. *mismatch-detection nuclease assay*) [156]. Dedama daug pastangų CRISPR-Cas9

metodo efektyvumui ir specifiškumui padidinti. Pavyzdžiui, mokslinio tyrimo metu nustatyta, kad šio metodo efektyvumas priklauso nuo gRNR GC sąsąto. Nustatyta, kad sistema redaguoja taikinį didžiausiu efektyvumu, jei GC sudaro apie 65 proc. gRNR sekos [157].

Taigi, CRISPR-Cas9 technologija leidžia žengti į naują medicinos erą, todėl šiuo metu sprendžiami su tuo susiję bioetiniai klausimai. Nepaisant to, tikimasi, kad artimoje ateityje šiuo tyrimo metodu bus efektyviai taisomi pažeisti genai ir gydomos genetinės ligos [134, 158].

### 2.3. Intelektinės negalios ir įgimtų anomalijų genetinė architektūra ir genomika

Pastaruosius dešimtmečius plėtojantis žmogaus genomo tyrimo technologijoms ir galimybėms jas eksploatuoti, buvo atlikta daug didelės apimties tyrimų, kurie padarė proveržį nustatant įvairių sveikatos sutrikimų ir ligų genetinę architektūrą [75, 89]. Viena svarbiausių socialinių ir ekonominių asmens sveikatos priežiūros problemų visame pasaulyje yra IN, kuri dažnai pasireiškia kartu su ĮA. Dėl tokios heterogeninės IN / ĮA prigimties šią patologiją lemiančių genetinių priežasčių nustatymas yra itin sudėtingas [1–4].

IN lemia įvairūs patogeniniai variantai: tiek taškiniai variantai, tiek chromosomų skaičiaus ir struktūros pokyčiai, kurie, manoma, sudaro apie 15 proc. visų atvejų. Panašus skaičius taip pat nustatomas patogeninių KSP. Nepaisant to, dažniausiai IN siejama su taškiniais variantais [60]. Atliekant paiešką žmogaus variantų duomenų bazėje *ClinVar* [8] pagal raktinius žodžius „*intellectual disability*“, nurodomi 10 092 patogeniniai variantai, siejami su IN. Dažniausi patogeniniai variantai yra rėmelio poslinkio (26 proc.), *missense* (25 proc.) bei *nonsense* (21 proc.) variantai. Rečiau nustatomi variantai nekoduojamoje DNR sekoje: sukirpimo vietose (7 proc.), 5' ir 3' netransliuojamose vietose (6 proc.), taip pat nekoduojančios RNR srityse (5 proc.) [9].

#### 2.3.1. Intelektinės negalios samprata ir klasifikacija

Nepaisant spartaus progreso medicinos ir sveikatos mokslų srityje, neuropsichiatrinių sutrikimų etiologija bei diagnostika vis dar kelia daug klausimų. Šių sutrikimų grupei priklauso intelektinė negalia (IN; angl. *Intellectual disability*), anksčiau vadinta protiniu atsilikimu. IN diagnozuojama asmenims iki 18 metų, kuriems būdingas intelekto

koeficientas (angl. *Intelligence quotient*; IQ) yra mažesnis nei 70 ir kartu pasireiškia kognityvinių funkcijų bei adaptyvaus elgesio sutrikimai [159, 160].

Pagrindiniai IN klasifikatoriai yra Pasaulio sveikatos organizacijos (PSO; angl. *World Health Organization*) sudaryta Tarptautinė statistinė ligų ir sveikatos sutrikimų klasifikacija (dešimtas leidimas, Australijos modifikacija; TLK-10-AM) ir Amerikos psichiatrų asociacijos patvirtintas Psichikos sutrikimų diagnostikos ir statistikos vadovas (penktasis leidimas; DSM-V) [159] (2.3 lentelė). Atsižvelgiant į klinikinius požymius yra skiriama sindrominė ir nesindrominė IN. Dažniausiai pasireiškia sindrominė IN, pavyzdžiui, Dauno (angl. *Down syndrome*; MIM #190685; ORPHA #870) ar lūžusios X chromosomos (angl. *Fragile X syndrome*; MIM #300624; ORPHA #908) sindromai, kuriems yra būdingi specifiniai neurofiziologiniai požymiai bei elgesio sutrikimai. Rečiau diagnozuojami nesindrominiai IN atvejai. Tokiems asmenims nepasireiškia kiti požymiai, išskyrus intelekto sutrikimą [161, 162].

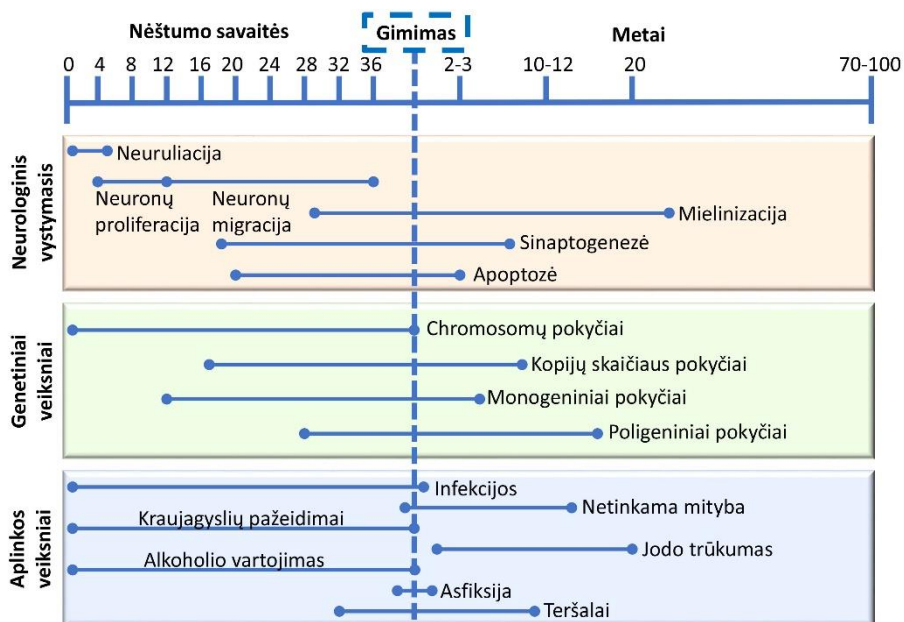
**2.3 lentelė.** IN laipsniai ir asmenų, turinčių IN, diagnozių statistika Lietuvoje pgl. SVEIDRA (2018)

IN laipsnis	Kodas pgl. TLK-10-AM	IQ	Sergančiųjų skaičius
Lengvas (angl. <i>Mild</i> )	F70	50–69	4 131 (44 proc.)
Vidutinis (angl. <i>Moderate</i> )	F71	35–49	2 350 (25 proc.)
Sunkus (angl. <i>Severe</i> )	F72	20–34	1 331 (14 proc.)
Gilus (angl. <i>Profound</i> )	F73	< 20	1 181 (13 proc.)
Nepatikslintas	F79	–	417 (4 proc.)

IN dažnis bendroje populiacijoje yra 1–3 proc. [10–12]. Ši patologija yra būdinga tiek vyrams, tiek moterims, tačiau vyriškosios lyties asmenims šis sutrikimas pasireiškia apie du kartus dažniau [112]. Toks pasiskirstymas aiškinamas tuo, kad daugiau nei 100 genų, siejamų su IN, yra X chromosomoje [112, 161, 163]. Higienos instituto Privalomojo sveikatos draudimo informacinės sistemos (SVEIDRA) duomenimis, 2018 metais Lietuvoje užregistruota 9 410 asmenų, kuriems diagnozuota IN. Daugumai visų atvejų (44 proc.) būdinga lengva IN (2.3 lentelė). Nors Lietuvos populiacijai būdingas mažesnis sergančiųjų skaičius (< 1 proc.), IN išlieka viena svarbiausių socialinių bei ekonominių asmens sveikatos priežiūros problemų ne tik Lietuvoje, bet ir visame pasaulyje. Skaičiuojama, jog išlaidos tyrimams bei konsultavimui dėl IN sudaro apie 5–10 proc. visų asmens sveikatos priežiūros sąnaudų [12, 164, 165].

### 2.3.2. Intelektinės negalios etiologija

IN etiologija yra heterogeninė, kadangi šią patologiją lemia tiek aplinkos, tiek genetiniai veiksniai. Šie veiksniai gali pažeisti įvairias organizmo sistemas, įskaitant centrinę nervų sistemą, kuri intensyviai vystosi prenataliniame laikotarpyje ir pirmaisiais gyvenimo metais. Todėl šiame laikotarpyje įvykę pažeidimai dažnai lemia neuropsichiatrinius sutrikimus, įskaitant IN (2.9 pav.) [166].

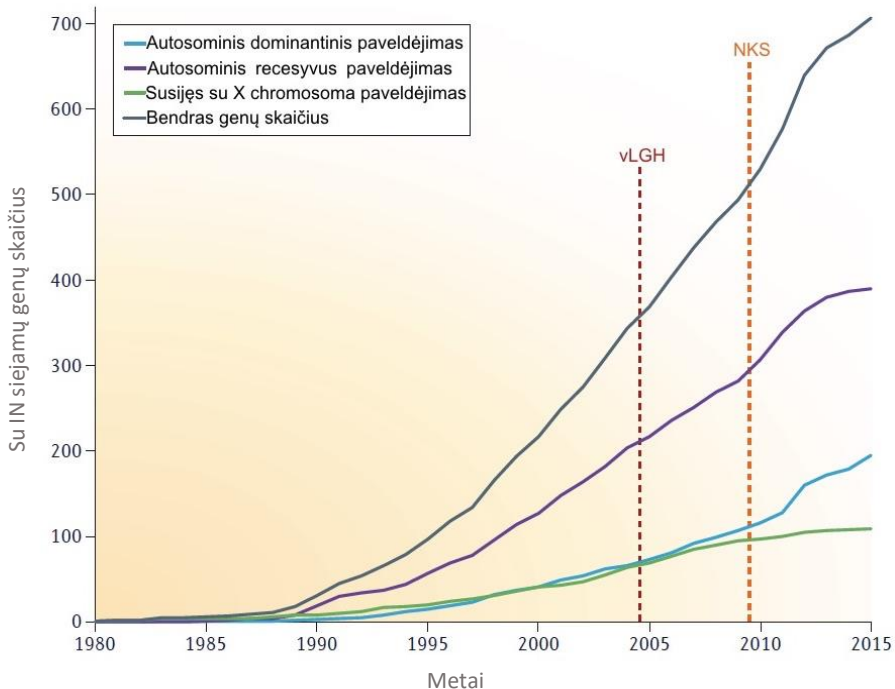


**2.9 pav.** IN lemiantys aplinkos ir genetiniai veiksniai priklausomai nuo neurologinio vystymosi stadijų. Schema adaptuota pagal mokslinę literatūrą [166]

Aplinkos veiksnių grupei priklauso įvairios infekcijos nėštumo metu ir po gimdymo: sepsio, toksoplazmozės, sifilio, raudonukės, citomegalo ar žmogaus imunodeficitu viruso infekcijos. Taip pat aplinkos veiksniams priskiriami teratogenai (pvz., alkoholis), metabolinės ligos (pvz., hiperbilirubinemija), netinkama mityba, jodo trūkumas bei komplikacijos gimdymo metu (pvz., smegenų pažeidimai, asfiksija) [167].

Manoma, kad 25–50 proc. visų IN atvejų lemia įvairios genetinės priežastys – nuo stambių (> 5 Mb) chromosomų skaičiaus bei struktūros pokyčių iki taškinių variantų specifiniuose genuose. Nors dažniausios monogeninės priežastys yra *de novo* sekos variantai, jos taip pat gali būti paveldėtos autosominiu dominantiniu, autosominiu recesyviu, susijusiu su

X chromosoma ar net mitochondrinio paveldėjimo būdu. Žinomos ne tik monogeninės, poligeninės, bet ir daugiaveiksniės ar net epigenetinės priežastys [167–169]. Šiuo metu yra patvirtinta apie 700 genų, lemiančių IN, tačiau tobulėjant tyrimo metodams ir technologijoms su šia patologija siejamų genų nustatoma vis daugiau (2.10 pav.) [170, 171]. Nepaisant to, apie 50 proc. visų IN atvejų etiologija lieka nežinoma [167].



**2.10 pav.** Naujų genų, siejamų su IN, nustatymas pagal paveldėjimo būdą ir naujos kartos sekoskaitos (NKS) ir vektorinės lyginamosios genomo hibridizacijos (vLGH) technologijas [170]

### 2.3.3. Sindrominė intelektinė negalia

Sindrominę IN turintiems asmenims dažnai yra būdingos gretutinės ligos ar sveikatos sutrikimai, pavyzdžiui, įgimtos anomalijos (IA; angl. *Congenital anomalies*), vystymosi atsilikimas (angl. *Developmental delay*), autizmo spektro sutrikimas (ASS; angl. *Autism spectrum disorder*) ar epilepsija, tačiau dažniausiai kartu su IN pasireiškia IA [172–174].

PSO IA apibūdina kaip patologinius klinikinius požymius, pasireiškiančius prenataliniame, perinataliniame, postnataliniame gyvenimo laikotarpyje arba vėliau gyvenime. Ši patologija yra neįgalumo ar net



ankstyvos mirties priežastis: pasaulyje kasmet miršta daugiau nei 300 tūkst. naujagimių [175, 176].

Sindrominiai IN ir ĮA atvejai dažniausiai būdingi chromosomų skaičiaus pokyčių nulemtiems sindromams: Dauno sindromui, sudarančiam 53 proc. visų atvejų, taip pat Turnerio (angl. *Turner*; ORPHA #881; 10 proc.), Edvardso (angl. *Edwards*; ORPHA #3380; 9 proc.) bei Patau (ORPHA #3378; 8 proc.) sindromams [177].

SVEIDRA duomenimis, 2018 metais Lietuvoje užregistruoti 54 096 asmenys, turintys įvairias ĮA. Lietuvos populiacijai daugiausiai būdingos įgimtos raumenų ir skeleto (41,6 proc.) bei kraujo apytakos (33,3 proc.) sistemų formavimosi ydos ir deformacijos (2.4 lentelė). Ankstesnio mokslinio tyrimo metu nustatyta, kad asmenims, turintiems IN, net 48,8 proc. visų atvejų taip pat būdinga bent viena ĮA, o trys ir daugiau ĮA stebimos rečiau (11,4 proc.). Dažniausia ĮA yra mikrocefalija, sudaranti 24 proc. visų anomalijų, o kitos ĮA stebimos rečiau – įgimta hidrocefalija (17,8 proc.), prieširdžių pertvaros defektas (10 proc.), nepatikslinti ausų dismorfiniai požymiai (7,6 proc.), klinodaktilija (7,3 proc.) ir hipertelorizmas (5,6 proc.) [32].

Nors Lietuvos populiacijoje nervų sistemos formavimosi ydos nustatomos rečiau nei kitos ydos (2.4 lentelė), žinomi net 307 genai, lemiantys IN ir nervų sistemos ĮA. Taip pat daug su IN siejamų genų lemia kitas ĮA: 146 genai siejami su IN ir mikrocefalija, 143 – su IN ir žemu ūgiu, 141 – su IN ir akių anomalijomis [174].

**2.4 lentelė.** Asmenų, turinčių ĮA, diagnozių statistika Lietuvoje pgl. SVEIDRA (2018)

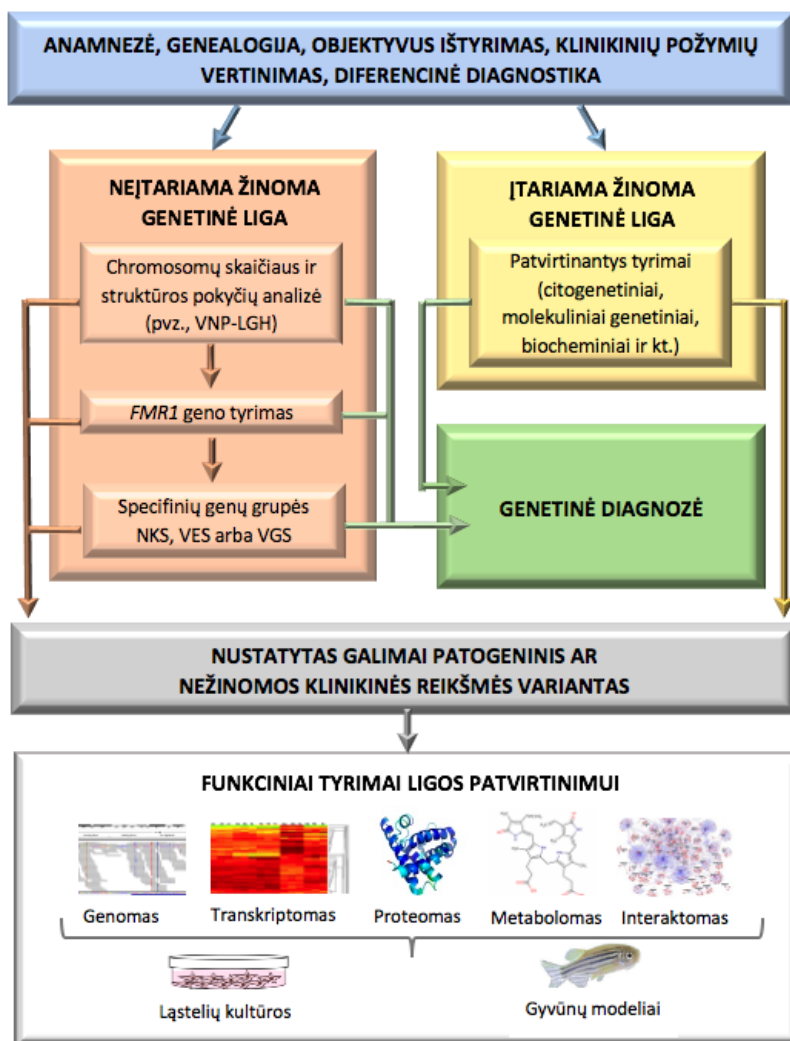
<b>Diagnozės pavadinimas</b>	<b>Kodas pgl. TLK-10-AM</b>	<b>Sergančiųjų skaičius</b>
Įgimtos nervų sistemos formavimosi ydos	Q00-Q07	1 659 (3,1 proc.)
Įgimtos akies, ausies, veido ir kaklo formavimosi ydos	Q10-Q18	1 910 (3,5 proc.)
Įgimtos kraujo apytakos sistemos formavimosi ydos	Q20-Q28	17 998 (33,3 proc.)
Įgimtos kvėpavimo sistemos formavimosi ydos	Q30-Q34	217 (0,4 proc.)
Lūpos ir gomurio defektai	Q35-Q37	221 (0,4 proc.)
Kitos įgimtos virškinimo sistemos formavimosi ydos	Q38-Q45	2 553 (4,7 proc.)
Įgimtos lyties organų formavimosi ydos	Q50-Q56	3 477 (6,4 proc.)
Įgimtos šlapimo organų sistemos formavimosi ydos	Q60-Q64	2 560 (4,7 proc.)
Įgimtos raumenų ir skeleto sistemos formavimosi ydos ir deformacijos	Q65-Q79	22 518 (41,6 proc.)
Chromosomų anomalijos, neklasifikuojamos kitur	Q90-Q99	983 (1,8 proc.)
	<b>Iš viso</b>	<b>54 096</b>

**2.3.4. Intelektinės negalios ir įgimtų anomalijų tyrimo gairės**

Dėl heterogeninės etiologijos ir klinikinių IN / ĮA požymių šią patologiją lemiančių genetinių priežasčių nustatymas yra itin sudėtingas [166]. Tačiau pastaraisiais dešimtmečiais yra nuolat tobulinami molekuliniai genetiniai ir funkciniai metodai, todėl IN / ĮA ištyrimo gairės, kurios buvo pasiūlytos 1997 metais, keičiasi [178–180].

Šiuo metu asmenims, turintiems IN / ĮA, yra rekomenduojamas etapinis ištyrimo planas (2.11 pav.) [13]. Šis diagnostinis tyrimas pradedamas nuo ligos anamnezės, genealogijos sudarymo bei klinikinių požymių ištyrimo. Atlikus diferencinę diagnostiką įtariamai genetinei ligai yra taikomi patvirtinantys tyrimai, pavyzdžiui, citogenetiniai ar molekuliniai genetiniai. Tačiau šio etapo diagnostinis efektyvumas yra nedidelis – diagnozė nustatoma ir patvirtinama tik 15–35 proc. tiriamųjų. Todėl dažnai yra atliekami išsamesni tyrimai: chromosomų skaičiaus ir struktūros pokyčių analizė naudojant, pavyzdžiui, VNP-LGH technologiją arba *FMRI* (MIM #300805) geno (CGG)<sub>n</sub> pasikartojimų skaičiaus tyrimas dėl lūžusios X chromosomos sindromo [13].

Pradėjus plačiau taikyti pažangias naujos kartos sekoskaitas (NKS; angl. *Next generation sequencing*) technologijas, asmenims, kuriems nepavyko nustatyti diagnozės, rekomenduojama atlikti variantų paiešką baltymus koduojančiose sekose taikant viso egzomo sekoskaitą (VES; angl. *Whole exome sequencing*) arba specifiniuose su IN / ĮA negalia siejamuose genuose, taikant specifinių genų grupės sekoskaitą (angl. *Gene panel sequencing*). Tam tikrais atvejais tikslinga atlikti viso genomo (VGS; angl. *Whole genome*) sekoskaitą [89, 181, 182].



**2.11 pav.** Asmenų, kuriems diagnozuota IN ir / ar ĮA, ištyrimo gairės. Schema adaptuota pagal mokslinę literatūrą [13, 30, 89, 115, 181, 182]

Vadovaujantis naujausiomis rekomendacijomis atrinktų variantų patogeniškumui ir įtakos fenotipui įvertinti, reikalingi funkciniai tyrimai ne tik genomo lygmeniu, bet ir transkriptomo, proteomo, metabolomo bei interaktomo [115]. Modeliuojant įvairių ligų, įskaitant IN / ĮA, požymius, vis plačiau sėkmingai naudojami ląstelių kultūrų bei įvairių gyvūnų, pavyzdžiui, zebražuvės (lot. *Danio rerio*) (2.11 pav.), pelės (lot. *Mus musculus*) ar žiurkės (lot. *Rattus norvegicus*), modeliai [30].

IN / ĮA diagnostikos efektyvumas sparčiai auga, todėl nenustačius diagnozės svarbu turėti išsaugotą tiriamo asmens biologinę medžiagą – DNR, RNR ir / arba ląstelių kultūrą, kuri gali būti pakartotinai iširta ateityje, atsiradus naujiems klinikiniams požymiams, patobulinus jau sukurtas arba panaudojus naujai atsiradusias tyrimo technologijas [13]. Pažymėtina, kad siūlomos ištyrimo gairės (2.11 pav.) yra tinkamos ne tik IN / ĮA etiopatogenezei tirti. Taikant panašią tyrimo strategiją, būtų galima įvertinti genetinių pokyčių patogenezės mechanizmus ir patvirtinti diagnozę kone kiekvienos heterogeninės prigimties paveldimos ligos kontekste.

### 3. MOKSLINIO DARBO METODIKA

#### 3.1. Tyrimo strategija ir tiriamieji

VU MF BMI ŽMGK mokslininkai jau kelis dešimtmečius analizuoja lietuvių genomą, vykdydami nacionalinius ir tarptautinius (pvz., LITGEN, CHERISH, PROGENET, UNIGENE) projektus. Naujos kartos sekoskaitos bei lyginamosios genomo hibridizacijos tyrimo metodais buvo sėkmingai ištirti asmenys (įskaitant tuos, kurie turi IN / IA) iš bendros Lietuvos populiacijos. Naudojant bioinformacinės analizės įrankius buvo nustatyti galimai patogeniniai arba nežinomos klinikinės reikšmės koduojančios ir nekoduojančios genomo sekos pokyčiai, tinkantys funkciniai genomo analizei.

Tarptautinis UNIGENE projektas suteikė pagrindą funkciniam tyrimams Lietuvoje, kadangi Šveicarijos partneriai variantų, siejamų su IN, ištyrimui taikė funkcinės genomikos metodus, o dalis jų pastaraisiais 2017–2020 metais buvo taikomi vykdant Lietuvos mokslo tarybos finansuojamą projektą „Intelektinės negalios ir įgimtų anomalijų genetinės architektūros identifikavimas“ (akronimas INGENES; sutarties Nr. S-MIP-17-19/LSS-150000-1179; vad. doc. dr. Eglė Preikšaitienė).

Disertacinis darbas atliktas vykdant INGENES projektą. INGENES biomedicininis tyrimas yra atliktas laikantis visų asmens duomenų apsaugos bei bioetikos reikalavimų. Mokslinis tyrimas patvirtintas Vilniaus regioninio biomedicininio tyrimų etikos komiteto (leidimo Nr. 158200-17-962-469; žr. I priedą), kaip to reikalauja Helsinkio deklaracija [183].

Šiame moksliniame darbe tirti lietuvių kilmės asmenys pagal nustatytus atrankos kriterijus. Pagrindiniai tiriamųjų atrankos kriterijai: (1) intelektinė negalia (IN), psichomotorinės raidos atsilikimas (PRA) ir / arba viena ar daugiau įgimtų anomalijų (IA); (2) tiriamajam nustatytas nežinomos ar iki šiol nepatvirtintos klinikinės reikšmės koduojančios ar nekoduojančios DNR sekos variantas ar kopijų skaičiaus pokytis. Naudojant ankstesnių projektų duomenis pagal užsibrėžtus atrankos kriterijus buvo atrinkti tiriamieji ( $n = 10$ ). Siekiant tyrimui gauti iRNR mėginius ir atlikti molekulinis bei funkcinis tyrimus, atrinkti tiriamieji buvo pakviesti dalyvauti INGENES projekte (3.1 lentelė). Genų raiškos tyrime taip pat tirti kontroliniai iRNR mėginiai, kurie buvo paimti iš atsitiktinių negiminingų neturinčių ligos požymių asmenų iš Lietuvos populiacijos. Keliais atvejais kontrolinę grupę sudarė tiriamo asmens tėvai, neturintys ligos požymių (žr. 3.3.5 skyrių). Tirtų asmenų fenotipų duomenys gauti iš medicininės dokumentacijos. Visi tyrime dalyvavę asmenys yra pasirašę Asmens informavimo ir Informuoto asmens

sutikimo formas. Tiriamųjų asmeninė informacija buvo konfidenciali ir saugiai užkoduota, todėl tyrimo metu buvo naudojami tik mėginių identifikaciniai numeriai (tiriamųjų žymėjimas: nuo #1 iki #10) (3.1 lentelė).

**3.1 lentelė.** Mokslinio darbo metu molekuliniais ir funkciniais tyrimo metodais ištirti asmenys (n = 10)

Tiriamieji	Varianto / KSP nustatymo ir patvirtinimo metodai	Nežinomos klinikinės reikšmės variantas / KSP	Genai	RNR gryninimo metodas	Molekuliniai ir / ar funkciniai tyrimo metodai	Pagrindiniai tiriamųjų klinikiniai požymiai		
						IN	PRA	IA
#1	Variantas nustatytas VES, patvirtintas Sangerio metodu	NM_020732.3: c.4986+2T>C	<i>ARID1B</i> MIM#614556	<i>Tempus™ Spin RNA Isolation Kit</i>	Sangerio sekoskaita kDNR mėginiui	+	+	Igimtos akių, veido ir smegenų anomalijos
#2	Variantas nustatytas Sangerio metodu	NM_000168.6: c.473+3A>T	<i>GLI3</i> MIM#165240	<i>Tempus™ Spin RNA Isolation Kit, RNeasy® Mini Kit</i>	Sangerio sekoskaita kDNR mėginiui; Fibroblastų kultivavimas	+	-	Igimtos skeleto anomalijos
#3	Variantas nustatytas VES, patvirtintas Sangerio metodu	NM_001042537.1: c.899+1G>A	<i>SLC9A6</i> MIM#300231	<i>Tempus™ Spin RNA Isolation Kit</i>	Sangerio sekoskaita kDNR mėginiui	+	+	Igimtos akių anomalijos
#4	Variantas nustatytas VES, patvirtintas Sangerio metodu	NM_017780.4: c.5535-1G>A	<i>CHD7</i> MIM#608892	<i>Tempus™ Spin RNA Isolation Kit</i>	Sangerio sekoskaita kDNR mėginiui	+	+	Igimta širdies yda, akių, smegenų ir skeleto anomalijos
#5	Variantas nustatytas atlikus genų grupės NKS, patvirtintas Sangerio metodu	NM_001024847.2: c.1600-2A>G	<i>TGFBR2</i> MIM#190182	<i>Tempus™ Spin RNA Isolation Kit</i>	kPGR; Sangerio sekoskaita kDNR mėginiui	-	-	Igimtos širdies ydos, skeleto anomalijos
#6	Variantas nustatytas atlikus genų grupės NKS, patvirtintas Sangerio metodu	NM_000057.2: c.2308-2A>G	<i>BLM</i> MIM#210900	<i>Tempus™ Spin RNA Isolation Kit</i>	kPGR; Sangerio sekoskaita kDNR mėginiui	-	+	-

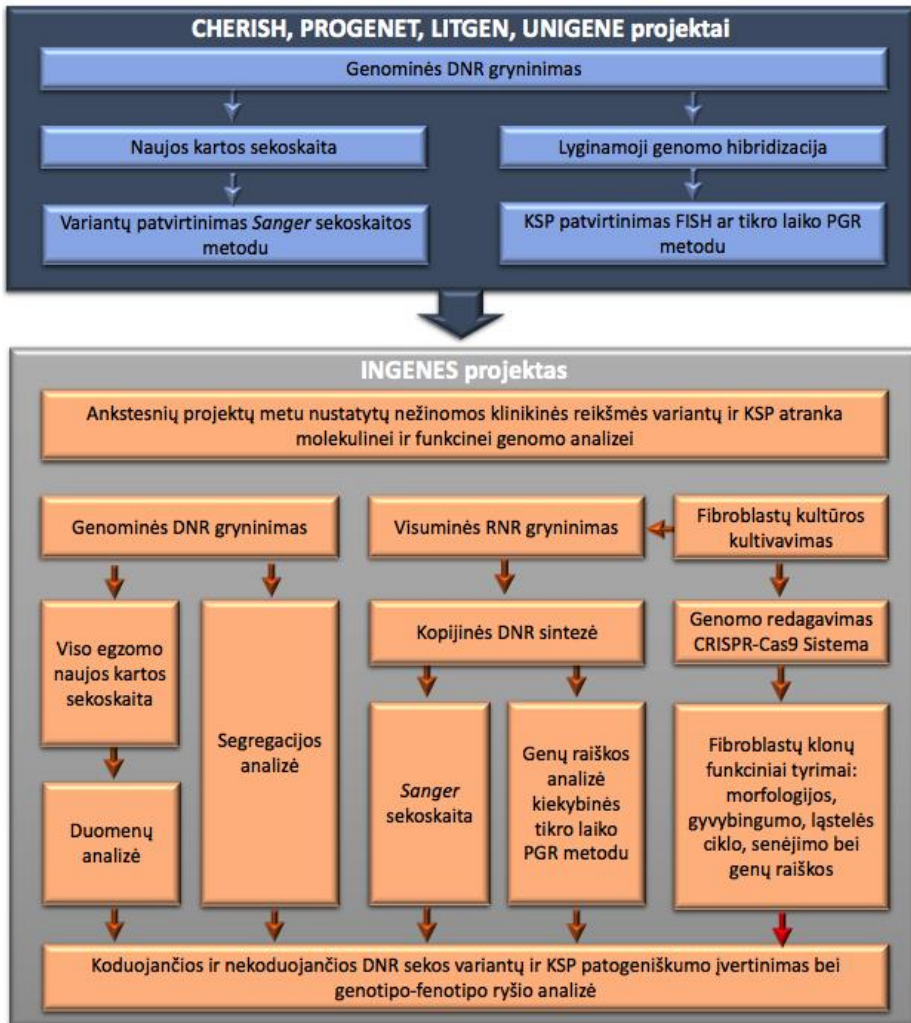
**3.1 lentelė (tęsinys).** Mokslinio darbo metu molekuliniais ir funkciniais tyrimo metodais ištirti asmenys (n = 10)

Tiria- mieji	Varianto / KSP nustatymo ir patvirtinimo metodai	Nežinomos klinikinės reikšmės variantai / KSP	Genai	RNR gryninimo metodas	Molekuliniai ir / ar funkciniai tyrimo metodai	Pagrindiniai tiriamųjų klinikiniai požymiai		
						IN	PRA	IA
#7	Variantas nustatytas atlikus genų grupės NKS, patvirtintas Sangerio metodu	NM_000070: c.1746-20C>G	<i>CAPN3</i> MIM#114240	<i>Tempus</i> <sup>TM</sup> <i>Spin RNA</i> <i>Isolation Kit</i>	Sangerio sekoskaita kDNR mėginiui	-	-	Igimtos raumenų ir skeleto anomalijos
#8	KSP nustatytas vLGH, patvirtintas <i>FISH</i> metodu	dup13q31.3	<i>GPC5</i> MIM#602446 <i>MIR17HG</i> MIM#609415	<i>PureLink</i> <sup>TM</sup> <i>Total RNA</i> <i>Blood Kit</i>	kPGR, VES	+	+	Igimta širdies yda, akių ir veido anomalijos
#9	KSP nustatytas vLGH, patvirtintas kPGR metodu	del4q21.22	<i>PRKG2</i> MIM#601591 <i>RASGEF1B</i> MIM#614532 <i>HNRNPD</i> MIM#601324 <i>HNRNPDL</i> MIM#607137 <i>ENOPH1</i>	<i>PureLink</i> <sup>TM</sup> <i>Total RNA</i> <i>Blood Kit</i>	kPGR	+	+	Igimtos veido ir skeleto anomalijos
#10	KSP nustatytas VNP-LGH, patvirtintas kPGR metodu	del12q24.21	<i>MED13L</i> MIM#608771 <i>RB1</i> (MIM #614041) <i>E2F1</i> (MIM #189971) <i>CCNC</i> (MIM #123838)	<i>Tempus</i> <sup>TM</sup> <i>Spin RNA</i> <i>Isolation Kit</i> , <i>RNeasy</i> <sup>®</sup> <i>Mini Kit</i>	kPGR; Sangerio sekoskaita kDNR mėginiui; CRISPR-Cas9 genomo redagavimas fibroblastuose	+	+	Igimtos veido ir skeleto anomalijos



### 3.2. Tyrimo eiga

Siekiant įgyvendinti užsibrėžtą disertacinio darbo tikslą ir užduotis, moksliniai tyrimai buvo atliekami pagal 3.1 paveiksle nurodytą tyrimo strategiją. Detaliau disertacijoje yra aprašomi tik tie tyrimo metodai (žr. 3.3 skyrių), kuriuos taikė disertacijos autorė rengdama mokslinį darbą. Autorės atliktų darbų indėlis pateiktas II priede.



3.1 pav. Mokslinio tyrimo eiga

### 3.3. Tyrimo metodai

#### 3.3.1. Genominės DNR gryninimas, švarumo nustatymas bei mėginių ilgalaikis saugojimas

Genominę DNR (gDNR) iš periferinio kraujo leukocitų išgrynino VU MF BMI ŽMGK darbuotojai. gDNR buvo gryninama iš periferinio kraujo leukocitų dviem metodais: 1) standartiniu fenolio-chloroformo-izoamilo alkoholio metodu [184]; 2) automatizuota sistema *TECAN Freedom EVO® 200* (*Tecan Schweiz AG*, Šveicarija) magnetinių dalelių prijungimo prie DNR metodu pagal gamintojo pateikiamus protokolus. DNR koncentracija ir švarumas buvo nustatomi spektrofotometru *NanoDrop® 2000* (*ThermoFisher Scientific*, JAV) vadovaujantis gamintojo vartotojo vadovu *NanoDrop® ND-1000 Spectrophotometer User Manual v.3.3* (*ThermoFisher Scientific*, JAV). Geros kokybės mėginiai buvo laikomi, jei absorbcijos santykis 260 nm ir 280 nm ilgio šviesos bangoje ( $A_{260/280}$ ) buvo daugiau nei 1,8 matavimo vieneto. Paruošti DNR mėginiai saugomi  $-20\text{ }^{\circ}\text{C}$  šaldiklyje.

#### 3.3.2. Fibroblastų kultūros kultivavimas bei ilgalaikis saugojimas

Fibroblastų ėminiai buvo paimti odos biopsijos procedūros metu, kurią atliko gydytojas chirurgas. Fibroblastų kultūros buvo sėjamos ir kultivuojamos dviejų dimensijų plastikiniuose indeliuose T25 (*TPP*, Šveicarija) mitybinėje terpėje, kurią sudarė 85,4 proc. *AmnioMAX C-100 Basal Medium* (*ThermoFisher Scientific*, JAV) bazinės terpės, 14,2 proc. *AmnioMAX C-100 Supplement* (*ThermoFisher Scientific*, JAV) priedo ir 0,4 proc. amfotericino B (angl. *Amphotericin B*, *Gibco*, Jungtinė Karalystė). Tyrimas buvo atliekamas pagal standartinį fibroblastų kultivavimo protokolą [185].

Fibroblastų kultūros tankiui pasiekus 80–90 proc. (angl. *Cell confluence*), ląstelės buvo užšaldomos 2 ml tūrio kriomėgintuvėliuose (*ThermoFisher Scientific*, JAV) paruoštoje šaldymo terpėje. Šaldymo terpė buvo ruošiamą naudojant 50 proc. mitybinės terpės, 33,5 proc. fetalinio veršelio serumo (angl. *Fetal bovine serum*; *Gibco*, Jungtinė Karalystė) bei 16,5 proc. dimetilsulfoksido (angl. *Dimethyl sulfoxide*; DMSO; *OriGen Biomedical*, JAV). Bendradarbiaujant su Nacionaliniu vėžio institutu (NVI), paruošti kriomėgintuvėliai pirmiausia buvo dedami į šaldymo dėžę *Mr. Frosty* (*ThermoFisher Scientific*, JAV). Siekiant užtikrinti, kad temperatūra kriomėgintuvėliuose mažėtų palaipsniui, t. y.  $1\text{ }^{\circ}\text{C}$  kas minutę, šaldymo dėžė

buvo dedama į šaldiklį (*Arctiko*, Danija), kuriame buvo –80 °C temperatūra. Po 24 val. kriomėgintuvėliai buvo perkeliama ilgalaikiam saugojimui į –150 °C šaldiklį (*Sanyo*, Japonija). Tuo pačiu principu tiriamojo #10 ir kontrolinio asmens mėginiai buvo saugomi VU GMC, su kuriuo bendradarbiauta redaguojant fibroblastų genomą CRISPR-Cas9 metodu ir atliekant funkcinius fibroblastų kultūros tyrimus (žr. 2.2.2.3 skyrių).

### 3.3.3. Visuminės RNR gryninimas

#### 3.3.3.1. Visuminės RNR gryninimas iš periferinio kraujo ėminių

Tiriamojo #8 ir #9 visuminė RNR (angl. *Total RNA*) buvo gryninama iš periferinio kraujo ėminių naudojant *PureLink™ Total RNA Blood Kit* (*Invitrogen*, JAV) rinkinį pagal gamintojo protokolą. Visų kitų šiame darbe aprašomų tiriamųjų visuminė RNR buvo gryninama iš periferinio kraujo ėminių, kurie buvo paimti į *Tempus™ Blood RNA Tube* (*ThermoFisher Scientific*, JAV) mėgintuvėlius su viduje esančiu komerciniu reagentu, užtikrinančiu RNR stabilumą. Visuminė RNR buvo gryninama pagal gamintojo protokolą naudojant *Tempus™ Spin RNA Isolation Kit* (*ThermoFisher Scientific*, JAV) rinkinį, kuris užtikrina išgrynintos RNR didelę išeią (6–25 µg RNR iš 3 ml kraujo ėminio) ir gerą kokybę ( $A_{260/280} > 1,9$ ). Taip pat šis rinkinys garantuoja itin mažą užteršimą gDNR (< 0,005 proc.). Tačiau, siekiant pašalinti galimus gDNR likučius, RNR gryninimo metu mėginiai buvo papildomai veikiami DNR nukleazės fermentu, naudojant *PureLink™ DNase Set* (*Invitrogen*, JAV) rinkinį pagal gamintojo protokolą.

Pažymėtina, kad molekulių ir funkcinių tyrimų metu tirtos taip pat negiminingų ir neturinčių ligos požymių kontrolinių asmenų visuminės RNR mėginiai. Siekiant užtikrinti tyrimo efektyvumą, atkartojamumą ir patikimumą, visuminė RNR buvo gryninama tuo pačiu metu ir tuo pačiu gryninimo metodu kaip ir tiriamų asmenų.

#### 3.3.3.2. Visuminės RNR gryninimas iš fibroblastų kultūros ėminių

Tiriamojo #2 ir #10 visuminė RNR buvo gryninama iš fibroblastų kultūros naudojant *RNeasy® Mini Kit* (*Qiagen*, Vokietija) rinkinį pagal gamintojo protokolą. Šis visuminės RNR gryninimo metodas taip pat buvo naudojamas CRISPR-Cas9 geno redagavimo eksperimente, gryninant

visuminę RNR iš neturinčio ligos požymio kontrolinio asmens fibroblastų kultūrų prieš ir po genomo redagavimo procedūros (žr. 3.3.8 skyrių).

Vadovaujantis gamintojo vartotojo vadovu *RNeasy® Mini Handbook* (Qiagen, Vokietija), rekomenduojama gryninti RNR iš ląstelių kultūros, kurioje yra apie  $3\text{--}4 \times 10^6$  ląstelių. Todėl RNR buvo gryninama, kai ląstelių kultūros, kultivuotos plastikiniame indelyje T25 (*TPP*, Šveicarija), tankis siekė  $> 90$  proc. (angl. *Cell confluence*) [186]. Vykdamas CRISPR-Cas9 genomo redagavimo eksperimentą, penkiuose (iš 11) kultivavimo indeliuose redaguotų fibroblastų dauginimosi efektyvumas buvo sumažėjęs, todėl ląstelių visuminė RNR išgryninta iš mažesnio kiekio ląstelių (apie 20–50 proc. ląstelių tankis).

### 3.3.3.3. Visuminės RNR koncentracijos ir švarumo nustatymas bei mėginių ilgalaikis saugojimas

Išgrynintos RNR koncentracija ir švarumas buvo matuojami spektrofotometru *NanoDrop® 2000* (*ThermoFisher Scientific*, JAV) vadovaujantis gamintojo vartotojo vadovu *NanoDrop® ND-1000 Spectrophotometer User Manual v.3.3* (*ThermoFisher Scientific*, JAV). Geros kokybės mėginiai buvo laikomi, jei absorbcijos santykis 260 nm ir 280 nm ilgio šviesos bangoje ( $A_{260/280}$ ) buvo daugiau nei du matavimo vienetai. Paruošti RNR mėginiai saugomi  $-80$  °C šaldiklyje.

### 3.3.4. Kopijinės DNR sintezė

iRNR būdingas jautrumas hidrolizės procesui ir mažesnis stabilumas, palyginti su DNR [15, 38, 48], todėl, siekiant to išvengti, visų tiriamųjų išgryninta visuminė RNR buvo naudojama kopijinės DNR (kDNR) sintezei, kuri toliau buvo tiriama Sangerio sekoskaitos ir / ar kiekybinės tikro laiko polimerazės grandininės reakcijos (kPGR; angl. *Quantitative real-time polymerase chain reaction*) metodais.

Kopijinė DNR buvo sintetinama atvirkštinės transkripcijos polimerazės grandininės reakcijos (AT-PGR; angl. *Reverse transcription polymerase chain reaction*) metodu, naudojant *High Capacity RNA-to-cDNA™ Kit* (*ThermoFisher Scientific*, JAV) rinkinį pagal gamintojo rekomendacijas. Ruošiant reakcijos mišinį (3.2 lentelė) ypač svarbu naudoti vienodą išgrynintos visuminės RNR kiekį, kadangi iRNR būdingas didelis koncentracijos variabilumas tarp skirtingų tiriamų mėginių. Todėl, siekiant užtikrinti AT-PGR efektyvumą, tikslumą ir atkartojamumą, priklausomai nuo

išgrynintos visuminės RNR koncentracijos į reakcijos mišinį buvo dedamas vienodas išgrynintos visuminės RNR kiekis – 500 ng, 250 ng arba 200 ng. Atsižvelgiant į gamintojo rekomendacijas, dedamos visuminės RNR kiekis buvo apskaičiuojamas taip, kad neviršytų 9 µl tūrio. Todėl, jei tiriamųjų ir kontrolinių mėginių išgrynintos visuminės RNR koncentracija buvo didesnė nei 56 ng/µl, reakcijai buvo naudojama 500 ng visuminės RNR. Retesniais atvejais, kai visuminės RNR koncentracija buvo mažesnė nei 55 ng/µl, reakcijai buvo naudojama 250 ng arba 200 ng visuminės RNR. Laikantis šios skaičiavimo strategijos reakcijos mišiniui pagaminti, dažniausiai buvo naudojama mažiau nei 9 µl visuminės RNR, dėl to likęs tūris buvo užpildomas vandeniu (be nukleazių) taip, kad bendras AT-PGR tūris būtų 20 µl. Tyrimo metu taip pat buvo naudojama neigiama atvirkštinės transkriptazės kontrolė (angl. *Negative reverse transcriptase control*; -RT), siekiant nustatyti, ar reakcija nebuvo užteršta DNR (3.2 lentelė).

**3.2 lentelė.** AT-PGR mišinio sudėtis

Medžiagos pavadinimas	Tūris (µl)	
	RT	-RT
2× RT BufferMix	10	10
20× RT EnzymeMix	1	–
Išgryninta visuminė RNR	Iki 9	Iki 9
Vanduo be nukleazių (jei taikoma)	Iki 20	Iki 20
<b>Mišinio tūris</b>	<b>20</b>	<b>20</b>

AT-PGR vykdoma *ProFlex PCR System* termociklerio kaitinimo bloke (*ThermoFisher Scientific*, JAV) gamintojo rekomenduojamomis sąlygomis: inkubacija 60 min. +37 °C, reakcijos sustabdymas 5 min. +95 °C. Susintetintos kDNR mėginiai saugomi –80 °C šaldiklyje.

### 3.3.5. Kiekybinė tikro laiko polimerazės grandininė reakcija

Naudojant oligonukleotidinius hidrolizės *TaqMan*<sup>TM</sup> žymenis (angl. *Probe*; *ThermoFisher Scientific*, JAV), kPGR metodu buvo iširta penkių tiriamųjų (#5, #6, #8, #9, #10) ir vieno kontrolinio asmens (genomo redagavimo eksperimente) 12 genų raiška (3.3 lentelė). kPGR eksperimentai buvo suplanuoti ir atliekami vadovaujantis MIQE rekomendacijomis ir gairėmis [130].

Siekiant išvengti genominės DNR priemaišų mėginyje, visi tyrimui parinkti komerciniai *TaqMan*<sup>TM</sup> žymenys buvo sukurti ant gretimų egzonų jungčių (3.3 lentelė). *MED13L* geno raiškos tyrimo metu komercinio žymens

1–2 egzono jungčiai nebuvo galima pasirinkti gamintojo internetinėje svetainėje [187], todėl pradmenys ir žymuo šiai genetinei sričiai buvo sukurti naudojant *NCBI Primer-Blast* [188] ir *Custom TaqMan® Probes tool* (*ThermoFisher Scientific*, JAV) [189] bioinformacinius įrankius (žr. III priedą, III.1 lentelę). Siekiant efektyvaus tiriamos srities atpažinimo, visi *TaqMan™* žymenys buvo žymėti 5'FAM (angl. *6-carboxyfluorescein*) fluoroforu bei 3'NFQ (angl. *nonfluorescent quencher*) slopikliu, kuris sąveikauja su DNR mažuoju grioviu (angl. *minor groove binder*, MGB).

### 3.3 lentelė. kPGR metodu tirti genai

Tiriamieji	Genai kandidatai	<i>TaqMan™</i> žymenų numeriai [187]	<i>TaqMan™</i> žymenų vieta [187]
#5	<i>TGFBR2</i>	Hs00234253_m1	7–8 egzono jungtis
		Hs00559660_m1	6–7 egzono jungtis
#6	<i>BLM</i>	Hs01119887_m1	19–20 egzono jungtis
#8	<i>GPC5</i>	Hs00270114_m1	3–4 egzono jungtis
		<i>PRKG2</i>	Hs00922440_m1
#9	<i>RASGEF1B</i>	Hs00959036_m1	12–13 egzono jungtis
	<i>HNRNPD</i>	Hs01086912_m1	2–3 egzono jungtis
	<i>HNRNPD</i>	Hs00943609_m1	1–2 egzono jungtis
	<i>ENOPH1</i>	Hs00368361_m1	1–2 egzono jungtis
#10	<i>MED13L</i>	Hs01573430_m1	3–4 egzono jungtis
		Hs01011103_m1	16–17 egzono jungtis
		Sukurtas naudojant <i>in silico</i> įrankius	1–2 egzono jungtis
	<i>RBI</i>	Hs00153108_m1	24–25 egzono jungtis
	<i>E2F1</i>	Hs00153451_m1	3–4 egzono jungtis
	<i>CCNC</i>	Hs01029307_m1	2–3 egzono jungtis

kPGR reakcijos mišiniui buvo naudojama tiriamųjų kDNR, vanduo be RNazės, *20x TaqMan™ Gene Expression Assays* (*ThermoFisher Scientific*, JAV) ir *2x TaqMan™ Gene Expression Master Mix* (*ThermoFisher Scientific*, JAV) pagal gamintojo rekomendacijas (3.4 lentelė).

### 3.4 lentelė. kPGR mišinio sudėtis

Medžiagos pavadinimas	Tūris (μl)	
	1 reakcija	NTC
20x TaqMan™ Gene Expression Assay	1	1
2x TaqMan™ Gene Expression Master Mix	10	10
kDNR (1–100 ng) <sup>2</sup>	4	–
Vanduo be nukleazių	5	9
<b>Mišinio tūris</b>	<b>20</b>	<b>20</b>

Siekiant užtikrinti rezultatų patikimumą, kiekvieno biologinio mėginio atveju buvo atlikta po 3 techninius pakartojimus. Taip pat, siekiant išvengti klaidingai teigiamų rezultatų (angl. *False-positive results*) dėl mėginio užteršimo ar pradmenų dimerų formavimosi, buvo naudojama neigiama kontrolė (angl. *No template control*; NTC). kPGR atlikta *Applied Biosystems 7900HT system (ThermoFisher Scientific, JAV)* termocikleryje pagal gamintojo rekomenduojamas standartines sąlygas. Duomenų analizė atlikta naudojant *SDS v.2.3 (Applied Biosystems, JAV)*, *ExpressionSuite v.1.1 (ThermoFisher Scientific, JAV)*, *qbase<sup>+</sup> (Biogazelle, Belgija)*, *Microsoft Excel 2010 (Microsoft Corporation, JAV)* programas bei *R v.3.4.0 [190]* statistinį paketą. Genų raiškos pokyčiai buvo vertinami naudojant  $\Delta\Delta Ct$  metodą (skaičiavimo formulė pateikta 2.2.2.1 skyriuje) ir normalizuojant pagal vieną arba dvi endogenines kontroles (angl. *Housekeeping genes*). *GPC5*, *PRKG2*, *RASGEF1B*, *HNRNPD*, *HNRNPDL* bei *ENOPHI* genų raiškos tyrimų metu buvo naudojama *ACTB* (kodas: Hs99999903\_m1; MIM #102630; *ThermoFisher Scientific, JAV*) endogeninė kontrolė, o *MED13L*, *TGFBR2*, *BLM*, *MED13L*, *RBI*, *E2F1* bei *CCNC* genų raiškos tyrimuose buvo naudojamos *ACTB* ir *GAPDH* (kodas: Hs99999905\_m1; MIM #102630; *ThermoFisher Scientific, JAV*) endogeninės kontrolės.

Tyrimų metu taip pat buvo naudojamos referentinės kontrolės. *GPC5* ir penkių genų: *PRKG2*, *RASGEF1B*, *HNRNPD*, *HNRNPDL* ir *ENOPHI* raiškos tyrimuose biologines kontrolines grupes sudarė atitinkamai #8 ir #9 tiriamųjų tėvai, neturintys ligos požymių. Atliekant *BLM* raiškos tyrimą, #6 tiriamojo mėginiai lyginti su asmenų ( $n = 4$ ), neturinčių ligos požymių, tarp jų – tiriamųjų tėvų, mėginiais. *MED13L* bei *TGFBR2* genų raiškos tyrimų metu buvo naudojamos vidinės (angl. *Internal control*) ir biologinės išorinės (angl. *External control*) kontrolės. *MED13L* tyrimo atveju vidinę kontrolę sudarė to paties geno dvi skirtingos genetinės sritys (1–2 bei 16–17 egzonų jungčių), o

<sup>2</sup> kPGR reakcijoje naudotos kDNR koncentracija priklausė nuo išgrynintos visuminės RNR išeigos (žr. 3.3.4 skyrių).

biologinę išorinę kontrolę sudarė ligos požymių neturintys trys sveiki asmenys ( $n = 3$ ). Panašiai buvo suplanuotas *TGFBR2* geno raiškos tyrimas: vidinę kontrolę sudarė to paties geno 6–7 egzonų jungties genetinė sritis, o išorinę kontrolę taip pat sudarė ligos požymių neturintys trys sveiki asmenys ( $n = 3$ ).

Lyginant su kontrolėmis, genų raiškos pokytis du ir daugiau kartų (angl. *Fold change*; FC) buvo vertinamas kaip padidėjusi arba sumažėjusi tiriamo geno raiška.

### 3.3.6. Sekoskaita Sangerio metodu

Siekiant patvirtinti nežinomos klinikinės reikšmės variantą, nustatytą NKS tyrimo metodu, bei įvertinti varianto segregaciją, Sangerio sekoskaitos metodu buvo ištirti keturių tiriamųjų (#1–#4) gDNR mėginiai. Siekiant patvirtinti sukirpimo varianto poveikį iRNR lygmeniu ir patikslinti vidugeninės *MEDI3L* delecijos dydį, šiuo metodu buvo ištirti aštuonių tiriamųjų (#1, #2, #3, #4, #5, #6, #7, #10) ir vieno kontrolinio asmens (genomo redagavimo eksperimente) kDNR mėginiai.

#### **Pradmenų kūrimas**

Pradmenys buvo sukurti naudojant *Primer3 v.0.4.0.* programą (*Whitehead Institute for Biomedical Research, JAV*) [191] arba bioinformacinį įrankį *Primer-BLAST* [188] vadovaujantis standartinėmis rekomendacijomis. Visų pirma, pradmenys buvo kuriami ne mažiau nei 50 nukleotidų atstumu iki tiriamo varianto vietos. Tinkamiausia pradmenų pora buvo pasirenkama atsižvelgiant į nespacificinių jungimosi vietų genome skaičių, G/C sąstatą (40–60 proc.), pradmenų komplementarumą bei lydimosi temperatūrų skirtumą ( $\leq 5$  °C) [192].

kDNR tyrimų metu itin svarbu pašalinti galimus gDNR likučius mėginyje. Nors komerciniai visuminės RNR gryninimo rinkiniai užtikrina itin mažą užteršimą gDNR, visų tyrimų metu mėginiai buvo ne tik papildomai veikiami komerciniu DNR nukleazės fermentu, bet taip pat tiesioginis (angl. *Forward*; F) arba grįžtamasis (angl. *Reverse*; R) pradmuo buvo kuriamas ant gretimų egzonų jungties.

Oligonukleotidinių pradmenų sekos ir PGR produktų ilgiai nurodyti III priede (III.1 lentelė).

#### **PGR ir reakcijos produkto vertinimas**

gDNR ir kDNR fragmentai buvo pagausinti atliekant polimerizacijos grandininę reakciją (PGR; angl. *Polymerase chain reaction*) naudojant  $2 \times$  *Phusion Flash High-Fidelity PCR MasterMix (ThermoFisher Scientific, JAV)*



mišinį, dejonizuotą vandenį, sukurtą pradmenų porą ir kDNR mėginį. Taip pat buvo naudojama NTC neigiama kontrolė (3.5 lentelė).

**3.5 lentelė.** PGR mišinio sudėtis

Medžiagos pavadinimas	Tūris (μl)	
	1 reakcija	NTC
Dejonizuotas vanduo	8	9
<i>2x Phusion Flash PCR MasterMix</i>	10	10
F (10 pmol/μl)	0,5	0,5
R (10 pmol/μl)	0,5	0,5
gDNR/kDNR <sup>3</sup>	1	–
<b>Mišinio tūris</b>	<b>20</b>	<b>20</b>

PGR buvo vykdoma *ProFlex PCR System* termociklerio kaitinimo bloke (*ThermoFisher Scientific*, JAV) nustatytais sąlygomis (3.6 lentelė).

**3.6 lentelė.** PGR termociklerio programos etapai ir sąlygos

Etapai	Temperatūra (°C)	Trukmė (s)	30 ciklų
Pradinė denatūracija	98	30	
Denatūracija	98	2	
Pradmenų prijungimas	64	10	
Sintezė	72	20	
Galutinė sintezė	72	60	
Programos pabaiga	4	∞	

Siekiant įvertinti PGR produkto specifiškumą, ilgį ir kiekį, buvo vykdoma elektroforezė 1,5 proc. agarozės (*Top Vision Agarose*, *ThermoFisher Scientific*, JAV) gelyje, kuris buvo ruošiamas pagal standartines gamintojo rekomendacijas. PGR produkto vizualizacijai santykiu 2:1 naudotas 3x dažas (*DNA LoadingDye*, *ThermoFisher Scientific*, JAV), o DNR molekulinės masės ir dydžio standartas (*100 bp GeneRuler*, *ThermoFisher Scientific*, JAV) naudotas produkto ilgiui įvertinti. Elektroforezė buvo vykdoma *EC 105* horizontalios elektroforezės sistemoje (*E-C Apparatus Corporation*, JAV). Agarozės gelis su PGR produktais buvo vizualizuojamas UV transiliumatoriuje (*E.A.S.Y. 442K*, *Herolab*, Vokietija),

<sup>3</sup> PGR reakcijoje naudota 100 ng/μl gDNR koncentracija, o naudotos kDNR koncentracija priklausė nuo išgrynintos visuminės RNR išėigos (žr. 3.3.4 skyrių).

o analizuojamas ir fotografuojamas *EasyWin 32* (*Herolab*, Vokietija) gelių analizavimo programa. Priklausomai nuo gautų rezultatų, keliais atvejais PGR sąlygos buvo optimizuojamos (III priedas, III.2 ir III.3 lentelės). Gavus kokybišką PGR produktą, toliau buvo vykdoma PGR produkto sekoskaita Sangerio metodu.

### **PGR produktų sekoskaita Sangerio metodu**

Pirmiausia, iš PGR reakcijos buvo išvalyti nepanaudoti nukleotidai ir pradmenys naudojant fermentinį *ExoI-FastAP* metodą. Šiai reakcijai naudojama Egzonukleazė I (angl. *Exonuclease I*; *ThermoFisher Scientific*, JAV) ir *FastAP™ Thermosensitive Alkaline Phosphatase* (*ThermoFisher Scientific*, JAV) šarminė fosfatazė pagal gamintojo rekomendacijas. Toliau buvo atliekama ciklinė sekoskaitos PGR naudojant *BigDye® Terminator v.3.1 Cycle Sequencing Kit* (*ThermoFisher Scientific*, JAV) rinkinį vadovaujantis gamintojo optimizuotu protokolu.

Sekoskaitos reakcijos produktas buvo valomas etanoliu ir natrio acetatu. Reakcijos mišinys (3.7 lentelė) išpilstomas po 80 µl į mėgintuvėlius, kuriuose yra 7 µl sekoskaitos PGR produkto. Mėginiai inkubuojami tamsoje 8 min., o po inkubacijos centrifuguojami 7 min. 13 000 aps./min. Mėgintuvėlius kruopščiai dekantavus, įpilama po 300 µl šalto etanolio (70 proc.) ir mėgintuvėliai purtomi 2 min. Mėgintuvėliai dar kartą centrifuguojami 7 min. 13 000 aps./min. ir kruopščiai dekantuojami. Taip paruošti mėginiai paliekami išdžiūti 20–30 min. kambario temperatūroje arba 5 min. +37 °C temperatūroje.

#### **3.7 lentelė.** Etanolio-natrio acetato mišinys vienam mėginiui

<b>Medžiagos pavadinimas</b>	<b>Tūris (µl)</b>
Dejonizuotas vanduo	14,5
Etanolis (96 proc.)	62,5
Natrio acetatas	3
<b>Mišinio tūris</b>	<b>80</b>

Išdžiovinus sekoskaitos produktus, vykdoma automatizuota kapiliarinė elektroforezė, kuri nustato tiesioginę nukleotidų seką. Šiam uždaviniui pasiekti liofilizuotos DNR mėginiai užpilami 8 µl *Hi-Di™* formamido (*Hi-Di™ Formamide*, *ThermoFisher Scientific*, JAV) ir kruopščiai pipetuojant pernešami į 96 šulinėlių sekoskaitos plokštelę (*Nerbe Plus*, *GmbH*, Vokietija). Plokštelė uždengiama gumine tarpine (*Nerbe Plus*, *GmbH*, Vokietija) ir nucentrifuguojama 1 min. 1 500 aps./min. *Universal 32* (*Andreas Hettich GmbH & Co*, Vokietija) centrifugoje. Kapiliarinė elektroforezė atliekama

genetiniu analizatoriumi *ABI PRISM<sup>®</sup> 3130xl Genetic Analyzer* (*ThermoFisher Scientific*, JAV). Šiuo metodu nuskaitytos sekos analizuotos *Sequence Analysis v.5.1* (*Thermo Fisher Scientific*, USA) ir *Chromas v.2.4.4* (*Technelysium Pty Ltd*) programomis. Naudojant *ExPASy LALIGN* [102] bioinformacinį įrankį, sekos buvo lyginamos su referentine tiriamo geno seka (*ARID1B* NM\_020732.3, *GLI3* NM\_000168.6, *SLC9A6* NM\_001042537.1, *CHD7* NM\_017780.4, *TGFBR2* NM\_001024847.2, *BLM* NM\_000057.2, *CAPN3* NM\_000070.3, *MED13L* NM\_015335.5), gauta iš *ENSEMBL* [95] genomo naršyklės.

### 3.3.7. Naujos kartos sekoskaita

INGENES projekto metu dr. Violeta Mikštienė, kuriai asistavo disertacijos autorė, doc. dr. Živilė Maldžienė bei dokt. Gunda Petraitytė, atliko viso egzomo sekoskaitos (VES) tyrimą 31 tiriamajam, turinčiam IN ir / ar JA. Tyrimo metu nustatyti nauji DNR sekos variantai, kurie buvo siejami su tiriamų asmenų fenotipais. Tačiau disertaciniame darbe pagal užsibrėžtą darbo tikslą ir uždavinius yra aprašomas tik vienas tiriamasis (#8), kuris buvo tiriamas taikant šią technologiją.

VES atlikta pagal gamintojo (*Illumina, Inc.*, JAV) tiekiamus reagentų ir instrumentų protokolus. Šiam tyrimui išgryninta DNR (žr. 3.3.1 skyrių) buvo skiedžiama 1× TE buferiniu tirpalu iki 5 ng/μl koncentracijos. DNR bibliotekos paruoštos naudojant 24 mėginiams skirtą rinkinį *Nextera Exome Kit* (*Illumina, Inc.*, JAV) pagal protokolą *Nextera DNA Exome Reference Guide* (*Illumina, Inc.*, JAV). Siekiant tiksliai išmatuoti DNR bibliotekų koncentraciją, buvo naudojamas rinkinys *Qubit dsDNA BR Assay kit* (*Invitrogen*, JAV) ir *Qubit v.2.0* fluorometras (*Invitrogen*, JAV) pagal gamintojo rekomendacijas. DNR fragmentų gausinimui ir klasterių paruošimui buvo naudojami rinkiniai *TruSeq PE Cluster Kit v.3-HS* (*Illumina, Inc.*, JAV) ir *TruSeq Dual Index Sequencing Primer Box, Paired End* (*Illumina, Inc.*, JAV). Klasteriai amplifikuoti *cBot* instrumentu (*Illumina, Inc.*, JAV) vadovaujantis gamintojo vartotojo vadovu *cBot System Guide v.0.2* (*Illumina, Inc.*, JAV). VES atlikta naudojant rinkinį *TruSeq SBS Kit v.3-HS* (*Illumina, Inc.*, JAV) ir genetinį analizatorių *HiScanSQ* (*Illumina, Inc.*, JAV) pagal gamintojo vartotojo vadovą *HiScanSQ System User Guide v.0.1* (*Illumina, Inc.*, JAV).

Gautus duomenis anotavo dr. Tautvydas Rančelis. Nuskaitytos sekos sulygiuotos BWA (angl. *Burrows-Wheeler Alignment tool*) [193] programa pagal žmogaus genomo referentinę seką, pateikiamą genomo naršyklėje UCSC (GRCh37/hg19) [194]. NKS analizės algoritmui naudotas įrankis

GATK (angl. *Genome Analysis Toolkit*) [195]. Nustatytų variantų anotacijai buvo naudojamas *ANNOVAR v.2018Apr16* [196] bei *VarAFT v.2.16* (angl. *Variant annotation and filter tool*) [197] įrankis.

Siekiant užtikrinti rezultatų patikimumą ir kokybę, galutinių duomenų vertinimas buvo atliekamas skirtingų INGENES projekto vykdytojų, tarp jų taip pat darbo autorės. Analizei buvo atrinkti variantai, kurių retojo alelio dažnis sveikų žmonių populiacijose yra  $\leq 0,05$ . Siekiant įvertinti variantų patogeniškumą buvo remtasi informacija *ANNOVAR* programa sukurtame dokumente, ACMG kriterijais [19], duomenų bazėmis bei moksline literatūra (IV priedas). Atrinkti patogeniniai ir / arba galimai patogeniniai sekos variantai buvo patikrinti BAM dokumentuose naudojant įrankį *IGV v.2.7.0* (angl. *Integrative Genomics Viewer*) [198].

### 3.3.8. Fibroblastų kultūros genomo redagavimas CRISPR-Cas9 technologija

CRISPR-Cas9 genomo redagavimo technologija pritaikyta *MED13L* genui nutildyti neturinio ligos požymių vyriškosios lyties asmens fibroblastų kultūroje. Pirmiausia buvo atšildytas saugomas kontrolinio asmens fibroblastų mėginys. Kriomėgintuvėlis išimtas iš skysto azoto šaldiklio ir nedelsiant įdėtas į +37 °C vandenį. Toliau mėginys buvo apdorojamas pagal standartinį fibroblastų kultivavimo protokolą [185].

Siekiant nutildyti *MED13L* geną kultivuojamoje fibroblastų kultūroje, naudota *MED13L* DNR sekai specifinė gRNR (Nr. CRISPR927871\_SGM; *TrueGuide*<sup>™</sup>, *Invitrogen*, JAV), Cas9 baltymas (*TrueCut*<sup>™</sup>, *Invitrogen*, JAV) ir transfekcijos reagentas (*Lipofectamine*<sup>™</sup>, *CRISPRMAX*<sup>™</sup>, *Invitrogen*, JAV) vadovaujantis gamintojo vartotojo vadovu *TrueGuide*<sup>™</sup> *Synthetic gRNA User Guide* (Nr. MAN0017058; *Invitrogen*, JAV). Siekiant įvertinti ląstelių užsėjimo tankio įtaką genomo redagavimo efektyvumui, kontrolinio asmens fibroblastai užsėti skirtingu tankiu ( $\sim 0,3 \times 10^6$ ) į tris 9,6 cm<sup>2</sup> skersmens kultivavimo indelius. Atsižvelgiant į gamintojo rekomendacijas, transfekcijos procedūra atlikta esant 30–70 proc. ląstelių konfluencijai.

Po dviejų redaguotų fibroblastų kultivavimo dienų vertintas genomo redagavimo efektyvumas naudojant *Primer-BLAST* [188] įrankiu sukurtus taikiniui specifinius PGR pradmenis (III priedas, III.1 lentelė) ir komercinį rinkinį (*GeneArt*<sup>™</sup> *Genomic Cleavage Detection Kit*, *Invitrogen*, JAV). Tyrimas atliktas vadovaujantis gamintojo vartotojo vadovu *GeneArt*<sup>™</sup> *Genomic Cleavage Detection Kit User Guide* (Nr. MAN0009849, *Invitrogen*, JAV). Siekiant užtikrinti efektyvią PGR amplifikaciją, pradmenys buvo kuriami atsižvelgiant į bendrąsias gamintojo rekomendacijas: 18–22 bp ilgis,

45–60 proc. GC sąstatas, produkto ilgis ~400–500 bp. Taip pat pradžioje kurti taip, kad prognozuojama kirpimo vieta nebūtų amplikono viduryje, kadangi priešingu atveju gelio elektroforezės vaizde nebūtų galimybės atskirti dviejų PGR produktų. Gautų PGR produktų specifiškumas įvertintas 1,5 proc. agarozės (*Top Vision Agarose, ThermoFisher Scientific, JAV*) gelyje (žr. 3.3.6 skyrių). Perkirpto tikslinio fragmento frakcija pagal agarozės gelyje išryškėjusias juostas įvertinta *ImageJ v.1.52t (National Institutes of Health, JAV)* programa. Genomo redagavimo efektyvumas apskaičiuotas pagal matematinę formulę:

$$\text{Efektyvumas (proc.)} = 1 - ((1 - \text{perkirpto tikslinio fragmento frakcija})^{1/2})$$

Įvertinus genomo redagavimo efektyvumą, redaguotų fibroblastų kultūros buvo klonuojamos skiedimo metodu (angl. *Limiting dilution cloning*). Praskiedus ląstelių kultūras iki 8 ląstelių viename mililitre, užsėta dešimt 96 šulinėlių kultivavimo indelių (iš viso 960 šulinėlių). Įvertinus pro invertuotą mikroskopą *Olympus CKX41 (Olympus Life Science, Japonija)*, atrinkti kultivavimo indeliai, kuriuose augo viena ląstelė. Šios atrinktos ląstelės toliau kultivuotos standartinėmis sąlygomis (žr. 3.3.2 skyrių) iki funkciniais tyrimams pakankamo ląstelių skaičiaus. Tačiau dėl *force majeure* aplinkybių (karantinas dėl SARS-CoV2 viruso) tyrimas buvo pertrauktas, užšaldžius 16-ka gyvybingų ląstelių klonų.

#### 3.3.8.1. Fibroblastų kultūrų funkciniai tyrimai

##### **Fibroblastų gyvybingumo tyrimas**

Fibroblastų skaičius ir gyvybingumas kultivuojamose kultūrose buvo vertintas naudojant *Bürker* skaičiavimo kamerą (angl. *Bürker counting chamber; Heinz-Herenz, Vokietija*), vadovaujantis gamintojo instrukcijomis.

##### **Fibroblastų morfologijos tyrimas**

Kultivuojamų fibroblastų morfologija tirta kiekvieno pasažo metu, ląsteles stebint pro invertuotą mikroskopą *Olympus CKX41 (Olympus Life Science, Japonija)*. Ląstelės fotografuotos ankstyvuose (III–IV) ir vėlyvuose (X–XV) pasažuose, naudojant ląstelių analizatorių *JuLI™ (NanoEnTek, Pietų Korėja)*.

## Fibroblastų senėjimo tyrimas

Ląstelės augintos T25 flakonuose. Fibroblastų senėjimas tirtas naudojant komercinį dažymo rinkinį *Senescence Cells Histochemical Staining Kit* (*Sigma-Aldrich*, JAV) pagal gamintojo rekomendacijas.

Kultivuojamų ląstelių senėjimas buvo įvertintas invertuotu mikroskopu *Nikon* (*Nikon*, Japonija). Dešimt kultivavimo indelio vietų nufotografuota *ImageJ v.1.52t* (*National Institutes of Health*, JAV) programa. Nuotraukose suskaičiavus bendrą ląstelių skaičių bei mėlynai nusidažiusių ląstelių skaičių, buvo santykinai įvertintas nusidažiusių ląstelių skaičius, kuris vėliau buvo išreikštas procentais.

## Fibroblastų ciklo tyrimas

Fibroblastai, kultivuoti mitybinėje terpėje *AmnioMAX* (*ThermoFisher Scientific*, JAV; žr. 3.3.2 skyrių), apdorojami pagal standartinį fibroblastų kultivavimo protokolą [185]. Gautos ląstelių suspensijos perpilamos į 1,5 ml talpos centrifuginius mėgintuvėlius ir centrifuguojama 10 min. 1 000 g. Nusiurbus supernatantą, fibroblastai suspenduojami 1 ml šalto PBS (*Gibco*, Jungtinė Karalystė) tirpale. Ląstelės dar kartą centrifuguojamos tomis pačiomis sąlygomis, suspenduojamos 100 µl PBS tirpale ir sulašinama 900 µl etanolio (96 proc.). Paruošti mėginiai laikomi 24 val. 4 °C temperatūroje.

Kitą dieną matuojamas ląstelių ciklas. Nucentrifugavus (5 min., 1 000 g) mėginius ir nusiurbus supernatantą, fibroblastai suspenduojami 1 ml PBS tirpale. Mėginiai dar kartą centrifuguojami tomis pačiomis sąlygomis. Nuosėdos suspenduojamos 500 µl propidžio jodido/RNazės (*BD Biosciences*, JAV) tirpale ir, apsaugant nuo šviesos, inkubuojama 15 min. kambario temperatūroje. Paruošti mėginiai perpilami į tėkmės citometro *BD FACSCalibur* (*BD Biosciences*, JAV) mėgintuvėlius ir analizuojami naudojant *Cell quest* programą.

### 3.3.9. Statistinė analizė

Statistinė duomenų analizė atlikta naudojant *SDS v.2.3* (*Applied Biosystems*, JAV), *ExpressionSuite v.1.1* (*ThermoFisher Scientific*, JAV), *qbase<sup>+</sup>* (*Biogazelle*, Belgija), *Microsoft Excel 2010* (*Microsoft Corporation*, JAV) programas bei *R v.3.4.0* [190] statistinį paketą. Atlikus statistinę analizę, gauti duomenys pateikti kaip aritmetinis vidurkis ± standartinis nuokrypis.

### 3.3.10. Bioinformacinė duomenų analizė

Siekiant įvertinti tiriamo varianto patogeniškumą bei įtaką baltymo struktūrai ir funkcijai buvo atliekama *in silico* analizė naudojant įvairias duomenų bazes, bioinformacinius įrankius ir analizuojant tarptautinę mokslinę literatūrą. Moksliniam darbui naudotos duomenų bazės ir bioinformaciniai įrankiai yra pateikti IV priede.

#### 4. TYRIMŲ REZULTATAI IR JŲ APTARIMAS

Pastaruosius dešimtmečius sparčiai tobulinami viso žmogaus genomo molekuliniai genetiniai tyrimai suteikė galimybę efektyviau ir išsamiau identifikuoti įvairių genetinių ligų, įskaitant intelektualinės negalios ir / ar įgimtų anomalijų (IN / IA), genetines priežastis. *ClinVar* [8] duomenimis, šiuo metu yra registruota: 10 092 (23 proc.) patogeniniai, 4 329 (10 proc.) galimai patogeniniai, 3 363 (8 proc.) nepatogeniniai, 4 915 (11 proc.) galimai nepatogeninių bei 20 522 (47 proc.) nežinomos klinikinės reikšmės variantai, kurie yra siejami su IN [9]. Todėl, nepaisant spartaus progreso, didžioji dalis šių genetinių pokyčių yra nežinomos ar vis dar nepatvirtintos klinikinės reikšmės.

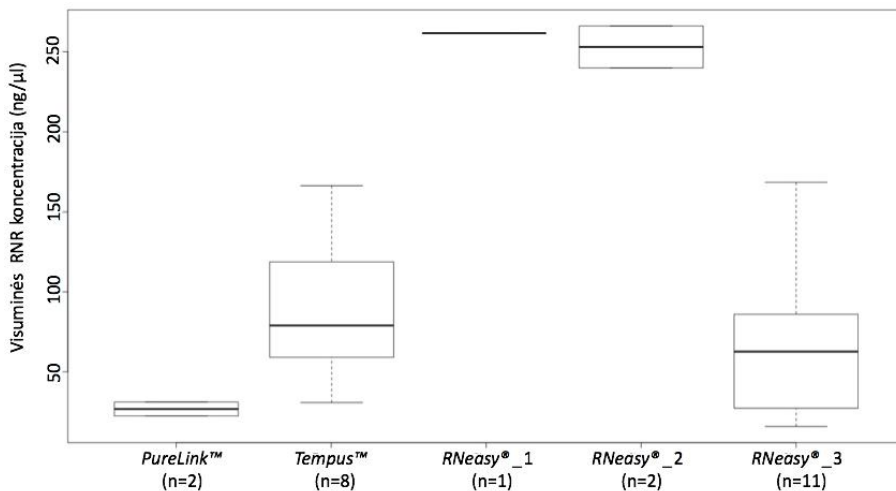
Siekiant įvertinti koduojančios ir nekoduojančios DNR sekos pokyčių patogeniškumą, naudojant ankstesnių mokslinių projektų (LITGEN, CHERISH, PROGENET, UNIGENE) duomenis buvo atrinkta dešimt tiriamųjų / tiriamųjų šeimų (#1–#10; 3.1 lentelė), kurioms pasireiškė intelektualinė negalia (IN), psichomotorinės raidos atsilikimas (PRA) ir / arba įgimtos anomalijos (IA). Taikant molekulinis bei funkcinis genomo tyrimo metodus, vykdant INGENES projektą, nežinomos klinikinės reikšmės DNR sekos variantai ir KSP ištirti iRNR lygmeniu.

iRNR būdingas jautrumas hidrolizės procesui ir mažesnis stabilumas, palyginti su DNR [15, 38, 48], todėl molekulių ir funkcinų iRNR tyrimų rezultatai priklauso nuo visuminės RNR gryninimo efektyvumo. Šiame moksliniame darbe buvo naudojami keli visuminės RNR gryninimo metodai.

Pirmiausia, dviejų tiriamųjų (#8 ir #9) visuminė RNR buvo išgryninta iš periferinio kraujo ėminių naudojant komercinį *PureLink™ Total RNA Blood Kit (Invitrogen, JAV)* rinkinį. Naudojant šį rinkinį, tiriamųjų išgrynintos visuminės RNR koncentracijos buvo nedidelės (22,4 ng/μl ir 31,1 ng/μl; žr. „*PureLink™*“ 4.1 pav.). Nors gautos visuminės RNR koncentracijos ir mėginių kokybė ( $A_{260/280} \sim 2$ ) buvo pakankama tolesniems tyrimams, dėl iRNR stabilumo ir mažos išgrynintos visuminės RNR išeigos parinktas ir pritaikytas kitas, *Tempus™ Spin RNA Isolation Kit (ThermoFisher Scientific, JAV)*, visuminės RNR gryninimo rinkinys. Naudojant šį rinkinį buvo išgryninta visuminė RNR iš visų kitų darbe aprašomų tiriamų asmenų periferinio veninio kraujo ėminių ( $n = 8$ ). Išgrynintos visuminės RNR koncentracijos (49,3 ng/μl – 166,5 ng/μl; žr. „*Tempus™*“ 4.1 pav.) ir mėginių kokybė ( $A_{260/280}: 2,04–2,17$ ) buvo geresnė, palyginti su rezultatais, gautais pirmuoju visuminės RNR gryninimo rinkiniu. Šiuo komerciniu rinkiniu išgrynintos visuminės RNR mėginiai taip pat pasižymėjo didesniu RNR stabilumu tiek gryninimo metu, tiek po ilgalaikio šaldymo.



Nors moksliniams tyrimams dažniausiai visuminė RNR gryninama iš periferinio veninio kraujo éminio, tam tikrų genų raiška periferiniame veniniame kraujyje yra labai maža arba iš viso nenustatoma. Nenustačius *GLI3* geno raiškos tiriamojo #2 periferinio veninio kraujo mėginyje, visuminė RNR buvo išgryninta iš fibroblastų kultūros naudojant *RNeasy® Mini Kit* (*Qiagen*, Vokietija) rinkinį (žr. „*Rneasy®\_1*“ 4.1 pav.). Šis visuminės RNR gryninimo metodas taip pat buvo naudojamas CRISPR-Cas9 eksperimente (žr. 4.2.3 skyrių), gryninant visuminę RNR iš neredaguotų ankstyvųjų pasažų (III–IV) tiriamojo #10 ir neturinčio ligos požymių kontrolinio asmens fibroblastų kultūros. Nustatyta visuminės RNR koncentracija atitinkamai 266,1 ng/μl ir 239,9 ng/μl (žr. „*Rneasy®\_2*“ 4.1 pav.). Redagavus kontrolinio asmens fibroblastų kultūrą, iš 11-kos redaguotų vėlyvųjų pasažų (X–XV) ląstelių klonų buvo išgryninta visuminė RNR ir įvertinta jos koncentracija (nuo 15,8 ng/μl iki 168,5 ng/μl; žr. „*Rneasy®\_3*“ 4.1 pav.). Dėl ilgalaikio kultivavimo ir / ar genomo redagavimo pasekmių redaguotose vėlyvųjų pasažų fibroblastų kultūrose buvo mažesnis skaičius ląstelių. Vertinant tiek ankstyvuosius, tiek vėlyvuosius kultivuojamų fibroblastų pasažus, nustatyta tiesioginė koreliacija tarp fibroblastų skaičiaus kultūroje ir išgrynintos visuminės RNR koncentracijos.



**4.1 pav.** Visuminės RNR koncentracijos<sup>4</sup> priklausomai nuo gryninimo metodų: *PureLink™ Total RNA Blood Kit* (*Invitrogen*, JAV); *Tempus™ Spin RNA*

<sup>4</sup> 4.1 paveiksle vaizduojami 10 tiriamųjų ir vieno kontrolinio asmens, tirtu CRISPR-Cas9 genomo redagavimo eksperimento metu, rezultatai. Neįtraukti genų raiškos tyrimuose tirtų kontrolinių asmenų visuminės RNR gryninimo rezultatai, kadangi genų raiškos tyrimų efektyvumui, atkartojamumui ir patikimumui užtikrinti visuminė RNR išgryninta tuo pačiu

*Isolation Kit (ThermoFisher Scientific, JAV); Rneasy® Mini Kit (Qiagen, Vokietija)*. Stačiakampė diagrama nubrėžta naudojant *R v.3.4.0* [190] statistikos paketą. Diagramos apatinė ir viršutinė vientisa linija žymi atitinkamai 25-ąjį ir 75-ąjį procentilį, o diagramos „ūsai“ nurodo mažiausią ir didžiausią koncentraciją. Horizontali paryškinta linija stačiakampyje žymi vidutinę koncentraciją

Tiriant iš išgrynintos visuminės RNR susintetintą kDNR, septyniems tiriamiesiems atitinkamai įvertintas *ARID1B*, *GLI3*, *SLC9A6*, *CHD7*, *TGFBR2*, *BLM* ar *CAPN3* genuose nustatytų neaiškios klinikinės reikšmės pre-iRNR sukirpimo variantų patogeniškumas; penkiems tiriamiesiems atitinkamai įvertintas *TGFBR2*, *BLM*, *GPC5*, *HNRNPD*, *HNRNPDL*, *ENOPH1*, *RASGEF1B*, *PRKG2*, *MEDI13L*, *RBI*, *E2F1* ar *CCNC* genų raiškos pokytis ir vienam tiriamajam patikslintas vidugeninės *MEDI13L* delecijos dydis. Tiriant nustatytos *MEDI13L* delecijos poveikį ląstelės lygmeniu, nauja CRISPR-Cas9 technologija taikyta šio geno nutildymo eksperimentui odos fibroblastų kultūroje. Siekiant įvertinti DNR sekos sukirpimo variantų ir KSP galimą poveikį baltymo lygmeniu bei suprasti sudėtingą ryšį tarp genotipo ir fenotipo, atlikta išsami *in silico* ir mokslinės literatūros analizė. Detalūs kiekvieno atlikto genetinio tyrimo rezultatai ir jų aptarimas pateikiami tolesniuose disertacijos skyriuose.

#### 4.1. Nekoduojančios DNR sekos sukirpimo variantų molekulinii ir funkcinių tyrimų rezultatai ir jų aptarimas

Remiantis *ClinVar* [8] duomenimis, 10 092 patogeniniai variantai yra siejami su IN [9]. Nekoduojančios DNR sekos sukirpimo vietų variantai sudaro apie 7 proc. visų patogeninių su IN siejamų variantų. Dažniausiai šie variantai lemia vieno ar, rečiau, kelių egzonzų iškritą. Taip pat šie patogeniniai variantai dažnai gali lemti kriptinį alternatyvųjų 5' ar 3' galo sukirpimą gretimame egzone arba introno (ar jo dalies) intarpą. Tačiau iki šiol nėra žinomas tikslus mechanizmas, koku būdu įvykus pokyčiui DNR sekoje pasirenkama alternatyvi kirpimo vieta [50, 52, 55]. Nors sukirpimo vietos variantai nėra tokia dažna IN / IA priežastis kaip taškiniai variantai ar KSP, šių variantų patogeniškumą taip pat reikalinga įvertinti. Šiuo tikslu būtina atlikti išsamią iRNR ir / ar kDNR analizę [48].

---

metu ir tuo pačiu gryninimo metodu kaip ir tiriamų asmenų. Dėl šios priežasties gautos koncentracijos nereikšmingai skyrėsi nuo tiriamųjų.

Išanalizavus ankstesnių projektų (LITGEN, CHERISH, PROGENET, UNIGENE) duomenis, septyniems tiriamiesiems (#1–#7; žr. 3.1 lentelę), turintiems IN / IA, atrinkti nežinomos klinikinės reikšmės sukirpimo vietos variantai, kurie buvo nustatyti VES ar genų grupės NKS technologijomis. Trijų genų (*ARID1B*, *GLI3*, *SLC9A6*) DNR sekos variantai nustatyti konservatyvioje 5' donorinėje sukirpimo vietoje, kitų trijų genų (*CHD7*, *TGFBR2*, *BLM*) variantai nustatyti konservatyvioje 3' akceptorinėje sukirpimo vietoje, o *CAPN3* geno variantas nustatytas 20 nt atstumu nuo akceptorinės srities. Keturiems tiriamiesiems (#1–#4) genominės DNR sekos variantai patvirtinti Sangerio metodu vykdant INGENES projektą, o kitų tiriamųjų variantai tuo pačiu metodu buvo patvirtinti anksčiau. Sangerio metodu taip pat atlikta variantų segregacijos analizė, kuri patvirtino variantų kilmę. Nustatyta, kad *ARID1B* ir *CHD7* genų variantai yra *de novo* kilmės, o *GLI3*, *SLC9A6*, *BLM* ir *CAPN3* DNR sekos variantai yra paveldėti iš vieno iš tėvų. *TGFBR2* c.1600-2A>G varianto kilmė nenustatyta, nes nebuvo galimybės ištirti tiriamojo #5 tėvų.

Atlikus mokslinės literatūros analizę, nustatyta, kad *ARID1B*, *GLI3*, *SLC9A6*, *CHD7* ir *TGFBR2* sukirpimo vietos variantai yra nauji ir iki šiol literatūroje neaprašyti. Taip pat šie variantai nėra registruoti *ClinVar* [8], *HGMD* [34], *dbSNP* [92], *dbVar* [93], 1 000 genomų projekto [6] bei *gnomAD* [96] duomenų bazėse. *BLM* c.2308-2A>G bei *CAPN3* c.1746-20C>G variantai mažesniu nei 1 proc. dažniu yra nustatomi skirtingose pasaulio populiacijose. Tačiau, *ClinVar* [8] duomenimis, iki šiol šių variantų patogeniškumas nėra patvirtintas. *BLM* c.2308-2A>G variantas dviejuose šaltiniuose yra klasifikuojamas kaip patogeninis ar galimai patogeninis. *CAPN3* c.1746-20C>G variantas trijuose šaltiniuose yra klasifikuojamas kaip nežinomos klinikinės reikšmės, o kituose – kaip patogeninis (1), nepatogeninis (1) ir galimai nepatogeninis (1).

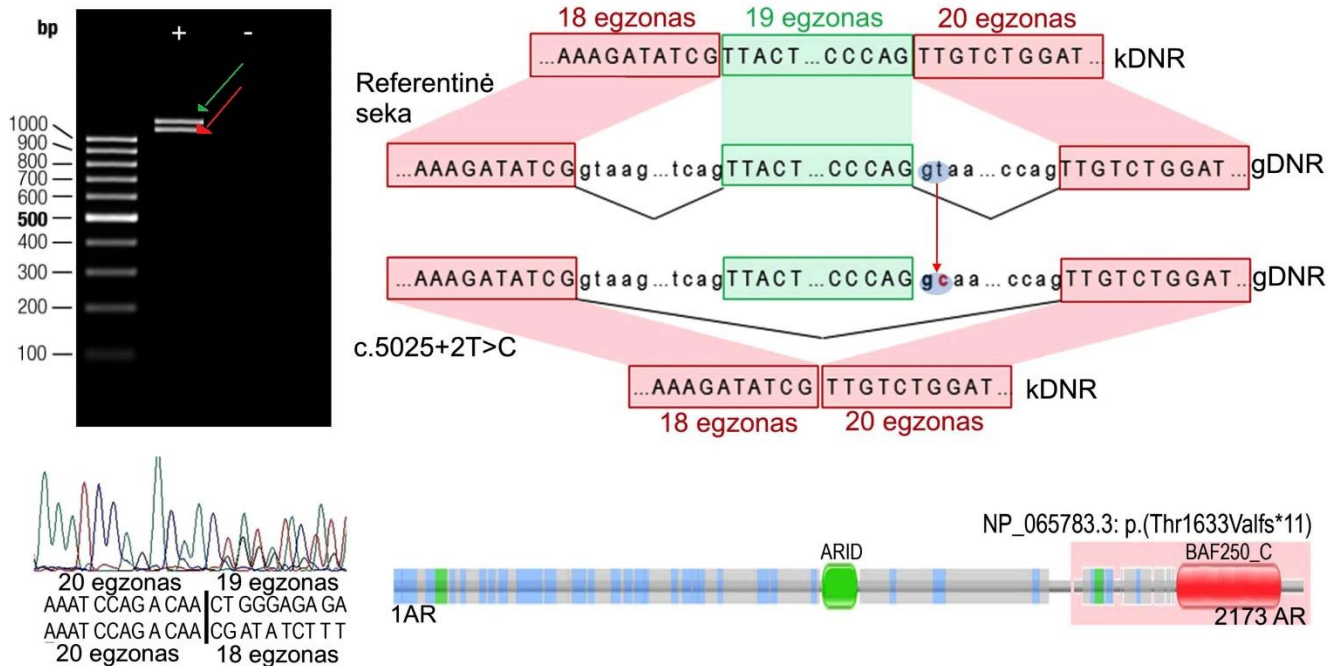
Naudojant bioinformacinius įrankius, įvertintas septynių variantų poveikis pre-iRNR sukirpimui. *MutationTaster2* [99] įrankiu šešių genų (*ARID1B*, *GLI3*, *SLC9A6*, *CHD7*, *TGFBR2*, *BLM*) sukirpimo variantai buvo įvertinti kaip patogeniniai, o *CAPN3* geno variantas priskirtas prie nepatogeninių DNR sekos variantų grupės. Nepaisant to, remiantis *Human Splicing Finder v.3.1* [101] prognozėmis, visi variantai galimai keičia pre-iRNR sukirpimo procesą. Nors kai kurių (pvz., *missense* ar *nonsense*) variantų atveju, atlikus išsamią *in silico* analizę, surenkama pakankamai įrodymų, pagrindžiančių varianto patogeniškumą ir jo įtaką fenotipui, sukirpimo variantų atveju tai yra tik tarpinis etapas. Todėl, siekiant tiksliai įvertinti šių variantų padarinius iRNR lygmeniu, būtina atlikti detalesnius molekulinis ir / ar funkcinis eksperimentus [82, 113].

#### 4.1.1. *ARID1B* c.5025+2T>C variantų tyrimas

Patogeniniai variantai šeštoje chromosomoje ilgajame petyje (6q25.3) esančiame *ARID1B* (MIM #614556) gene yra viena dažniausių sindrominės ir nesindrominės IN priežasčių [170, 199]. *HGMD* [34] duomenimis, šiuo metu registruojama 116 patogeninių *ARID1B* DNR sekos variantų, iš kurių dažniausi yra delecijos / insercijos (39,7 proc.) bei *missense* / *nonsense* (29,3 proc.) tipo variantai, o apie pusę visų patogeninių *ARID1B* variantų yra siejama su autosominiu dominantiniu būdu paveldimu pirmojo tipo Coffino–Siris sindromu (MIM #135900; ORPHA #1465).

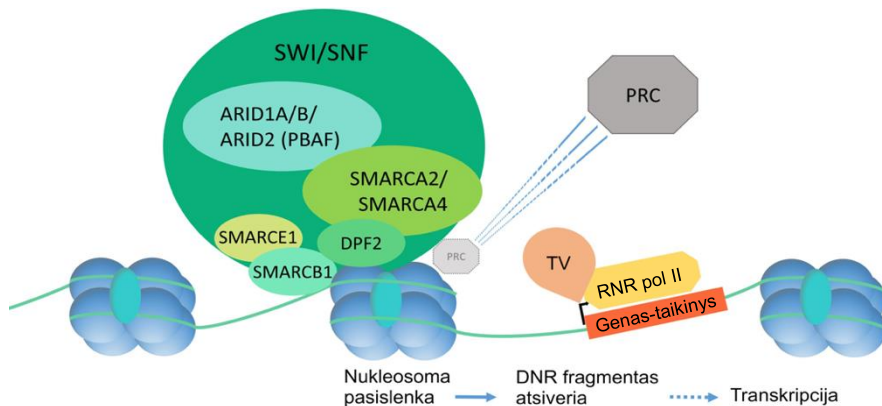
Genetinio konsultavimo metu 14 metų amžiaus moteriškosios lyties tiriamam asmeniui #1, kuriam pasireiškė sunki IN, psichomotorinis raidos atsilikimas bei įvairus spektras IA: žemas ūgis, grubūs veido bruožai, regėjimo sutrikimai, galvos smegenų pakitimai, hipertrichozė ir kt., buvo įtariamas Coffino–Siris (CSS) sindromas. Moksliniais tyrimais nustatyta, kad šį sindromą lemia genų, koduojančių SWI/SNF (angl. *SWItch/Sucrose Non-Fermentable*) komplekso subvienetus, patogeniniai variantai. Priklausomai nuo pažeisto geno, išskiriama net 11-ka CSS tipų. Todėl, atsižvelgiant į heterogeninę CSS prigimtį, VES yra tinkamiausias šią patologiją lemiančių genetinių priežasčių nustatymo metodas.

Atlikus išsamią tiriamojo #1 gDNR mėginio VES duomenų analizę, nustatytas naujas heterozigotinis 5' donorinės sukirpimo vietos NC\_000006.12(NM\_020732.3):c.5025+2T>C variantas 19-ame *ARID1B* introne. Tiriamojo #1 ir jo tėvų gDNR segregacijos tyrimas Sangerio metodu patvirtino varianto *de novo* kilmę. Siekiant įvertinti sukirpimo vietos varianto poveikį pre-iRNR sukirpimo mechanizmui, atliktas tiriamojo #1 kDNR tyrimas Sangerio metodu (4.2 pav. A, B). Tyrimo metu patvirtinta *ARID1B* geno 19-o egzono iškrita (4.2 pav. C). *In silico* dėl 131 bp ilgio 19-ojo egzono nebuvimo kDNR susidaro skaitymo rėmelio poslinkis ir priešlaikinis baigmės kodonas (NP\_065783.3:p.(Thr1633Valfs\*11)). Dėl šios priežasties, manoma, yra sintetinas trumpesnis ARID1B (UniProtKB #Q8NFD5) baltymas, neturintis BAF250\_C domeno (4.2 pav. D).



**4.2 pav.** Tiriamojo #1 kDNR mėginio tyrimas Sangerio metodu: **A)** PGR produkto vizualizacija agarozės gelyje. Elektroforezės pirmame takelyje (žymima „+“) vizualizuojamas 1162 bp ilgio laukinio tipo alelio (žalia rodyklė) kDNR fragmentas bei 1031 bp ilgio pakitusio alelio (raudona rodyklė) kDNR fragmentas, o antrame takelyje (žymima „-“) vizualizuojama neigiama PGR kontrolė. Gelio elektroforezės rezultatai vertinti pagal 100 bp DNR molekulinės masės ir dydžio standartą; **B)** Tiriamojo #1 PGR produkto chromatograma; **C)** Sukirpimo c.5025+2T>C varianto sutrikdyto pre-iRNR sukirpimo proceso rezultato schema; **D)** Linijinė ARID1B baltymo (UniProtKB #Q8NFD5) struktūra su apytiksliai nurodyta prognozuojamo NP\_065783.3:p.(Thr1633Valfs\*11) pokyčio vieta. Schema adaptuota pagal *Pfam* [120] duomenų bazę

ARID1, kurio yra žinomos dvi izoformos – ARID1A ir ARID1B, yra vienas pagrindinių SWI/SNF, kitaip vadinamo BAF (angl. *Brahma-Associated Factor*), komplekso subvienetų [200]. Šis sudėtingas baltymų kompleksas dalyvauja nuo ATP priklausomo chromatinio struktūros pertvarkymo procese. SWI/SNF tiesiogiai sąveikaujant su PRC (angl. *Polycomb Repressive Complex 1*) ir naudojant ATP energiją, šalina nukleosomas nuo DNR grandinės. Tokiu būdu II tipo RNR polimerazė ir transkripcijos veiksniai jungiasi prie taikinio ir aktyvuoja transkripcijos procesą (4.3 pav.) [199]. Be to, SWI/SNF gali keisti genų raišką sąveikaujant su histonų acetiltransferaze, kuri acetilina trečiąjį (H3) ir ketvirtąjį (H4) histoną [201].



**4.3 pav.** SWI/SNF baltymų komplekso reguliuojama genų raiška. Adaptuota pagal mokslinę literatūrą [199]

SWI/SNF kompleksas sąveikauja su DNR dėl ARID1B baltymo ARID domeno, kuris atpažįsta ir jungiasi prie AT turtingos DNR sekos [199]. Kitas funkcinis požūriu svarbus ARID1B domenas yra šio baltymo C gale esantis BAF250. Yanas su bendradarbiais (2008), užslopinę (angl. *Knock-out*) BAF250 domeno raišką pelių EKL, nustatė padidėjusią šių ląstelių diferenciaciją ir sumažėjusį atsinaujinimo potencialą. Todėl šis domenas yra ypač svarbus ankstyvos embriogenezės metu [202]. Tiek BAF250, tiek ARID domenas yra ypač svarbus įvairių audinių funkcijoms palaikyti. *GTE*x portalo [118] duomenimis, *ARID1B* geno raiška didžiausia yra kraujo, odos, skydliaukės, kiaušidžių, šlapimo pūslės, smegenų audiniuose bei embrioninėse kamieninėse ląstelėse. Jungas su kolegomis (2017) tirdami modelines *Arid1b*<sup>+/-</sup> peles įrodė, kad šio geno haplonepakankamumas turi įtakos nervų sistemos neurotransmisijos procesui, tarpneuroninių pirmtakų proliferacijai bei GABAerginių tarpneuronų kiekiui smegenų žievėje. Taip pat tyrimo metu buvo stebimi *Arid1b*<sup>+/-</sup> pelių kognityvinių funkcijų bei

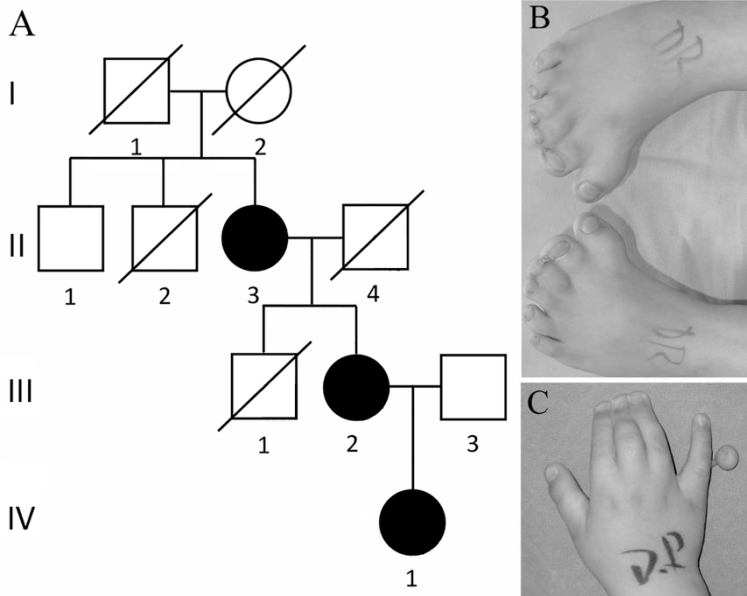
adaptyvaus elgesio sutrikimai, kurie yra analogiški CSS sergantiems asmenims [200].

Remiantis ankstesnių tyrimų rezultatais, galima teigti, kad patogeniniai variantai *ARID1B* gene neabejotinai nulemia neuropsichiatrinis klinikinius požymius, kurie pasireiškia tiriamajam #1. Šiam tiriamajam nustatytas sukirpimo vietos c.5025+2T>C variantas sutrikdo *ARID1B* pre-iRNR sukirpimo procesą. Atsižvelgiant į nustatyto c.5025+2T>C varianto reikšmę pre-iRNR sukirpimui ir tiriamajam #1 pasireiškiančius klinikinius požymius, šis *ARID1B* variantas yra patogeninis ir lemia pirmojo tipo CSS sindromą.

#### 4.1.2. *GLI3* c. 473+3A>T varianto tyrimas

Polidaktilija, kuri gali būti tiek sindrominė, tiek nesindrominė, yra viena dažniausių plaštakų ir pėdų įgimtų anomalijų. Priklausomai nuo papildomo piršto atsiradimo vietos, klasifikuojama postaksialinė polidaktilija bei preaksialinė polidaktilija, kurios pagal Temtamy ir McKusicko (1978) apibrėžtą klasifikacijos schemą papildomai skirstomos į kelis tipus [203]. IV tipo preaksialinė polidaktilija (PPD4; MIM #174700; ORPHA #93338), kitaip vadinama polisindaktilija, kartu pasireiškiant su kitais klinikiniais požymiais yra būdinga keliems sindromams, tarp jų GCPS (angl. *Greig Cephalopolysyndactyly syndrome*; MIM #175700; ORPHA #380) bei PHS (angl. *Pallister-Hall syndrome*; MIM #146510; ORPHA #672). Šiuos autosominiu dominantiniu būdu paveldimus sindromus nulemia septintos chromosomos trumpajame petyje (7p14.1) esančio *GLI3* (angl. *Gli-kruppel Family Member 3*; MIM #165240) geno patogeniniai variantai.

GCPS būdingi požymiai – plaštakų ir pėdų pirštų PPD4 (4.4 pav. B, C) bei specifinės kraniofacialinės srities anomalijos, genetinio konsultavimo metu buvo stebimos #2 šeimos IV-1 ir III-2 nariams. PPD4 taip pat buvo diagnozuota šios šeimos II-3 nariui (4.4 pav. A).

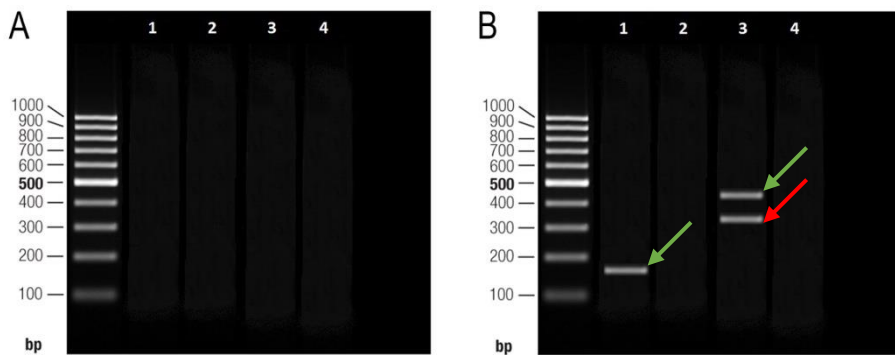


**4.4 pav.** A) Tiriamosios #2 šeimos genealoginis medis. Juodos spalvos apskritimai žymi sukirpimo c.473+3A>T variantu *GLI3* gene paveldėjimą; B) Tiriamojo IV-1 pėdų pirštų polisindaktilija; C) Tiriamojo IV-1 plaštakos pirštų polisindaktilija (gyd. D. Radzevičiaus nuotraukos)

Sangerio sekoskaitos metodu ištyrus IV-1 tiriamojo DNR mėginį, nustatyti du heterozigotiniai variantai *GLI3* gene: 5' sukirpimo vietos NC\_000007.14(NM\_000168.6):c.473+3A>T variantas ir *missense* tipo NM\_000168.6:c.2179G>A (NP\_000159.3:p.(Gly727Arg); rs121917710) variantas. Atlikus variantų segregacijos analizę, sukirpimo vietos c.473+3A>T variantas patvirtintas taip pat III-2 ir II-3 tiriamos šeimos nariams. Tiriant c.2179G>A variantu segregaciją, nustatyta, kad IV-1 tiriamasis paveldėjo šį variantą iš neturinio ligos požymių tėvo (III-3). Anksčiau c.2179G>A variantą, paveldimą keturių kartų šeimoje, aprašė Radhakrishna su bendradarbiais (1999). Mokslininkai nustatė c.2179G>A variantą aštuoniems tiriamiesiems, kuriems pasireiškė postaksialinė polidaktilija A ir B tipo, o pirštų sindaktilija nebuvo stebima [204]. Tačiau, remiantis naujesniais 1 000 genomų projekto [6] ir *gnomAD* [96] duomenų bazių duomenimis, šis variantas yra nustatomas sveikiems asmenims, atitinkamai 0,003 ir 0,006 dažniu. Dėl šios priežasties tikėtina, kad tiriamajai šeimai #2 būdingi klinikiniai požymiai yra nulemti tik *GLI3* sukirpimo c.473+3A>T variantu.



Siekiant įvertinti anksčiau literatūroje neaprašyto sukirpimo c.473+3A>T varianto patogeniškumą, atliktas II-3 tiriamajam (toliau tiriamasis #2) kDNR molekulinis tyrimas Sangerio metodu. Pirmiausia tirta kDNR, susintetinta pagal iRNR, kuri buvo išgryninta iš periferinio veninio kraujo mėginio. Atlikus PGR ir įvertinus gelio elektroforezės vaizdą, nenustatyta jokio PGR produkto nei tiriamajam #2, nei kontroliniam asmeniui (4.5 pav. A). *GTEx* portalo [118] duomenimis, *GLI3* geno raiška kraujo mėginyje yra itin maža (TPM = 0,03), todėl galėjo būti nepakankama molekuliniam genetiniam tyrimui atlikti. Dėl šios priežasties toliau tirta kDNR, susintetinta pagal iRNR, kuri buvo išgryninta iš tiriamojo #2 fibroblastų kultūros. Pakartojus PGR ir įvertinus gelio elektroforezės vaizdą, sėkmingai nustatytas 166 bp ilgio PGR produktas (4.5 pav. B). Remiantis šiuo rezultatu, iškelta prielaida, kad reakcijoje amplifikavosi tik laukinio tipo alelio (angl. *wild-type allele*) PGR produktas. Dėl šios priežasties PGR pakartota sukūrus naujus pradmenis, apimančius ilgesnį DNR fragmentą (419 bp; III priedas, III.1 lentelė). Išanalizavus gelio elektroforezės vaizdą, nustatyti du PGR produktai, kurių ilgis įvertintas apie 419 bp ir 313 bp (4.5 pav. B).

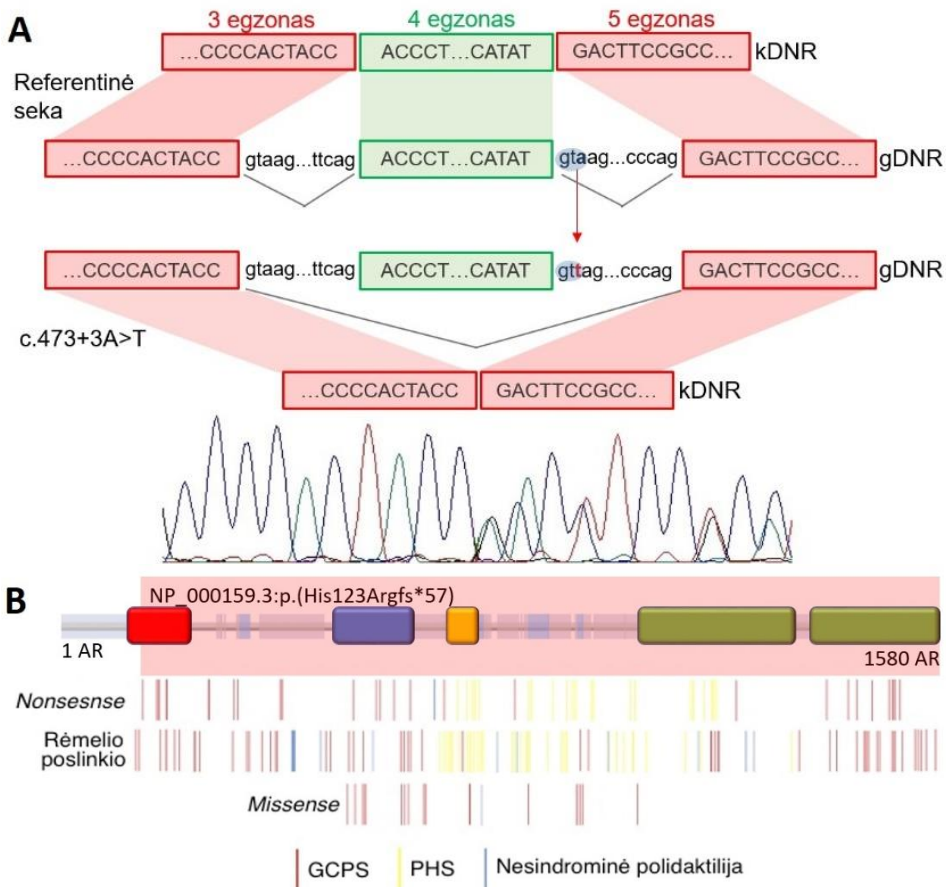


**4.5 pav.** Tiriamojo #2 kDNR mėginio PGR produktų vizualizacija agarozės gelyje. Elektroforezės rezultatai vertinti pagal 100 bp DNR molekulinės masės ir dydžio standartą: **A)** Tiriamojo #2 kDNR, susintetintos pagal iRNR, kuri buvo išgryninta iš periferinio veninio kraujo mėginio, tyrimas. Tiriamojo #2 PGR produktas vizualizuojamas gelio pirmame ir antrame takeliuose, o trečiame ir ketvirtame takeliuose vizualizuojamas kontrolinio asmens kDNR tyrimas. PGR produktų negauta; **B)** kDNR, susintetintos pagal iRNR, kuri buvo išgryninta iš fibroblastų kultūros, tyrimas. Tiriamojo #2 laukinio tipo alelio (žalia rodyklė) 166 bp ilgio DNR fragmentas vizualizuojamas gelio pirmame takelyje. Tiriamojo #2 laukinio tipo alelio 419 bp ilgio DNR fragmentas bei pakitusio alelio (raudona rodyklė) 313 bp ilgio DNR fragmentas vizualizuojamas gelio trečiame takelyje. Neigiama PGR kontrolė atspindima antrame ir ketvirtame takeliuose

Sangerio metodu ištyrus gautus PGR produktus, nustatyta, kad dėl 5<sup>4</sup> sukirpimo vietos c.473+3A>T varianto sutrikdomas pre-iRNR sukirpimo procesas iškrentant ketvirtam egzoniui, kurio ilgis yra 106 bp (4.6 pav. A). Bioinformaciniais įrankiais analizuojant sutrikdyto pre-iRNR sukirpimo proceso įtaką baltymo lygmeniui, nustatyta, kad dėl sukirpimo varianto susidaro 57-ių aminorūgščių skaitymo rėmelio poslinkis ir priešlaikinis baigmės kodonas. Dėl šios priežasties tikėtina, kad yra sintetinamas trumpesnis NP\_000159.3:p.(His123Argfs\*57) baltymas arba inicijuojamas iRNR skaidymas. Remiantis moksline literatūra ir *Pfam* [104] bei *UniProt* [103] duomenų bazėmis (4.6 pav. B), galima teigti, kad trumpesnis GLI3 (*UniProtKB* #P10071) baltymas neturi slopinančiojo domeno (AR<sup>5</sup> 106-AR 263), cinko pirštelių (angl. *Zinc-finger*) domeno (AR 462-AR 645), proteolitinės skaidymo vietos (AR 703-AR 740), *transaktyvinančiojo* domeno 2 (AR 1044-AR1322) bei *transaktyvinančiojo* domeno 1 (AR 1376-AR 1580) [205].

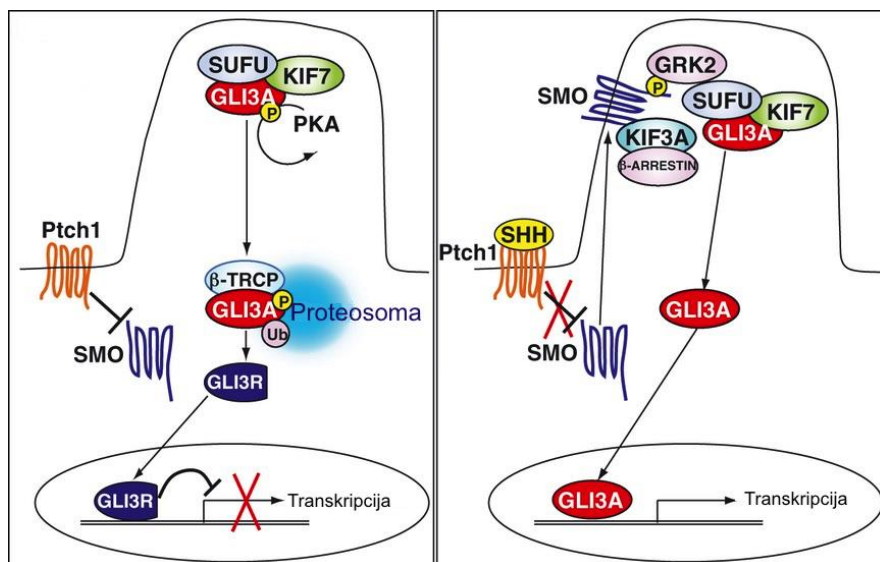
---

<sup>5</sup> AR – aminorūgštis.



**4.6 pav. A)** Tiriamojo #2 kDNR mėginio chromatograma ir sukirpimo c.473+3A>T varianto sutrikdyto pre-iRNR sukirpimo proceso rezultato schema; **B)** Linijinė GLI3 baltymo (UniProtKB #P10071) struktūra su apytiksliai nurodyta prognozuojamo NP\_000159.3: p.(His123Argfs\*57) pokyčio vieta ir pagrindiniais baltymo domenais. Raudonu stačiakampiu vaizduojamas slopinantis domenas (AR 106-AR 263), mėlynu – cinko pirštelių domenas (AR 462-AR 645), geltonu – proteolitinė skaidymo vieta (AR 703-AR 740), žaliu – transaktyvinantis domenas 2 (AR 1044-AR1322) bei transaktyvinantis domenas 1 (AR 1376-AR 1580). Žemiau pavaizduoti dažniausi patogeniniai *GLI3* missense, nonsense ir rėmelio poslinkio variantai, lemiantys skirtingus fenotipus: GCPS, PHS ir nesindrominę polidaktiliją. Schema sudaryta remiantis *Pfam* duomenų baze [104] ir mokslinė literatūra [205, 206]

GLI3 baltymas – tai transkripcijos veiksnys, kuris dalyvaujant SHH (angl. *Sonic Hedgehog Signaling Molecule*) signaliniame kelyje turi dvigubą poveikį susijusių genų raiškai dėl slopinančio ir *transaktyvuojančių* struktūrinių domenų. Tačiau moksliniais tyrimais įrodyta, kad balansas tarp slopinančio ir aktyvuojančių domenų yra ypač svarbus ankstyvos embriogenezės metu formuojantis galūnių pirštams (4.7 pav.) [207, 208]. Modelinių *Gli3*<sup>+/-</sup> pelių tyrimai patvirtina šį teiginį, kadangi šioms heterozigotinėms pelėms yra stebima polidaktilija ir kraniofacialinių anomalijų požymiai, kurie būdingi GCPS bei PHS fenotipui [209].



**4.7 pav.** GLI3 dalyvavimas SHH signaliniame kelyje. **A)** Specifiniai baltymai sąveikaudami su aktyvinančiu GLI3 (žymima GLI3A) baltymu jį fosforilina. Fosforilintas GLI3A yra degraduojamas proteosomoje, taip sudarant trumpesnę slopinančią GLI3 baltymo formą (žymima GLI3R), kuri slopina genų, atsakingų už galūnių pirštų formavimąsi, raišką; **B)** SHH sąveikaujant su specifiniais baltymais, GLI3 sudaro baltymų kompleksą, kuris išvengia proteolitinio skaidymo taip palaikant GLI3 aktyvia forma [208]

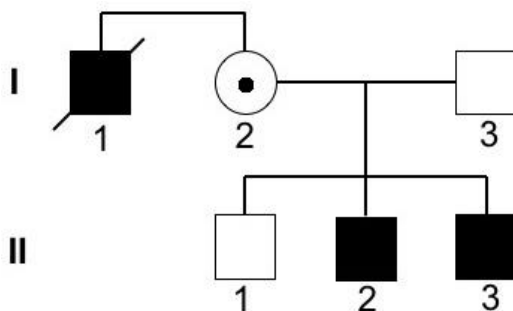
Patogeniniai *GLI3* variantai sutrikdo balansą tarp aktyvuojančiojo ir slopinančiojo *GLI3* transkripto, todėl sutrikdoma SHH signaliniame kelyje veikiančių genų veikla. *HGMD* [34] duomenimis, šiuo metu registruota daugiau nei 200 patogeninių variantų *GLI3* gene. Didžioji dalis šių genetinių variantų lemia sindrominę polidaktiliją – GCPS. Manoma, kad nuo variantų vietos *GLI3* geno srityje priklauso skirtingi fenotipai – GCPS, PHS ar nesindrominė polidaktilija. GCPS yra siejamas su DNR sekos variantais, kurie

yra *GLI3* geno pradžioje (iki 1 998 nt) ir gale (nuo 3481 nt). PHS lemia geno vidurio patogeniniai variantai, kurie nepažeidžia represuojančiojo domeno (4.6 pav. B). Nepaisant to, patogeniniai variantai ne visada tiksliai koreliuoja su klinikiniais požymiais, todėl iki šiol nėra žinomas skirtingus fenotipus nulemiantis mechanizmas [206, 210].

Atsižvelgiant į tiriamajam #2 pasireiškiančius klinikinius požymius ir gautus genetinių tyrimų rezultatus, naujas sukirpimo c.473+3A>T variantas, sutrikdantis pre-iRNR sukirpimo procesą, yra laikytinas patogeniniu ir lemiančiu GCPS sindromą.

#### 4.1.3. *SLC9A6* c.899+1G>A varianto tyrimas

IN yra būdinga tiek vyrams, tiek moterims, tačiau, mokslinės literatūros duomenimis, vyriškosios lyties asmenims šis sutrikimas pasireiškia apie du kartus dažniau [112]. Toks pasiskirstymas aiškinamas tuo, kad daugiau nei 100 genų, siejamų su IN, yra X chromosomoje, o naujų genų šioje chromosomoje identifikuojama vis daugiau [112, 161, 163]. Vienas iš tokių su X chromosoma susijusių genetinių susirgimų buvo įtariamas tiriamajai šeimai #3, kurioje vyriškosios lyties asmenims (I-1, II-2, II-3) pasireiškė IN, psichomotorinės raidos sutrikimas, mažas ūgis, mikrocefalija, traukuliai ir kitos IA (4.8 pav.).

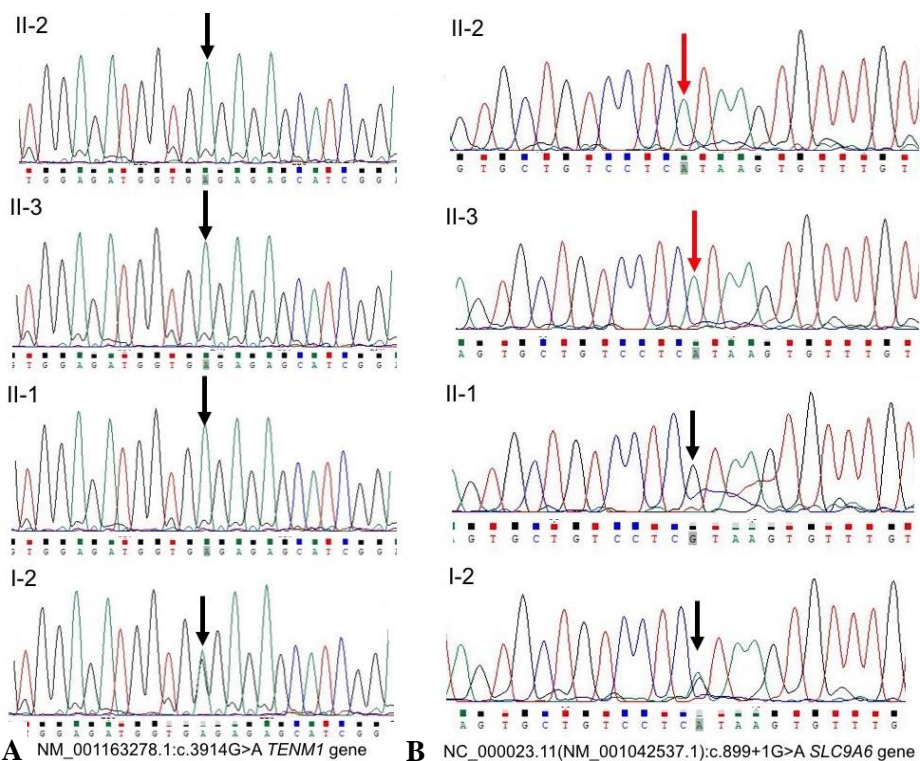


**4.8 pav.** Tiriamosios šeimos #3 genealoginis medis. Juodos spalvos kvadratai žymi vyriškosios lyties asmenis, turinčius ligos požymius. Juodos spalvos perbrauktas kvadratas žymi vyriškosios lyties asmenį, turėjusį ligos požymius. Tašku žymėtas apskritimas žymi su X chromosoma susijusios ligos moteriškosios lyties nešiotąją

Išanalizavus VES duomenis, II-2 ir II-3 šeimos #3 nariams X chromosomoje nustatyti trys variantai – *TENM1* NM\_001163278.1:c.3914G>A, *TAF9B* NM\_015975.4:c.510T>A ir *SLC9A6* NC\_000023.11(NM\_001042537.1):c.899+1G>A. Nors *TENM1* (MIM

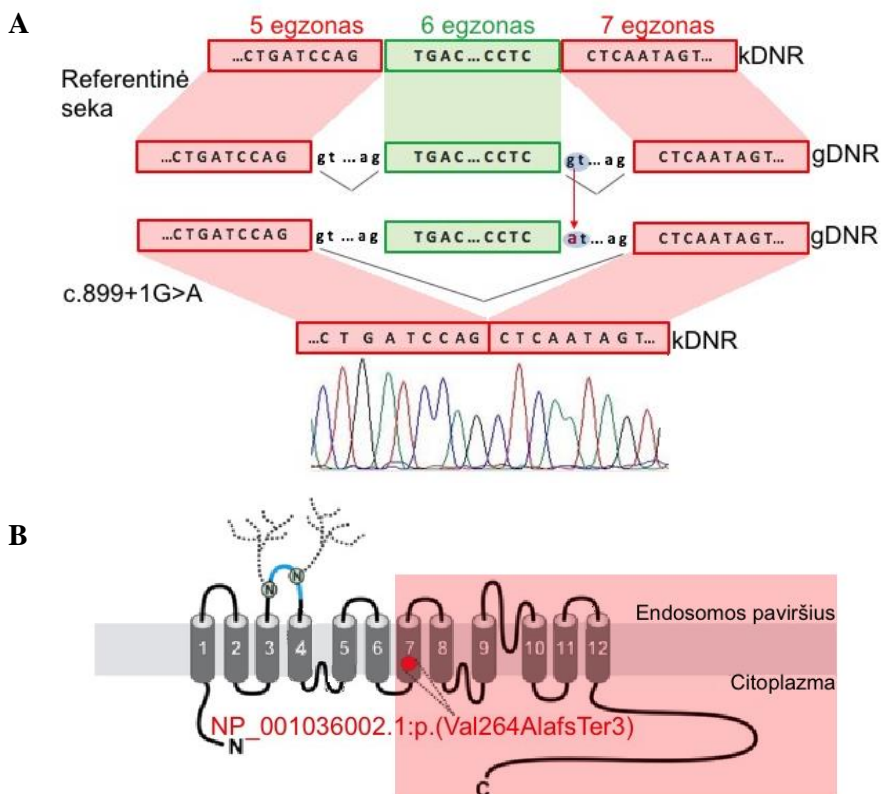
#300588) ir *TAF9B* (MIM #300754) iki šiol nėra siejami su jokia pagal Mendelio dėsnius paveldima liga, abiejų genų raiška yra intensyvi smegenų ląstelėse, o šių genų išjungimas modeliniuose organizmuose nulemia neuropsichiatrinius sutrikimus [211, 212]. Patogeniniai variantai *SLC9A6* (MIM #300231) gene siejami su Christiansono tipo sindromine IN (MIM #300243; ORPHA #85278), kuri charakterizuojama progresuojančia IN, sunkiais psichomotorinės raidos sutrikimais, mikrocefalija, hipotonija, įvairaus tipo traukuliais ir IĀ. Šis su X chromosoma susijęs dominantiniu būdu paveldimas sindromas pirmą kartą buvo aprašytas prieš 20 metų [213]. Nepaisant to, *HGMD* [34] duomenimis, iki šiol nustatyta tik 18-ka patogeninių *SLC9A6* variantų, iš kurių trys yra sukirpimo vietos variantai.

Siekiant įvertinti trijų nežinomos klinikinės reikšmės variantų segregaciją šeimoje #3, atlikta dviejų sergančių probandų (II-2, II-3), neturinčio ligos požymių sibso (II-1) bei motinos (I-2) gDNR mėginių Sangerio sekoskaita. Tiriant *TAF9B* geną, *missense* tipo variantas nebuvo nustatytas, kadangi visiems tiriamiems šeimos nariams, įskaitant probandus, tiriamoje pozicijoje nustatytas nukleotidas, atitinkantis žmogaus referentinės sekos (NM\_015975.4) duomenis. Toks rezultatas gali būti aiškinamas tuo, kad NKS metu *TAF9B* geno sekos *missense* tipo variantas galėjo būti neteisingai prilygiuotas dėl genome esančių panašios sekos genų. Analizuojant *TENM1* geno Sangerio sekoskaitos rezultatus, *missense* tipo variantas nustatytas II-2, II-3 bei II-1 tiriamiesiems hemizigotinėje būklėje bei I-2 tiriamojo heterozigotinėje būklėje (4.9 pav. A). Iš I-2 (motina) paveldėtas variantas taip pat būdingas neturinčiam ligos požymių sibsiui II-1, todėl šis genetinis pokytis negali paaiškinti ligos patogenezės. Analizuojant *SLC9A6* genetinio tyrimo rezultatus, nustatyta, kad sukirpimo vietos variantas nėra būdingas II-1, tačiau II-2 ir II-3 paveldėjo šį variantą iš I-2, kuri yra c.899+1G>A varianto nešioja (4.9 pav. B). Iškelta prielaida, kad 5' donorinės sukirpimo vietos variantas galimai nulėmė II-2 ir II-3 tiriamųjų, kuriems būdingi identiški klinikiniai požymiai, sindrominės IN patogenezės mechanizmą.



**4.9 pav.** Tiriamosios šeimos #3 gDNR sekos variantų Sangerio sekoskaitos rezultatai. **A)** NM\_001163278.1:c.3914G>A variantas nustatytas *TENM1* gene; **B)** NC\_000023.11(NM\_001042537.1):c.899+1G>A variantas nustatytas *SLC9A6* gene. Juoda rodyklė žymi varianto vietą. Raudona rodyklė žymi galimai patogeninį sukirpimo vietos variantą, kuris toliau tirtas kDNR lygmeniu

Iškelta hipotezė dėl *SLC9A6* gene nustatyto c.899+1G>A varianto patogeniškumo buvo tikrinama Sangerio sekoskaitos metodu tiriant vieno iš probandų (II-2; toliau tiriamasis #3) kDNR mėginį. Nustatyta, kad 5' donorinės sukirpimo vietos variantas lemia 6-ojo egzono iškritą (4.10 pav. A). *In silico* dėl prarasto 6-ojo egzono *SLC9A6* gene susidaro skaitymo rėmelio poslinkis (NP\_001036002.1:p.(Val264Alafs\*3)). Dėl susidariusio priešlaikinio baigmės kodono, tikėtina, yra sintetinamas pakitusios funkcijos trumpesnis SLC9A6 (UniProtKB #Q9258; (4.10 pav. B) baltymas arba vyksta NMD inicijuojamas šios iRNR skaidymas.



**4.10 pav.** A) Tiriamąjo #3 kDNR mėginio chromatograma ir c.899+1G>A varianto pre-iRNR sukirpimo proceso rezultato schema; B) SLC9A6 baltymo (UniProtKB #Q9258) struktūros schema endosomos membranoje su apytiksliai nurodyta prognozuojamo NP\_001036002.1:p.(Val264Alafs\*3) pokyčio vieta. Cilindrinėmis struktūromis pažymėti transmembraniniai domenai, dalyvaujantys Na<sup>+</sup> bei H<sup>+</sup> jonų pernašoje. Schema adaptuota pagal mokslinę literatūrą [214, 215]

*GTEx* portalu [118] duomenimis, *SLC9A6* intensyviausia raiška yra smegenų audiniuose. Nepaisant to, šis baltymas taip pat yra svarbus palaikant kitų audinių, pavyzdžiui, inkstų ir žarnyno, veiklą. *SLC9A6* koduojamas produktas yra vienas iš devynių monovalentinių Na<sup>+</sup> ir H<sup>+</sup> jonų nešiklių (angl. *Na<sup>+</sup>/H<sup>+</sup> exchanger*; NHE) šeimos baltymų, todėl kitaip vadinamas NHE6. Šis baltymas, daugiausia randamas ląstelių endosomų ir mitochondrijų membranose, yra atsakingas už Na<sup>+</sup> bei H<sup>+</sup> jonų pernašą priešinga kryptimi tarp organelės ir ląstelės citoplazmos [216]. NHE6 baltymą sudaro 12-a transmembraninių spiralių, iš kurių šešta ir septinta yra būdingos visiems NHE šeimos baltymams [216, 217]. Remiantis Wakabayashi ir kt. (2000) sukurtu



modeliu, prognozuojamas NP\_001036002.1:p.(Val264Alafs\*3) pokytis yra konservatyvioje septintoje transmembraninėje spiralėje (4.10 pav. B). Todėl tikėtina, kad tiriamajam #3 nustatytas sukirpimo vietos c.899+1G>A variantas *SLC9A6* gene sutrikdo Na<sup>+</sup> ir H<sup>+</sup> jonų pernašą ir tokiu būdu nulemia Christiansono tipo sindrominės IN fenotipą. Tikslus mechanizmas, galintis paaiškinti genotipo ir fenotipo ryšį, iki šiol nėra žinomas. Tačiau Ouyang su bendradarbiais (2013), tirdami modelines peles, nustatė, kad patogeniniai variantai *Slc9a6* sutrikdo nervinių ląstelių aksonų ir dendritų šakojimąsi, kuris yra būtinas tinkamai neuronų veiklai užtikrinti. Tikėtina, kad sutrikus šiai veiklai vystosi kognityvinių funkcijų bei adaptyvaus elgesio sutrikimai [218].

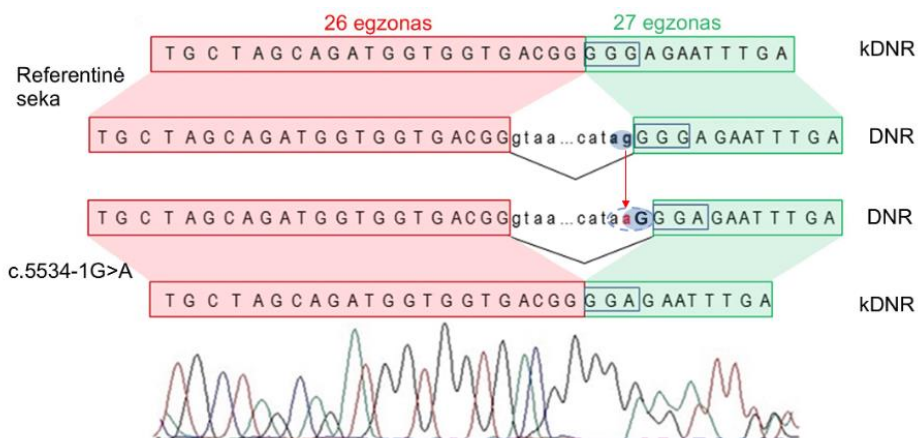
Atsižvelgiant į sukirpimo vietos c.899+1G>A varianto sutrikdomą pre-irNR sukirpimo mechanizmą ir baltymo prognozuojamus struktūrinius pokyčius, tiriamajam #3 nustatytas *SLC9A6* genetinis pokytis yra patogeninis ir siejamas su Christiansono tipo sindromine intelektine negalia.

#### 4.1.4. *CHD7* c.5535-1G>A varianto tyrimas

CHARGE sindromas (MIM #214800; ORPHA #138), kuris genetinio konsultavimo metu buvo įtariamas moteriškosios lyties tiriamajam asmeniui #4, charakterizuojamas akies koloboma, širdies yda, choanų atrezija, psichomotorinės raidos atsilikimu, genitalijų ĮA, ausų ĮA ir / ar kurtumu (angl. *Coloboma of the eye, heart anomaly, choanal atresia, retardation of mental and somatic development, genital abnormalities, ear abnormalities and/or deafness*) [219]. Dalis šių klinikinių požymių būdinga taip pat MFDGA sindromui (angl. *Mandibulofacial dysostosis Guion-Almeida type*; MIM #610536; ORPHA #79113), Kabuki sindromui 1 (MIM #147920; ORPHA #2322) bei hipogonadotropiniam hipogonadizmui (MIM #612370; ORPHA #432), kuriuos atitinkamai lemia genetiniai variantai *EFTUD2* (MIM #603892), *KMT2D* (MIM #602113) bei *CHD7* (MIM #608892) genuose. O autosominiu dominantiniu būdu paveldimą CHARGE sindromą lemia patogeniniai DNR sekos variantai *SEMA3E* (MIM #608166) ir *CHD7* genuose [220–223]. Manoma, kad aštuntos chromosomos ilgajame petyje (8q12) esančio *CHD7* geno heterozigotiniai variantai lemia apie 70 proc. visų CHARGE sindromo atvejų [224]. Specifinės *CHD7* duomenų bazės [225] duomenimis, šiuo metu yra registruojamas 801 genetinis variantas *CHD7* gene, o iš jų 686 yra patogeniniai arba nežinomos klinikinės reikšmės [226]. Remiantis *HGMD* [34] duomenų baze, iš viso skaičiuojami 737 patogeniniai variantai, iš kurių dažniausi yra *missense / nonsense* (43 proc.), insercijos / delecijos (39,2 proc.) bei sukirpimo variantai (12,5 proc.).

Atlikus išsamią 16 metų amžiaus tiriamojo #4 VES duomenų analizę, nustatytas nežinomos klinikinės reikšmės heterozigotinis sukirpimo vietos variantas NC\_000008.11(NM\_017780.4):c.5535-1G>A, kuris yra *CHD7* 26-ojo introno akceptorinėje srityje, bei patogeninis heterozigotinis variantas NM\_004004.5:c.167delT, NP\_003995.2:p.(Leu56Argfs\*26), rs80338942 *GJB2* gene. Patogeninių ar galimai patogeninių variantų *EFTUD2*, *KMT2D* ar *SEMA3E* genuose nenustatyta.

Siekiant įvertinti sukirpimo varianto c.5535-1G>A patogeniškumą ir galimą poveikį tiriamam fenotipui, atlikta tiriamojo #4 kDNR sekos analizė Sangerio sekoskaitos metodu. Nustatyta, kad sukirpimo vietos variantas aktyvuoja alternatyvią 3' akceptorinę sukirpimo vietą, kuri nuo pirminės vietos yra nutolusi tik vienu nukleotidu. Tokiu būdu paskutinis 26-ojo introno A nukleotidas ir pirmasis 27-ojo egzono G nukleotidas tampa alternatyvia 3' sukirpimo vieta (4.11 pav.).

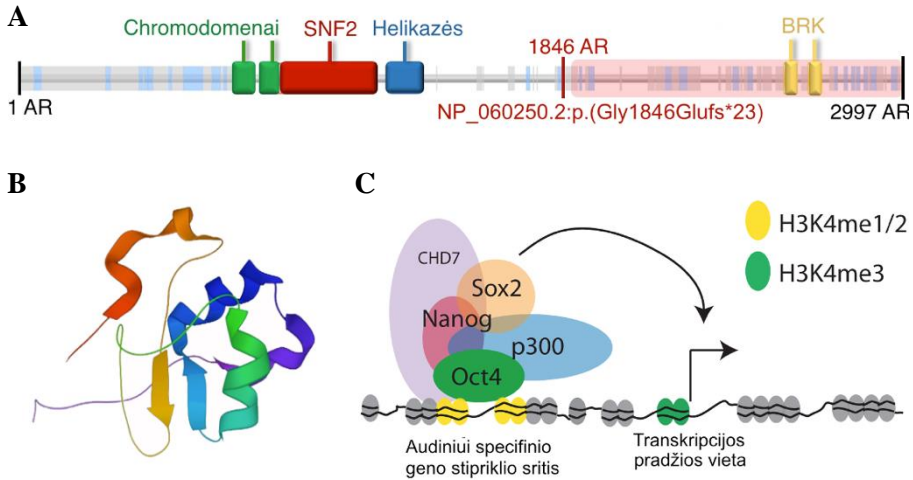


**4.11 pav.** Tiriamojo #4 kDNR mėginio chromatograma ir aktyvuojamos alternatyviosios 3' akceptorinės sukirpimo vietos *CHD7* 26-ame introne proceso schema

*ENSEMBL* [95] duomenų bazės duomenimis, apie 46,52 proc. (271 628/583 835) žmogaus referentinės sekos egzonų prasideda G nukleotidu. Taip pat, remiantis moksline literatūra, gana dažnai akceptorinėje srityje -1 pozicijoje G nukleotidas keičiamas į A. McDaniellas su kolegomis (2006) aprašė *NOTCH2* (MIM # 600275) geno sukirpimo variantą c.5930-1G>A, kai pirmasis 33-iojo egzono nukleotidas yra G. Panašiai Pasmooij ir kt. (2007), Mukhopadhyay ir kt. (2006), Gelincik ir kt. (2015), Simpsonas ir kt. (2007) aprašė akceptorinės vietos variantus -1 pozicijoje, kai gretimame egzone pirmasis nukleotidas yra G [227–231]. Visais atvejais šie sukirpimo variantai lėmė egzono iškritą. Anksčiau yra

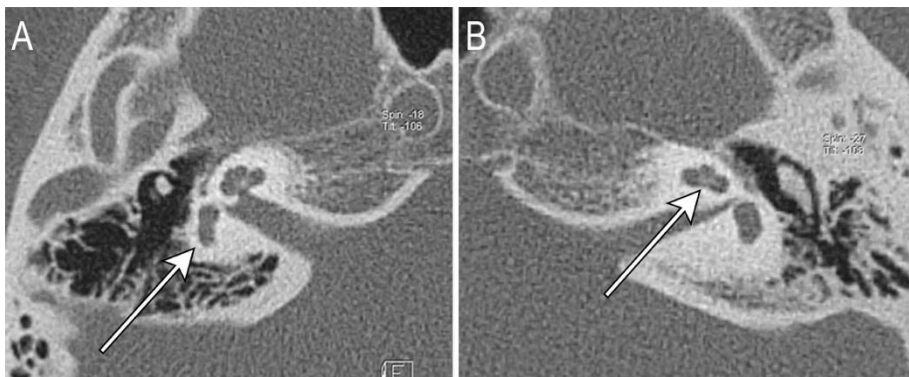
aprašytas tik vienas *ATM* (MIM #607585) geno patogeninis variantas c.4437-1G>A, kuris lėmė kriptinės alternatyvios sukirpimo vietos susidarymą vienu nukleotidu toliau nuo pirminės vietos, panašiai kaip tiriamojo #4 atveju [53]. Iki šiol nėra žinomas tikslus sukirpimo mechanizmas, nulemiantis kirpimo vietos pasirinkimą [50, 52, 55]. Manoma, kad tam gali turėti įtakos genomo vieta, ESST ir ESSL tankis genome ar net pre-iRNR antrinės struktūros [48].

*In silico* įvertinus sukirpimo varianto c.5535-1G>A poveikį baltymo lygmeniu, nustatyta, kad susidaro 23-jų aminorūgščių skaitymo rėmelio poslinkis NP\_060250.2:p.(Gly1846Glufs\*23). Dėl susidariusio priešlaikinio baigmės kodono, tikėtina, sintetinamas trumpesnis CHD7 baltymas (angl. *Chromodomain Helicase DNA-Binding Protein 7*; UniProtKB #Q9P2D1) arba inicijuojamas iRNR skaidymas. CHD7 baltymą sudaro 2 997 aminorūgštys, o pagrindinės šio baltymo struktūrinės dalys yra du chromodomenai, SNF2 ir helikazės domenai bei du BRK domenai (4.12 pav. A) [232, 233]. Pastarieji du BRK domenai (4.12 pav. B) dėl tiriamo sukirpimo varianto nesusidaro. Moksliniais tyrimais yra nustatyta, kad BRK domenai reguliuoja genų raišką ir chromatino organizacijos procesą, sąveikaujant su aukštesniųjų eukariotų chromatino komponentais (pvz., CTCF baltymu) [234]. Schnetzas su bendradarbiais (2009) nustatė, kad CHD7 kartu su kofaktoriais (angl. *Cofactor*), pavyzdžiui, P300, OCT4, SOX2 ir NANOG, reguliuoja audiniui specifinių genų raišką. Baltymų kompleksas atpažįsta reguliacinio elemento, vadinamojo stipriklio, trečiojo histono ketvirtoje pozicijoje esantį (mono- arba di-) metilintą liziną (H3K4) ir, sudarant chromatino kilpą, priartina stipriklį prie transkripcijos pradžios vietos (angl. *transcription start site*; 4.12 pav. C). Manoma, kad tokiu būdu reguliuojama audiniui specifinių genų raiška yra ypač svarbi embrioninių kamieninių ląstelių proliferacijai ir diferenciacijai, o sutrikus šiam mechanizmui nulemiami CHARGE sindromui būdingi klinikiniai požymiai [232, 235]. Šį mechanizmą patvirtina *Chd7*<sup>+/-</sup> modelinių pelių tyrimai, kurių metu hemizigotinės pelės pasižymėjo klinikiniais požymiais, siejamais su CHARGE sindromu [236–238].



**4.12 pav. A)** Linijinė CHD7 baltymo (UniProtKB #Q9P2D1) struktūra su apytiksliai nurodyta prognozuojamo NP\_060250.2:p.(Gly1846Glu) pokyčio vieta. Schema adaptuota pagal *Pfam* duomenų bazę [104]; **B)** BRK domeno trimatės struktūros schema, sukurta *LiteMol* [239] įrankiu; **C)** CHD7 reguliuojama audiniui specifinių genų raiška. Schema adaptuota pagal mokslinę literatūrą [235]

Nors tiriamajam #4 pasireiškia CHARGE sindromui būdingas klausos sutrikimas (4.13 pav.), ši patologija gali būti siejama su patogeniniais DNR sekos variantais tiek *CHD7*, tiek *GJB2* gene.



**4.13 pav.** Tiriamojo #4 kaukolės laikinojo kaulo kompiuterinės tomografijos vaizdai. **A)** Pusapvalių kanalų aplazija; **B)** Nervinio kanalo atrezija

Patogeniniai *GJB2* geno variantai dažniausiai yra siejami su autosominiu recesyviu būdu paveldimu klausos sutrikimu (MIM #220290). Anksčiau rs80338942 variantas keletą kartų buvo nustatytas tiriamiesiems

asmenims iš bendros Lietuvos populiacijos, kuriems pasireiškė įvairaus laipsnio klausos sutrikimai ar net kurtumas. Nustatyta, kad homozigotinis arba sudėtinis heterozigotinis variantas lemia skaitymo rėmelio poslinkį ir priešlaikinį baigmės kodoną, dėl kurio susidaro sutrumpėjęs baltymas, nulemiantis klausos defektus. Tačiau yra patvirtinta, kad *GJB2* variantas heterozigotinėje būklėje yra nepakankamas paaiškinti klausos sutrikimo potogenezę [240]. Todėl tiriamasis #4 yra šio *GJB2* patogeninio varianto nešiotojas, o jam pasireiškiantis klausos sutrikimas yra aiškinamas sukirpimo variantu, kuris buvo nustatytas *CHD7* gene.

Remiantis išanalizuota literatūra, gautais genetinio tyrimo rezultatais ir tiriamojo #4 klinikinio fenotipu, c.5535-1G>A sukirpimo vietos variantas yra laikytinas patogeniniu ir lemiančiu dalinį arba visišką *CHD7* geno haplonepakankamumą, dėl kurio tiriamajam #4 diagnozuotas CHARGE sindromas.

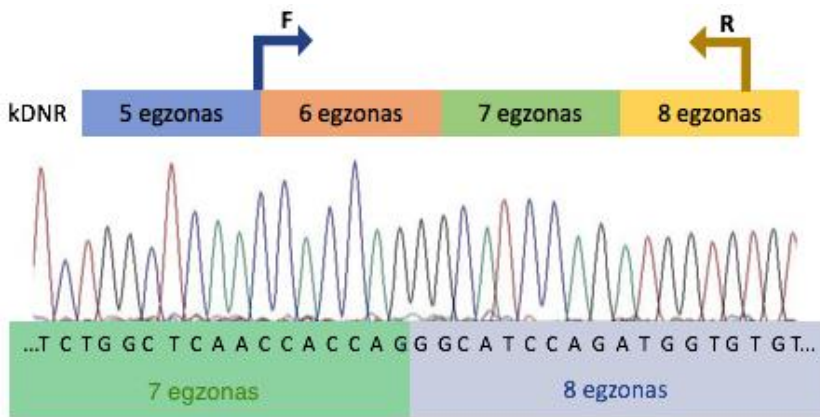
#### 4.1.5. *TGFBR2* c.1600-2A>G varianto tyrimas

Baltymai, veikiantys TGF- $\beta$  (angl. *Transforming growth factor beta*) signaliniame kelyje, reguliuoja daugelį ląstelės funkcionavimui būtinų procesų, įskaitant augimo, diferenciacijos, apoptozės bei homeostazės. Šiuos baltymus koduojančių genų patogeniniai variantai lemia autosominiu dominantiniu būdu paveldimus Marfano, Shprintzeno–Goldbergo bei Loeyso–Dietzo sindromus (LDS), kurie charakterizuojami skeleto, akių, plaučių, širdies ir kraujagyslių sistemų pažeidimais (4.1 lentelė) [241]. Dėl pasireiškiančios įgimtos skeleto deformacijos bei įgimtų širdies defektų vienas šių sindromų buvo įtartas 21 metų amžiaus tiriamajam #5.

**4.1 lentelė.** TGF- $\beta$  signaliniame kelyje veikiantys genai ir siejami sindromai pagal *OMIM* [91]

Sindromas	MIM	Genas	MIM
Marfano	#154700	<i>FBN1</i> (angl. <i>Fibrillin 1</i> )	#134797
Shprintzeno–Goldbergo	#182212	<i>SKI</i> (angl. <i>SKI proto-oncogene</i> )	#164780
LDS1	#609192	<i>TGFB1</i> (angl. <i>Transforming Growth Factor-Beta Receptor, Type I</i> )	#190181
LDS2	#610168	<i>TGFB2</i> (angl. <i>Transforming Growth Factor-Beta Receptor, Type II</i> )	#190182
LDS3	#613795	<i>SMAD3</i> (angl. <i>SMAD Family Member 3</i> )	#603109
LDS4	#614816	<i>TGFB2</i> (angl. <i>Transforming Growth Factor-Beta, Type II</i> )	#190220
LDS5	#615582	<i>TGFB3</i> (angl. <i>Transforming Growth Factor-Beta, Type III</i> )	#190230

NKS genų grupės tyrimo technologija ištyrus tiriamojo #5 gDNR mėginį, nustatytas galimai patogeninis sukirpimo vietos NC\_000003.12(NM\_001024847.2):c.1600-2A>G variantas *TGFBR2* geno heterozigotinėje būklėje. Šis 3' akceptorinės sukirpimo vietos variantas yra *TGFBR2* septintame introne, o aštuntas egzonas yra paskutinis tiriamo geno transkripto egzonas. Atsižvelgiant į ankstesnių tyrimų rezultatus, iškelta hipotezė, kad c.1600-2A>G variantas gali lemti kriptinės alternatyvios sukirpimo vietos susidarymą arba septinto egzono iškritą. Sukūrus oligonukleotidinius pradmenis, apimančius penkto–šešto egzonų jungties vietą, septintą ir dalį aštunto egzono (III priedas, III.1 lentelė), Sangerio metodu nusekvenotas tik nepakitusio alelio kDNR sekos fragmentas (4.14 pav.).



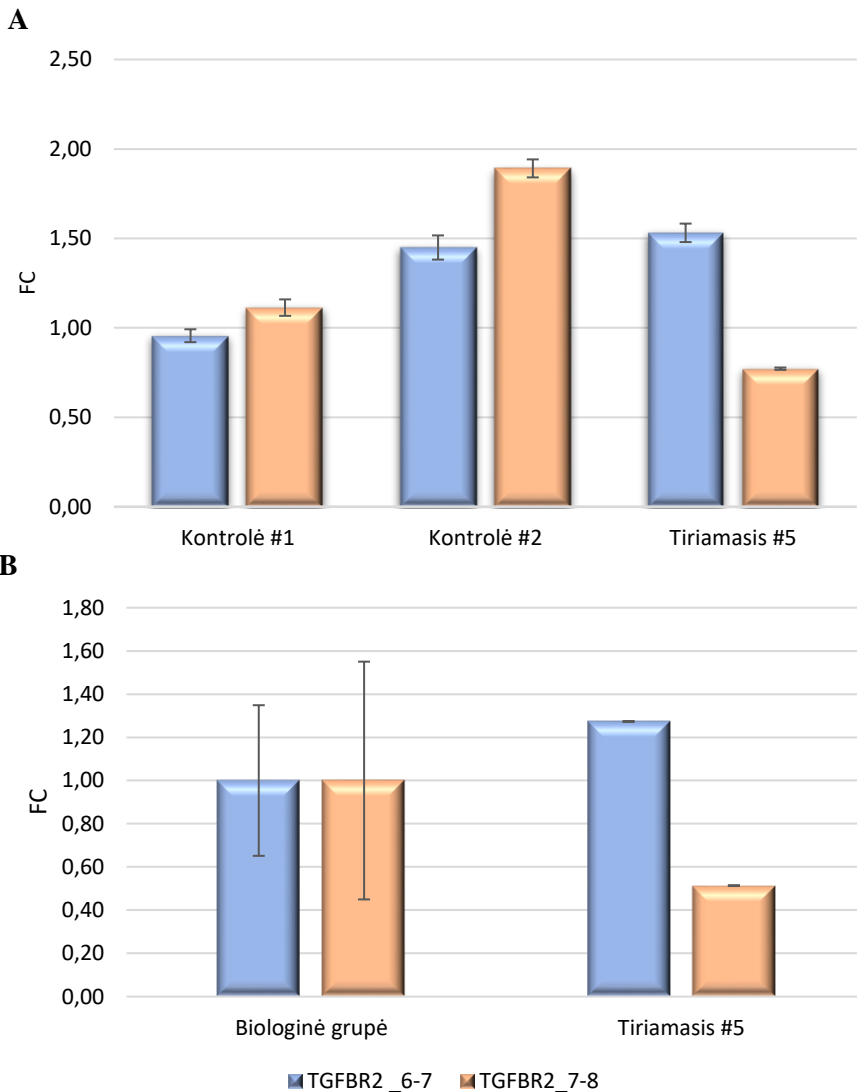
**4.14 pav.** Tiriamojo #5 kDNR mėginio chromatograma. Schemoje nurodytais oligonukleotidiniais pradmenimis (F ir R) atlikus Sangerio sekoskaitą, nusekvenotas tik nepakitusio *TGFBR2* geno alelio fragmentas

Remiantis šiuo rezultatu, iškelta nauja hipotezė, kad c.1600-2A>G variantas gali lemti aštuntojo egzono iškritą, kadangi sukurtais pradmenimis nepavyktų nusekvenuoti kDNR fragmento be egzono, kuriame jungiasi grįžtamasis oligonukleotidinis pradmuo. Dėl šios priežasties atlikta kPGR naudojant oligonukleotidinius *TaqMan*<sup>™</sup> žymenis c.1600-2A>G varianto poveikiui pre-iRNR sukirpimo procesui įvertinti. Hipotezei dėl galimai iškritusio aštunto egzono patikrinti buvo naudojami žymenys, kurie hibridizuojasi skirtingose *TGFBR2* geno srityse – 6–7 ir 7–8 egzonų jungties vietose. Tiriamoji 6–7 egzonų jungties sritis buvo naudojama kaip vidinė kontrolė, pagal kurią nustatyta, kad 7–8 egzonų jungties vietos fluorescencijos intensyvumas yra apie du kartus mažesnis (FC=1,99, 4.15 pav. A). Taip pat atlikta tiriamojo #5 ir trijų neturinčių ligos požymių negiminingų asmenų (kontrolė #1–#3) mėginių lyginamoji analizė. Normalizuojant duomenis,

vieno iš kontrolinio asmens rezultatai reikšmingai išsiskyrė iš tiriamos imties, todėl galutinė lyginamoji analizė atlikta lyginant tiriamąjį #5 ir biologinę grupę, kurią sudarė kontrolė #1 bei kontrolė #3<sup>6</sup>. Nustatyta 1,96 karto mažesnė *TGFBR2* 7–8 egzono jungties vietos raiška ir normali 6–7 egzono jungties vietos raiška (4.15 pav. B). Remiantis gautais rezultatais, sukirpimo vietos c.1600-2A>G variantas lemia aštuntojo egzono iškritą, galimai kartu apimant introninę sritį. Siekiant patvirtinti tiksliai sutrikusio pre-iRNR sukirpimo proceso vietas, Sangerio sekoskaitos tyrimas pakartotas, sukūrus naują oligonukleotidinį grįžtamąjį pradmenį 3' netransliuojamoje srityje (3'UTR). Išanalizavus gautą tyrimo rezultatą, nustatytas tik nepakitęs *TGFBR2* geno alelio fragmentas. Atsižvelgiant į gautus geno raiškos ir Sangerio sekoskaitos tyrimų rezultatus, tikėtina, kad dėl sukirpimo vietos c.1600-2A>G varianto iškrenta ne tik aštuntas egzonas, bet ir 3'UTR. Laimeris su kolegomis (2006) nustatė, kad patogeniniai variantai *CINH* (MIM #606860) geno 3'UTR lemia koduojamo baltymo haplonepakankamumą, dėl kurio sergama paveldima angioedema (MIM #106100) [242]. Todėl manytina, kad c.1600-2A>G variantas lemia *TGFBR2* baltymo (UniProtKB #P37173) haplonepakankamumą.

---

<sup>6</sup> *TGFBR2* geno raiškos tyrime eliminuotas 29 metų amžiaus moteriškosios lyties kontrolinis asmuo. Atlikus geno raiškos tyrimą, nustatyta, kad šio kontrolinio asmens *TGFBR2* raiška yra reikšmingai sumažėjusi lyginant tiek su tiriamojo, tiek su kitais kontroliniais mėginiais. Nutarta šį mėginį eliminuoti iš analizės, nes priešingu atveju reikėtų atlikti papildomus tyrimus, kurie galėtų paaiškinti kontrolinio asmens kraujyje sumažėjusios *TGFBR2* raiškos priežastį. Pažymėtina, kad šiame geno raiškos tyrime naudota papildoma vidinė kontrolė. Todėl vienos vidinės kontrolės ir dviejų išorinių biologinių kontrolių užtenka įvertinti genų raiškos pokytį ir patikrinti iškeltą hipotezę.

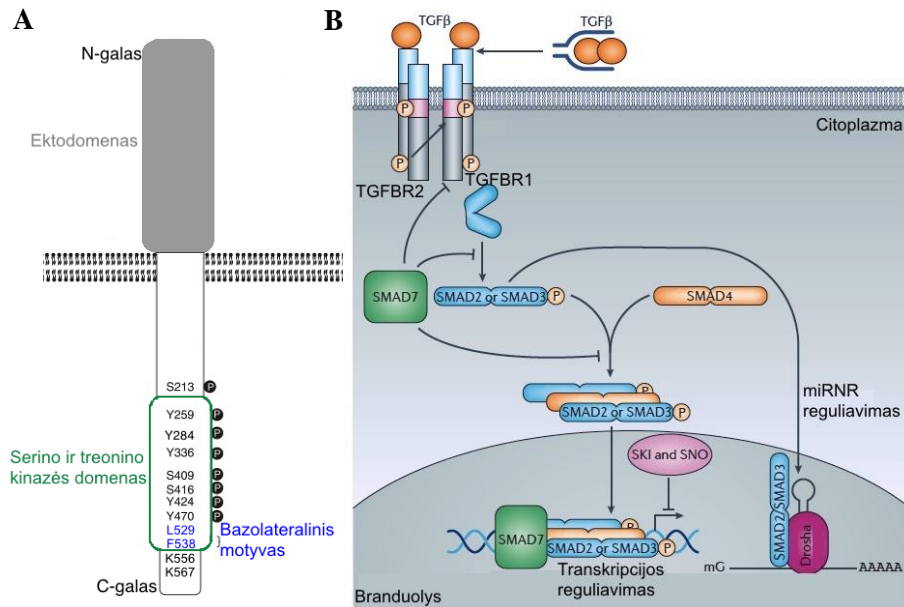


**4.15 pav. A)** Lyginamuoju  $\Delta\Delta C_t$  metodu įvertintas *TGFBR2* raiškos pokytis (FC) tarp tiriamojo #5, dviejų neturinčių ligos požymių negiminingų asmenų (kontrolė #1 ir #3; biologinė grupė) bei vidinės kontrolės (TGFBR2\_6-7). Histogramų „ūšiai“ žymi techninių pakartojimų vidurkio SN. **B)** Tiriamojo #5 *TGFBR2* raiškos pokytis pagal biologinę grupę (n = 2). Histogramų „ūšiai“ žymi biologinės grupės vidurkio SN. Histogramos nubrėžtos naudojant *Microsoft Excel 2010* (*Microsoft Corporation, JAV*) statistinį paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems, JAV*) ir *qbase<sup>+</sup>* (*Biogazelle, Belgija*) programa

*TGFBR2* geno produktas yra viduląstelinis TGF- $\beta$  II tipo receptorius (4.16 pav. A), kuris atsakingas už signalo perdavimą ląstelės citozolyje.



Pirmiausia, TGF- $\beta$  augimo veiksniai, prisijungę prie II tipo receptoriaus N-gale esančio ektodomeno, inicijuoja receptoriaus dimerizaciją. Šie homodimerai sąveikauja su TGF- $\beta$  I tipo receptoriaus homodimerais ir sudaro stabilią heterotetramerinę baltymų struktūrą. Baltymų kompleksas aktyvuojamas II tipo receptoriaus C-gale esančiu serino ir treonino kinazės domenams fosforilinus I tipo receptorių. Tuo pačiu principu I tipo receptorių toliau aktyvuoja SMAD baltymų signalinio kelio kaskadą, kuris reguliuoja miRNR brendimą ir iRNR transkripciją (4.16 pav. B) [243]. *TGFBR2* koduojamas baltymas yra svarbus ne tik aktyvuojant TGF- $\beta$  signalinį kelią. II tipo receptoriaus gale yra bazolateralinis motyvas (angl. *Basolateral Motif*), kuris nukreipia receptorių į poliarizuotų epitelinių ląstelių bazolateralinę dalį [244].



**4.16 pav. A)** TGFBR2 (UniProtKB #P37173) baltymo struktūros ląstelės membranoje schema. Juodi apskritimai, pažymėti P raide, žymi fosforilinimo sritis; **B)** TGFBR2 signalo perdavimo TGF- $\beta$  kelyje schema. Schemas adaptuotos pagal mokslinę literatūrą [243, 244]

Moksliniais modelinių organizmų tyrimais patvirtinta, kad TGF- $\beta$  signalinio kelio sutrikimai yra siejami su įvairių audinių, įskaitant širdies ir kraujagyslių, pažeidimais. *HGMD* [34] duomenimis, žinomi 127 patogeniniai *TGFBR2* variantai, o *TGFBR1* gene registruojama dvigubai mažiau kliniškai reikšmingų variantų. Didžioji dalis visų aprašytų patogeninių variantų, siejamų su LDS, – heterozigotiniai *missense* variantai, kurie yra evoliuciškai

konservatyviuose I ir II tipo receptoriaus serino ir treonino kinazės domenuose. Gallo su bendradarbiais (2014), tirdami modelines *Tgfb1*<sup>M318R/+</sup> ir *Tgfb2*<sup>G357W/+</sup> peles, nustatė, kad aortos sienelėje kaupiasi fosforilintas SMAD2, o kitų baltymų, veikiančių TGF-β signaliniame kelyje, raiška yra reikšmingai padidėjusi. Dėl sutrikusios genų raiškos šioms pelėms pasireiškė aortos aneurizma ir skeleto deformacijos, kurios taip pat yra būdingos tiek pirmojo, tiek antrojo tipo LDS [241, 245].

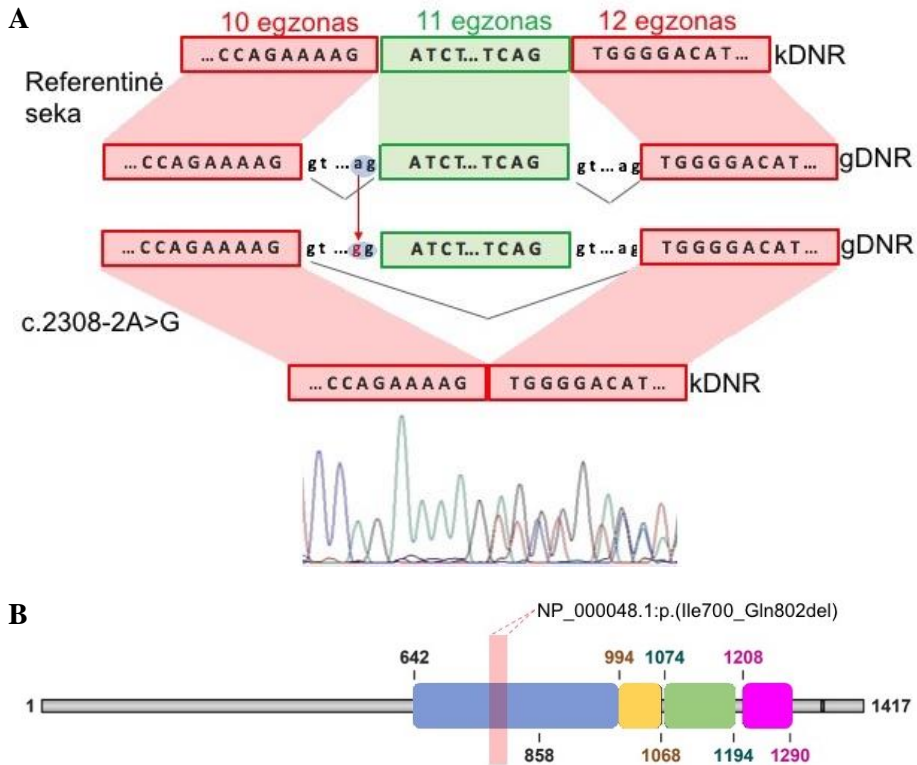
Sukirpimo vietos c.1600-2A>G variantas lemia *TGFBR2* aštuntojo egzono ir 3'UTR iškritą, todėl, remiantis ACMG [20] rekomendacijomis, variantas klasifikuotinas patogeniniu. Baltymo lygmeniu tikėtina, kad c.1600-2A>G variantas lemia II tipo receptoriaus haplonepakankamumą. Atsižvelgiant į tiriamojo #5 nustatytą genotipą ir jam pasireiškiančius klinikinius požymius, patogeninis *TGFBR2* variantas lemia antrojo tipo Loeyso–Dietzo sindromą.

#### 4.1.6. *BLM* c.2308-2A>G varianto tyrimas

Anemija, kitaip vadinama mažakraujyste, pasireiškia eritrocitų skaičiaus ir hemoglobino koncentracijos sumažėjimu kraujyje. Ši patologinė organizmo būseną būdinga ne tik sergant onkologinėmis ar autoimuninėmis ligomis, bet taip pat ir genetinėmis, pavyzdžiui, Bloomo sindromu (OMIM #210900; ORPHA #125). Šį autosominiu recesyviu būdu paveldimą sindromą lemia homozigotiniai arba sudėtiniai heterozigotiniai DNR sekos variantai *BLM* (MIM #604610) gene. *HGMD* [34] duomenimis, retesniais atvejais, kurie sudaro apie 14 proc. visų žinomų atvejų, *BLM* patogeniniai variantai yra siejami su kitais susirgimais (pvz., krūties ar kasos vėžys).

Pagal klinikinius požymius Bloomo sindromas buvo įtariamas labai smulkaus kūno sudėjimo 29 metų amžiaus moteriškosios lyties asmeniui #6, kuriam pasireiškė lengvi psichomotorinės raidos sutrikimai, pancitopenija bei lengvo laipsnio aplastinė anemija su paroksizminės naktinės hemoglobinurijos (angl. *Paroxysmal nocturnal hemoglobinuria*, PNH) klonu. Tiriamojo #6 gDNR mėginys ištirtas specifinių genų grupės NKS tyrimo technologija. Atlikus sekoskaitos duomenų analizę, nustatytas galimai patogeninis sukirpimo vietos NC\_000015.10(NM\_000057.2):c.2308-2A>G (rs1248548542) variantas *BLM* geno heterozigotinėje būklėje. Tiriant kDNR mėginį Sangerio metodu, nustatyta, kad šis 3' akceptorinės sukirpimo vietos variantas sutrikdo pre-iRNR sukirpimo procesą pašalinant 11-ąjį egzoną (4.17 pav. A). Kadangi šiuo metodu buvo tirti taip pat tiriamojo #6 tėvų kDNR mėginiai, paaiškėjo, kad variantas yra paveldėtas iš tėvo. Bioinformaciniais įrankiais įvertinus tiriamo varianto pasekmes baltymo lygmeniu, nustatyta,

kad dėl c.2308-2A>G varianto formuojasi galimai trumpesnis BLM baltymas (UniProtKB #P54132; NP\_000048.1:p.(Ile700\_Gln802del)) (4.17 pav. B).

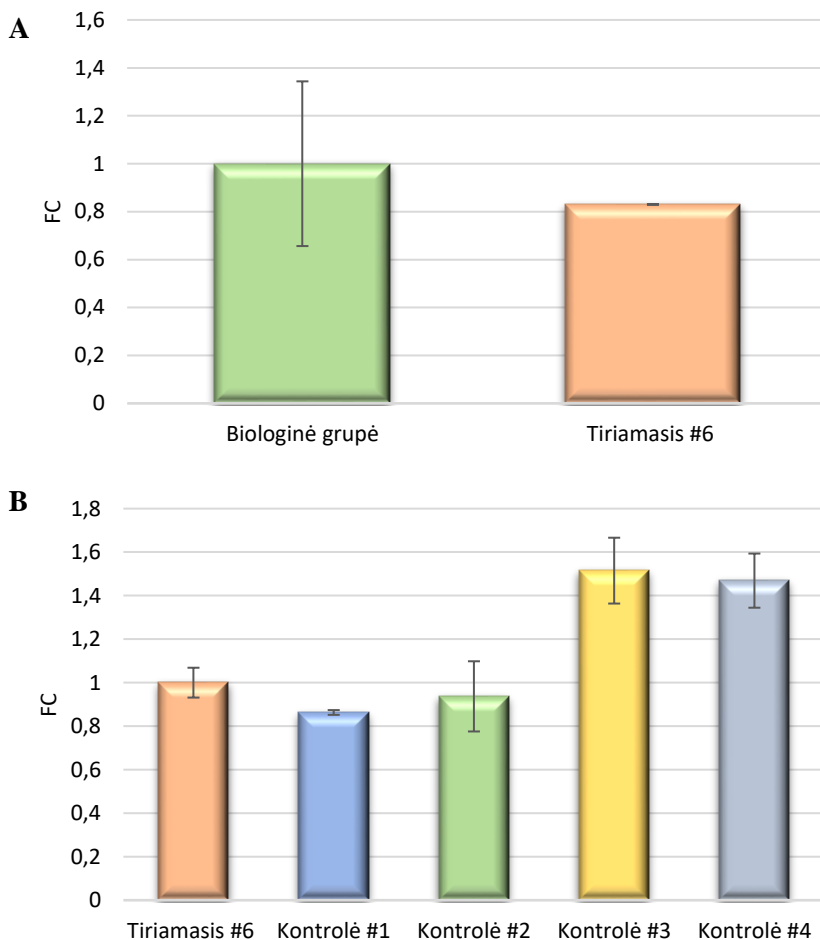


**4.17 pav. A)** Tiriomojo #6 kDNR mėginio chromatograma ir 3' akceptorinės sukirpimo vietos c.2308-2A>G varianto pre-iRNR sukirpimo proceso rezultato schema; **B)** Linijinė BLM baltymo (UniProtKB #P54132) struktūra su apytiksliai nurodyta prognozuojamo NP\_000048.1:p.(Ile700\_Gln802del) pokyčio vieta ir pagrindiniais funkciniais ir struktūriniais baltymo domenais. Mėlynu stačiakampiu vaizduojamas ATPazės domenas, geltonu – ZFN, žaliu – RQC, violetiniu – HRDC. Schema adaptuota pagal mokslinę literatūrą [246]

*BLM* koduoja RecQ tipo DNR helikazę, todėl šis genas dar vadinamas *RECQ3* (angl. *RecQ Protein-like 3*). RecQ tipo baltymai priklauso helikazių šeimai, kurios nariams, turintiems struktūrinį RQC (angl. *RecQ C-terminal*) ir HRDC (angl. *Helicase and RNaseD C-terminal*) domeną, būdingas ypač didelis konservatyvumas tiek bakterijų, tiek žmogaus ląstelėse. RecQ tipo DNR helikazės, prisijungusios prie ATPazės domeno, hidrolizuoja ATP molekulę. Šie baltymai, naudojant ATP kaip energijos šaltinį, reguliuoja DNR replikaciją, rekombinaciją ir reparaciją bei dalyvauja pažaidų taisymo

mechanizme [246, 247]. Tiriama NP\_000048.1:p.(Ile700\_Gln802del) pokyčio vieta yra BLM baltymo ATPazės domene (4.17 pav. B), todėl tikėtina, kad dėl 3' akceptorinės sukirpimo vietos varianto sutrinka baltymo ATP-azinis aktyvumas. Nepaisant to, nustatytas tik vienas patogeninis variantas, kurio nepakanka autosominio recesyvaus *Bloom* sindromo patogenezėi paaiškinti. Dėl šios priežasties *BLM* kDNR seka buvo tiriama detaliau.

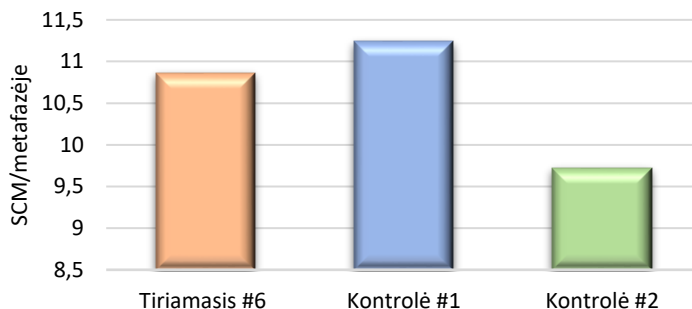
Iškelta hipotezė dėl galimai sutrikusios *BLM* geno raiškos. Kiekybinės PGR metodu atliktas šio geno raiškos tyrimas. Santykinis  $\Delta\Delta C_t$  metodu įvertintas tiriamojo #6 *BLM* geno raiškos pokytis (FC) pagal referentinius mėginius. Palyginti su biologine grupe ( $n = 4$ ), nustatyta, kad tiriamojo #6 *BLM* raiška yra 1,2 karto mažesnė, tačiau atitinkanti normos ribas (4.18 pav. A). Vertinant FC individualiai tarp tiriamojo #6 ir kiekvieno kontrolinio mėginio, nustatyta, kad geno raiškos pokytis ( $0,86 \leq FC \leq 1,51$ ) atitinka normą (4.18 pav. B).



**4.18 pav. A)** Lyginamuoju  $\Delta\Delta C_t$  metodu įvertintas *BLM* raiškos pokytis (FC) tiriamojo #6 ir biologinės grupės ( $n = 4$ ) periferinio veninio kraujo mėginiuose. Histogramų „ūšai“ žymi biologinės grupės vidurkio standartinį nuokrypį (SN; angl. *Standard deviation*); **B)** Tiriamojo #6 *BLM* FC pagal kontrolinius mėginius. Žymėjimas: kontrolė #1 – tiriamojo motina, kontrolė #2 – tiriamojo tėvas, kontrolė #3 ir #4 – neturintys ligos požymių negiminingi priešingų lyčių asmenys. Histogramų „ūšai“ žymi techninių pakartojimų vidurkio SN. Histogramos nubrėžtos naudojant *Microsoft Excel 2010* (*Microsoft Corporation*, JAV) statistinį paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems*, JAV) programa

Remiantis literatūra, Bloomo sindromą turintiems asmenims dažnai būdingas padidėjęs genomo nestabilumas dėl sutrikusios RecQ tipo DNR helikazės veiklos [247]. Todėl, bendradarbiaujant su VU GMC Botanikos ir

genetikos katedra, tiriamajam #6 ir jo tėvams diferenciniu dažymo metodu atliktas<sup>7</sup> seserinių chromatidžių mainų (SCM; angl. *Sister chromatid exchange*) tyrimas. Tačiau visų SCM metafazėje skaičius nustatytas kontrolinio lygio (SCM < 14; 4.19 pav.).



**4.19 pav.** Tiriamojo #6 ir kontrolinių asmenų seserinių chromatidžių mainų tyrimo rezultatai. Žymėjimas: kontrolė #1 – tiriamojo motina, kontrolė #2 – tiriamojo tėvas

Atsižvelgiant į c.2308-2A>G lemiamą pre-iRNR sukirpimo proceso pokytį, funkciškai svarbų baltymo domeną koduojančioje sekoje, *BLM* geno variantas pagal ACMG [20] rekomendacijas galėtų būti klasifikuojamas kaip patogeninis. Visgi šis variantas heterozigotinėje būklėje negali paaiškinti autosominiu recesyviu būdu paveldimo Bloomo sindromo patogenezės tiriamojo #6 atveju. Todėl šio varianto patogeniškumui pagrįsti reikėtų papildomų įrodymų, kurie galėtų būti gauti atlikus funkcinius tyrimus modelinėse sistemose.

#### 4.1.7. *CAPN3* c.1746-20C>G varianto tyrimas

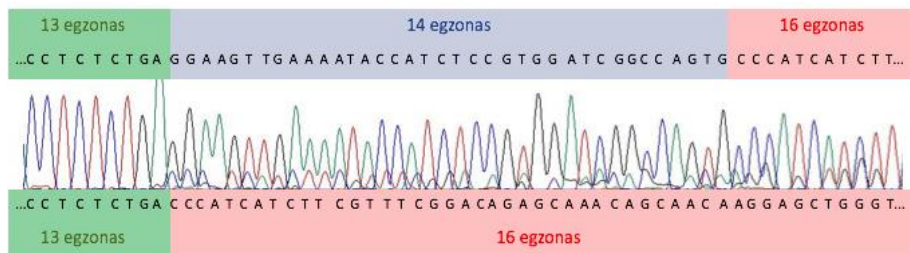
Galūnių ir juosmens raumenų distrofijos (angl. *Limb girdle muscular dystrophies*; LGMD) – tai autosominiu recesyviu ar autosominiu dominantiniu būdu paveldimų raumenų sutrikimų grupė, kuriai būdingas genetinis heterozigotiškumas. Priklausomai nuo pažeisto geno, klasifikuojama mažiausiai 14-ka LGMD tipų. Daugelyje šalių labiausiai paplitusi yra autosominiu recesyviu būdu paveldima antrojo A tipo LGMD (LGMD2A; MIM #253600; ORPHA #267), kurią lemia homozigotiniai arba sudėtiniai heterozigotiniai variantai 15 chromosomos (15q15.1) *CAPN3*

<sup>7</sup> SCM tyrimą atliko doc. dr. Veronika Dedonytė.

(MIM #114240) gene [248]. *HGMD* [34] ir *LOVD* [249] duomenimis, šiuo metu žinoma apie 400–500 patogeninių *CAPN3* variantų. Nors didžioji dalis šių variantų yra koduojančioje DNR sekoje, apie 15–25 proc. variantų yra nekoduojančioje geno dalyje [248].

Dėl pasireiškiančios raumenų patologijos 43 metų amžiaus tiriamajam #7 atliktas specifinių genų grupės NKS tyrimas. Nustatyti du heterozigotiniai *CAPN3* variantai. Ketvirtame egzone nustatytas patogeninis NM\_000070.2:c.598\_612del (rs72750383) variantas, kuris anksčiau buvo patvirtintas asmenims, kuriems pasireiškė LGMD požymiai (PMID #9452114, #25135358, #18854869, #16141003). 13-ame introne nustatytas nežinomos klinikinės reikšmės NC\_000015.10(NM\_000070.2):c.1746-20C>G variantas. Segregacijos analizė Sangerio sekoskaitos metodu patvirtino variantus sudėtinėje heterozigotinėje būklėje, sukirpimo variantas paveldėtas iš motinos, o 15 nt delecija – iš tėvo.

Siekiant įvertinti c.1746-20C>G varianto padarinius iRNR lygmeniu, atlikta tiriamojo #7 kDNR analizė Sangerio metodu. Nustatyti du pakitę transkriptai. Pagal *CAPN3* NM\_000070.2 referentinę seką, viename transkripte nustatyta 15-ojo egzono iškrita, o kitame – 14-ojo ir 15-ojo egzono iškrita (4.20 pav.).



**4.20 pav.** Tiriamojo #7 kDNR mėginio chromatograma

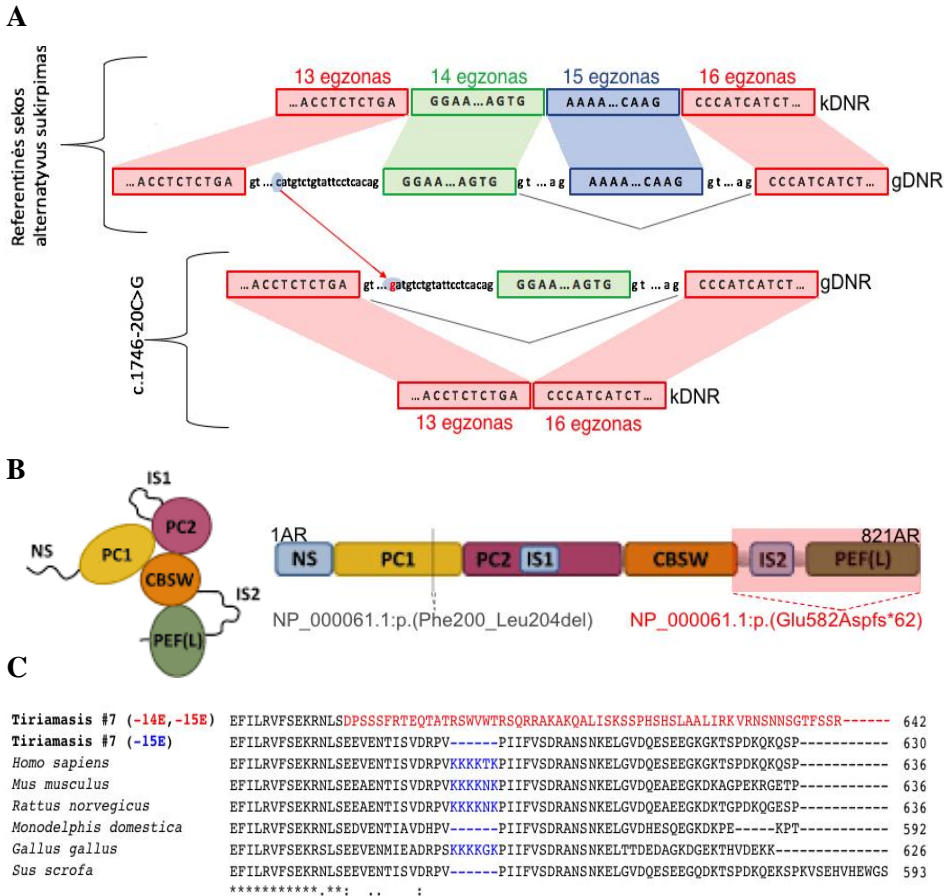
Remiantis moksline literatūra, galima teigti, kad retesniais atvejais patogeniniai sukirpimo variantai gali lemti dviejų ar daugiau pakitusių transkriptų atsiradimą tam tikrame audinyje. Pavyzdžiui, Nascimbeni su kolegomis (2010) tyrė *CAPN3* sukirpimo vietas c.1524+1G>C varianto poveikį iRNR lygmeniu. Tyrimo metu nustatyta, kad dėl sukirpimo vietos varianto susidaro du skirtingi transkriptai. Vienas iš tiriamų transkriptų susidarė praradus dalį 11-ojo egzono, o kitame transkripte nustatytas dalies 11-ojo introno tarpas į gretimą egzoną [248]. Kita vertus, alternatyvus pre-iRNR sukirpimo procesas yra normalus reiškinys, užtikrinantis baltymų įvairovę ląstelėje. Pavyzdžiui, dėl alternatyvaus sukirpimo mechanizmo susidaro net 46 skirtingi *NF1* transkriptai [43, 44]. Taip pat Blázquez su

bendradarbiais (2008) 26-ių tiriamųjų kraujo mėginiuose nustatė keturis skirtingus *CAPN3* koduojamus iRNR transkriptus. Visi ištirti kraujo transkriptai neturėjo 15-ojo egzono, o tarpusavyje skyrėsi 6-ojo ir 16-ojo egzonų kombinacija. Palyginę c.1782+1072G>C varianto poveikį pre-iRNR sukirpimui skirtinguose kraujo ir raumens bioptato ėminiuose, mokslininkai nustatė 100 bp ilgio 14-ojo introno intarpą tarp 14-ojo ir 16-ojo egzono kraujo transkripte, o raumens transkripte šis introno intarpas nustatytas tarp 14-ojo ir 15-ojo egzono [250]. Remiantis šiuo rezultatu tikėtina, kad žmogaus kraujo ląstelėse vyksta alternatyvių, neturinčių 15-ojo egzono, *CAPN3* koduojamų iRNR transkriptų raiška. Todėl manytina, kad sukirpimo c.1746-20C>G variantas tiriamojo #7 kDNR mėginyje lemia tik 14-ojo egzono iškritą, o 15-asis egzonas yra prarandamas dėl ląstelėms būdingo įprasto alternatyvaus sukirpimo proceso (4.21 pav. A).

*CAPN3* geno produktas yra kalcio ( $\text{Ca}^{2+}$ ) aktyvuojama cisteino proteazė, kuri priklauso kalpainų (angl. *Calpains*) superšeimai. Funkciniu požiūriu kalpainai, įskaitant *CAPN3*, yra ypač svarbūs ne tik hidrolizuojant įvairius baltymus, bet ir reguliuojant ląstelės ciklo, apoptozės ir diferenciacijos procesus. Visi kalpainai turi du aktyvųjų baltymo centrą formuojančius proteolitinius šerdies domenų (angl. *Proteolytic core domains*; PC), CBSW (angl. *Calpain-type  $\beta$ -sandwich*) domeną bei baltymo homodimerizaciją užtikrinantį EF-rankos motyvo kilpos (angl. *Penta EF-hand loop*; PEFL) domeną. *CAPN3* (UniProtKB #P20807) baltymas, skirtingai nei kiti kalpainų šeimos nariai, papildomai turi tris specifines įsiterpusias sekas (angl. *insertion sequences*; IS1, IS2, NS) (4.21 pav. B). Manoma, kad dėl šių unikalių sekų *CAPN3* yra mažiau stabilus ir pasižymi didesne savaiminio suardymo (angl. *autodegradation*) geba. Be to, moksliniais tyrimais pagrįsta, kad IS2 seka, kurios dalį koduoja *CAPN3* 15 egzonas, perduoda branduolio lokalizacijos signalą ir užtikrina sąveiką su raumenų *TTN* baltymu [250]. *GTEx* portalo duomenimis, intensyvi *TTN* (MIM #188840) raiška vyksta tik skeleto ir širdies raumenyse, o kraujyje ir kituose audiniuose ji yra labai neryški. Kadangi žmogaus periferiniame veniniame kraujyje nėra *TTN* baltymo poreikio, tikėtina, kad vykstant evoliucijai įvyko alternatyvus *CAPN3* sukirpimas, kuris lėmė 15 egzono, koduojančio tik šešias aminorūgštis, iškritą. Šis pokytis nesukelia skaitymo rėmelio poslinkio, todėl raumenų ir kraujo *CAPN3* baltymo struktūra skiriasi nežymiai. Be to, konservatyvumo analizė, apėmusi žmogų ir penkias atsitiktinai parinktas evoliuciškai nutolusias rūšis, atskleidė, kad kelių organizmų, pavyzdžiui, *Monodelphis Domestica* ir *Sus Scrofa*, *CAPN3* kDNR taip pat nėra 15-ojo egzono. Priešingai, 14 egzono, kurio iškritą lėmė tiriamas



sukirpimo c.1746-20C>G variantas, koduojama sritis yra labai konservatyvi ir būdinga visiems analizuotiems organizmams (4.21 pav. C).



**4.21 pav. A)** Sukirpimo c.1746-20C>G varianto sutrikdyto pre-iRNR sukirpimo proceso rezultato schema; **B)** CAPN3 (UniProtKB #P20807) baltymo struktūra su apytiksliai nurodyta prognozuojamų NP\_000061.1:p.(Phe200\_Leu204del) ir NP\_000061.1:p.(Glu582Aspfs\*62) pokyčių vieta. Schema adaptuota pagal mokslinę literatūrą [251]; **C)** CAPN3 baltymo sekos konservatyvumo analizė, apėmusi šešias skirtingas rūšis. Sukirpimo varianto nulemtas 62-ų aminorūgščių skaitymo rėmelio poslinkis pažymėtas raudona spalva, 15-ojo egzono koduojamos aminorūgštys pažymėtos mėlyna spalva. Konservatyvios sekos vietos pažymėtos žvaigždute (\*), konservatyvūs variantai – dvitaškiu (:), pusiau konservatyvūs variantai – tašku (.), nepažymėtos vietos reiškia nekonservatyvias

sekos vietas<sup>8</sup>. Konservatyvumo analizė atlikta *Clustal Omega* [252] bioinformaciniu įrankiu

Abu tiriamajam #7 nustatyti *CAPN3* variantai yra funkciniuose baltymo domenuose. Ketvirtojo *CAPN3* egzono 15 nt delecija lemia penkių aminorūgščių iškritą (NP\_000061.1:p.(Phe200\_Leu204del)) iš aktyvųjį baltymo centrą formuojančio proteolitinio šerdies domeno (PC1). *In silico* sukirpimo vietos variantas lemia skaitymo rėmelio poslinkį ir priešlaikinio baigmės kodono susiformavimą. Prognozuojamas NP\_000061.1:p.(Glu582Aspfs\*62) pokytis apima CBSW domeno galą, įsiterpusią antrąją seką (IS2) bei EF-rankos motyvo kilpos domeną (4.21 pav. B). Todėl tikėtina, kad nustatyti variantai pažeidžia *CAPN3* funkciją. Nors *CAPN3* geno funkcijos praradimą (angl. *Loss-of-function variants*) nulemiantys variantai yra nustatomi sergantiesiems LGMD2A, iki šiol tikslus patogenezės mechanizmas nėra žinomas. Tačiau moksliniais modelių gyvūnų tyrimais patvirtinta, kad išjungus *CAPN3* geną sutrikdoma specifinių genų raiška (pvz., *NCX3*) ir  $Ca^{2+}$  reguliacijos procesas, atsiranda mitochondrijų pažeidimų bei sukeliamas oksidacinis stresas, o visa tai sutrikdo normalią raumenų veiklą [251].

Nustatytas sukirpimo vietos c.1746-20C>G variantas sutrikdo pre-iRNR sukirpimą, todėl yra klasifikuojamas patogeniniu. Tiriamajam #7 nustatytas *CAPN3* geno sudėtinis heterozigotinis genotipas nulemia autosominiu recesyviu būdu paveldimą antrojo A tipo LGMD.

#### 4.2. Kopijų skaičiaus pokyčių molekulinų ir funkcinių tyrimų rezultatai ir jų aptarimas

Koduojančios ir nekoduojančios DNR sekos KSP, ypač delecijos ir duplikacijos, yra viena dažniausių IN / ĮA priežasčių [9, 60]. Trims tiriamiesiems (#8, #9, #10) ankstesnių projektų metu (CHERISH, PROGENET, UNIGENE) atlikus viso genomo molekulinio kariotipavimo tyrimą, nustatyti trys nežinomos klinikinės reikšmės KSP: 13q31.3 duplikacija, 4q21 delecija ir 12q24.21 (vidugeninė *MED13L*) delecija. Kitų patogeninių ar galimai patogeninių KSP, sietinų su tiriamiesiems asmenims pasireiškiančiais klinikiniais požymiais, nenustatyta. Išanalizavus mokslinę

---

<sup>8</sup> Nors tirtos baltymo sekos yra konservatyvios, dėl sukirpimo varianto nulemtos skaitymo rėmelio poslinkio *Clustal Omega* bioinformacinis įrankis šias sritis žymi pusiau konservatyviomis ar nekonservatyviomis.

literatūrą, įvertinta, kad visi nustatyti nežinomos klinikinės reikšmės KSP yra unikalūs ir iki šiol neaprašyti. DGV [94] duomenimis, šių, moksliniame darbe nagrinėjamų, genominių sričių KSP nėra būdingi sveikų žmonių populiacijai. Todėl siekiant įvertinti 13q31.3 duplikacijos, 4q21 delecijos bei vidugeninės *MED13L* delecijos patogeniškumą ir ryšį tarp genotipo ir fenotipo, atliktas jų klinikinis, molekulinis bei funkcinis apibūdinimas.

#### 4.2.1. Molekulinis, funkcinis ir klinikinis 13q31.3 duplikacijos apibūdinimas

Molekulinio kariotipavimo vLGH (1M skiriamosios gebos, *Agilent Technologies*, JAV) [253] metodu buvo iširta 19 metų moteris, tiriamasis #8 (*DECIPHER* #366736), kuriai pasireiškė IN (IQ = 49), psichomotorinės raidos vėlavimas, makrocefalija, aukštas ūgis, veido ir skeleto anomalijos. Ištyrus tiriamąją #8 DNR mėginį buvo nustatyta intersticinė 13 chromosomos duplikacija (13q31.3), kuri pagal referentinį žmogaus genomą (GRCh37/hg19) apėmė 844 153 bp sritį. Tyrimas patvirtintas FISH metodu [253], taip pat šiuo metodu nustatyta *de novo* duplikacijos kilmė. Išanalizavus genominę sritį, nustatyta, kad 13q31.3 duplikacija apėmė *MIR17HG* (MIM #609415) geną koduojantį miR-17~92 klasterį, pagal kurį transkribuojamos šešios specifinės miRNR: miR17, miR18A, miR19A, miR20A, miR19B1 ir miR92A1. Šios miRNR molekulės yra ypač svarbios reguliuojant genų raišką. Pavyzdžiui, patvirtinta, kad miR17 ir miR20A tiesiogiai reguliuoja nuo ciklinų priklausomų kinazių slopiklių *CDKN1A* (MIM #116899) ir *CDKN1C* (MIM # 600856) raišką. Taip pat yra reguliuojami genai, kurie svarbūs apoptozės, diferenciacijos bei angiogenezės procesui [254].

De Pontualis su kolegomis (2011) nustatė, kad 13q31.3 hemizigotinė delecija lemia autosominiu dominantiniu būdu paveldimą antrojo tipo Feingoldo sindromą (MIM #614326). Mokslininkų tirtos 13q31.3 delecijos sritys apėmė *MIR17HG* bei visą arba dalį *GPC5* (MIM #602446) geno. Funkcinių miR-17~92<sup>+/-</sup> pelių tyrimų metu nustatyta, kad heterozigotinėms pelėms pasireiškia II tipo Feingoldo sindromui būdingi klinikiniai požymiai: mikrocefalija, žemas ūgis ir skeleto anomalijos. Nors šis tyrimas patvirtina miR-17~92 klasterio įtaką skeleto vystymuisi, iki šiol buvo keliami hipotezė, kad *GPC5* geno pokyčiai gali prisidėti nulemiant II tipo Feingoldo sindromui būdingus kliniskus požymius [255]. Šią hipotezę patvirtina ankstesni modelinių pelių tyrimai, kurių metu nustatyta, kad sutrikusi *Gpc5* raiška turi įtakos pelių centrinės nervų sistemos, inkstų bei galūnių formavimuisi. Taip pat įrodyta, kad *Gpc5* geno koduojamas ląstelės paviršiaus heparano sulfato

proteoglikanas yra itin svarbus reguliuojant ląstelės augimą, dalijimąsi ir diferenciaciją [256].

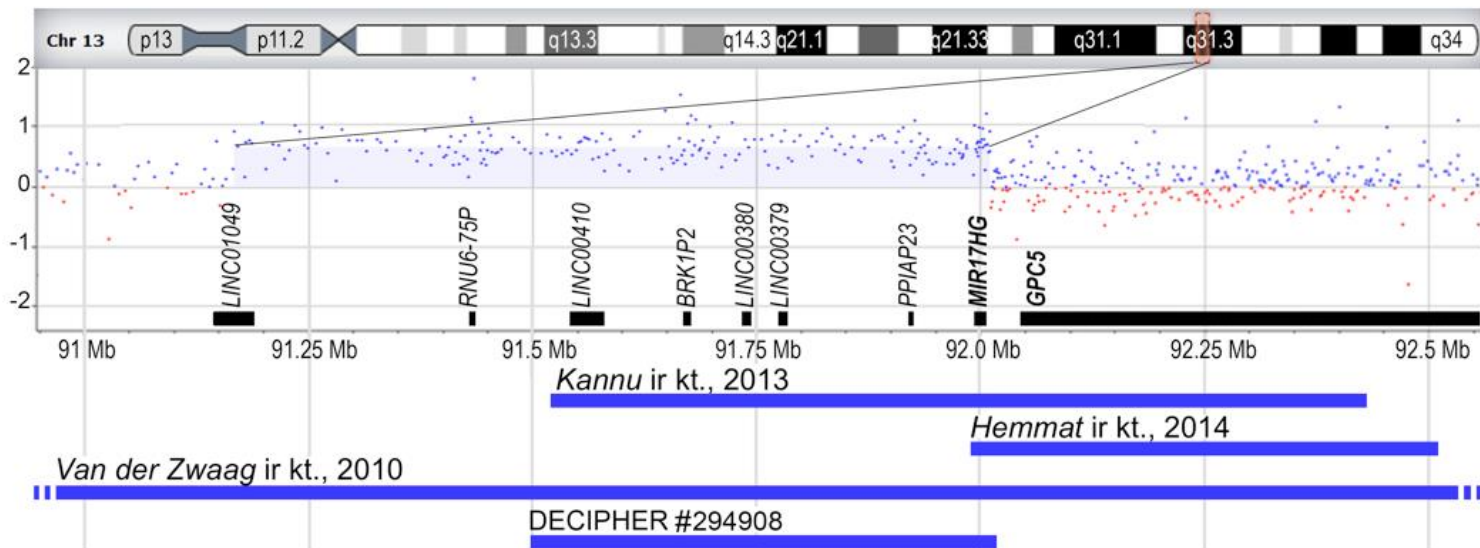
Remiantis išanalizuota literatūra ir duomenų bazėmis, yra aprašytos dvi panašaus dydžio (516 Kb ir 845,7 Kb) 13q31.3 srities duplikacijos, apimančios *MIR17HG* ir dalį *GPC5* geno [257, 258]. Šios duplikacijos buvo nustatytos penkiems tiriamiesiems iš dviejų skirtingų šeimų, kuriems, panašiai kaip tiriamajam #8, pasireiškė psichomotorinės raidos vėlavimas, makrocefalija, augimo sutrikimai bei skeleto anomalijos (4.22 pav., 4.2 lentelė). Svarbu pažymėti, kad dalis fenotipinių požymių, t. y. makrocefalija, aukštas ūgis, pirštų polidaktilija, kuriuos nulemia 13q31.3 srities duplikacija, yra priešingi požymiams: mikrocefalija, žemas ūgis, pirštų sindaktilija, kurie yra stebimi esant II tipo Feingoldo sindromui, nulemtam 13q31.3 srities delecijos. Kadangi duplikacijos sritys apėmė tiek *MIR17HG*, tiek *GPC5* geną, autoriai pritarė De Pontualio ir kt. (2011) iškeltai hipotezei, kad galimai abu genai nulemia tiriamus klinikinius požymius. *DECIPHER* duomenų bazėje [108] yra registruotas kitas tiriamasis (#294908), kuriam nustatyta 521 Kb dydžio duplikacija, apimanti *MIR17HG*, tačiau klinikinis fenotipas yra nežinomas. Mokslinėje literatūroje yra aprašytos kelios didesnės persidengiančios 13q31.3 srities duplikacijos, tačiau tik dviejų duplikacijų, nustatytų vLGH metodu, yra nurodomos tikslios KSP sritys. Vienu atveju genotipo ir fenotipo ryšio analizė yra neinformatyvi dėl duplikacijos dydžio (22.5 Mb) [259]. Kitu atveju negalima įvertinti duplikacijos (5.58 Mb) įtakos kognityvinių funkcijų vystymuisi dėl tiriamojo amžiaus (27 mėn.; 4.22 pav., 4.2 lentelė) [260].

**4.2 lentelė.** Analizuotų asmenų, kuriems nustatyta 13q31.3 duplikacija, klinikinių požymių santrauka

Pagrindiniai klinikiniai požymiai	Tiriamasis #8	Kannu ir kt. (2013)		Hemmat ir kt. (2014)			Van der Zwaag ir kt. (2010)
		Probandas	Motina	Probandas	Sibasas	Motina	
Dup13q31.3 sritis (GRCh37/hg19)	91 166 748– 92 010 901 bp	91 520– 92 430 Kb	91,520– 92 430 Kb	91 989 261– 92 504 857 bp	91 989 261– 92 504 857 bp	91 989 261– 92 504 857 bp	90 867 465– 96 454 223 bp
Duplikacijos dydis	844 Kb	845,7 Kb	845,7 Kb	516 Kb	516 Kb	516 Kb	5.59 Mb
Amžius	17 metų	4,5 metų	41 metų	9 metų	11 metų	34 metų	27 mėn.
Lytis	Mot.	Vyr.	Mot.	Vyr.	Vyr.	Mot.	Vyr.
Gimimo svoris (procentilis)	75	25	ND <sup>9</sup>	Norma	ND	ND	> 97
Galvos apimtis	Makrocefalija	Makrocefalija	Makrocefalija	Makrocefalija	Makrocefalija	Makrocefalija	Makrocefalija
Ūgis (procentilis)	> 97	> 97	> 97	Žemas ūgis	Žemas ūgis	Žemas ūgis	Norma
IN	+	ND	ND	ND	ND	+	ND
PRA <sup>10</sup>	+	+	ND	+	+	ND	Norma
Akys	Žvairumas, astigmatizmas	Telekantas	ND	Hipertelorizmas	Hipertelorizmas	Hipertelorizmas	Telekantas
Igimta širdies yda	+	+	ND	ND	ND	ND	ND
Pirštų ir / ar skeleto anomalijos	Arachnodaktilija, <i>pectus excavatum</i> , skoliozė	Polidaktilija	Polidaktilija	Brachidaktilija, klinodaktilija	Brachidaktilija, klinodaktilija	Brachidaktilija, klinodaktilija	Polidaktilija

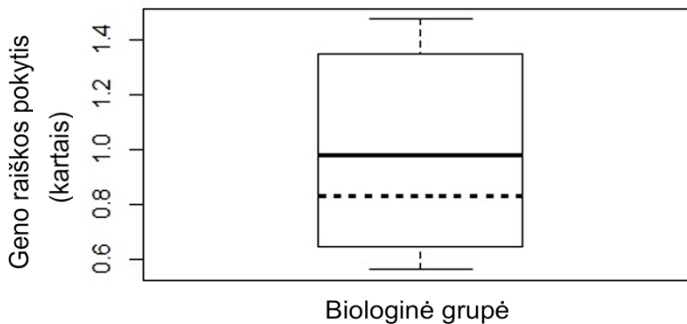
<sup>9</sup> ND – nėra duomenų.

<sup>10</sup> PRA – psichomotorinės raidos atsilikimas.



**4.22 pav.** 13 chromosomos ideograma, tiriamojo #8 vLGH tyrimo rezultatai 13q31.3(91 166 748-92 010 901)×3 *dn* (GRCh37/hg19) bei genų, patenkančių į duplikacijos sritį, schema. Tiriamojo #8 13q31.3 duplikacijos sritis pažymėta šviesiai mėlynu stačiakampiu, *OMIM* genai paryškinti, o mėlynos linijos žymi pagal mokslinę literatūrą [257, 258, 260] ir *DECIPHER* [108] duomenų bazę persidengiančių 13q31.3 duplikacijų sritis

Tiriamąjį #8 nustatyta 844 Kb dydžio duplikacija be geno kandidato *MIR17HG* apėmė taip pat tris *BRK1P2*, *PPIAP23* ir *RNU6-75P* pseudogenus bei keturis ilgas nekoduojančias RNR koduojančius genus (*LINC01049*, *LINC00410*, *LINC00380* ir *LINC00379*), tačiau neapėmė *GPC5* geno (4.22 pav.). Pažymėtina, kad atstumas, skiriantis nustatytos duplikacijos sritį ir *GPC5* geną, yra apie 40 Kb. Remiantis *GeneCards* duomenų baze [261], tiriamoje duplikacijos srityje yra mažiausiai vienas *GPC5* geną reguliuojantis elementas (Nr. GH13J091346). Be to, moksliniais tyrimais įrodyta, kad KSP turi įtakos gretimų genų raiškai, o poveikis gali būti stebimas visos chromosomos ar net transkriptomo mastu [262–264]. Todėl siekiant patikrinti De Pontualio ir kt. (2011), Kannu ir kt. (2013) bei Hemmato ir kt. (2014) hipotezę dėl galimų dviejų *MIR17HG* ir *GPC5* genų kandidatų bei norint įvertinti, ar *GPC5* geno raiška nebuvo sutrikdyta, atlikta *GPC5* geno raiškos analizė kPGR metodu. Duomenų analizė atlikta santykiniu  $\Delta\Delta Ct$  metodu, lyginant tiriamojo #8 *GPC5* geno raiškos pokytį (FC) su biologinės grupės, kurią sudarė neturintys požymių keturi sveiki asmenys (tarp jų tiriamojo #8 tėvai), mėginiais. Lyginant su referentiniais mėginiais, nustatyta, kad *GPC5* raiška atitinka normą ( $0,56 \leq FC \leq 1,47$ ) (4.23 pav.).



**4.23 pav.** Stačiakampė diagrama atspindi *GPC5* geno raiškos pokytį (FC) tarp tiriamojo #8 ir biologinės grupės ( $n = 4$ ) kraujo mėginių. Horizontali paryškinta linija stačiakampyje žymi biologinės grupės vidutinę vertę ( $FC = 1$ ), pagal kurią nustatyta, kiek kartų pakito *GPC5* raiška ( $FC = 0,83$ ; žymima punktyrine linija) tiriamojo #8 kraujo mėginyje. Diagramos apatinė ir viršutinė vientisa linija žymi atitinkamai 25-ąją ir 75-ąją procentilį, o diagramos „ūsai“ nurodo referentinės biologinės grupės mažiausią ir didžiausią FC reikšmes. Stačiakampė diagrama nubrėžta naudojant *R v.3.4.0* [190] statistikos paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems*, JAV) ir *ExpressionSuite v.1.1* (*ThermoFisher Scientific*, JAV) programa

Dalis tiriamojo #8 klinikinių požymių (aukštas ūgis, veido asimetrija, įgimta širdies yda, arachnodaktilija, skoliozė) taip pat būdinga autosominiu dominantiniu būdu paveldimam Marfano sindromui (MIM #154700), kurį nulemia DNR sekos variantai *FBNI* (MIM #134797) gene. Siekiant įvertinti, ar tiriamojo #8 genome nėra patogeninių variantų, atlikta VES. Atlikus išsamią gautų duomenų bioinformacinę analizę, nenustatyta jokių patogeninių ar galimai patogeninių DNR sekos variantų, kurie galėtų būti siejami su tiriamojo #8 fenotipu.

Remiantis klinicine tiriamojo #8 informacija, genetinių tyrimų rezultatais bei išanalizuota mokslinė literatūra ir duomenų bazėmis, galima teigti, kad miR-17~92 klasterio duplikacija, kuri, *DGV* [94] duomenimis, nėra būdinga sveikiesiems asmenims, yra patogeninė ir lemia naują sindromą, charakterizuojamą požymiais, iš dalies priešingais antrojo tipo Feingoldo sindromui.

#### 4.2.2. Molekulinis, funkcinis ir klinikinis 4q21.22 delecijos apibūdinimas

vLGH (105K skiriamosios gebos, *Agilent Technologies*, JAV) [253] metodu ištirtas 16 metų amžiaus moteriškosios lyties tiriamasis #9, kuriam būdinga IN, psichomotorinės raidos atsilikimas (PRA), kalbos sutrikimai, žemas ūgis, nutukimas, hipotonija bei dismorfiniai požymiai (pvz., veido anomalijos, skersinė delno raukšlė, brachidaktilija). Ketvirtosios chromosomos ilgajame petyje (4q21.22) nustatyta *de novo* delecija, kuri patvirtinta kPGR metodu [253]. Pagal referentinį žmogaus genomą (GRCh36/hg18) 4q21.22 delecija apėmė 824 053 bp sritį.

Remiantis *DGV* [94] duomenų baze, intersticinėje 4q21.22 delecijos srityje yra baltymą koduojantys genai (*HNRNPD* (MIM#601324), *HNRNPD* (MIM #607137), *ENOPH1*, *TMEM150C* (MIM #617292), *SCD5* (MIM #608370), *SEC31A* (MIM #610257), *THAP9* (MIM #612537)), ilgą nekoduojančią RNR koduojantis *LINC00575* genas bei miRNR koduojantis *MIR575* genas. Delecija taip pat apėmė dalį *LIN54* (MIM #613367) geno. Nors 4q21 delecijos srityje nustatoma daug baltymus koduojančių genų, tik keli iš jų turi patvirtintą poveikį žmogaus fenotipui. Moksliniais tyrimais nustatyta, kad *SEC31A* lemia autosominiu recesyviu būdu paveldimą nervų sistemos vystymosi sutrikimą (MIM #618651), o *HNRNPD*, *HNRNPD* ir *ENOPH1* yra siejami su 4q21 delecijos sindromu (MIM #613509).

Šiuo metu žinoma daugiau nei 50 klinikinių atvejų, kai yra nustatytos persidengiančios 4q21 srities delecijos: iš jų 31 atvejis yra registruotas *DECIPHER* duomenų bazėje [108], o kita dalis yra aprašyta mokslinėje literatūroje [265–274]. Siekiant įvertinti ryšį tarp genotipo ir fenotipo, buvo



analizuojami tik tie publikuoti 4q21 delecijos atvejai, kurių buvo nurodomos tikslios KSP sritys (4.3 lentelė, 4.24 pav.).

Bonnet su kolegomis (2010) ištyrė devynis negiminingus asmenis, kuriems, kaip ir tiriamajam #9, pasireiškė sunkūs psichomotorinės raidos defektai (įskaitant kalbinės raidos vėlavimą ar nebuvimą), augimo sutrikimai, raumenų hipotonija bei veido anomalijos. Nors visiems tiriamiesiems buvo nustatyta skirtingo dydžio 4q21 delecija, apibrėžta persidengianti KSP 1,37 Mb dydžio sritis, kuri apėmė penkis kandidatinius genus: *HNRNPD*, *HNRNPD*, *ENOPH1*, *RASGEF1B* (MIM #614532) ir *PRKG2* (MIM#601591). Pastarasis genas yra siejamas su žemu ūgiu [265], o kiti genai, manoma, lemia psichomotorinės raidos atsilikimą, hipotoniją, veido anomalijas ir kitus dismorfinius požymius [266, 267, 270].

*DECIPHER* duomenų bazėje [108] bei mokslinėje literatūroje yra aprašyta keliasdešimt persidengiančių 4q21 srities delecijų, tačiau 29 analizuotų asmenų klinikinis fenotipo aprašymas yra gana informatyvus genotipo ir fenotipo ryšio analizei [265–274]. Nors 4q21 delecijos sindromas yra būdingas tiek vyrams, tiek moterims, visgi dažniau (63,3 proc.) pasireiškia moteriškosios lyties asmenims. Didžioji dauguma (25/30) visų analizuotų atvejų kilmė yra *de novo*, o likusių penkių tiriamųjų paveldėjimo būdas yra nežinomas. Visiems tiriamiesiems pasireiškė psichomotorinės raidos atsilikimas. Taip pat dažnai stebimos veido įgimtos anomalijos (90 proc.), žemas ūgis (83 proc.), kalbinės raidos vėlavimas (80 proc.) bei hipotonija (70 proc.). Smegenų pažeidimai (43,3 proc.) bei hipertelorizmas (33,3 proc.) pasireiškia rečiau (4.3 lentelė).

Augimo sutrikimas – tai vienas pagrindinių 4q21 delecijos klinikinių požymių, kuris siejamas su *PRKG2* haplonepakankamumu. *PRKG2* koduoja II tipo nuo cGMP (angl. *Cyclic Guanine Monophosphate*) priklausomą proteinkinazę, kuri jungiasi su ląstelių paviršiaus baltymais ir reguliuoja įvairius procesus, pavyzdžiui, renino sekreciją ar kaulų augimą. Siekdami įvertinti šio geno poveikį kaulų formavimosi procesui, Miyazawa su bendradarbiais (2002) tyrė *Prkg2*<sup>-/-</sup> pelių linijas, kurioms dėl sutrikusios endochondralinės osifikacijos postnataliniame laikotarpyje buvo stebimas nykštukinis (angl. *dwarfism*) fenotipas [275].

Palyginti su tiriamojo #9 824 Kb dydžio 4q21 delecija, mokslinėje literatūroje yra aprašyta tik viena mažesnė persidengianti delecija (761 Kb), kuri panašiai kaip tiriamojo #9 atveju, neapėmė *PRKG2* geno [266]. Zarrei ir kt. (2017), Yano ir kt. (2015), Bhoj ir kt. (2013), Tsang ir kt. (2012) bei *DECIPHER* #271211 ir #401283 tiriamiesiems taip pat nustatytos persidengiančios delecijos, kurios neapėmė su tiriamųjų žemu ūgiu siejamo *PRKG2* geno (4.24 pav.). Nors šių tiriamųjų genome yra nepakitęs *PRKG2*

kopijų skaičius, visgi daugumai (5/8; 62,5 proc.) pasireiškė augimo sutrikimai. Svarbu pažymėti, kad augimo sutrikimai nepasireiškė tik vienai septynerių metų tiriamajai (1/8; 12,5 proc.), o kitais atvejais (2/8; 25 proc.) antropometriniai duomenys nebuvo pateikti [266, 267, 269–271].

**4.3 lentelė.** Analizuotų asmenų, kuriems nustatyta 4q21 delecija, klinikinių požymių santrauka

Analizuoti asmenys	Del4q21 sritis (bp) GRCh36/hg18	Dydis	Paveldėjimas	Lytis	Pagrindiniai klinikiniai požymiai, būdingi 4q21 delecijos sindromui						
					PRA	Kalbos sutrikimai	Žemas ūgis	Hipotonija	Smegenų pažeidimai	Veido anomalijos	Hipertelorizmas
<b>Tiriamasis #9</b>	83 273 844– 84 097 897	824 Kb	<i>De novo</i>	Mot.	+	+	+	+	–	+	–
Bonnet ir kt., 2010	A 80 926 692– 84 215 893	3,2 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	+
	B 82 228 875– 86 711 060	4,5 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	+
	C 78 995 025– 84 459 549	5,5 Mb	<i>De novo</i>	Vyr.	+	+	+	+	+	+	–
	D 80 293 554– 86 615 654	6,3 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	+
	E 76 949 115– 83 601 083	6,6 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	+
	F 77 174 381– 86 773 353	9,6 Mb	<i>De novo</i>	Vyr.	+	+	+	+	+	+	+
	G 80 293 554– 86 615 654	9,7 Mb	<i>De novo</i>	Vyr.	+	+	+	+	+	+	–
	H 70 547 247– 84 066 553	13,5 Mb	<i>De novo</i>	Vyr.	+	+	+	+	+	+	+
	I 81 706 999– 96 867 539	15,1 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	–

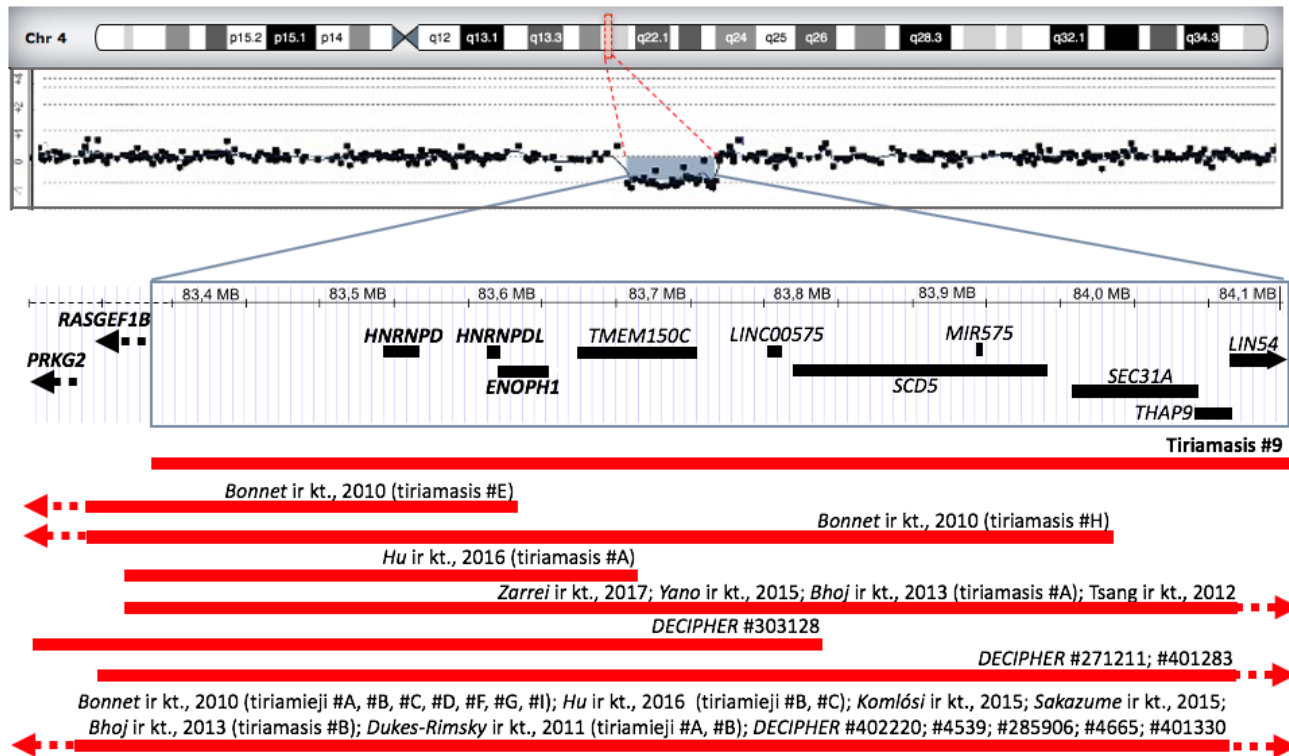
**4.3 lentelė (tęsinys).** Analizuotų asmenų, kuriems nustatyta 4q21 delecija, klinikinių požymių santrauka

Analizuoti asmenys	Del4q21 sritis (bp) GRCh36/hg18	Dydis	Pavel-dėjimas	Lytis	Pagrindiniai klinikiniai požymiai, būdingi 4q21 delecijos sindromui						
					PRA	Kalbos sutrikimai	Žemas ūgis	Hipotonija	Smegenų pažeidimai	Veido anomalijos	Hipertelorizmas
Zarrei ir kt., 2017	82 909 269– 84 260 359	1,3 Mb	<i>De novo</i>	Vyr.	+	+	ND <sup>11</sup>	ND	+	ND	–
Hu ir kt., 2016	A 82 942 185– 83 703 439	761 Kb	<i>De novo</i>	Vyr.	+	+	+	+	–	ND	–
	B 81 872 297– 85 234 249	3,36 Mb	<i>De novo</i>	Mot.	+	+	+	+	–	+	–
	C 80 044 702– 100 378 930	20,33 Mb	<i>De novo</i>	Vyr.	+	+	+	+	–	+	–
Komlósi ir kt., 2015	81 628 004– 86 480 977	4,85 Mb	<i>De novo</i>	Mot.	+	+	+	+	–	+	+
Yano ir kt., 2015	82 900 791– 84 408 390	1,6 Mb	ND	Vyr.	+	ND	+	+	–	+	+
Sakazume ir kt., 2015	78 396 380– 91 106 330	12,7 Mb	ND	Mot.	+	ND	+	+	ND	+	ND
Bhoj ir kt., 2013	A 83 061 747– 85 090 749	2,0 Mb	<i>De novo</i>	Mot.	+	+	–	–	+	+	ND
	B 82 048 194– 85 469 787	3,4 Mb	<i>De novo</i>	Vyr.	+	+	+	ND	ND	+	ND

<sup>11</sup> ND – nėra duomenų.

**4.3 lentelė (tęsinys).** Analizuotų asmenų, kuriems nustatyta 4q21 delecija, klinikinių požymių santrauka

Analizuoti asmenys	Del4q21 sritis (bp) GRCh36/hg18	Dydis	Paveldėjimas	Lytis	Pagrindiniai klinikiniai požymiai, būdingi 4q21 delecijos sindromui						
					PRA	Kalbos sutrikimai	Žemas ūgis	Hipotonija	Smegenų pažeidimai	Veido anomalijos	Hipertelorizmas
Tsang ir kt., 2012	82 429 949– 91 434 337	9 Mb	<i>De novo</i>	Mot.	+	+	+	+	ND	+	–
Dukes-Rimsky ir kt., 2011	A 81 055 150– 86 169 140	5,1 Mb	<i>De novo</i>	Mot.	+	+	+	+	+	+	+
	B 81 658 220– 87 569 120	5,9 Mb	ND	Mot.	+	+	+	ND	ND	+	ND
<i>DECIPHER</i> #303128	82 318 467– 83 827 820	1,51 Mb	<i>De novo</i>	Mot.	+	+	ND	ND	ND	+	ND
<i>DECIPHER</i> #271211	83 052 214– 84 868 876	1,82 Mb	<i>De novo</i>	Vyr.	+	+	ND	+	ND	+	ND
<i>DECIPHER</i> #402220	82 642 606– 84 626 781	1,98 Mb	<i>De novo</i>	Mot.	+	ND	+	+	ND	+	ND
<i>DECIPHER</i> #401283	83 321 610– 85 419 369	2,1 Mb	ND	Vyr.	+	ND	+	ND	ND	+	ND
<i>DECIPHER</i> #4539	82 301 439– 85 425 757	3,12 Mb	<i>De novo</i>	Mot.	+	+	ND	+	ND	+	ND
<i>DECIPHER</i> #285906	82 941 985– 86 371 992	3,43 Mb	ND	Mot.	+	ND	+	ND	+	+	+
<i>DECIPHER</i> #4665	82 229 075– 86 710 919	4,48 Mb	<i>De novo</i>	Mot.	+	+	+	ND	ND	ND	ND
<i>DECIPHER</i> #401330	80 149 484– 85 340 729	5,19 Mb	<i>De novo</i>	Mot.	+	ND	+	ND	ND	+	ND



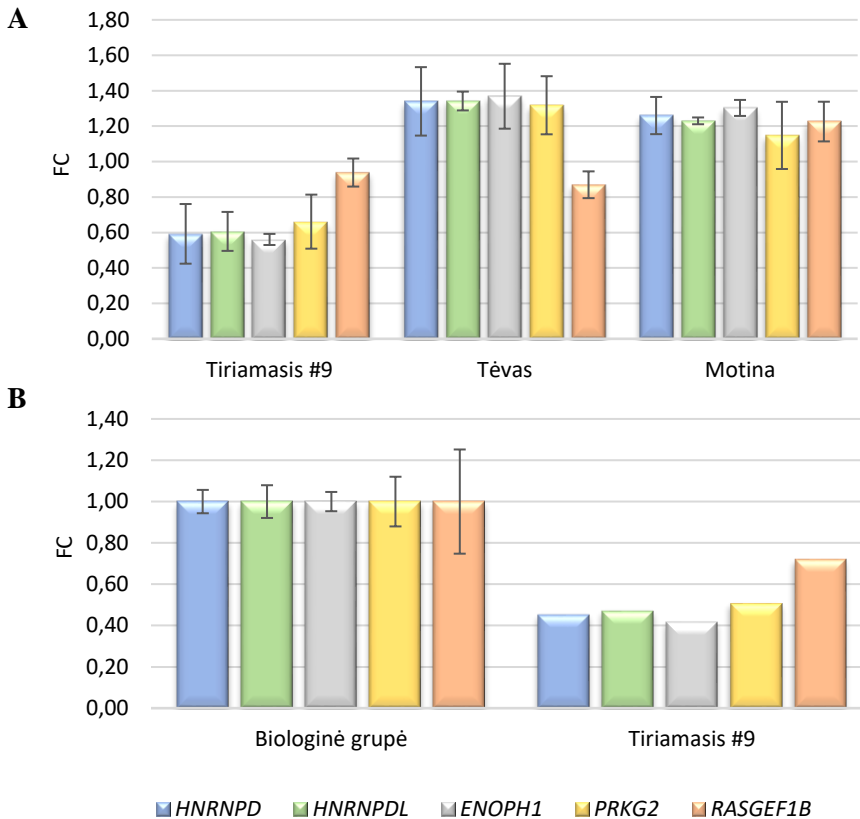
**4.24 pav.** 4 chromosomos ideograma, tiriamojo #9 vLGH tyrimo rezultatai 4q21.22(83 273 844-84 097 897)×1 dn (GRCh36/hg18) bei delecijos srities genų schema. Tiriamojo #9 4q21 delecijos sritis pažymėta pilku stačiakampiu, kandidatiniai genai paryškinti, o raudonos linijos žymi pagal mokslinę literatūrą [265–274] persidengiančių 4q21 delecijų sritis

Siekiant patikrinti mokslinėje literatūroje keliamą hipotezę dėl penkių kandidatinių *HNRNPD*, *HNRNPD*, *ENOPHI*, *PRKG2* bei *RASGEF1B* genų, siejamų su 4q21 delecijos sindromu, atlikta genų raiškos analizė kPGR metodu. Nepaisant to, kad tiriamojo #9 genome vLGH metodu nustatytas normalus *PRKG2* (chr4: 82 228 861-82 355 242, hg18) ir *RASGEF1B* (chr4: 82 566 571-83 184 421, hg18) kopijų skaičius, šių genų raiška gali būti sutrikdyta, kadangi KSP poveikis gali būti stebimas visos chromosomos mastu, o ypač gretimiems genams [262–264]. Be to, atstumas tarp tiriamos delecijos trūkio vietos ir *PRKG2*, *RASGEF1B* genų yra atitinkamai apie 918 Kb ir 89 Kb. Taip pat net 62,5 proc. analizuotų asmenų, kurių 4q21 srities delecijos neapėmė *PRKG2* geno, siejamo su žemu ūgiu, visgi pasireiškė augimo sutrikimai. Atkreiptinas dėmesys, kad tiriamos delecijos srityje yra miR575, kuri sąveikauja su įvairiais iRNR taikiniiais ir neigiamai reguliuoja jų raišką.

Atsižvelgiant į tai, kas išdėstyta, atlikta lyginamoji *HNRNPD*, *HNRNPD*, *ENOPHI*, *PRKG2* bei *RASGEF1B* genų raiškos analizė, apimanti tiriamojo #9 periferinio veninio kraujo mėginį ir jo sveikų tėvų referentinius mėginius (4.25 pav. A). Palyginti su biologinės grupės, kurią sudarė tiriamojo #9 tėvai, genų raiškos rezultatais, nustatyta, kad *HNRNPD*, *HNRNPD* ir *ENOPHI* genų, esančių 4q21 delecijos srityje, raiška tiriamojo #9 kraujo ląstelėse yra reikšmingai sumažėjusi, kadangi nustatytas didesnis nei du kartus genų raiškos pokytis (*HNRNPD* FC = 2,19; *HNRNPD* FC = 2,12; *ENOPHI* FC = 2,38). Gretimojo *RASGEF1B* raiška atitinka normą (FC = 1,39), o, tiriant *PRKG2*, nustatyta reikšmingai sumažėjusi (FC = 2) *PRKG2* raiška (4.25 pav. B). Siekiant patvirtinti *PRKG2* raiškos rezultatą, svarbu pakartoti tyrimą, padidinus referentinės biologinės grupės imtį. Taip pat, remiantis *GTE*x portalu [118] duomenimis, šio geno raiška yra maža tiriamų asmenų kraujyje (TPM<sup>12</sup> = 0,03), todėl patikimiau būtų tirti šio geno raišką, išgryninus visuminę RNR iš kito audinio, pavyzdžiui, odos fibroblastų, kuriems būdinga didesnė *PRKG2* raiška (TPM = 2,9). Pakartotinis tyrimas nebuvo atliktas, kadangi tiriamojo #9 išgrynintos visuminės RNR kiekio neužteko pakartoti kPGR, o šeima nesutiko odos biopsijos ar pakartotinei kraujo paėmimo procedūrai.

---

<sup>12</sup> TPM (angl. *Transcripts per million*) – 1 000 000 RNR transkriptų viename mėginyje.



**4.25 pav. A)** Lyginamuoju  $\Delta\Delta C_t$  metodu įvertintas *HNRNPD*, *HNRNPDL*, *ENOPH1*, *PRKG2* bei *RASGEF1B* genų raiškos pokytis (FC) tiriamojo #9 ir jo tėvų periferinio veninio kraujo mėginiuose. Histogramų „ūšai“ žymi techninių pakartojimų vidurkio SN. **B)** Tiriamojo #9 penkių genų raiškos pokytis pagal biologinę grupę (n = 2). Histogramų „ūšai“ žymi biologinės grupės vidurkio SN. Histogramos nubrėžtos naudojant *Microsoft Excel 2010* (*Microsoft Corporation*, JAV) statistinį paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems*, JAV) ir *qbase+* (*Biogazelle*, Belgija) programa

Be tirtų genų kandidatų, 4q21 delecijos srityje taip pat yra *MIR575* koduojama miRNR, kuri, *RegNetwork* [276] duomenų bazės duomenimis, reguliuoja 190 skirtingų genų raišką. *Phenoanalyzer* [110] bioinformaciniu įrankiu išanalizavus pateiktą genų sąrašą, nustatyta, kad nei vienas analizuotas genas nėra reikšmingai susijęs su tiriamojo #9 klinikiu fenotipu. Taip pat analizuojant mokslinę literatūrą ir duomenų bazes, nerasta jokio galimo ryšio tarp šios miRNR ir penkių kandidatinių genų, ypač *RASGEF1B* ir *PRKG2*, kuriems nustatytas nepakitęs kopijų skaičius tiriamojo #9 genome. Dėl šių priežasčių miRNR nebuvo toliau tirta funkciniais genetiniais tyrimais.

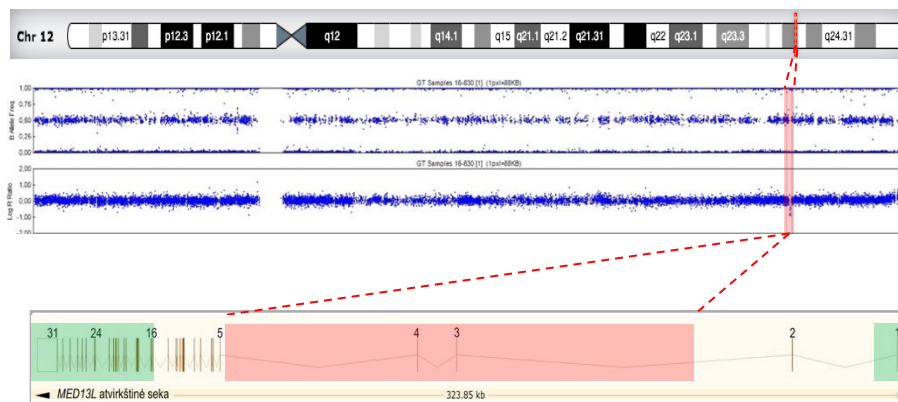


Nors *PRKG2* geno kopijų skaičius tiriamojo #9 genome yra normalus, dėl sumažėjusios raiškos yra tikėtinas šio geno haplonepakankamumas, kuriuo gali būti aiškinamas probando žemas ūgis. kPGR metodu tiriamojo #9 tirtame mėginyje nustatyta normali *RASGEF1B* raiška suteikė galimybę susiaurinti anksčiau su 4q21delecijos sindromu siejamą genetinę sritį. Todėl, remiantis išanalizuota moksline literatūra ir duomenų bazėmis bei tiriamojo #9 kPGR rezultatais, galima teigti, kad patogeninė 4q21 delecija, apimanti *HNRNPD*, *HNRNPD*, *ENOPH1* ir *PRKG2*, lemia 4q21 delecijos sindromą.

#### 4.2.3. Molekulinis, funkcinis ir klinikinis 12q24.21 delecijos apibūdinimas

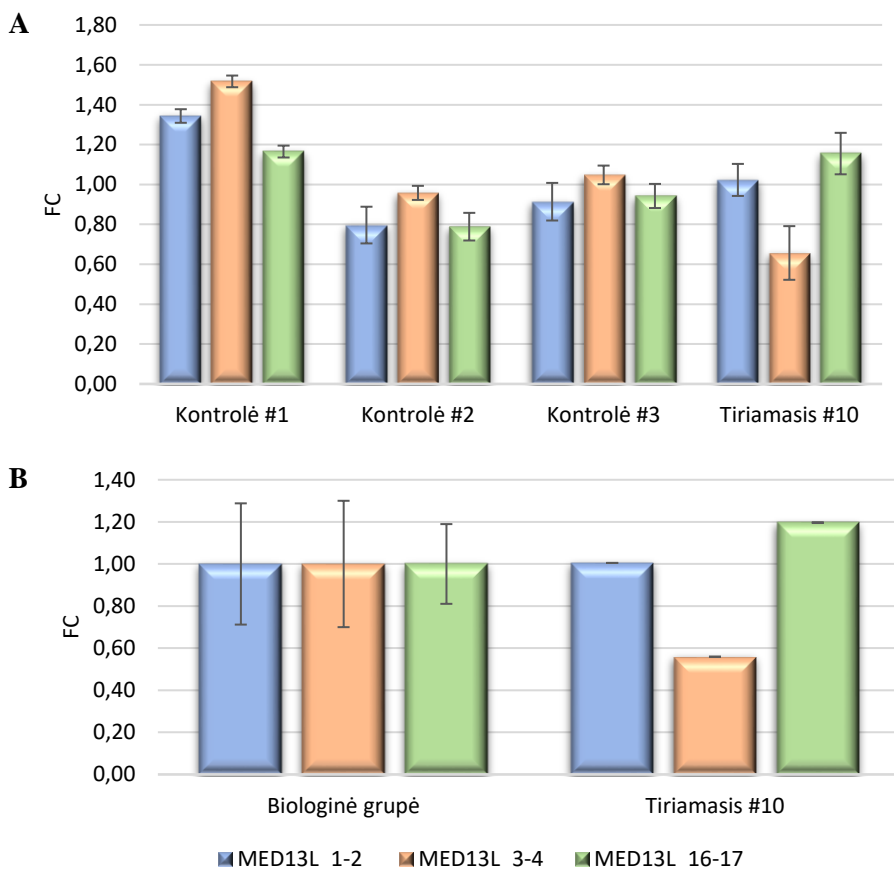
Intelektinė negalia su specifiniais veido bruožais su ar be širdies ydų (angl. *Intellectual disability and distinctive facial features with or without cardiac defects*; MIM #616789; ORPHA #369891) – tai autosominiu dominantiniu būdu paveldimas sutrikimas, kuriam būdinga IN, psichomotorinės, ypač kalbinės, raidos vėlavimas, mikrocefalija, mažosios veido anomalijos bei galimos įvairios širdies ydos. Šios patologijos, dažniau vadinamos *MEDI3L* haplonepakankamumo sindromu, priežastis yra baltymo funkcijos praradimą lemiantys heterozigotiniai variantai *MEDI3L* (MIM #608771) gene, kuris yra 12 chromosomos ilgajame petyje (12q24.21).

Tiriamajam #10 (5 metų amžiaus), kuriam pasireiškė *MEDI3L* haplonepakankamumo sindromui būdingi klinikiniai požymiai, VNP-LGH metodu [277] nustatyta galimai patogeninė vidugeninė *MEDI3L* geno delecija. Pagal referentinį žmogaus genomą (GRCh37/hg19) 12q24.21 delecija apėmė 97 880 bp sritį, kurioje yra *MEDI3L* geno 3 ir 4 egzonus (4.26 pav.). Atlikus 3–5 egzonų kiekybinį tyrimą kPGR [253] metodu tiriamojo asmens gDNR mėginyje, patvirtintas KSP tiriamojo #10 genome. Šis tyrimas atskleidė, kad dėl šiai sričiai specifinių zondu trūkumo VNP-LGH rinkinyje buvo nustatyta trumpesnė delecija, nei yra iš tiesų, kadangi ne tik 3 ir 4 egzonus, bet ir 5 egzono kopijų skaičius tiriamojo DNR mėginyje sumažėjęs. Ištyrus tiriamojo #10 tėvus, nustatyta šios delecijos *de novo* kilmė.



**4.26 pav.** 12 chromosomos ideograma, tiriamojo #10 VNP-LGH tyrimo rezultatai 12q24.21(116 523 305-116 621 185)x1 *dn* (GRCh37/hg19) bei vidugeninės *MED13L* delecijos schema. Tiriamojo #10 12q24.21 delecijos sritis pažymėta raudonu stačiakampiu, o 1 ir 16–31 egzoniai, kurių kopijų skaičius tiriamojo #10 genome nustatytas nepakitęs, pažymėti žaliais stačiakampiais

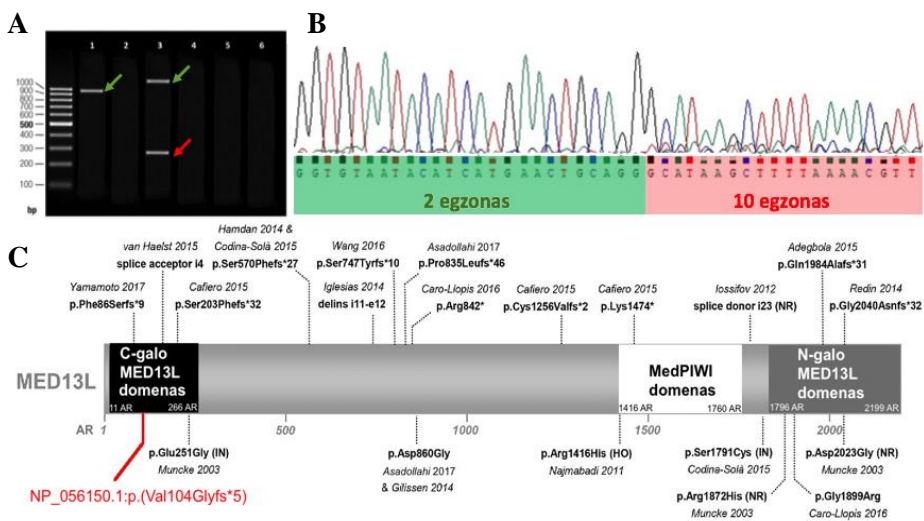
Nepaisant atliktų VNP-LGH ir kPGR tyrimų, 2–3 bei 5–16 *MED13L* geno egzonų sritys buvo neinformatyvios (4.26 pav.). Siekiant nustatyti tikslią KSP sritį ir įvertinti šio pokyčio padarinius iRNR lygmeniu, atliktas tiriamojo #10 *MED13L* raiškos tyrimas lyginamuoju kPGR metodu. Tyrime buvo naudojami oligonukleotidiniai *TaqMan*<sup>™</sup> žymenys, kurie hibridizuojasi skirtingose *MED13L* geno srityse – 1–2, 3–4 ir 16–17 egzonų jungties vietose. Kadangi VNP-LGH tyrimu buvo nustatytos dvi *MED13L* 1, 2 bei 16–31 egzonų kopijos, 1–2 ir 16–17 egzonų jungties sritys atliko vidinės kontrolės funkciją, remiantis jų tyrimo rezultatais nustatyta, kad 3–4 egzonų jungties fragmento fluorescencijos intensyvumas yra apie du kartus mažesnis (4.27 pav. A). Taip pat atlikta tiriamojo #10 ir trijų neturinčių ligos požymių negiminingų asmenų (kontrolė #1–#3; biologinė grupė) mėginių lyginamoji analizė. Nustatyta 1,79 karto mažesnė *MED13L* 3–4 egzonų jungties vietos raiška, o 1–2 bei 16–17 egzonų jungties vietų raiška buvo nepakitusi (4.27 pav. B).



**4.27 pav. A)** Lyginamuoju  $\Delta\Delta C_t$  metodu įvertintas *MED13L* raiškos pokytis (FC) tarp tiriamojo #10, trijų neturinčių ligos požymių negiminingų asmenų (kontrolės #1–#3; biologinė grupė) bei vidinės kontrolės (*MED13L\_1-2* ir *MED13L\_16-17*). Histogramų „ūšai“ žymi techninių pakartojimų vidurkio SN. **B)** Tiriamojo #10 *MED13L* raiškos pokytis pagal biologinę grupę (n = 3). Histogramų „ūšai“ žymi biologinės grupės vidurkio SN. Histogramos nubrėžtos naudojant *Microsoft Excel 2010* (*Microsoft Corporation, JAV*) statistinį paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems, JAV*) ir *qbase+* (*Biogazelle, Belgija*) programa

*MED13L* koduojamos srities delecijos dydis patikslintas, atlikus tiriamojo #10 kDNR mėginio analizę Sangerio sekoskaitos metodu. Tiesioginis oligonukleotidinis pradmuo buvo kuriamas antrojo egzono srityje, o grįžtamasis oligonukleotidinis pradmuo – 8, 10 ir 15 egzonų srityse. Nesant delecijos, naudojant šiuos pradmenis, amplifikuojasi DNR sekos fragmentai, kurių ilgis 953 bp (2–8 egzonų amplikonas), 1 262 bp (2–10 egzonų

amplikonas) bei 2 524 bp (2–15 egzonų amplikonas) (III priedas, III.1 lentelė). Tyrimas pradėtas, naudojant pradmenis, amplifikuojančius trumpiausią fragmentą. Atlikus PGR ir įvertinus gelio elektroforezės vaizdą, nustatytas vienas 953 bp ilgio PGR produktas (4.28 pav. A). Remiantis šiuo rezultatu, iškelta prielaida, kad reakcijoje amplifikavosi tik laukinio tipo alelio (angl. *wild-type allele*) kDNR fragmentas. Dėl šios priežasties atlikta papildoma PGR, naudojant pradmenis, apimančius ilgesnius kDNR fragmentus (1 262 bp ir 2 524 bp; III priedas, III.1 lentelė). PGR reakcijoje, kurioje naudoti 2 ir 10 egzonams specifiniai pradmenys, amplifikavosi du PGR produktai, kurių apytikslis ilgis įvertintas apie 1 262 bp ir 292 bp (4.28 pav. A). Ilgesniojo 2 524 bp ilgio fragmento, apimančio *MED13L* geno 2–15 egzonus, amplifikuoti nepavyko. Amplifikuoti PGR produktai ištirti Sangerio sekoskaitos metodu. Nustatyta, kad 12q24.21 delecija apėmė koduojančią sritį nuo 3 iki 9 egzono imtinai (4.28 pav. B). *In silico* NG\_023366.1(NM\_015335.5):c.(310+1\_311-1)(1280+1\_1281-1)del pokytis lemia skaitymo rėmelio poslinkį ir priešlaikinio baigmės kodono susidarymą (NP\_056150.1:p.(Val104Glyfs\*5)) (4.28 pav. C).



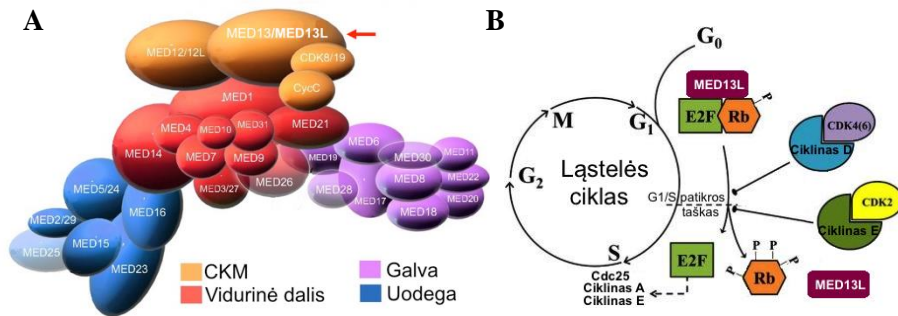
**4.28 pav. A)** Tiriamojo #10 kDNR mėginio PGR produkto vizualizacija agarozės gelyje. Elektroforezės pirmame takelyje vizualizuojamas 953 bp ilgio laukinio tipo alelio (žalia rodyklė) kDNR fragmentas. Trečiame takelyje vizualizuojamas laukinio tipo alelio 1 262 bp ilgio kDNR fragmentas bei pakitusio alelio (raudona rodyklė) 292 bp ilgio kDNR fragmentas. Penktas takelis atitinka 2 524 bp ilgio fragmento PGR. Neigiama PGR kontrolė vaizduojama 2, 4 ir 6 takelyje. Rezultatai vertinti pagal 100 bp DNR molekulinės masės ir dydžio standartą; **B)** Tiriamojo #10 PGR produkto chromatograma; **C)** Linijinė *MED13L* baltymo

(UniProtKB #Q71F56) struktūra su apytiksliai nurodyta prognozuojamo NP\_056150.1:p.(Val104Glyfs\*5) pokyčio vieta ir literatūroje aprašytų pokyčių sritimis. Schema adaptuota pagal *Pfam* [120] duomenų bazę ir mokslinę publikaciją [278]

Remiantis moksline literatūra, *MED13L* koduojamo baltymo funkcijos praradimą lemiantys skaitymo rėmelio poslinkio ar *nonsense* tipo variantai sudaro didžiąją dalį (76 proc.) visų iki šiol šiame gene nustatytų genetinių pokyčių [279]. *HGMD* [34] duomenų bazėje įtraukti 28 patogeniniai *MED13L* pokyčiai. *ClinVar* [8] duomenimis, tokių patogeninių variantų ir KSP yra registruojama daugiau – 91, iš kurių didžiąją dalį sudaro šio geno delecijos (36 proc.) ir duplikacijos (20 proc.). Nors *MED13L* geno delecija yra dažniausiai šiame gene nustatomas genetinis pokytis, tiriamajam #10 nustatyta nauja vidugeninė delecija, apimanti 3–9 egzonus, kuri anksčiau literatūroje nebuvo aprašyta.

*MED13L* koduojamas produktas (angl. *Mediator Complex Subunit 13-Like*) yra didelio baltymų komplekso, vadinamo mediatoriumi (angl. *Mediator complex*), subvienetas (UniProtKB #Q71F56; 4.29 pav. A). Mediatoriaus komplekso struktūra yra labai konservatyvi. Išskiriamos trys pagrindinės struktūrinės dalys: galva (angl. *head*), vidurinė dalis (angl. *middle*), uodega (angl. *tail*) bei CDK8 kinazės modulis (CKM). Mediatoriaus kinazės dalį sudaro CDK8, ciklinas C bei *MED13L*, *MED13*, *MED12L*, *MED12* koduojami baltymai (4.29 pav. A). Moksliniai tyrimai rodo, jog mediatoriaus *MED13L* ir jo paralogo *MED13* koduojami subvienetai užtikrina tiesioginę sąveiką tarp MedPIWI domeno (4.28 pav. C) ir antrojo tipo RNR polimerazės, todėl yra ypač svarbūs reguliuojant genų raišką [280]. Utami su bendradarbiais (2014) tyrė hemizigotinės būklės *MED13L* geno poveikį įvairiose modelinėse sistemose. Pirmiausia, išjungę *MED13L* ortologą (*med13b*) zebražuvėje, mokslininkai nustatė sutrikusią nervinių skiauterinių ląstelių (angl. *Neural crest cells*) migraciją, kuri vėlesnėse zebražuvų vystymosi stadijose nulėmė kremzlinių struktūrų deformacijas. Taip pat mokslininkai tyrė *in vitro* šio geno poveikį žmogaus neuronuose, kurie buvo sukurti transformuojant iš EKL kilusius nervinių ląstelių pirmtakus. Šiose ląstelėse *MED13L* išjungtas panaudojus siRNR technologiją, o, atlikę viso transkriptomo analizę, mokslininkai nustatė sutrikusią net 1 117 genų raišką. Daugelis šių genų dalyvauja Wnt (angl. *Wingless*, *Int-1*) ir FGF (angl. *Fibroblast growth factors*) signaliniuose keliuose, kurie intensyviai veikia reguliuojant įvairius viduląstelinis ir tarpląstelinis procesus ankstyvose embriono vystymosi stadijose, o ypač formuojant galūnių skeletą bei nervų sistemą [281]. Angusas ir Nevinsas (2012), tirdami žmogaus vėžinius

audinius, nustatė, kad *MED13L* yra svarbus Rb/E2F signalinio kelio, reguliuojančio ląstelės G1/S (arba R) patikros tašką, kofaktorius. Baltymas Rb (angl. *Retinoblastoma*) yra žinomas kaip svarbus vėžio slopiklis, kuris jungiasi su E2F transkripcijos veiksniu specifinių genų promotoriuose ir slopina jų raišką. Ciklino ir nuo ciklino priklausomos kinazės (CDK) kompleksas fosforilina Rb ir pakeičia baltymo konformaciją taip, kad E2F transkripcijos veiksnyms negali sąveikauti su Rb. Todėl laisvas E2F aktyvuoja genų, koduojančių daugelį ląstelės ciklui svarbių baltymų (pvz., ciklinas A), raišką (4.29 pav. B). Mokslinio tyrimo metu nustatyta, kad *MED13L* sąveikaujant su RB/E2F baltymų kompleksu kartu slopinama ciklino A raiška, nuo kurio priklauso ne tik tolesnė ląstelės ciklo veikla, bet ir ląstelės senėjimo procesas [282].



**4.29 pav. A)** Mediatoriaus komplekso schema [280]. Raudona rodyklė žymi MED13L ir jo paralogo MED13 subvienetą; **B)** Rb/E2F signalinio kelio schema, kuri adaptuota pagal mokslinę literatūrą [282, 283]

Manoma, kad patogeninių *MED13L* variantų sutrikdyti Wnt, FGF bei Rb/E2F signaliniai keliai nulemia *MED13L* haplonepakankamumo sindromą. Iki šiol literatūroje yra aprašyti 69 asmenys [279], kuriems buvo nustatyti patogeniniai pokyčiai *MED13L* gene. *DECIPHER* [108] duomenų bazėje registruotos devynios patogeninės *MED13L* delecijos ir trys duplikacijos (4.4 lentelė). Tačiau ne visi paskelbti KSP yra tinkami genotipo ir fenotipo ryšio analizei. Tiriamųjų, kuriems nustatyta 275,59 Kb dydžio delecija (#294797) ir 18,71 Mb dydžio duplikacija (#370352), nėra žinomas klinikinis fenotipas. Kito tiriamojo (#285997) genotipo ir fenotipo analizė yra neinformatyvi dėl delecijos dydžio (101,30 Mb). *MED13L* haplonepakankamumo sindromui būdingi įvairaus spektro klinikiniai požymiai, tačiau visiems tiriamiesiems, įskaitant tiriamąjį #10, pasireiškia įvairaus laipsnio IN ir psichomotorinės, įskaitant kalbinės, raidos sutrikimas. Kiti fenotipiniai požymiai – raumenų

hipotonija, smegenų pažeidimai, akių, veido ir galūnių anomalijos diagnozuojamos ne visiems, bet daugumai tiriamųjų (37–76 proc.). Įgimtos širdies ydos (IŠY), kurios mokslinėje literatūroje yra dažnai pabrėžiamos kaip vienas pagrindinių *MED13L* haplonepakankamumo sindromo požymių, yra nustatomos tik penktadaliui analizuotų asmenų (4.4 lentelė).

**4.4 lentelė.** Analizuojamų asmenų, kuriems nustatyti patogeniniai *MED13L* variantai ar KSP, klinikinių požymių santrauka

Analizuoti asmenys	IN	PRA	Kalbinės raišos atsilikimas	Dismorfiniai veido požymiai	Galūnių anomalijos	Hipotonija	Akių anomalijos	Smegenų pažeidimai	IŠY	Autizmas
<b>Tiriamasis #10</b>	+	+	+	–	+	–	+	–	–	–
<b>Literatūroje aprašyti tiriamieji (n = 69) [108]</b>	69/69	69/69	68/69	50/67	31/69	46/66	23/69	26/58	12/64	16/69
<b>12q24.21 delecija pgl. DECIPHER</b>	#264232	ND	+	ND	+	ND	ND	+	ND	ND
	#260719	ND	+	ND	ND	ND	+	ND	ND	ND
	#271765	ND	+	ND	+	ND	ND	ND	ND	ND
	#271595	+	+	ND	+	ND	ND	ND	ND	+
	#285418	+	ND	ND	ND	ND	ND	ND	ND	ND
	#283932	ND	+	+	+	ND	ND	+	ND	ND
	#277527	ND	+	+	+	ND	ND	+	ND	ND
<b>12q24.21 duplikacija pgl. DECIPHER</b>	#296364	+	+	+	ND	ND	ND	ND	ND	ND
	#296365	ND	+	ND	+	ND	ND	ND	ND	ND
<b>Bendras skaičius</b>	72/72	78/78	72/73	56/74	35/73	46/67	27/73	27/60	12/65	17/71
<b>Bendras skaičius (proc.)</b>	100 proc.	100 proc.	99 proc.	76 proc.	48 proc.	69 proc.	37 proc.	45 proc.	18 proc.	23 proc.



Nors moksliniais tyrimais yra pagrįsta, kad patogeniniai *MEDI3L* variantai lemia *MEDI3L* haplonepakankamumo sindromą, tikslus *MEDI3L* genetinio pokyčio nulemtos patologijos etiopatogenezės mechanizmas ir poveikis žmogaus ląstelių lygmeniu nėra visapusiškai ištirtas. Šiame tyrime, siekiant įvertinti nustatyto *MEDI3L* KSP lemiamą poveikį ląstelių morfologijai, gyvybingumui, senėjimui, ląstelių ciklui bei susijusių genų raiškai, *MEDI3L* genui nutildyti kontrolinio asmens odos fibroblastų kultūroje buvo naudota CRISPR-Cas9 technologija. Siekiant įgyvendinti šį uždavinį, užsibrėžta palyginti kontrolinio asmens redaguotos ir neredaguotos fibroblastų kultūros ypatybes su tiriamojo #10 nemodifikuotos ląstelių kultūros savybėmis. Pirmiausia, siekiant įvertinti fibroblastų užsėjimo tankio įtaką genomo redagavimo efektyvumui, kontrolinio asmens fibroblastai buvo užsėti skirtingu tankiu ( $\sim 0,3 \times 10^6$ ) į tris kultivavimo indelius. Apytiksliai įvertinus vienodomis sąlygomis kultivuotų ląstelių tankį, trijuose kultivavimo indeliuose nustatyta maždaug 50 proc., 60 proc. ir 65 proc. konfluencija. Visuose trijuose indeliuose kultivuotų ląstelių skaičius buvo pakankamas transfekcijos procedūrai atlikti. Nors ląstelių tankis trijuose kultivavimo indeliuose skyrėsi, *MEDI3L* taikinio kirpimo efektyvumas apskaičiuotas panašus, atitinkamai – 44 proc., 43 proc. ir 43 proc.

Lyginamajai ir funkicinei ląstelių kultūrų analizei reikalingos iš vienos ląstelės išaugintos kultūros, todėl apie 44 proc. efektyvumu modifikuota fibroblastų kultūra praskiesta iki vienos ląstelės kultivavimo indelyje. Iš viso kultivavimo indeliuose užsėta 960 šulinėlių. Atrinkus šulinėlius, kuriuose augo vienos ląstelės klonai, iš viso kultivuoti 25 gyvybingi ląstelių klonai iki funkciniais tyrimams pakankamo ląstelių kiekio. Tačiau dėl *force majeure* aplinkybių (karantinas dėl SARS-CoV2 viruso sukeltos COVID-19 infekcijos) tyrimas buvo pertrauktas, užšaldžius 16-ka<sup>13</sup> fibroblastų kultūrų. Pasibaigus karantinui Lietuvos Respublikos teritorijoje, kontrolinio asmens redaguotos fibroblastų kultūros bei tiriamojo #10 neredaguota fibroblastų kultūra buvo atšildytos ir toliau kultivuojamos.

Redaguotas ląstelių kultūras kultivuojant iš vienos ląstelės neišvengiama ilgalaikio ląstelių klonų kultivavimo ir su tuo susijusių pasekmių: morfologijos pakitimų, gyvybingumo sumažėjimo, ląstelės ciklo pažeidimų ir ląstelių senėjimo proceso [284]. Kultivuojant redaguotas kontrolinio asmens fibroblastų kultūras, pastebėta, kad dalies ląstelių kultūrų morfologija pradėjo kisti anksčiau. Tokiose morfologiškai besiskiriančių

---

<sup>13</sup> Devyniose fibroblastų kultūrose ląstelių skaičius buvo nepakankamas šaldymo procedūrai atlikti.

ląstelių kultūrose fibroblastai dauginosi akivaizdžiai lėčiau ir mažėjo jų gyvybingumas. Dėl sulėtėjusios ląstelių proliferacijos ir gyvybingumo ne visuose kultivavimo indeliuose buvo pakankamas ląstelių skaičius visuminės RNR gryninimo procedūrai<sup>14</sup>. Dėl šios priežasties visuminė RNR išgryninta iš 11-kos kontrolinio asmens redaguotų fibroblastų kultūrų ir tiriamojo #10 neredaguotos ankstyvojo pasažo fibroblastų kultūros.

Sangerio sekoskaitos metodu tiriant pagal iRNR susintetintą kDNR, šešiose kontrolinio asmens redaguotose fibroblastų kultūrose nustatyti pokyčiai *MED13L* gene, o likusiose penkiose redaguotose fibroblastų kultūrose jokių *MED13L* variantų nenustatyta. Trijose fibroblastų kultūrose (žymima: #A–#C) nustatyta heterozigotinė antrojo egzono iškrita, o vienoje ląstelių kultūroje (#D) nustatyta dviejų G nukleotidų iškrita heterozigotinėje būklėje. Kitose dviejose fibroblastų kultūrose (#E, #F) abiejuose aleliuose nustatyti skirtingi genetiniai pokyčiai: antrojo egzono iškrita, dviejų G nukleotidų iškrita bei G nukleotido intarpas. *In silico* įvertinus šių pokyčių prognozuojamus padarinius baltymo lygmeniu, nustatyta, kad visais atvejais susidaro skaitymo rėmelio poslinkis ir priešlaikinis baigmės kodonas, lemiantys *MED13L* haplonepakankamumą ląstelėje (4.5 lentelė).

**4.5 lentelė.** Sangerio sekoskaitos metodu nustatyti *MED13L* variantai po genomo redagavimo CRISPR-Cas9 technologija bei *in silico* prognozuojamas šių variantų poveikis baltymo lygmeniu

Redaguota ląstelių kultūra	Geno-tipas	<i>MED13L</i> redagavimo pasekmės kDNR lygmeniu	<i>In silico</i> prognozuojamas poveikis baltymo lygmeniu
#A	Heterozigotinis	NG_023366.1(NM_015335.5): c.(72+1_73-1)_(310+1_311-1)del	NP_056150.1: p.(Ala25Cysfs*13)
#B	Heterozigotinis	NG_023366.1(NM_015335.5): c.(72+1_73-1)_(310+1_311-1)del	NP_056150.1: p.(Ala25Cysfs*13)
#C	Heterozigotinis	NG_023366.1(NM_015335.5): c.(72+1_73-1)_(310+1_311-1)del	NP_056150.1: p.(Ala25Cysfs*13)
#D	Heterozigotinis	NM_015335.5:c.95_96del	NP_056150.1: p.(Trp32Serfs*1)
#E	Skirtingi pokyčiai abiejuose aleliuose	NG_023366.1(NM_015335.5): c.(72+1_73-1)_(310+1_311-1)del	NP_056150.1: p.(Ala25Cysfs*13)
		NM_015335.5:c.95_96del	NP_056150.1: p.(Trp32Serfs*1)

<sup>14</sup> Vadovaujantis gamintojo vartotojo vadovu *RNeasy® Mini Handbook* (Qiagen, Vokietija) visuminės RNR gryninimui rekomenduojamas fibroblastų skaičius yra 3–4 x 10<sup>6</sup> ląstelių.

#F	Skirtingi pokyčiai	NM_015335.5:c.95_96del	NP_056150.1: p.(Trp32Serfs*1)
	abiejuose aleliuose	NM_015335.5:c.96_97insG	NP_056150.1: p.(Arg33Alafs*1)

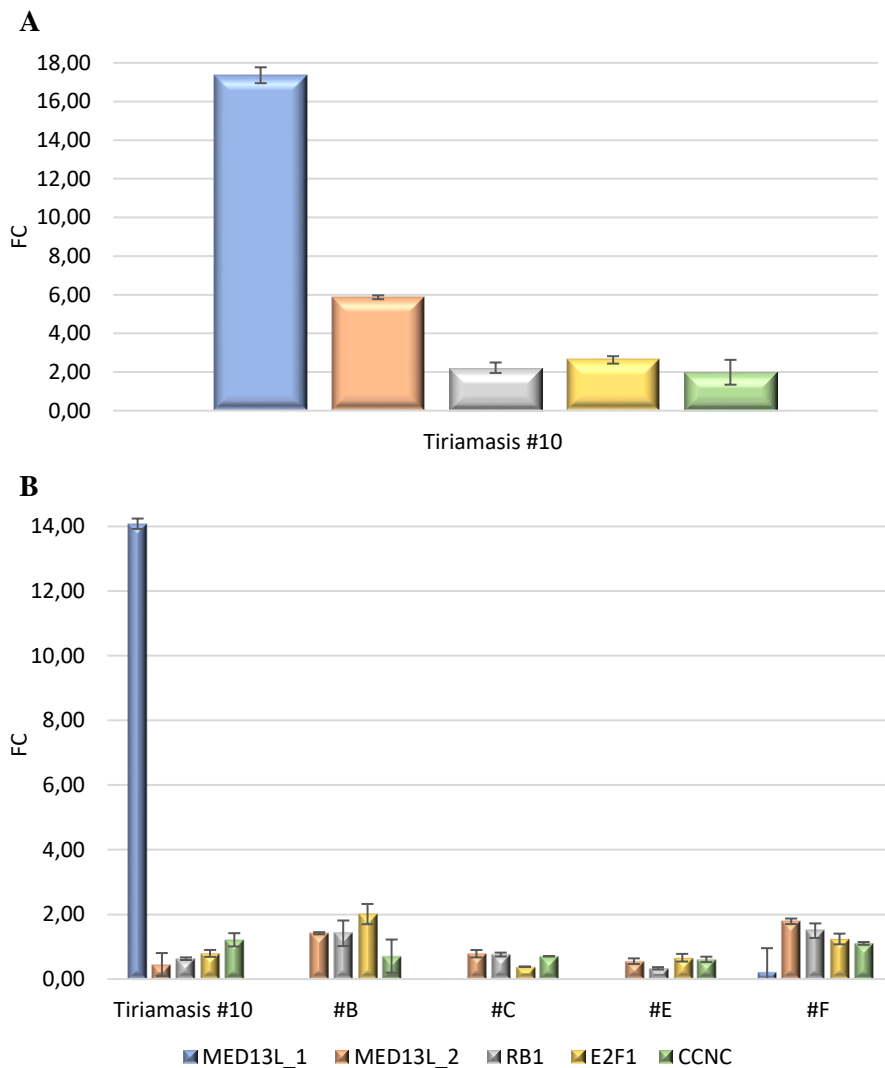
Siekiant įvertinti nustatytų *MED13L* variantų įtaką ląstelės ciklą valdančių *RBI* (MIM #614041), *E2F1* (MIM #189971) ir *CCNC* (MIM #123838) genų raiškai, atlikta lyginamoji genų raiškos analizė kPGR metodu. Pirmiausia įvertinta šių genų raiška tiriamojo #10 kraujyje. Lyginant su vidine kontrole<sup>15</sup> (1–2 egzonų jungtis; *MED13L\_1*), 3–4 egzonų jungties (*MED13L\_2*) vietos fluorescencijos intensyvumas yra apie 2,9 karto mažesnis. Ląstelės ciklą valdančių *RBI*, *E2F1* ir *CCNC* genų raiška nustatyta taip pat mažesnė, atitinkamai 7,8, 6,6 ir 8,7 karto (4.30 pav. A). *GTE*x portalo [118] duomenimis, 755 ištirtų asmenų kraujyje ląstelės ciklą valdančių genų (*RBI* TPM = 1,9; *E2F1* TPM = 4,2; *CCNC* TPM = 2,5) raiška yra atitinkamai 3,10, 1,4 ir 2,4 karto mažesnė, palyginti su *MED13L* (TPM=5,9) geno raiška. Tačiau, remiantis atlikto kPGR tyrimo rezultatais, tiriamojo #10 kraujyje analizuojamų ląstelės ciklą valdančių genų raiška yra ne tik 6,6–8,7 karto mažesnė, palyginti su vidine *MED13L* kontrole, bet ir apie du kartus mažesnė nei 3–4 egzonų jungties vietos (4.30 pav. A), kurios delecija patvirtinta kPGR ir Sangerio sekoskaitos metodais, raiška. *GTE*x portale [118] taip pat nurodoma, kad žmogaus fibroblastuose analizuojamų keturių genų raiška yra žymiai didesnė nei kraujo mėginiuose (*MED13L* TPM = 28,9; *RBI* TPM = 38,1; *E2F1* TPM = 9,1; *CCNC* TPM = 29,8). Todėl ta pati kPGR tyrimo strategija taikyta siekiant įvertinti šių genų raiškos pokytį tiriamojo #10 bei kontrolinio asmens redaguotose fibroblastų kultūrose (#B, #C, #E, #F<sup>16</sup>). Palyginti su vidine kontrole (*MED13L\_1*), 3–4 egzonų jungties (*MED13L\_2*) vietos fluorescencijos intensyvumas yra apie 32 kartus mažesnis, o ląstelės ciklą valdančių *RBI*, *E2F1* ir *CCNC* genų raiška yra daugiau nei 7 kartus

<sup>15</sup> Genų raiškos tyrime taip pat tirta biologinė kontrolė – vienodomis sąlygomis kultivuota neredaguota neturinio ligos požymio asmens fibroblastų kultūra. Atlikus genų raiškos tyrimą, nustatyta, kad šio kontrolinio asmens *MED13L* raiška yra reikšmingai sumažėjusi, lyginant su tiriamojo #10 mėginiu. Dėl šios priežasties šis mėginys eliminuotas iš analizės, o lyginamoji analizė atlikta pagal vidinę kontrolę (*MED13L* 1–2 egzonų jungties vieta), kurios nepakitusi raiška tiek tiriamojo, tiek biologinės grupės (n = 3) mėginiuose patvirtinta pirmo *MED13L* geno raiškos tyrimo metu (4.27 pav. B).

<sup>16</sup> Genų raiška netirta #A ir #D redaguotose fibroblastų kultūrose, kadangi išgrynintos visuminės RNR koncentracijos buvo per mažos, atitinkamai 20 ng/μl ir 30,6 ng/μl. AT-PGR mišinyje turi būti naudojamas vienodas išgrynintos visuminės RNR kiekis, siekiant užtikrinti genų raiškos tyrimo efektyvumą, tikslumą ir atkartojamumą.

mažesnė tiek tiriamojo #10, tiek kontrolinio asmens redaguotose fibroblastų kultūrose (4.30 pav. B).

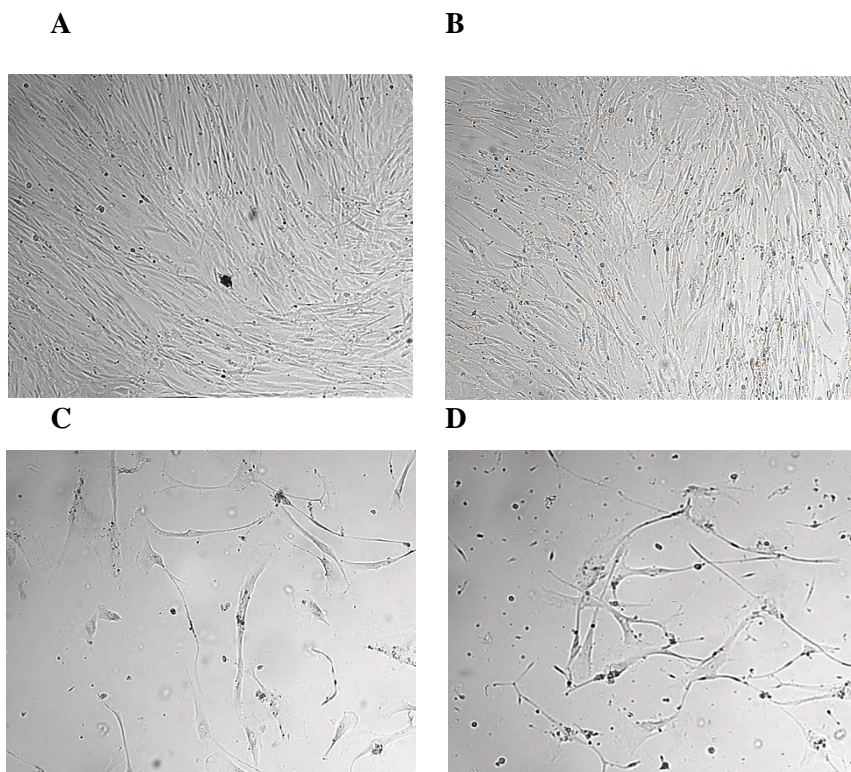
Dviejuose fibroblastų kultūrose (#E, #F) abiejuose aleliuose nustatyti skirtingi antrojo egzono genetiniai pokyčiai. Prognozuojama, kad šie variantai lemia skaitymo rėmelio poslinkį ir priešlaikinį baigmės kodoną abiejuose aleliuose, todėl 3–31 *MED13L* egzono raiška neturėtų būti stebima. Atliktame kPGR tyrime užfiksuotas 1–2 egzono jungties vietos fluorescencijos intensyvumas #F fibroblastų kultūroje bei abiejose kultūrose nustatyta 3–4 egzono jungties vietos raiška (4.30 pav. B). Dėl šios priežasties tikėtina, kad #E ir #F fibroblastų kultūroje buvo kultivuojami dviejų redaguotų ląstelių klonai, kurie yra netinkami tolesnei funkcinei analizei.



**4.30 pav. A)** Lyginamuoju  $\Delta\Delta C_t$  metodu pagal vidinę kontrolę (MED13L\_1) įvertintas *MED13L*, *RBI*, *E2F1* ir *CCNC* raiškos pokytis (FC) tiriamojo #10 periferinio veninio kraujo mėginyje. Histogramų „ūsai“ žymi techninių pakartojimų vidurkio SN. **B)** *MED13L*, *RBI*, *E2F1* ir *CCNC* raiškos pokytis tiriamojo #10 ir redaguotose (#B, #C, #E, #F) fibroblastų kultūrose. Histogramų „ūsai“ žymi techninių pakartojimų vidurkio SN. Histogramos nubrėžtos naudojant *Microsoft Excel 2010* (*Microsoft Corporation*, JAV) statistinį paketą pagal rezultatus, apdorotus *SDS v.2.3* (*Applied Biosystems*, JAV) ir *qbase<sup>+</sup>* (*Biogazelle*, Belgija) programa

Sangerio sekoskaitos ir kPGR metodu įvertinus *MED13L* variantus tiek tiriamojo #10, tiek kontrolinio asmens fibroblastų kultūrose, toliau atlikta lyginamoji šių ląstelių kultūrų funkcinė analizė, kadangi ląstelių morfologijos, gyvybingumo, senėjimo ir ciklo pokyčiams gali turėti įtakos ne tik ilgalaikis kultivavimas, bet ir genų, atsakingų už ląstelės ciklą, raiškos pokyčiai [282, 285].

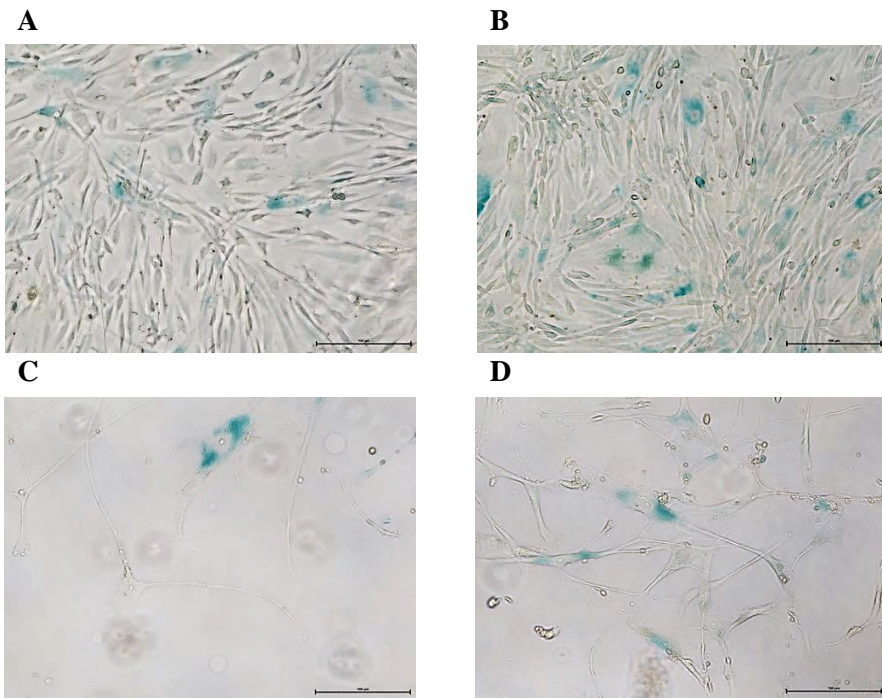
Kultivuojant fibroblastų kultūras stebėta ląstelių morfologija kiekviename pasaže. Prieš atliekant genomo redagavimo eksperimentą buvo palyginta ankstyvųjų pasažų (~III–IV) tiriamojo #10 ir kontrolinio asmens neredaguotų fibroblastų kultūrų morfologija. Įvertinta, kad abiejose fibroblastų kultūrose dauguma ląstelių buvo nepakitusios taisyklingos verpstės formos (4.31 pav. A). Tolesnio kultivavimo metu, nagrinėjant tiriamojo #10 neredaguotą fibroblastų kultūrą bei kontrolinio asmens redaguotas fibroblastų kultūras, pastebėta, kad su kiekvienu vėlyvesniu pasažu daugėja padidėjusių ir netaisyklingos formos ląstelių. Tokios pakitusios morfologijos fibroblastai kultūroje neformuoja tankių monosluoksnių plotų (4.31 pav. B). Kultivavimo pabaigoje vėlyvųjų pasažų (~X–XV) fibroblastų kultūrose tokių ląstelių buvo didžioji dauguma.



**4.31 pav.** Ankstyvojo (~III–IV) ir vėlyvojo (~X–XV) pasažo fibroblastų kultūrų morfologijos tyrimas invertuotu mikroskopu *Olympus CKX41* (*Olympus Life Science*, Japonija; padidinimas 4x, mastelis 500  $\mu\text{m}$ ): **A**) Reprezentatyvi ankstyvojo pasažo kontrolinio asmens neredaguotos fibroblastų kultūros nuotrauka; **B**) Reprezentatyvi ankstyvojo pasažo tiriamojo #10 neredaguotos fibroblastų kultūros nuotrauka; **C**) Reprezentatyvi kontrolinio asmens redaguotos vėlyvojo pasažo fibroblastų kultūros, kurioje nenustatyta *MED13L* variantų, nuotrauka; **D**) Reprezentatyvi kontrolinio asmens redaguotos vėlyvojo pasažo fibroblastų kultūros, kurioje nustatytas *MED13L* variantas heterozigotinėje būklėje, nuotrauka

Ląstelių padidėjimas, suplokštėjimas ir įgyta netaisyklinga forma yra ląstelių senėjimo požymis [285]. Siekiant įvertinti *MED13L* haplonepakankamumo įtaką fibroblastų senėjimui, įvertintas senstančioms ląstelėms būdingas  $\beta$ -galaktozidazės aktyvumas tiriamojo #10 neredaguotoje fibroblastų kultūroje bei kontrolinio asmens redaguotoje ir neredaguotoje fibroblastų kultūroje. Fermentas  $\beta$ -galaktozidazė skaldo dažą X-gal (5-bromo-4-chloro-3-indolil- $\beta$ -D-galaktopiranozidas), susiformuoja mėlynos spalvos skaidymo produktai, todėl senstančios ląstelės nusidažo mėlynai [286]. Įvertinus nusidažiusių ląstelių ir bendrą ląstelių skaičių, nustatyta, kad

kontrolinio asmens neredaguotoje ankstyvojo pasažo fibroblastų kultūroje senstančių ląstelių yra apie 10 proc. Nepaisant to, kad tiriamojo #10 ankstyvojo pasažo fibroblastų kultūroje tik pavienės ląstelės buvo morfologiškai pakitusios, atlikus senėjimo tyrimą nustatyta beveik 3 kartus daugiau (27,6 proc.) senstančių ląstelių, palyginti su kontroline neredaguota fibroblastų kultūra (4.32 pav. A, B). Tiriant kontrolinio asmens vėlyvųjų pasažų fibroblastų kultūras po CRISPR-Cas9 genomo redagavimo eksperimento nustatyta, kad fibroblastų kultūroje, kurioje nenustatyta *MED13L* variantų, yra apie 55,2 proc. senstančių ląstelių, o fibroblastų kultūroje, kurioje nustatytas *MED13L* variantas heterozigotinėje būklėje, senstančios ląstelės sudarė 76,8 proc. (4.32 pav. C, D).



**4.32 pav.** Tiriamojo #10 bei kontrolinio asmens redaguotos ir neredaguotos fibroblastų kultūrų senėjimo tyrimas invertuotu mikroskopu (*Nikon*, Japonija; padidinimas 40x, mastelis 100  $\mu\text{m}$ ): **A**) Reprezentatyvi ankstyvojo pasažo kontrolinio asmens neredaguotos fibroblastų kultūros nuotrauka; **B**) Reprezentatyvi ankstyvojo pasažo tiriamojo #10 neredaguotos fibroblastų kultūros nuotrauka; **C**) Reprezentatyvi kontrolinio asmens redaguotos vėlyvojo pasažo fibroblastų kultūros, kurioje nenustatyta *MED13L* variantų, nuotrauka. **D**) Reprezentatyvi kontrolinio asmens redaguotos vėlyvojo pasažo fibroblastų kultūros, kurioje nustatytas *MED13L* variantas heterozigotinėje būklėje, nuotrauka. Senstančios ląstelės yra pakitusios morfologijos ir nusidažiusios žalsvai-mėlynai

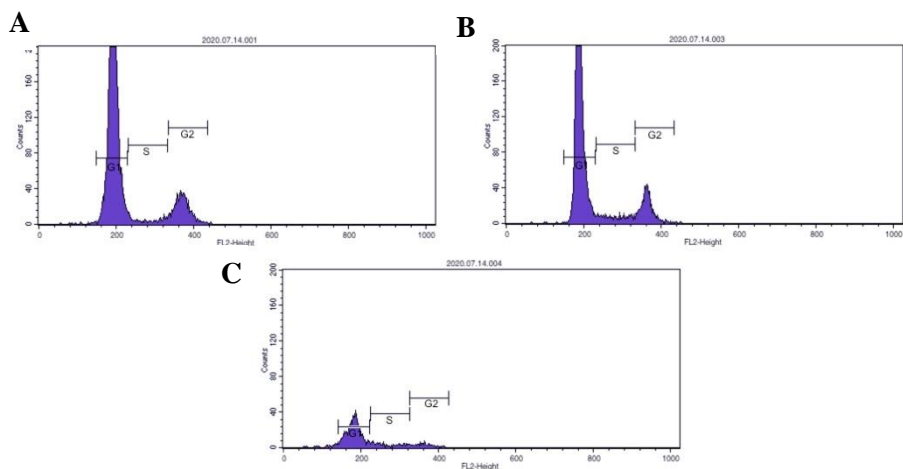
Ilgalaikis ląstelių kultivavimas turi įtakos ląstelių gyvybingumui, kadangi susidaro dideli ląstelių agregatai, kurie slopina ląstelių proliferaciją ir spartina apoptozės procesą [287]. Prieš genomo redagavimo eksperimentą kontrolinio asmens fibroblastų kultūros gyvybingumas siekė apie 96 proc., o tiriamojo #10 buvo kiek mažesnis – 89 proc. Įvertinus redaguotų vėlyvųjų pasażų fibroblastų kultūrų gyvybingumą, nustatyta, kad ilgalaikio kultivavimo metu ląstelių gyvybingumas sumažėjo visose tirtose ląstelių kultūrose. Be to, nustatyta, kad redaguotoje fibroblastų kultūroje, turinčioje heterozigotinį *MEDI3L* variantą, lemiantį haplonepakankamumą, yra apie du kartus daugiau žuvusių ląstelių nei fibroblastų kultūroje be *MEDI3L* varianto. Todėl tikėtina, kad tiek ilgalaikis kultivavimas, tiek heterozigotiniai *MEDI3L* geno variantai lemia fibroblastų gyvybingumo mažėjimą.

Galiausiai, vertinti ankstyvųjų pasażų tiriamojo #10 ir kontrolinio asmens neredaguotų ir vėlyvųjų pasażų redaguotų fibroblastų ciklo pokyčiai. Tiek tiriamojo #10, tiek kontrolinio asmens fibroblastai ankstyvuosiuose pasażuose buvo gyvybingi ir efektyviai dauginosi, todėl tirtas 10 000 ląstelių ciklas. Nustatyta, kad neredaguotoje kontrolinėje kultūroje dauguma ląstelių (82 proc.) efektyviai augo ir vykdė medžiagų apykaitą, nes buvo G1 (angl. *growth 1*) fazėje. Dalis ląstelių (3 proc.) buvo S (angl. *synthesis*) fazėje, kurioje dvigubėjo ląstelių genetinė medžiaga, o likusi dalis ląstelių (15 proc.) aptikta G2 (angl. *growth 2*) fazėje (4.33 pav. A). Panašus nustatytas tiriamojo #10 fibroblastų pasiskirstymas pagal fazes: G1 – 83 proc., S – 5 proc., G2 – 12 proc. (4.33 pav. B). Dėl sumažėjusios ląstelių proliferacijos ir gyvybingumo vėlyvuosiuose pasażuose redaguotoje fibroblastų kultūroje, kurioje nustatytas *MEDI3L* variantas heterozigotinėje būklėje<sup>17</sup>, įvertintas 2 235 ląstelių ciklas. Nustatyta, kad 79 proc. ląstelių buvo G1 fazėje, 11 proc. ląstelių – S fazėje, o 10 proc. – G2 fazėje (4.33 pav. C). Kadangi nenustatyta reikšmingo pokyčio tarp kontrolinės ir tiriamojo #10 ankstyvųjų pasażų fibroblastų kultūrų, tikėtina, kad heterozigotiniai *MEDI3L* variantai neturi reikšmingo įtakos ląstelės ciklui.

---

<sup>17</sup> Kitų kultūrų ląstelių skaičius buvo per mažas ląstelės ciklui įvertinti.





**4.33 pav.** Ląstelės ciklo analizės rezultatai. **A)** Kontrolinio asmens neredaguotos ankstyvojo pasažo fibroblastų kultūros; **B)** Tiriamojo #10 ankstyvojo pasažo fibroblastų kultūros; **C)** Kontrolinio asmens redaguotos vėlyvojo pasažo fibroblastų kultūros

Remiantis klinikiniais tiriamojo #10 požymiais ir atliktų molekulinį bei funkcinių *MED13L* geno tyrimų rezultatais, nustatyta NG\_023366.1(NM\_015335.5):c.(310+1\_311-1)\_(1280+1\_1281-1)del delecija yra patogeninė ir nulemia *MED13L* haplonepakankamumo sindromą. Sėkmingai redagavus kontrolinio asmens fibroblastų kultūrą, įvertinta šio KSP funkcinė reikšmė ląstelės lygmeniu.

#### 4.3. Rezultatų apibendrinimas ir praktinės rekomendacijos

Šio mokslinio darbo tikslas – molekuliniais ir funkciniais iRNR tyrimo metodais įvertinti DNR sekos sukirpimo vietos variantų ir kopijų skaičiaus pokyčių patogeniškumą bei nustatyti galimą reikšmę IN / IA etiopatogenezei. Siekiant užsibrėžto darbo tikslo, nustatytos septynių sukirpimo variantų (#1–#7) ir trijų KSP (#8–#10) pasekmės iRNR lygmeniu bei įvertintas jų patogeniškumas. Taip pat atliktas šių genetinių pokyčių klinikinis, molekulinis ir / ar funkcinis apibūdinimas.

*ARID1B*, *GLI3*, *SLC9A6*, *CHD7* ir *TGFBR2* sukirpimo vietos variantai yra neaprašyti literatūroje ir nėra registruojami *ClinVar* [8], *HGMD* [34], *dbSNP* [92], *dbVar* [93], 1 000 genomų projekto [6] bei *gnomAD* [96] duomenų bazėse. Nors kitų dviejų *BLM* ir *CAPN3* genų sukirpimo variantai yra nustatomi skirtingose pasaulio populiacijose, *ClinVar* [8] duomenimis, iki šiol jų patogeniškumas taip pat nebuvo patvirtintas.

Sukirpimo variantai gali turėti skirtingą poveikį pre-iRNR [48]. Iširti *ARID1B*, *GLI3*, *SLC9A6*, *BLM* ir *CAPN3* genų sukirpimo variantai lėmė vieno egzono iškritą. Sukirpimo variantas *TGFBR2* gene lėmė paskutinio egzono ir 3'UTR iškritą. Sukirpimo variantai 5' donorinėje vietoje (*ARID1B* c.5025+2T>C, *GLI3* c.473+3A>T, *SLC9A6* c.899+1G>A) lėmė egzono, esančio prieš variantą, iškritą, o sukirpimo variantai 3' akceptorinėje vietoje (*TGFBR2* c.1600-2A>G, *BLM* c.2308-2A>G) lėmė egzono už varianto iškritą. Mokslinės literatūros duomenimis, DNR sekos variantai, kurie yra nutolę nuo konservatyvių sukirpimo vietų, dažniau lemia introno ar jo dalies intarpą [36, 48]. Visgi iširtas c.1746-20C>G sukirpimo variantas *CAPN3* gene lėmė 14-ojo egzono iškritą. Kitas neįprastas rezultatas – sukirpimo vietos c.5535-1G>A varianto *CHD7* gene aktyvuojama kriptinė alternatyvi 3' akceptorinė sukirpimo vieta, kuri nuo pirminės vietos yra nutolusi tik vienu nukleotidu (4.6 lentelė). Iki šiol nėra žinomas tikslus mechanizmas, koku būdu įvykus pokyčiui DNR sekoje pasirenkama alternatyvi kirpimo vieta [50, 52, 55]. Manoma, kad tam gali turėti įtakos genomo vieta, egzoninė sukirpimo proceso stipriklio / slopiklio vieta ar net pre-iRNR antrinės struktūros [48].

Atlikus išsamią *in silico* analizę, nustatyta, kad *ARID1B* c.5025+2T>C, *GLI3* c.473+3A>T, *SLC9A6* c.899+1G>A, *CHD7* c.5535-1G>A ir *CAPN3* c.1746-20C>G variantai lėmė skaitymo rėmelio poslinkį ir priešlaikinio baigmės kodono susidarymą (4.6 lentelė). Baltymų lygmeniu šie pokyčiai lemia baltymo sutrumpėjimą (angl. *Protein truncation*) arba haplonepakankamumą dėl NMD (angl. *Nonsense-mediated decay*) inicijuojamo iRNR skaidymo. *Posttranskripciniu* lygmeniu veikiant šiam mechanizmui, specifinis egzonų jungties baltymų kompleksas atpažįsta priešlaikinį baigmės kodoną, jei jis yra apie 200 nt atstumu nutolęs nuo pradžios kodono ir maždaug 50–55 nt atstumu nutolęs nuo paskutinės egzonų poros jungties vietos. Tokiu atveju šis kompleksas, sąveikaujant su papildomais specifiniais baltymais, inicijuoja iRNR, turinčios priešlaikinį baigmės kodoną, skaidymo procesą [40]. Pažymėtina, kad dalis transkriptų gali išvengti šio mechanizmo, tokiu atveju nustatomas mažesnis kiekis pakitusių transkriptų. Tais atvejais, kai NMD proceso metu atpažįstami ir suardomi visi pakitę transkriptai, pagrindinis molekulinės patogenezės mechanizmas yra baltymo haplonepakankamumas. Kita vertus, jei NMD visiškai išvengiama, sutrumpėjusio transkripto raiška gali turėti neigiamos įtakos baltymo funkcijai. *ARID1B* sukirpimo varianto nulemtas priešlaikinis baigmės kodonas nėra nutolęs 50–55 nt atstumu nuo paskutinės egzonų poros jungties vietos, todėl NMD šio baigmės kodono negali atpažinti. Keturių kitų genų (*GLI3*, *SLC9A6*, *CHD7*, *CAPN3*) sukirpimo variantų nulemti priešlaikiniai baigmės kodonai atitinka NMD atpažinimo sąlygas, todėl iRNR

skaidymas galėtų būti inicijuojamas. Sangerio sekoskaitos tyrimo metu šių genų transkriptai buvo ištirti, todėl tikėtina, kad NMD proceso dalinai arba visiškai išvengta. Remiantis šiais rezultatais, tikėtina, kad *ARID1B*, *GLI3*, *SLC9A6*, *CHD7* ir *CAPN3* genų variantai lemia baltymo sutrumpėjimą, o ne haplonepakankamumą. Kiti tirti *TGFBR2* ir *BLM* sukirpimo variantai nesudaro skaitymo rėmelio poslinkio (4.6 lentelė). *BLM* c.2308-2A>G sukirpimo vietos variantas lemia dalies funkciškai svarbaus baltymo domeno praradimą, todėl *BLM* funkcija gali būti sutrikusi. Sukirpimo vietos c.1600-2A>G variantas *TGFBR2* gene lemia paskutinio egzono ir 3'UTR praradimą, todėl tikėtina, kad *TGFBR2* haplonepakankamumas nulėmė ligos etiopatogenezę.

Analizuojant skirtingų, atsitiktine tvarka parinktų, evoliuciškai nutolusių rūšių *ARID1B*, *GLI3*, *SLC9A6*, *CHD7*, *TGFBR2*, *BLM* ir *CAPN3* aminorūgščių sekas, nustatyta, kad pokyčiai įvyko ypač konservatyviose srityse, todėl šios sritys gali būti svarbios funkcinio požiūriu (4.34 pav., 4.21 pav. C).

**4.6 lentelė.** Nekoduojančios DNR sekos sukirpimo variantų molekulinį ir / ar funkcinių tyrimų apibendrinti rezultatai

Tiriamasis	Sukirpimo variantas	Genas	Genotipas	Varianto kilmė	Nustatytas poveikis iRNR lygmeniu	<i>In silico</i> prognozuojamas poveikis baltymo lygmeniu	Patvirtinta diagnozė	Varianto patogeniškumas
#1	NM_020732.3: c.4986+2T>C	ARID1B MIM#614556	Heterozigotinis	<i>De novo</i>	19 egzono iškrita	NP_065783.3: p.(Thr1633Valfs*11)	Coffino–Siris sindromas 1 (MIM #135900)	Patogeninis
#2	NM_000168.6: c.473+3A>T	GLI3 MIM#165240	Heterozigotinis	Paveldėtas trijose kartose	4 egzono iškrita	NP_000159.3: p.(His123Argfs*57)	Greigo cefalopolisindaktilija (MIM #175700)	Patogeninis
#3	NM_001042537.1: c.899+1G>A	SLC9A6 MIM#300231	Hemizigotinis	Paveldėtas iš mamos	6 egzono iškrita	NP_001036002.1: p.(Val264Alafs*3)	Christiansono tipo sindrominė IN (MIM #300243)	Patogeninis
#4	NM_017780.4: c.5535-1G>A	CHD7 MIM#608892	Heterozigotinis	<i>De novo</i>	Alternatyvi akceptorinė sukirpimo vieta	NP_060250.2: p.(Gly1846Gluufs*23)	CHARGE sindromas (MIM #214800)	Patogeninis
#5	NM_001024847.2: c.1600-2A>G	TGFBR2 MIM#190182	Heterozigotinis	ND	8 egzono ir 3'UTR iškrita	TGFBR2 (UniProtKB #P37173) haplonepakankamumas	Loeyso–Dietzo sindromas 2 (MIM #610168)	Patogeninis
#6	NM_000057.2: c.2308-2A>G, rs1248548542	BLM MIM#210900	Heterozigotinis	Paveldėtas iš tėvo	11 egzono iškrita	NP_000048.1: p.(Ile700_Gln802del)	-	Galimai patogeninis
#7	NM_000070.3: c.1746-20C>G, rs201892814	CAPN3 MIM#114240	Heterozigotinis	Paveldėtas iš mamos	14 egzono iškrita	NP_000061.1: p.(Glu582Aspfs*62)	Galūnių ir juosmens raumenų distrofija 2 (MIM #253600)	Patogeninis

<b>Tiriamasis #1</b>	FPPGSVEASQPVLKQRRKITSKDIVVWISRTFSRVL-----	1643
<i>Homo sapiens</i>	FPPGSVEASQPVLKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1707
<i>Mus musculus</i>	FPPGSVEASQPVILKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1662
<i>Rattus norvegicus</i>	FPPGSVEASQPVLKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1165
<i>Monodelphis domestica</i>	FPPGSVEASQPVLKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1530
<i>Gallus gallus</i>	FPPGSVEASQPVLKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	959
<i>Sus scrofa</i>	FPPGSVEASQPVLKQRRKITSKDIVTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1737
<i>Danio rerio</i>	YPLDSVEATQPQLKRRRLTSKDIGTPEAWRVMSLKSGLLAESTWALDTINILLYDDST	1529
<i>Latimeria chalumnae</i>	FPPGSVEATPPVLKARRKITSKDIGAPEAWRVMSLKSGLLAESTWALDTINILLYDDSS	950
	:* .****: * * * * :***** . : . *	
<b>Tiriamasis #2</b>	EPHYRLPPYLVALRIRTCPSLGSPTHTPLLLPSLPSALHIPTLIPTWTISAPCTAAHRSP-----	179
<i>Homo sapiens</i>	EPHYHPPHLFPFAPHPP-VPIDARHHEGRYHYDPSPIPLHMTSALSSSPTYDLPFFIRISPHRNPTAAS	186
<i>Mus musculus</i>	EPHYHPPHLFPFAPHPP-VPIDARHHEGRYHYDPSPIPLHMTSALSSSPTYDLPFFIRISPHRNPTAAS	186
<i>Rattus norvegicus</i>	EPHYHPPHLFPFAPHPP-VPIDARHHEGRYHYDPSPIPLHMTSALSSSPTYDLPFFIRISPHRNPTAAS	186
<i>Gallus gallus</i>	DPHYHPPHLFPFAPHPP-VPIDARHHEGRYHYEPSPIPLHMTSALSSSPTYDLPFFIRISPHRNPTAAS	186
<i>Sus scrofa</i>	EPHYHPPHLFPFAPHPP-VPIDARHHEGRYHYDPSPIPLHMTSALSSSPTYDLPFFIRISPHRNPTAAS	186
<i>Danio rerio</i>	DPHYTAPQFFPAPHPP-VPIDRRHTQGRYIYEPSVPSVLSHMPALAAASPTFSDISLRISPHRNPTAAS	187
<i>Latimeria chalumnae</i>	DPHYHPPHLFPFAPHPP-VPIEARHHEGRYHYEPSPIQLHVSATIGNPAYDIPFFIRISPHRNPTAAS	79
	:*** * : * : * * ** ** : : : * : ** *	
<b>Tiriamasis #3</b>	VTGQLAGDFYFTDCLLFGAIVSATDPAQ-----	266
<i>Homo sapiens</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	246
<i>Mus musculus</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	299
<i>Rattus norvegicus</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	318
<i>Monodelphis domestica</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	299
<i>Gallus gallus</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	407
<i>Sus scrofa</i>	VTGQLAGDFYFTDCLLFGAIVSATDPVTVLAIFHELQVDVLYALLFGESVLNDAVAIVL	246
<i>Danio rerio</i>	IIGQLGGDFYFTDCLLFGAIVSATDPVTVLAIFNELKVDVLYALMFGESVLNDAVAIVL	280
<i>Latimeria chalumnae</i>	QISQLGGDFYFTDCLLFGAIVSATDPVTVLAIFNELQVDVLYALLFGESVLNDAVAIVL	301
	.**.*:*****:*****:*****.	
<b>Tiriamasis #4</b>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGENLEKMKTONINQPEHRSKMK-----	1874
<i>Homo sapiens</i>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RTPFKDEIDEFA	1872
<i>Mus musculus</i>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RAPFKDEIDEFA	1862
<i>Rattus norvegicus</i>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RAPFKDEIDEFA	1861
<i>Monodelphis domestica</i>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RTPFKDEIDEFA	1872
<i>Gallus gallus</i>	EKYNMRADSTLFCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RTPFKDEIDEFA	1873
<i>Sus scrofa</i>	EKYNMRADPALCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RTPFKDEIDEFA	1865
<i>Latimeria chalumnae</i>	EKYNVRADPTLFCFLERVGMPPDAKAIAAEQRTDMLADGGDGEFDREDEDEPEYKPT--RTPFKDEMDEFA	1856
	*****:** :*****:*****:*****:** : : : : * : *	
<b>Tiriamasis #5</b>	PSFWLNHQ-----	533
<i>Homo sapiens</i>	PSFWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HLDRLSGRSCSCEEKIPEDGSLNNTK--	592
<i>Mus musculus</i>	PSFWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HPERLSGRSCSCEEKIPEDGSLNNTK--	592
<i>Rattus norvegicus</i>	PSFWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HLDRLSGRSCSCEEKIPEDGSLNNTK--	567
<i>Monodelphis domestica</i>	PNTWLSHRGIQMVCETLIECWHDHDPPEARLTAQCVAERFDQLG-DLDRLSGRSCSCEEKIPEDCSLSPTK--	564
<i>Gallus gallus</i>	PSSWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HHDKLSGRSCSCEEKIPEDGSLNNTK--	671
<i>Sus scrofa</i>	PSSWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HLDRLSGRSCSCEEKIPEDGSLNNTK--	564
<i>Danio rerio</i>	PSSWMKHQGVAAVCATINECWHDHDPPEARLTAQCVAERFNEMDDDLKSLSCSCEEKIPEDCAVSVSDDK	556
<i>Latimeria chalumnae</i>	SSWLNHQIQMVCETLIECWHDHDPPEARLTAQCVAERFSELE-HLDRLSGRSCSCEEKIPEDCESDGGK--	585
	. * : * :	
<b>Tiriamasis #6</b>	SEATNIYLQLSKKDP I IKLLYVTPFK-----	W 770
<i>Homo sapiens</i>	SEATNIYLQLSKKDP I IKLLYVTPFKICASNRLISTLENLYERKLLARFVIDEACHVSQW	803
<i>Mus musculus</i>	SEAAIYLQLSKKDP I IKLLYVTPFKVCASNRLISTLENLYERKLLARFVIDEACHVSQW	814
<i>Rattus norvegicus</i>	SEAAIYLQLSKKDP I IKLLYVTPFKVCASNRLISTLENLYERKLLARFVIDEACHVSQW	794
<i>Monodelphis domestica</i>	SEAAIYLQLSKKDP I IKLLYVTPFKVCASNRLISTLENLYERKLLARFVIDEACHVSQW	837
<i>Gallus gallus</i>	ADASKTYMQLSKKDP I IKLLYVTPFKVCASNRLISALENLYDRKLLARFVIDEACHVSQW	771
<i>Sus scrofa</i>	SEATS IYLQLSKKDP I IKLLYVTPFKVCASNRLISTLENLYERKLLARFVIDEACHVSQW	815
<i>Danio rerio</i>	SEAARIYMLSRKDP I IKLLYVTPFKVCASGRMISALQNLVYERKLLARFVIDEACHVSQW	814
<i>Latimeria chalumnae</i>	AEAGSIYMLSKKDP I IKLLYVTPFKVCASNRLISTLENLYERKMLTCFVIDEACHVSQW	832
	:* * : * : * : * : * : * : * : * : *	

**4.34 pav.** ARID1B (tiriamasis #1), GLI3 (tiriamasis #2), SLC9A6 (tiriamasis #3), CHD7 (tiriamasis #4), TGFBR2 (tiriamasis #5), BLM (tiriamasis #6) baltymo sekos konservatyvumo analizė lyginant skirtingas rūšis. Ištirtų sukirpimo variantų nulemtas aminorūgščių skaitymo rėmelio poslinkis pažymėtas raudona spalva.

Konservatyvios sekos vietos pažymėtos žvaigždute (\*), konservatyvūs variantai – dvitaškiu (:), pusiau konservatyvūs variantai – tašku (.), nepažymėtos vietos reiškia nekonservatyvias sekas vietas<sup>18</sup>. Konservatyvumo analizė atlikta *Clustal Omega* [252] bioinformaciniu įrankiu

Atsižvelgiant į *ARID1B* c.5025+2T>C, *GLI3* c.473+3A>T, *SLC9A6* c.899+1G>A, *CHD7* c.5535-1G>A, *CAPN3* c.1746-20C>G ir *TGFBR2* c.1600-2A>G lemiamus pre-iRNR sukirpimo proceso, prognozuojamus baltymo struktūros pokyčius, konservatyvumo analizės rezultatus ir tiriamiesiems pasireiškiančius klinikinius požymius, šie variantai pagal ACMG [20] rekomendacijas klasifikuotini kaip patogeniniai (4.6 lentelė). *BLM* nustatyto c.2308-2A>G varianto nepakanka įtarto autosominio recesyvaus Bloomo sindromo patogenezėi paaiškinti. Atlikus SCM ir *BLM* geno raiškos analizę, nenustatyta jokių kitų *BLM* geno pokyčių, todėl hipotezė dėl sudėtinio heterozigotinio varianto, lemiančio Bloomo sindromą, nepatvirtinta. Nors c.2308-2A>G variantas nelemia tiriamojo #6 klinikinio fenotipo, atsižvelgiant į varianto dažnį (MAF < 0,01) ir nustatytą galimą funkcinį poveikį konservatyvioje baltymo srityje, pagal ACMG [20] rekomendacijas šis DNR sekos variantas klasifikuotinas kaip galimai patogeninis (4.6 lentelė).

Šiame moksliniame darbe ištirta *de novo* kilmės 13q31.3 duplikacija, 4q21 delecija bei vidugeninė *MED13L* delecija. Atlikus šių KSP klinikinį, molekulinį ir / ar funkcinį apibūdinimą, ne tik patvirtintas šių genetinių pokyčių patogeniškumas, bet ir gauti unikalūs, iki šiol literatūroje neaprašyti duomenys.

Iki šiol buvo manoma, kad *GPC5* kartu su *MIR17HG* galimai lemia asmenų su persidengiančia duplikacija 13q31.3 fenotipus [257, 258]. Tačiau ši hipotezė buvo paneigta, patvirtinus, kad *GPC5* geno raiška ir kopijų skaičius tiriamojo #8 yra nepakitęs, o pasireiškiančius klinikinius požymius lemia greta *GPC5* geno esančio *MIR17HG* geno duplikacija. Šiame tyrime išanalizuotais duomenimis, galimai apibūdintas naujas *MIR17HG* geno duplikacijos sindromas.

Literatūros duomenimis, 4q21 delecijos sindromą lemia penkių *HNRNPD*, *HNRNPD*, *ENOPH1*, *RASGEF1B* ir *PRKG2* genų delecija. vLGH metodu ištyrus tiriamojo #9 gDNR, nustatyta viena mažiausių iki šiol

---

<sup>18</sup> Nors tirtos baltymų sekos yra konservatyvios, dėl sukirpimų variantų nulemta skaitymo rėmelio poslinkio *Clustal Omega* bioinformacinis įrankis šias sritis žymi pusiau konservatyviomis ar nekonservatyviomis.

aprašytų šios srities delecijų, kuri apėmė *HNRNPD*, *HNRNPD*L ir *ENOPH1* geną. Nors, tiriamajam #9 nustatytas nepakitęs *RASGEF1B* ir *PRKG2* kopijų skaičius genome, lyginamuoju kPGR metodu ištyrus gretimų dviejų genų raišką, nustatyta, kad *RASGEF1B* raiška yra nepakitusi, o *PRKG2* – sumažėjusi. Tai suteikė galimybę sumažinti galimų kandidatinių genų skaičių iki keturių: *HNRNPD*, *HNRNPD*L, *ENOPH1*, *PRKG2*.

Tiriamajam #10 VNP-LGH tyrimu 12q24.21 srityje nustatyta viduginė *MED13L* delecija. Kiekybinės PGR ir Sangerio sekoskaitos metodu ištyrus tiriamojo #10 kDNR, patikslintas delecijos dydis – NG\_023366.1(NM\_015335.5):c.(310+1\_311-1)\_(1280+1\_1281-1)del. *In silico* prognozuojama, kad šis pokytis lemia skaitymo rėmelio poslinkį ir priešlaikinio baigmės kodono susidarymą (NP\_056150.1:p.(Val104Glyfs\*5)). Susidaręs priešlaikinis baigmės kodonas yra pakankamu atstumu nutolęs nuo transkripto pradžios ir paskutinės egzono poros jungties vietos, todėl ši sritis galėtų būti atpažinta NMD mechanizmo ir inicijuotas iRNR suardymas. Tačiau dėl tos pačios priežasties kaip sukirpimo variantų atveju tikėtina, iRNR suardymo iš dalies arba visiškai išvengta. Nors kPGR ir Sangerio sekoskaitos tyrimų užteko pagrįsti nustatyto KSP patogeniškumą ir jo ryšį su fenotipu, molekulinis patogenezės mechanizmas iki šiol nėra iki galo žinomas. Siekiant įvertinti šio KSP pasekmes ląstelės lygmeniu, atliktas neturinčio ligos požymių kontrolinio asmens fibroblastų kultūros genomo redagavimo tyrimas, kuriame *MED13L* geno kirpimo efektyvumas nustatytas apie 44 proc. Sangerio sekoskaitos metodu ištyrus 11-ka redaguotų fibroblastų kultūrų, keturiuose kultūrose nustatyti heterozigotiniai variantai *MED13L* gene. Jokių homozigotinių *MED13L* variantų nenustatyta. Kultivuojant redaguotas fibroblastų kultūras, buvo pastebėta, kad dalies kultūrų morfologija pasikeitė akivaizdžiai anksčiau, taip pat sulėtėjo ląstelių proliferacija ir sumažėjo gyvybingumas. Dėl šios priežasties nepavyko padauginti šių ląstelių iki funkciniais tyrimams reikiamo skaičiaus. Galimai šiose fibroblastų kultūrose buvo homozigotiniai *MED13L* variantai, kurie yra letalūs ląstelėi. Siekiant patvirtinti iškeltą prielaidą, reikėtų taikyti kitokią tyrimo strategiją, pavyzdžiui, iš karto po genomo redagavimo iš vienos ląstelės išgryninti DNR, ją pagausinti WGA (angl. *whole genome amplification*) metodu ir iširti taikant Sangerio sekoskaitos ir / ar NKS technologijas. Viso genomo ir / ar transkriptomo tyrimas NKS metodu taip pat rekomenduotinas, siekiant tiksliau įvertinti atlikto genomo redagavimo pasekmes, kadangi CRISPR-Cas9 technologijai yra būdingas kirpimas šalutinėse vietose (angl. *Off-target*). Nepaisant to, šiame tyrime sėkmingai įvaldytas genomo redagavimo metodas fundamentinių tyrimų tikslais. Gauti

unikalūs tyrimo rezultatai, kurie papildė žinias apie tiriamo KSP funkcinę reikšmę.

Šioje disertacijoje, taikant kPGR ir Sangerio sekoskaitos tyrimo technologijas bei redagavus ląstelių kultūros genomą CRISPR-Cas9 metodu, įvertinti dešimties koduojančios ir nekoduojančios DNR sekos sukirpimo vietos variantų ir KSP patogeniškumas iRNR lygmeniu. Atlikus išsamią genotipo ir fenotipo ryšio analizę, patvirtinta klinikinė diagnozė devyniems tiriamiesiems / šeimos nariams, turintiems IN, PRA ir / arba vieną ar daugiau IA. Šie molekuliniai ir funkciniai iRNR tyrimai suteikė galimybę ne tik suprasti IN / IA etiopatogenezės mechanizmus, bet ir patobulinti molekulinės diagnostikos strategiją (žr. 2.3.4 skyrių; 2.11 pav.). Atkreiptinas dėmesys, kad parinkta žmogaus genomo koduojančios ir nekoduojančios DNR sekos tyrimo strategija yra tinkama ne tik IN / IA etiopatogenezei tirti. Taikant panašią tyrimo strategiją, būtų galima įvertinti genetinių pokyčių patogenezės mechanizmus ir patvirtinti diagnozę kone kiekvienos heterogeninės prigimties paveldimos ligos kontekste. Tiksliai nustatyta diagnozė yra svarbi tiek konsultuojančiam gydytojui, tiek tiriamiesiems ir jų šeimoms, kadangi tokiu būdu suteikiama galimybė parinkti tinkamą gydymą, pateikti išankstines prognozes bei įvertinti pasikartojimo riziką kitų nėštumų metu.

Haruki Murakami (g. 1949) yra pasakęs: „*Tęstinumas – tai pastovaus ritmo išlaikymas. Tai būtina ilgalaikio darbo sąlyga*“, tad, siekiant suprasti išanalizuotų patogeninių DNR sekos variantų ir KSP įtaką baltymo struktūrai ir funkcijai proteomo, metabolomo ar net interaktomo lygmeniu bei papildyti žinias apie ryšį tarp genotipo ir fenotipo, tikslinga tęsti šiuos tyrimus. Tokių tyrimų pavyzdžiai yra baltymų analizė *Western blot* metodu ar modeliųjų sistemų eksperimentai CRISPR-Cas9 metodu. Šie tyrimai ne tik patvirtintų *in silico* analizės rezultatus, bet ir padėtų suprasti NMD inicijuojamą iRNR skaidymo procesą, baltymų tarpusavio sąveikas bei kitus svarbius molekulinis procesus. Taip pat atsižvelgiant į tai, kad intensyviausia analizuoto *SLC9A6* geno raiška yra nervinėse ląstelėse, rekomenduotina fibroblastų transformacija į nervinių ląstelių liniją, po kurios tikslinga atlikti šių ląstelių funkcinis tyrimus. Kadangi *BLM* sukirpimo variantas heterozigotinėje būklėje negali paaiškinti Bloomo sindromo patogenezės, reikalingi detalesni tiriamojo #6 tyrimai. *HGMD* [34] duomenimis, *BLM* patogeniniai variantai lemia ne tik Bloomo sindromą, bet taip pat gali būti krūties ar kasos vėžio priežastis. Dėl šios priežasties rekomenduotina toliau stebėti tiriamojo #6 klinikinius požymius, kurie gali pasireikšti vėlesniame amžiuje. Tiriamajam #6 pasireiškiantys klinikiniai požymiai būdingi taip pat ligoms (pvz., Fanconi anemija), kurias lemia kitų genų patogeniniai variantai. Atsižvelgiant į tai, kad tiriamajam #6, naudojant NKS technologiją, iki šiol



ištirta tik specifinių genų grupė, tikslinga atlikti VGS, kadangi tiriamojo genų grupėje galėjo nebūti tikslinio geno. Galiausiai, siekiant patvirtinti su 4q21 delecijos sindromu siejamus keturis kandidatinius genus, rekomenduotina pakartoti tiriamojo #9 kPGR, padidinus referentinės biologinės grupės imtį. Taip pat tikslinga atlikti funkcinius modelinių sistemų tyrimus, siekiant suprasti molekulinį mechanizmą, dėl kurio sumažėja *PRKG2* raiška.

## IŠVADOS

1. Taikant bioinformacinius žmogaus genomo sekos analizės įrankius, pagal nustatytus atrankos kriterijus atrinkti septynių *ARID1B*, *GLI3*, *SLC9A6*, *CHD7*, *TGFBR2*, *BLM* ir *CAPN3* genų sukirpimo variantai ir trijų 13q31.3, 4q21.22, 12q24.21 genetinių sričių kopijų skaičiaus pokyčiai, kurių klinikinė reikšmė buvo nežinoma ar nepatvirtinta.
2. Siekiant įvertinti atrinktų DNR sekos pokyčių patogeniškumą informacinės RNR lygmeniu, taikant kPGR ir Sangerio sekoskaitos metodus iširti dešimties tiriamų asmenų kopijinės DNR mėginiai.
3. Nežinomos ar nepatvirtintos klinikinės reikšmės DNR sekos pokyčiai klasifikuoti kaip patogeniniai (9) arba galimai patogeniniai (1).
4. *In silico* *ARID1B* c.5025+2T>C, *GLI3* c.473+3A>T, *SLC9A6* c.899+1G>A, *CHD7* c.5535-1G>A, *CAPN3* c.1746-20C>G ir *MEDI3L* NG\_023366.1(NM\_015335.5):c.(310+1\_311-1)\_(1280+1\_1281-1)del lemia skaitymo rėmelio poslinkį ir priešlaikinio baigmės kodono susidarymą, *BLM* c.2308-2A>G sukirpimo variantas lemia dalies funkciškai svarbaus baltymo domeno praradimą, *TGFBR2* c.1600-2A>G sukirpimo variantas lemia koduojamo baltymo haplonepakankumą.
5. CRISPR-Cas9 metodu redagavus kontrolinio asmens fibroblastų kultūrą 44 proc. efektyvumu, nustatyta, kad patogeniniai *MEDI3L* variantai heterozigotinėje būklėje mažina fibroblastų kultūros gyvybingumą, spartina senėjimo procesą ir slopina ląstelės ciklą valdančių genų raišką.
6. Atliktas *ARID1B*, *GLI3*, *SLC9A6*, *CHD7*, *TGFBR2*, *BLM* ir *CAPN3* genų nekoduojančios DNR sekos sukirpimo variantų bei 13q31.3 duplikacijos, 4q21 delecijos ir vidugeninės *MEDI3L* delecijos molekulinis, funkcinis ir klinikinis apibūdinimas, kuris padėjo suprasti sudėtingą ryšį tarp genotipo ir fenotipo:
  - 6.1. *ARID1B*, *GLI3*, *SLC9A6* sukirpimo variantai 5' donorinėje vietoje lemia atitinkamai pirmojo tipo Coffino–Siris sindromą, Greigo cefalopolisindaktiliją, Christiansono tipo sindrominę intelektinę negalią. *CHD7* ir *TGFBR2* sukirpimo variantai 3' akceptorinėje vietoje lemia atitinkamai CHARGE sindromą ir antrojo tipo Loeyso–Dietzo sindromą. *CAPN3* c.1746-20C>G sukirpimo variantas sudėtinėje heterozigotinėje būklėje lemia antrojo tipo galūnių ir juosmens raumenų distrofiją;
  - 6.2. *BLM* 3' akceptorinės sukirpimo vietos c.2308-2A>G varianto heterozigotinėje būklėje nepakanka paaiškinti autosominiu recesyviu būdu paveldimo Bloomo sindromo patogenezės;

- 6.3. Tryliktos chromosomos ilgojo peties duplikacija galimai lemia naują *MIR17HG* duplikacijos sindromą, kuris pasižymi klinikiniais požymiais, iš dalies priešingais antrojo tipo Feingoldo sindromui, nulemtam šios srities delecijos;
- 6.4. Remiantis nepakitusiu *RASGEF1B* kopijų skaičiumi ir nepakitusia šio geno raiška tiriamojo genome, susiaurinta anksčiau su 4q21delecijos sindromu siejama genetinė sritis, kuri apima keturis genus: *HNRNPD*, *HNRNPD*, *ENOPH1*, *PRKG2*;
- 6.5. Dvyliktos chromosomos ilgojo peties vidugeninė *MED13L* delecija lemia *MED13L* haplonepakankamumo sindromą.

## LITERATŪROS SĄRAŠAS

1. Collins FS, Lander ES, Rogers J, Waterson RH. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431:931–45. doi:10.1038/nature03001.
2. Rehm HL, Fowler DM. Keeping up with the genomes: scaling genomic variant interpretation. *Genome Med*. 2020;12:5. doi:10.1186/s13073-019-0700-4.
3. Feingold EA, Good PJ, Guyer MS, Kamholz S, Liefer L, Wetterstrand K, et al. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* (80- ). 2004;306:636–40. doi:10.1126/science.1105136.
4. Tragante V, Moore JH, Asselbergs FW. The ENCODE project and perspectives on pathways. *Genet Epidemiol*. 2014;38:275–80. doi:10.1002/gepi.21802.
5. Gloss BS, Dinger ME. Realizing the significance of noncoding functionality in clinical genomics. *Exp Mol Med*. 2018;50:97. doi:10.1038/s12276-018-0087-0.
6. Altshuler DL, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061–73. doi:10.1038/nature09534.
7. Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65. doi:10.1038/nature11632.
8. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42:D980–5. doi:10.1093/nar/gkt1113.
9. ClinVar. ClinVar (NCBI). <https://www.ncbi.nlm.nih.gov/clinvar>. Accessed 27 Apr 2020.
10. Hagberg B, Kyllerman M. Epidemiology of mental retardation—A Swedish survey. *Brain Dev*. 1983;5:441–9. doi:10.1016/S0387-7604(83)80072-2.
11. Association AP, others. *Diagnostic and Statistical Manual of Mental Disorders*. 4th text revision ed. Washington, DC Am Psychiatr Assoc. 2000.
12. Moeschler JB, Shevell M. *Comprehensive Evaluation of the Child With Intellectual Disability or Global Developmental Delays*. *Pediatrics*. 2014;134:e903–18. doi:10.1542/peds.2014-1839.
13. Preikšaitienė E, Kasnauskienė J, Utkus A, Kučinskas V. Asmenų su intelektine negalia genetinio ištyrimo gairės. *Neurol Semin*. 2012; 16:283–

- 8.
14. Bell CJ, Dinwiddie DL, Miller NA, Hateley SL, Ganusova EE, Mudge J, et al. Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing. *Sci Transl Med*. 2011;3:65ra4-65ra4. doi:10.1126/scitranslmed.3001756.
15. Brooker RJ. *Genetics: Analysis & Principles*. 3 edition. McGraw-Hill; 2009.
16. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. *Nature*. 2006;444:444–54. doi:10.1038/nature05329.
17. Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. *Nat Rev Genet*. 2015;16:172–83. doi:10.1038/nrg3871.
18. Pearson CE, Edamura KN, Cleary JD. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet*. 2005;6:729–42. doi:10.1038/nrg1689.
19. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet Med*. 2008;10:294–300. doi:10.1097/GIM.0b013e31816b5cae.
20. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–23. doi:10.1038/gim.2015.30.
21. Karki R, Pandya D, Elston RC, Ferlini C. Defining “mutation” and “polymorphism” in the era of personal genomics. *BMC Med Genomics*. 2015;8:37. doi:10.1186/s12920-015-0115-z.
22. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med*. 2011;13:680–5. doi:10.1097/GIM.0b013e3182217a3a.
23. Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature*. 2010;464:713–20. doi:10.1038/nature08979.
24. Choudhuri S. Small noncoding RNAs: Biogenesis, function, and emerging significance in toxicology. *J Biochem Mol Toxicol*. 2010;24:195–216. doi:10.1002/jbt.20325.
25. Palazzo AF, Gregory TR. The Case for Junk DNA. *PLoS Genet*.

- 2014;10:e1004351. doi:10.1371/journal.pgen.1004351.
26. Green D, Dalmay T, Chapman T. Microguards and micromessengers of the genome. *Heredity* (Edinb). 2016;116:125–34. doi:10.1038/hdy.2015.84.
  27. Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518:337–43. doi:10.1038/nature13835.
  28. Guenther CA, Tasic B, Luo L, Bedell MA, Kingsley DM. A molecular basis for classic blond hair color in Europeans. *Nat Genet*. 2014;46:748–52. doi:10.1038/ng.2991.
  29. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet*. 2006;7:85–97. doi:10.1038/nrg1767.
  30. Tărlungeanu DC, Novarino G. Genomics in neurodevelopmental disorders: an avenue to personalized medicine. *Exp Mol Med*. 2018;50:100. doi:10.1038/s12276-018-0129-7.
  31. Bridges CB. The Bar “gene” a duplication. *Science* (80- ). 1936;83:210–1.
  32. Preikšaitienė E. Asmenų, turinčių intelektinę negalią, genomo struktūrinių pokyčių įtakos fenomui įvertinimas. Vilniaus universitetas; 2013.
  33. Rosenberg C. Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J Med Genet*. 2005;43:180–6. doi:10.1136/jmg.2005.032268.
  34. Krawczak M, Ball E V., Stenson P, Cooper DN. HGMD: The Human Gene Mutation Database. *Bioinformatics: Databases and Systems*. 2006;:99–104. <http://www.hgmd.cf.ac.uk/ac/index.php>. Accessed 3 Feb 2019.
  35. Hgmd.cf.ac.uk. HGMD® home page. 2015. <http://www.hgmd.cf.ac.uk/ac/index.php>. Accessed 27 Apr 2020.
  36. Kasnauskienė J. Viso žmogaus genomo analizės metodai. Vilniaus universitetas; 2014.
  37. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*. 2018;19:770–88. doi:10.1038/s41576-018-0059-1.
  38. Strachan T, Read A. *Human Molecular Genetics*. 4 edition. Garland Science; 2010.
  39. DeBoever C, Tanigawa Y, Lindholm ME, McInnes G, Lavertu A, Ingelsson E, et al. Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. *Nat Commun*. 2018;9:1612. doi:10.1038/s41467-018-03910-9.

40. Hug N, Longman D, Cáceres JF. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* 2016;44:1483–95. doi:10.1093/nar/gkw010.
41. Harper AR, Nayee S, Topol EJ. Protective alleles and modifier variants in human health and disease. *Nat Rev Genet.* 2015;16:689–701. doi:10.1038/nrg4017.
42. Gilbert W. Why genes in pieces? *Nature.* 1978;271:501. doi:10.1038/271501a0.
43. Park E, Pan Z, Zhang Z, Lin L, Xing Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. *Am J Hum Genet.* 2018;102:11–26. doi:10.1016/j.ajhg.2017.11.002.
44. Vandenbroucke I, Callens T, De Paepe A, Messiaen L. Complex splicing generates great diversity in human NFI transcripts. *BMC Genomics.* 2002;3:13. doi:10.1186/1471-2164-3-13.
45. Sakharkar MK, Chow VTK, Kanguane P. Distributions of exons and introns in the human genome. *In Silico Biol.* 2004;4:387–93. <http://www.ncbi.nlm.nih.gov/pubmed/15217358>.
46. Watanabe T, Kimura A, Kuroyanagi H. Alternative Splicing Regulator RBM20 and Cardiomyopathy. *Front Mol Biosci.* 2018;5. doi:10.3389/fmolb.2018.00105.
47. Saravana RKM. Characterization of Small Conductance Ca<sup>2+</sup> activated K<sup>+</sup> Channel 2 Isoforms in Mouse Brain. University of Göttingen; 2007.
48. Abramowicz A, Gos M. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet.* 2018;59:253–68. doi:10.1007/s13353-018-0444-7.
49. Juan WC, Roca X, Ong ST. Identification of cis-Acting Elements and Splicing Factors Involved in the Regulation of BIM Pre-mRNA Splicing. *PLoS One.* 2014;9:e95210. doi:10.1371/journal.pone.0095210.
50. Krawczak M, Reiss J, Cooper D. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. *Hum Genet.* 1992;90:41–54. doi:10.1007/BF00210743.
51. Krawczak M, Thomas NST, Hundrieser B, Mort M, Wittig M, Hampe J, et al. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat.* 2007;28:150–8. doi:10.1002/humu.20400.
52. Saez B, Walter MJ, Graubert TA. Splicing factor gene mutations in hematologic malignancies. *Blood.* 2017;129:1260–9. doi:10.1182/blood-2016-10-692400.
53. Eng L, Coutinho G, Nahas S, Yeo G, Tanouye R, Babaei M, et al. Nonclassical Splicing Mutations in the Coding and Noncoding Regions of

- the ATM Gene: Maximum Entropy Estimates of Splice Junction Strengths. *Hum Mutat.* 2004;23:67–76. doi:10.1002/humu.10295.
54. Moon SD, Park JH, Kim EM, Kim JH, Han JH, Yoo SJ, et al. A novel IVS2-1G> A mutation causes aberrant splicing of the HRPT2 gene in a family with hyperparathyroidism-jaw tumor syndrome. *J Clin Endocrinol Metab.* 2005;90:878–83. doi:10.1210/jc.2004-0991.
  55. Caminsky NG, Mucaki EJ, Rogan PK. Interpretation of mRNA splicing mutations in genetic disease: review of the literature and guidelines for information-theoretical analysis. *F1000Research.* 2015;3:282. doi:10.12688/f1000research.5654.2.
  56. Sobczyńska-Tomaszewska A, Ołtarzewski M, Czerska K, Wertheim-Tysarowska K, Sands D, Walkowiak J, et al. Newborn screening for cystic fibrosis: Polish 4 years' experience with CFTR sequencing strategy. *Eur J Hum Genet.* 2013;21:391–6. doi:10.1038/ejhg.2012.180.
  57. Sanz DJ, Hollywood JA, Scallan MF, Harrison PT. Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA. *PLoS One.* 2017;12:e0184009. doi:10.1371/journal.pone.0184009.
  58. Anczuków O, Rosenberg AZ, Akerman M, Das S, Zhan L, Karni R, et al. The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. *Nat Struct Mol Biol.* 2012;19:220–8. doi:10.1038/nsmb.2207.
  59. Rivas MA, Pirinen M, Conrad DF, Lek M, Tsang EK, Karczewski KJ, et al. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* (80- ). 2015;348:666–9. doi:10.1126/science.1261877.
  60. Zahir F, Friedman J. The impact of array genomic hybridization on mental retardation research: a review of current technologies and their clinical utility. *Clin Genet.* 2007;72:271–87. doi:10.1111/j.1399-0004.2007.00847.x.
  61. Conner KJ, Hartl LD. *A Primer of Ecological Genetics.* Sinauer Associates; 2004.
  62. Kučinskas V. *Genomo įvairovė: lietuviai Europoje. Švytėjimas;* 2004.
  63. Kucinskas V, Uktveryte I. Genetic Variation and Genomic Origin of Lithuanians. *Ifeh 12Th World Congr Environ Heal.* 2012;:7–11.
  64. Uktverytė I. Lietuvos etnolingvistinių grupių genetinės struktūros analizė remiantis informatyviais genomo žymenimis. *Vilniaus universitetas;* 2014.
  65. Rančelis T. Patogeninių genomo variantų ir jų genų, lemiančių autosominės recesyvias ligas, įvairovės analizė, panaudojant viso egzomo sekoskaitą. *Vilniaus universitetas;* 2016.



66. Pranckėnienė L. De novo mutacijų intensyvumo įvertinimas bendrosios Lietuvos populiacijos ir asmenų su intelektine negalia egzomuose. Vilniaus universitetas; 2019.
67. Čiuladaitė Ž. Nesubalansuotų chromosominių struktūros pokyčių paplitimas žmogaus genome ir jų reikšmė intelektinės negalios etiopatogenezei. Daktaro disertacija. Vilniaus universitetas; 2013.
68. Jorde LB, Wooding SP. Genetic variation, classification and “race.” *Nat Genet.* 2004;36:S28–33. doi:10.1038/ng1435.
69. Burokienė N, Karčiauskaitė D, Kasiulevičius V, Kučinskas V, Kučinskienė ZA. Širdies ir kraujagyslių ligų sąsajos su aterosklerozės rizikos veiksniais ir biocheminiais žymenimis Lietuvos populiacijoje. *Lab Med.* 2015;17:47–52.
70. Domarkienė I. Lietuvos populiacijos genetinės struktūros tyrimas remiantis vieno nukleotido polimorfizmų asociacijos su liga analize. Vilniaus universitetas; 2014.
71. Ambrozaitytė L. Genominių veiksnių įtaka daugiaveiksnių etiologijos lūpos ir (arba) gomurio nesuaugimams lietuvių pacientų grupėje. Vilniaus universitetas; 2011.
72. Mikštienė V. Įgimto / paveldimo klausos sutrikimo genomika: įtaka patogenezei ir fenotipinei išraiškai Lietuvos populiacijoje. Vilniaus universitetas; 2017.
73. Preiksaitienė E, Molytė A, Kasnauskienė J, Čiuladaitė Z, Utkus A, Patsalis PC, et al. Considering specific clinical features as evidence of pathogenic copy number variants. *J Appl Genet.* 2014;55:189–96. doi:10.1007/s13353-014-0197-x.
74. Apie InGenes projektą. <http://www.ingenef.vu.lt/index.php/lt/apie-3>. Accessed 2 Jan 2020.
75. Nowakowska B. Clinical interpretation of copy number variants in the human genome. *J Appl Genet.* 2017;58:449–57. doi:10.1007/s13353-017-0407-4.
76. Lu X, Shaw CA, Patel A, Li J, Cooper ML, Wells WR, et al. Clinical Implementation of Chromosomal Microarray Analysis: Summary of 2513 Postnatal Cases. *PLoS One.* 2007;2:e327. doi:10.1371/journal.pone.0000327.
77. Shaffer LG, Kashork CD, Saleki R, Rorem E, Sundin K, Ballif BC, et al. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. *J Pediatr.* 2006;149:98-102.e5. doi:10.1016/j.jpeds.2006.02.006.
78. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci.* 1977;74:5463–7.

doi:10.1073/pnas.74.12.5463.

79. Sanger F. Determination of Nucleotide Sequences in DNA. *Biosci Rep.* 2004;24:237–53. doi:10.1007/s10540-005-2733-8.
80. Hoff KJ. The effect of sequencing errors on metagenomic gene prediction. *BMC Genomics.* 2009;10:520. doi:10.1186/1471-2164-10-520.
81. Codina-Solà M, Rodríguez-Santiago B, Homs A, Santoyo J, Rigau M, Aznar-Lain G, et al. Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. *Mol Autism.* 2015;6:21. doi:10.1186/s13229-015-0017-0.
82. Jamuar SS, Hekha., Tan E-C. Clinical application of next-generation sequencing for Mendelian diseases. *Hum Genomics.* 2015;9:10. doi:10.1186/s40246-015-0031-5.
83. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet.* 2012;20:490–7. doi:10.1038/ejhg.2011.258.
84. Sun Y, Ruivenkamp CAL, Hoffer MJV, Vrijenhoek T, Kriek M, van Asperen CJ, et al. Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome? *Hum Mutat.* 2015;36:648–55. doi:10.1002/humu.22783.
85. Becich M, Santana-Santos L, Gullapalli R, Desai K, Kant J. Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics. *J Pathol Inform.* 2012;3:40. doi:10.4103/2153-3539.103013.
86. Mu W, Lu H-M, Chen J, Li S, Elliott AM. Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *J Mol Diagnostics.* 2016;18:923–32. doi:10.1016/j.jmoldx.2016.07.006.
87. Cano-Gamez E, Trynka G. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front Genet.* 2020;11:1–21.
88. Paul DS, Soranzo N, Beck S. Functional interpretation of non-coding sequence variation: Concepts and challenges. *BioEssays.* 2014;36:191–9. doi:10.1002/bies.201300126.
89. Harripaul R, Noor A, Ayub M, Vincent JB. The Use of Next-Generation Sequencing for Research and Diagnostics for Intellectual Disability. *Cold Spring Harb Perspect Med.* 2017;7:a026864. doi:10.1101/cshperspect.a026864.
90. Guillaume JC. PubMed. In: *The Dictionary of Genomics, Transcriptomics and Proteomics.* Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2015. p. 1–1. doi:10.1002/9783527678679.dg10319.
91. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online

- Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 2005;33 DATABASE ISS.:D514–7. doi:10.1093/nar/gki033.
92. Sherry ST. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29:308–11. doi:10.1093/nar/29.1.308.
  93. Lappalainen I, Lopez J, Skipper L, Hefferon T, Spalding JD, Garner J, et al. dbVar and DGVA: public archives for genomic structural variation. *Nucleic Acids Res.* 2012;41:D936–41. doi:10.1093/nar/gks1213.
  94. MacDonald JR, Ziman R, Yuen RKC, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 2014;42:D986–92. doi:10.1093/nar/gkt958.
  95. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. *Nucleic Acids Res.* 2002;30:38–41. doi:10.1093/nar/30.1.38.
  96. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581:434–43. doi:10.1038/s41586-020-2308-7.
  97. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012;40:W452–7. doi:10.1093/nar/gks539.
  98. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7:248–9. doi:10.1038/nmeth0410-248.
  99. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014;11:361–2. doi:10.1038/nmeth.2890.
  100. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One.* 2012;7:e46688. doi:10.1371/journal.pone.0046688.
  101. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37:e67–e67. doi:10.1093/nar/gkp215.
  102. Gasteiger E. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003;31:3784–8. doi:10.1093/nar/gkg563.
  103. Bateman A. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 2019;47:D506–15. doi:10.1093/nar/gky1049.

104. Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res.* 2014;42:D222–30. doi:10.1093/nar/gkt1223.
105. Sun H, Yu G. New insights into the pathogenicity of non-synonymous variants through multi-level analysis. *Sci Rep.* 2019;9:1667. doi:10.1038/s41598-018-38189-9.
106. Miosge LA, Field MA, Sontani Y, Cho V, Johnson S, Palkova A, et al. Comparison of predicted and actual consequences of missense mutations. *Proc Natl Acad Sci.* 2015;112:E5189–98. doi:10.1073/pnas.1511585112.
107. Ernst C, Hahnen E, Engel C, Nothnagel M, Weber J, Schmutzler RK, et al. Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. *BMC Med Genomics.* 2018;11:35. doi:10.1186/s12920-018-0353-y.
108. Firth H V., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet.* 2009;84:524–33. doi:10.1016/j.ajhg.2009.03.010.
109. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med.* 2019;25:60–4. doi:10.1038/s41591-018-0279-0.
110. Yang H, Robinson PN, Wang K. Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. *Nat Methods.* 2015;12:841–3. doi:10.1038/nmeth.3484.
111. Köhler S, Carmody L, Vasilevsky N, Jacobsen JOB, Danis D, Gourdiine J-P, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* 2019;47:D1018–27. doi:10.1093/nar/gky1105.
112. Patel DR, Apple R, Kanungo S, Akkal A. Intellectual disability: definitions, evaluation and principles of treatment. *Pediatr Med.* 2018;1:11–11. doi:10.21037/pm.2018.12.02.
113. Rodenburg RJ. The functional genomics laboratory: functional validation of genetic variants. *J Inherit Metab Dis.* 2018;41:297–307. doi:10.1007/s10545-018-0146-7.
114. Bunnik EM, Le Roch KG. An Introduction to Functional Genomics and Systems Biology. *Adv Wound Care.* 2013;2:490–8. doi:10.1089/wound.2012.0379.
115. Sun Y V., Hu Y-J. Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. In: *Advances in Genetics.* 2016. p. 147–90. doi:10.1016/bs.adgen.2015.11.004.

116. Huo Y, Li S, Liu J, Li X, Luo X-J. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun.* 2019;10:670. doi:10.1038/s41467-019-08666-4.
117. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* 2005;6:279–84. doi:10.1038/sj.gene.6364190.
118. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013;45:580–5. doi:10.1038/ng.2653.
119. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med.* 2017;9:eaal5209. doi:10.1126/scitranslmed.aal5209.
120. Pellegrini M, Ferrari R. Epigenetic analysis: ChIP-chip and ChIP-seq. In: *Methods in Molecular Biology.* 2012. p. 377–87. doi:10.1007/978-1-61779-400-1\_25.
121. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* 2008;582:1977–86. doi:10.1016/j.febslet.2008.03.004.
122. Parnas D, Linial M. Highly sensitive ELISA-based assay for quantifying protein levels in neuronal cultures. *Brain Res Protoc.* 1998;2:333–8. doi:10.1016/S1385-299X(98)00009-9.
123. Yang P-C, Mahmood T. Western blot: Technique, theory, and trouble shooting. *N Am J Med Sci.* 2012;4:429–34. doi:10.4103/1947-2714.100998.
124. Garg U, Dasouki M. Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: Clinical and laboratory aspects. *Clin Biochem.* 2006;39:315–32. doi:10.1016/j.clinbiochem.2005.12.009.
125. Gowda GAN, Djukovic D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. In: *Methods in Molecular Biology.* 2014. p. 3–12. doi:10.1007/978-1-4939-1258-2\_1.
126. Maccarrone G, Bonfiglio JJ, Silberstein S, Turck CW, Martins-De-Souza D. Characterization of a protein interactome by Co-Immunoprecipitation and shotgun mass spectrometry. In: *Methods in Molecular Biology.* 2017. p. 223–34. doi:10.1007/978-1-4939-6730-8\_19.
127. Walker DG, Whetzel AM, Serrano G, Sue LI, Lue L-F, Beach TG. Characterization of RNA isolated from eighteen different human tissues: results from a rapid human autopsy program. *Cell Tissue Bank.* 2016;17:361–75. doi:10.1007/s10561-016-9555-8.
128. Freeman WM, Walker SJ, Vrana KE. Quantitative RT-PCR: Pitfalls and

- potential. *Biotechniques*. 1999;26:112–25. doi:10.2144/99261rv01.
129. Rao X, Huang X, Zhou Z, Lin X. An improvement of the  $2^{-\Delta\Delta CT}$  method for quantitative real-time polymerase chain reaction data analysis. *Biostat Bioinforma Biomath*. 2013;3:71–85. <http://www.ncbi.nlm.nih.gov/pubmed/25558171>.
130. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin Chem*. 2009;55:611–22. doi:10.1373/clinchem.2008.112797.
131. Piazza V, Bartke A, Miquet J, Sotelo A. Analysis of Different Approaches for the Selection of Reference Genes in RT-qPCR Experiments: A Case Study in Skeletal Muscle of Growing Mice. *Int J Mol Sci*. 2017;18:1060. doi:10.3390/ijms18051060.
132. Van Haelst MM, Monroe GR, Duran K, Van Binsbergen E, Breur JM, Giltay JC, et al. Further confirmation of the MED13L haploinsufficiency syndrome. *Eur J Hum Genet*. 2015;23:135–8. doi:10.1038/ejhg.2014.69.
133. Fusco C, Morlino S, Micale L, Ferraris A, Grammatico P, Castori M. Characterization of Two Novel Intronic Variants Affecting Splicing in FBN1-Related Disorders. *Genes (Basel)*. 2019;10:442. doi:10.3390/genes10060442.
134. Bonjoch L, Mur P, Arnau-Collell C, Vargas-Parra G, Shamloo B, Franch-Expósito S, et al. Approaches to functionally validate candidate genetic variants involved in colorectal cancer predisposition. *Mol Aspects Med*. 2019;69:27–40. doi:10.1016/j.mam.2019.03.004.
135. Meneely P. *Genetic Analysis: Genes, Genomes, and Networks in Eukaryotes*. 2 edition. Oxford University Press; 2014.
136. Puliti A, Caridi G, Ravazzolo R, Ghiggeri GM. Teaching molecular genetics: chapter 4—positional cloning of genetic disorders. *Pediatr Nephrol*. 2007;22:2023–9. doi:10.1007/s00467-007-0548-5.
137. Wortmann SB, Ziętkiewicz S, Kousi M, Szklarczyk R, Haack TB, Gersting SW, et al. CLPB Mutations Cause 3-Methylglutaconic Aciduria, Progressive Brain Atrophy, Intellectual Disability, Congenital Neutropenia, Cataracts, Movement Disorder. *Am J Hum Genet*. 2015;96:245–57. doi:10.1016/j.ajhg.2014.12.013.
138. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. *Nature*. 2017;550:204–13. doi:10.1038/nature24277.
139. Fernandes IR, Russo FB, Pignatari GC, Evangelinellis MM, Tavorali S, Muotri AR, et al. Fibroblast sources: Where can we get them? *Cytotechnology*. 2016;68:223–8. doi:10.1007/s10616-014-9771-7.

140. Wu D, Hamilton B, Martin C, Gao Y, Ye M, Yao S. Generation of Induced Pluripotent Stem Cells by Reprogramming Human Fibroblasts with the Stemgent Human TF Lentivirus Set. *J Vis Exp*. 2009. doi:10.3791/1553.
141. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 2006;126:663–76. doi:10.1016/j.cell.2006.07.024.
142. Romero-Moya D, Santos-Ocaña C, Castaño J, Garrabou G, Rodríguez-Gómez JA, Ruiz-Bonilla V, et al. Genetic Rescue of Mitochondrial and Skeletal Muscle Impairment in an Induced Pluripotent Stem Cells Model of Coenzyme Q 10 Deficiency. *Stem Cells*. 2017;35:1687–703. doi:10.1002/stem.2634.
143. Scherer S, Davis RW. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci*. 1979;76:4951–5. doi:10.1073/pnas.76.10.4951.
144. Addgene. CRISPR 101: A Desktop Resource. 2 edition. 2017.
145. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res*. 2011.
146. Method of the Year 2011. *Nat Methods*. 2012;9:1–1. doi:10.1038/nmeth.1852.
147. Addgene | Homepage. <https://www.addgene.org/>. Accessed 31 Jan 2020.
148. Behrmann L, McComb S, Aguadé-Gorgorió J, Huang Y, Hermann M, Pelczar P, et al. Efficient Generation of Multi-gene Knockout Cell Lines and Patient-derived Xenografts Using Multi-colored Lenti-CRISPR-Cas9. *Bio-Protocol*. 2017;7. doi:10.21769/bioprotoc.2222.
149. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*. 2012;109:E2579-86.
150. Giuliano CJ, Lin A, Girish V, Sheltzer JM. Generating Single Cell-Derived Knockout Clones in Mammalian Cells with CRISPR/Cas9. *Curr Protoc Mol Biol*. 2019;128. doi:10.1002/cpmb.100.
151. Gueneau L, Fish RJ, Shamseldin HE, Voisin N, Tran Mau-Them F, Preiksaitiene E, et al. KIAA1109 Variants Are Associated with a Severe Disorder of Brain Development and Arthrogryposis. *Am J Hum Genet*. 2018;102:116–32. doi:10.1016/j.ajhg.2017.12.002.
152. ClinicalTrials.gov. <https://clinicaltrials.gov/>. Accessed 31 Jan 2020.
153. Demirci S, Leonard A, Haro-Mora JJ, Uchida N, Tisdale JF. CRISPR/Cas9 for Sickle Cell Disease: Applications, Future Possibilities, and Challenges. In: *Advances in Experimental Medicine and Biology*.

2019. p. 37–52. doi:10.1007/5584\_2018\_331.
154. Ledford H. CRISPR treatment inserted directly into the body for first time. *Nature*. 2020;579:185. doi:10.1038/d41586-020-00655-8.
155. Parfitt DA, Lane A, Ramsden CM, Carr A-JF, Munro PM, Jovanovic K, et al. Identification and Correction of Mechanisms Underlying Inherited Blindness in Human iPSC-Derived Optic Cups. *Cell Stem Cell*. 2016;18:769–81. doi:10.1016/j.stem.2016.03.021.
156. Wu X, Kriz AJ, Sharp PA. Target specificity of the CRISPR-Cas9 system. *Quant Biol*. 2014;2:59–70. doi:10.1007/s40484-014-0030-x.
157. Ren F, Ren C, Zhang Z, Duan W, Lecourieux D, Li S, et al. Efficiency Optimization of CRISPR/Cas9-Mediated Targeted Mutagenesis in Grape. *Front Plant Sci*. 2019;10. doi:10.3389/fpls.2019.00612.
158. Ormond KE, Mortlock DP, Scholes DT, Bombard Y, Brody LC, Faucett WA, et al. Human Germline Genome Editing. *Am J Hum Genet*. 2017;101:167–76. doi:10.1016/j.ajhg.2017.06.012.
159. American Psychiatric Association. DSM-5 Diagnostic Classification. In: *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association; 2013. doi:10.1176/appi.books.9780890425596.x00DiagnosticClassification.
160. Surís A, Holliday R, North C. The Evolution of the Classification of Psychiatric Disorders. *Behav Sci (Basel)*. 2016;6:5. doi:10.3390/bs6010005.
161. Kaufman L, Ayub M, Vincent JB. The genetic basis of non-syndromic intellectual disability: a review. *J Neurodev Disord*. 2010;2:182–209. doi:10.1007/s11689-010-9055-2.
162. Lee S, Rudd S, Gratten J, Visscher PM, Prins JB, Dawson PA. Gene networks associated with non-syndromic intellectual disability. *J Neurogenet*. 2018;32:6–14. doi:10.1080/01677063.2017.1404058.
163. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, et al. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet*. 2009;41:535–43. doi:10.1038/ng.367.
164. Queirós FC, Wehby GL, Halpern CT. Developmental Disabilities and Socioeconomic Outcomes in Young Adulthood. *Public Health Rep*. 2015;130:213–21. doi:10.1177/003335491513000308.
165. Polder JJ, Meerding WJ, Bonneux L, Maas PJ van der. Healthcare costs of intellectual disability in the Netherlands: a cost-of-illness perspective. *J Intellect Disabil Res*. 2002;46:168–78. doi:10.1046/j.1365-2788.2002.00384.x.
166. Chiurazzi P, Pirozzi F. Advances in understanding – genetic basis of



- intellectual disability. *F1000Research*. 2016;5:599. doi:10.12688/f1000research.7134.1.
167. Milani D, Ronzoni L, Esposito S. Genetic Advances in Intellectual Disability. *J Pediatr Genet*. 2015;04:125–7. doi:10.1055/s-0035-1564438.
168. Gilissen C, Hahir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014;511:344–7. doi:10.1038/nature13394.
169. Redin C, Gérard B, Lauer J, Herenger Y, Muller J, Quartier A, et al. Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J Med Genet*. 2014;51:724–36. doi:10.1136/jmedgenet-2014-102554.
170. Vissers LELM, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet*. 2016;17:9–18. doi:10.1038/nrg3999.
171. Harripaul R, Vasli N, Mikhailov A, Rafiq MA, Mittal K, Windpassinger C, et al. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. *Mol Psychiatry*. 2018;23:973–84. doi:10.1038/mp.2017.60.
172. Brett M, McPherson J, Zang ZJ, Lai A, Tan E-S, Ng I, et al. Massively Parallel Sequencing of Patients with Intellectual Disability, Congenital Anomalies and/or Autism Spectrum Disorders with a Targeted Gene Panel. *PLoS One*. 2014;9:e93409. doi:10.1371/journal.pone.0093409.
173. Marrus N, Hall L. Intellectual Disability and Language Disorder. *Child Adolesc Psychiatr Clin N Am*. 2017;26:539–54. doi:10.1016/j.chc.2017.03.001.
174. Kochinke K, Zweier C, Nijhof B, Fenckova M, Cizek P, Honti F, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet*. 2016;98:149–64. doi:10.1016/j.ajhg.2015.11.024.
175. Modell B, Darlison MW, Malherbe H, Moorthie S, Blencowe H, Mahaini R, et al. Congenital disorders: epidemiological methods for answering calls for action. *J Community Genet*. 2018;9:335–40. doi:10.1007/s12687-018-0390-4.
176. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: A meta-analysis of population-based studies. *Res Dev Disabil*. 2011;32:419–36. doi:10.1016/j.ridd.2010.12.018.
177. Egbe A, Ho D, Lee S, Uppu S, Srivastava S. Prevalence of congenital anomalies in newborns with congenital heart disease diagnosis. *Ann*

- Pediatr Cardiol.* 2014;7:86. doi:10.4103/0974-2069.132474.
178. Curry CJ, Stevenson RE, Aughton D, Byrne J, Carey JC, Cassidy S, et al. Evaluation of mental retardation: Recommendations of a consensus conference. *Am J Med Genet.* 1997;72:468–77. doi:10.1002/(SICI)1096-8628(19971112)72:4<468::AID-AJMG18>3.0.CO;2-P.
179. Gijsbers AC, Lew JY, Bosch CA, Schuurs-Hoeijmakers JH, van Haeringen A, den Hollander NS, et al. A new diagnostic workflow for patients with mental retardation and/or multiple congenital abnormalities: test arrays first. *Eur J Hum Genet.* 2009;17:1394–402. doi:10.1038/ejhg.2009.74.
180. Baker K, Raymond FL, Bass N. Genetic investigation for adults with intellectual disability. *Curr Opin Neurol.* 2012;25:150–8. doi:10.1097/WCO.0b013e328351820e.
181. Mithyantha R, Kneen R, McCann E, Gladstone M. Current evidence-based recommendations on investigating children with global developmental delay. *Arch Dis Child.* 2017;102:1071–6. doi:10.1136/archdischild-2016-311271.
182. Malinowski J, Miller DT, Demmer L, Gannon J, Pereira EM, Schroeder MC, et al. Systematic evidence-based review: outcomes from exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability. *Genet Med.* 2020;22:986–1004. doi:10.1038/s41436-020-0771-z.
183. World Medical Association Declaration of Helsinki. *JAMA.* 2013;310:2191–4. doi:10.1001/jama.2013.281053.
184. Javadi A, Shamaei M, Ziazi LM, Pourabdollah M, Dorudinia A, Seyedmehdi SM, et al. Qualification Study of Two Genomic DNA Extraction Methods in Different Clinical Samples. *Tanaffos.* 2014;13:41–7.
185. Villegas J, McPhaul M. Establishment and culture of human skin fibroblasts. *Curr Protoc Mol Biol.* 2005;Chapter 28. doi:10.1002/0471142727.mb2803s71.
186. Useful Numbers for Cell Culture | Thermo Fisher Scientific - LT. <https://www.thermofisher.com/lt/en/home/references/gibco-cell-culture-basics/cell-culture-protocols/cell-culture-useful-numbers.html>. Accessed 4 Feb 2020.
187. ThermoFisher. TaqMan Search. [https://www.thermofisher.com/order/genome-database/?pearUXVerSuffix=pearUX2&elcanoForm=true#!/genotyping/assays/genotyping\\_all/?keyword=\\*&searchMethod=keyword](https://www.thermofisher.com/order/genome-database/?pearUXVerSuffix=pearUX2&elcanoForm=true#!/genotyping/assays/genotyping_all/?keyword=*&searchMethod=keyword). Accessed 4 Feb 2020.

188. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134. doi:10.1186/1471-2105-13-134.
189. Applied Biosystems TaqMan Probes - create and order custom oligos - Life Technologies. <https://www.thermofisher.com/order/custom-oligo/custom-taqman-probes>. Accessed 4 Feb 2020.
190. R Development Core Team R. R: A Language and Environment for Statistical Computing. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. doi:10.1007/978-3-540-74686-7.
191. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40:e115–e115. doi:10.1093/nar/gks596.
192. Dieffenbach CW, Lowe TM, Dveksler GS. General concepts for PCR primer design. *Genome Res*. 1993;3:S30–7. doi:10.1101/gr.3.3.S30.
193. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–60. doi:10.1093/bioinformatics/btp324.
194. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. *Genome Res*. 2002;12:996–1006. doi:10.1101/gr.229102.
195. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–303. doi:10.1101/gr.107524.110.
196. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164–e164. doi:10.1093/nar/gkq603.
197. Desvignes J-P, Bartoli M, Delague V, Krahn M, Miltgen M, Bérout C, et al. VarAFT: a variant annotation and filtration system for human next generation sequencing data. *Nucleic Acids Res*. 2018;46:W545–53. doi:10.1093/nar/gky471.
198. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–6. doi:10.1038/nbt.1754.
199. Bögershausen N, Wollnik B. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci*. 2018;11. doi:10.3389/fnmol.2018.00252.
200. Jung E-M, Moffat JJ, Liu J, Dravid SM, Gurusurthy CB, Kim W-Y. *Arid1b* haploinsufficiency disrupts cortical interneuron development and

- mouse behavior. *Nat Neurosci.* 2017;20:1694–707. doi:10.1038/s41593-017-0013-0.
201. Hassan AH, Neely KE, Workman JL. Histone acetyltransferase complexes stabilize SWI/SNF binding to promoter nucleosomes. *Cell.* 2001;104:817–27. doi:10.1016/S0092-8674(01)00279-3.
202. Yan Z, Wang Z, Sharova L, Sharov AA, Ling C, Piao Y, et al. BAF250B-Associated SWI/SNF Chromatin-Remodeling Complex Is Required to Maintain Undifferentiated Mouse Embryonic Stem Cells. *Stem Cells.* 2008;26:1155–65.
203. Temtamy SA, McKusick VA. The genetics of hand malformations. *Birth Defects Orig Artic Ser.* 1978;14:i–xviii, 1–619. doi:10.1097/00006534-198001000-00022.
204. Radhakrishna U, Bornholdt D, Scott HS, Patel UC, Rossier C, Engel H, et al. The Phenotypic Spectrum of GLI3 Morphopathies Includes Autosomal Dominant Preaxial Polydactyly Type-IV and Postaxial Polydactyly Type-A/B; No Phenotype Prediction from the Position of GLI3 Mutations. *Am J Hum Genet.* 1999;65:645–55. doi:10.1086/302557.
205. Ito S, Kitazawa R, Haraguchi R, Kondo T, Ouchi A, Ueda Y, et al. Novel GLI3 variant causing overlapped Greig cephalopolysyndactyly syndrome (GCPS) and Pallister-Hall syndrome (PHS) phenotype with agenesis of gallbladder and pancreas. *Diagn Pathol.* 2018;13:1. doi:10.1186/s13000-017-0682-8.
206. Ni F, Han G, Guo R, Cui H, Wang B, Li Q. A Novel Frameshift Mutation of GLI3 Causes Isolated Postaxial Polydactyly. *Ann Plast Surg.* 2019;82:570–3.
207. Litingtung Y, Li Y, Fallon JF, Chiang C. Shh and Gli3 are dispensable for limb skeleton formation but regulate digit number and identity. *Nature.* 2002;418:979–83.
208. Suzuki T. How is digit identity determined during limb development? *Development Growth and Differentiation.* 2013;55:130–8.
209. Quinn ME, Haaning A, Ware SM. Preaxial polydactyly caused by Gli3 haploinsufficiency is rescued by Zic3 loss of function in mice. *Hum Mol Genet.* 2012;21:1888–96.
210. Johnston JJ, Sapp JC, Turner JT, Amor D, Aftimos S, Aleck KA, et al. Molecular analysis expands the spectrum of phenotypes associated with GLI3 mutations. *Hum Mutat.* 2010;31:1142–54.
211. Herrera FJ, Yamaguchi T, Roelink H, Tjian R. Core promoter factor TAF9B regulates neuronal gene expression. *Elife.* 2014;3. doi:10.7554/elife.02559.
212. Cheung A, Trevers KE, Reyes-Corral M, Antinucci P, Hindges R.

- Expression and Roles of Teneurins in Zebrafish. *Front Neurosci.* 2019;13. doi:10.3389/fnins.2019.00158.
213. Christianson AL, Stevenson RE, Van der Meyden CH, Pelsler J, Theron FW, Van Rensburg PL, et al. X linked severe mental retardation, craniofacial dysmorphology, epilepsy, ophthalmoplegia, and cerebellar atrophy in a large South African kindred is localised to Xq24-q27. *J Med Genet.* 1999;36:759–66.
214. Ilie A, Gao AYL, Reid J, Boucher A, McEwan C, Barrière H, et al. A Christianson syndrome-linked deletion mutation ( $\Delta$  287ES288) in SLC9A6 disrupts recycling endosomal function and elicits neurodegeneration and cell death. *Mol Neurodegener.* 2016;11.
215. Wakabayashi S, Pang T, Su X, Shigekawa M. A novel topology model of the human Na<sup>+</sup>/H<sup>+</sup> exchanger isoform 1. *J Biol Chem.* 2000;275:7942–9.
216. Orłowski J, Grinstein S. Na<sup>+</sup>/H<sup>+</sup> exchangers of mammalian cells. *J Biol Chem.* 1997;272:22373–6. doi:10.1074/jbc.272.36.22373.
217. Donowitz M, Ming Tse C, Fuster D. SLC9/NHE gene family, a plasma membrane and organellar family of Na<sup>+</sup>/H<sup>+</sup> exchangers. *Mol Aspects Med.* 2013;34:236–51. doi:10.1016/j.mam.2012.05.001.
218. Ouyang Q, Lizarraga SB, Schmidt M, Yang U, Gong J, Ellisor D, et al. Christianson Syndrome Protein NHE6 Modulates TrkB Endosomal Signaling Required for Neuronal Circuit Development. *Neuron.* 2013;80:97–112.
219. Källén K, Robert E, Mastroiacovo P, Castilla EE, Källén B. CHARGE association in newborns: A registry-based study. *Teratology.* 1999;60:334–43.
220. Martin DM, Sheldon S, Gorski JL. CHARGE association With choanal atresia and inner ear hypoplasia in a child with a de novo chromosome translocation t(2;7)(p14;q21.11). *Am J Med Genet.* 2001;99:115–9.
221. Lalani SR. SEMA3E mutation in a patient with CHARGE syndrome. *J Med Genet.* 2004;41:e94–e94. doi:10.1136/jmg.2003.017640.
222. Luquetti D V., Hing A V., Rieder MJ, Nickerson DA, Turner EH, Smith J, et al. “Mandibulofacial dysostosis with microcephaly” caused by EFTUD2 mutations: Expanding the phenotype. *Am J Med Genet Part A.* 2013;161:108–13.
223. Sakata S, Okada S, Aoyama K, Hara K, Tani C, Kagawa R, et al. Individual clinically diagnosed with CHARGE syndrome but with a mutation in KMT2D, a gene associated with Kabuki syndrome: A case report. *Front Genet.* 2017;8:1–6. doi:10.3389/fgene.2017.00210.
224. Issekutz KA, Graham JM, Prasad C, Smith IM, Blake KD. An

- epidemiological analysis of CHARGE syndrome: Preliminary results from a Canadian study. *Am J Med Genet.* 2005;133 A:309–17. doi:10.1002/ajmg.a.30560.
225. Janssen N, Bergman JEH, Swertz MA, Tranebjaerg L, Lodahl M, Schoots J, et al. Mutation update on the CHD7 gene involved in CHARGE syndrome. *Hum Mutat.* 2012;33:1149–60. doi:10.1002/humu.22086.
226. CHD7 database. <https://chd7.org/>. Accessed 23 Mar 2020.
227. McDaniell R, Warthen DM, Sanchez-Lara PA, Pai A, Krantz ID, Piccoli DA, et al. NOTCH2 Mutations Cause Alagille Syndrome, a Heterogeneous Disorder of the Notch Signaling Pathway. *Am J Hum Genet.* 2006;79:169–73. doi:10.1086/505332.
228. Pasmooij AMG, Pas HH, Bolling MC, Jonkman MF. Revertant mosaicism in junctional epidermolysis bullosa due to multiple correcting second-site mutations in LAMB3. *J Clin Invest.* 2007;117:1240–8.
229. Mukhopadhyay A, Nikopoulos K, Maugeri A, De Brouwer APM, Van Nouhuys CE, Boon CJF, et al. Erosive vitreoretinopathy and wagner disease are caused by intronic mutations in CSPG2/Versican that result in an imbalance of splice variants. *Investig Ophthalmol Vis Sci.* 2006;47:3565–72. doi:10.1167/iovs.06-0141.
230. Gelincik A, Demir S, Olgaç M, Karaman V, Toksoy G, Çolakoğlu B, et al. Idiopathic angioedema with F12 mutation: is it a new entity? *Ann Allergy, Asthma Immunol.* 2015;114:154–6. doi:10.1016/j.anai.2014.11.018.
231. Simpson MA, Hsu R, Keir LS, Hao J, Sivapalan G, Ernst LM, et al. Mutations in FAM20C Are Associated with Lethal Osteosclerotic Bone Dysplasia (Raine Syndrome), Highlighting a Crucial Molecule in Bone Development. *Am J Hum Genet.* 2007;81:906–12. doi:10.1086/522240.
232. Schnetz MP, Bartels CF, Shastri K, Balasubramanian D, Zentner GE, Balaji R, et al. Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. *Genome Res.* 2009;19:590–601.
233. Brehm A, Tufteland KR, Aasland R, Becker PB. The many colours of chromodomains. *BioEssays.* 2004;26:133–40. doi:10.1002/bies.10392.
234. Allen MD, Religa TL, Freund SMV, Bycroft M. Solution Structure of the BRK Domains from CHD7. *J Mol Biol.* 2007;371:1135–40. doi:10.1016/j.jmb.2007.06.007.
235. Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Filipe Pereira C, Fisher AG, et al. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet.* 2010;6:1–15.
236. Vissers LELM, van Ravenswaaij CMA, Admiraal R, Hurst JA, de Vries BBA, Janssen IM, et al. Mutations in a new member of the chromodomain

- gene family cause CHARGE syndrome. *Nat Genet.* 2004;36:955–7. doi:10.1038/ng1407.
237. Bosman EA, Penn AC, Ambrose JC, Kettleborough R, Stemple DL, Steel KP. Multiple mutations in mouse *Chd7* provide models for CHARGE syndrome. *Hum Mol Genet.* 2005;14:3463–76. doi:10.1093/hmg/ddi375.
238. Ogier JM, Arhatari BD, Carpinelli MR, McColl BK, Wilson MA, Burt RA. An intronic mutation in *Chd7* creates a cryptic splice site, causing aberrant splicing in a mouse model of CHARGE syndrome. *Sci Rep.* 2018;8:5482. doi:10.1038/s41598-018-23856-8.
239. Sehnal D, Deshpande M, Vařeková RS, Mir S, Berka K, Midlik A, et al. LiteMol suite: Interactive web-based visualization of large-scale macromolecular structure data. *Nature Methods.* 2017;14:1121–2.
240. Mikstiene V, Jakaitiene A, Byckova J, Gradauskiene E, Preiksaitiene E, Burnyte B, et al. The high frequency of GJB2 gene mutation c.313\_326del14 suggests its possible origin in ancestors of Lithuanian population. *BMC Genet.* 2016;17:45. doi:10.1186/s12863-016-0354-9.
241. Takeda N, Hara H, Fujiwara T, Kanaya T, Maemura S, Komuro I. TGF- $\beta$  signaling-related genes and thoracic aortic aneurysms and dissections. *International Journal of Molecular Sciences.* 2018;19.
242. Laimer M, Klausegger A, Aberer W, Oender K, Steinhuber M, Lanschuetzer CM, et al. Haploinsufficiency due to deletion within the 3'-UTR of C1-INH-gene associated with hereditary angioedema. *Genet Med.* 2006;8:249–54.
243. Akhurst RJ, Hata A. Targeting the TGF $\beta$  signalling pathway in disease. *Nature Reviews Drug Discovery.* 2012;11:790–811.
244. Heldin CH, Moustakas A. Signaling receptors for TGF- $\beta$  family members. *Cold Spring Harb Perspect Biol.* 2016;8.
245. Gallo EM, Loch DC, Habashi JP, Calderon JF, Chen Y, Bedja D, et al. Angiotensin II-dependent TGF- $\beta$  signaling contributes to Loeys-Dietz syndrome vascular pathogenesis. *J Clin Invest.* 2014;124:448–60.
246. Kitano K. Structural mechanisms of human RecQ helicases WRN and BLM. *Front Genet.* 2014;5 OCT. doi:10.3389/fgene.2014.00366.
247. Cobb JA, Bjergbaek L. RecQ helicases: Lessons from model organisms. *Nucleic Acids Research.* 2006;34:4106–14.
248. Nascimbeni AC, Fanin M, Tasca E, Angelini C. Transcriptional and translational effects of intronic CAPN3 gene mutations. *Hum Mutat.* 2010;31.
249. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: The next generation in gene variant databases. *Hum Mutat.* 2011;32:557–63.

250. Blázquez L, Azpitarte M, Sáenz A, Goicoechea M, Otaegui D, Ferrer X, et al. Characterization of novel CAPN3 isoforms in white blood cells: An alternative approach for limb-girdle muscular dystrophy 2A diagnosis. *Neurogenetics*. 2008;9:173–82.
251. Lasa-Elgarresta J, Mosqueira-Martín L, Naldaiz-Gastesi N, Sáenz A, López de Munain A, Vallejo-Illarramendi A. Calcium Mechanisms in Limb-Girdle Muscular Dystrophy with CAPN3 Mutations. *Int J Mol Sci*. 2019;20. doi:10.3390/ijms20184548.
252. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7:539. doi:10.1038/msb.2011.75.
253. Preikšaitienė E, Ambrozaitytė L, Maldžienė Ž, Morkūnienė A, Cimbalistienė L, Rančelis T, et al. Identification of genetic causes of congenital neurodevelopmental disorders using genome wide molecular technologies. *Acta medica Litu*. 2016;23.
254. Concepcion CP, Bonetti C, Ventura A. The MicroRNA-17-92 family of MicroRNA clusters in development and disease. *Cancer Journal (United States)*. 2012;18:262–7.
255. De Pontual L, Yao E, Callier P, Faivre L, Drouin V, Cariou S, et al. Germline deletion of the miR-17~92 cluster causes skeletal and growth defects in humans. *Nat Genet*. 2011;43:1026–31.
256. Saunders S, Paine-Saunders S, Lander AD. Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb, and brain. *Dev Biol*. 1997;190:78–93.
257. Kannu P, Campos-Xavier AB, Hull D, Martinet D, Ballhausen D, Bonafé L. Post-axial polydactyly type A2, overgrowth and autistic traits associated with a chromosome 13q31.3 microduplication encompassing miR-17-92 and GPC5. *Eur J Med Genet*. 2013;56:452–7.
258. Hemmat M, Rumble MJ, Mahon LW, Strom CM, Anguiano A, Talai M, et al. Short stature, digit anomalies and dysmorphic facial features are associated with the duplication of miR-17 ~ 92 cluster. *Mol Cytogenet*. 2014;7:1–20.
259. Jønch AE, Larsen LG, Pouplier S, Nielsen K, Brøndum-Nielsen K, Tümer Z. Partial duplication of 13q31.3-q34 and deletion of 13q34 associated with diaphragmatic hernia as a sole malformation in a fetus. *Am J Med Genet Part A*. 2012;158 A:2302–8.
260. van der Zwaag PA, Dijkhuizen T, Gerssen-Schoorl KBJ, Colijn AW, Broens PMA, Flapper BCT, et al. An interstitial duplication of chromosome 13q31.3q32.1 further delineates the critical region for



- postaxial polydactyly type A2. *Eur J Med Genet.* 2010;53:45–9.
261. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinforma.* 2016;54. doi:10.1002/cpbi.5.
262. Reymond A, Henrichsen CN, Harewood L, Merla G. Side effects of genome structural changes. *Curr Opin Genet Dev.* 2007;17:381–6. doi:10.1016/j.gde.2007.08.009.
263. Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot MT, et al. Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes. *Am J Hum Genet.* 2006;79:332–41.
264. Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, et al. Segmental copy number variation shapes tissue transcriptomes. *Nat Genet.* 2009;41:424–9.
265. Bonnet C, Andrieux J, Béri-Dexheimer M, Leheup B B, Boute O, Manouvrier S, et al. Microdeletion at chromosome 4q21 defines a new emerging syndrome with marked growth restriction, mental retardation and absent or severely delayed speech. *J Med Genet.* 2010;47:377–84. doi:10.1136/jmg.2009.071902.
266. Hu X, Chen X, Wu B, Soler IM, Chen S, Shen Y. Further defining the critical genes for the 4q21 microdeletion disorder. *Am J Med Genet Part A.* 2017;173:120–5.
267. Zarrei M, Merico D, Kellam B, Engchuan W, Scriver T, Jokhan R, et al. A de novo deletion in a boy with cerebral palsy suggests a refined critical region for the 4q21.22 microdeletion syndrome. *Am J Med Genet Part A.* 2017;173:1287–93.
268. Komlósi K, Duga B, Hadzsiev K, Czakó M, Kosztolányi G, Fogarasi A, et al. Phenotypic variability in a Hungarian patient with the 4q21 microdeletion syndrome. *Mol Cytogenet.* 2015;8.
269. Yano S, Mcnamara M, Halbach S, Waggoner D. 4Q21 Microdeletion in a Patient With Epilepsy and Brain Malformations. *American Journal of Medical Genetics, Part A.* 2015;167:1409–13.
270. Bhoj E, Halbach S, McDonald-McGinn D, Tan C, Lande R, Waggoner D, et al. Expanding the spectrum of microdeletion 4q21 syndrome: A partial phenotype with incomplete deletion of the minimal critical region and a new association with cleft palate and pierre robin sequence. *Am J Med Genet Part A.* 2013;161:2327–33.
271. Tsang E, Rupps R, Mcgillivray B, Eydoux P, Marra M, Arbour L, et al. Life-history chronicle for a patient with the recently described chromosome 4q21 microdeletion syndrome. *Am J Med Genet Part A.*

- 2012;158 A:2606–9.
272. Dukes-Rimsky L, Guzauskas GF, Holden KR, Griggs R, Ladd S, del Carmen Montoya M, et al. Microdeletion at 4q21.3 is associated with intellectual disability, dysmorphic facies, hypotonia, and short stature. *Am J Med Genet Part A*. 2011;155:2146–53.
273. Sakazume S, Kido Y, Murakami N, Matsubara T, Numabe H. Additional patients with 4q deletion: Severe growth delay and polycystic kidney disease associated with 4q21q22 loss. *Pediatr Int*. 2015;57:880–3.
274. Harada N, Nagai T, Shimokawa O, Niikawa N, Matsumoto N. A 4q21-q22 deletion in a girl with severe growth retardation. *Clin Genet*. 2002;61:226–8. doi:10.1034/j.1399-0004.2002.610311.x.
275. Miyazawa T, Ogawa Y, Chusho H, Yasoda A, Tamura N, Komatsu Y, et al. Cyclic GMP-dependent protein kinase II plays a critical role in C-type natriuretic peptide-mediated endochondral ossification. *Endocrinology*. 2002;143:3604–10.
276. Liu Z-P, Wu C, Miao H, Wu H. RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. *Database*. 2015;2015:bav095. doi:10.1093/database/bav095.
277. Aleksuniene B, Preiksaitiene E, Morkuniene A, Ambrozaityte L, Utkus A. A de novo 1q22q23.1 Interstitial Microdeletion in a Girl with Intellectual Disability and Multiple Congenital Anomalies Including Congenital Heart Defect. *Cytogenet Genome Res*. 2018;154:6–11.
278. Asadollahi R, Zweier M, Gogoll L, Schiffmann R, Sticht H, Steindl K, et al. Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. *Eur J Med Genet*. 2017;60:451–64.
279. Tørring PM, Larsen MJ, Brasch-Andersen C, Krogh LN, Kibæk M, Laulund L, et al. Is MED13L-related intellectual disability a recognizable syndrome? *Eur J Med Genet*. 2019;62:129–36.
280. Yin JW, Wang G. The Mediator complex: A master coordinator of transcription and cell lineage development. *Dev*. 2014;141:977–87.
281. Utami KH, Winata CL, Hillmer AM, Aksoy I, Long HT, Liany H, et al. Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. *Hum Mutat*. 2014;35:1311–20. doi:10.1002/humu.22636.
282. Angus SP, Nevins JR. A role for Mediator complex subunit MED13L in Rb/E2F-induced growth arrest. *Oncogene*. 2012;31:4709–17.
283. Biggar K, Storey K. Perspectives in Cell Cycle Regulation: Lessons from an Anoxic Vertebrate. *Curr Genomics*. 2009;10:573–84.
284. Kundrotas G, Gasperskaja E, Slapsyte G, Gudleviciene Z, Krasko J,

- Stumbryte A, et al. Identity, proliferation capacity, genomic stability and novel senescence markers of mesenchymal stem cells isolated from low volume of human bone marrow. *Oncotarget*. 2016;7:10788–802.
285. SETHE S, SCUTT A, STOLZING A. Aging of mesenchymal stem cells. *Ageing Res Rev*. 2006;5:91–116. doi:10.1016/j.arr.2005.10.001.
286. Wagner W, Bork S, Lepperdinger G, Joussem S, Ma N, Strunk D, et al. How to track cellular aging of mesenchymal stromal cells? *Aging (Albany NY)*. 2010;2:224–30. doi:10.18632/aging.100136.
287. Ho JH, Chen YF, Ma WH, Tseng TC, Chen MH, Lee OK. Cell contact accelerates replicative senescence of human mesenchymal stem cells independent of telomere shortening and p53 activation: Roles of ras and oxidative stress. *Cell Transplant*. 2011;20:1209–20.
288. Home - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene>. Accessed 17 Jan 2020.
289. Home - Genome - NCBI. <https://www.ncbi.nlm.nih.gov/genome/>. Accessed 17 Jan 2020.
290. Petryszak R, Keays M, Tang YA, Fonseca NA, Barrera E, Burdett T, et al. Expression Atlas update - An integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res*. 2016;44:D746–52.
291. Rath A, Olry A, Dhombres F, Brandt MM, Urbero B, Ayme S. Representation of rare diseases in health information systems: The orphanet approach to serve a wide range of end users. *Hum Mutat*. 2012;33:803–8. doi:10.1002/humu.22078.

## PRIEDAI

### I PRIEDAS. Vilniaus regioninio biomedicininų tyrimų etikos komiteto leidimas atlikti INGENES biomedicininį tyrimą



#### VILNIAUS UNIVERSITETO MEDICINOS FAKULTETAS

Viešoji įstaiga, Universiteto g. 3, LT-01513 Vilnius, tel. (8 5) 268 7001, faks. (8 5) 272 8646, el. p. [infor@cr.vu.lt](mailto:infor@cr.vu.lt).  
Duomenys kaupiami ir saugomi Juridinių asmenų registre, kodas 211950810.  
Fakulteto duomenys: M.K. Čiurlionio g. 21/27, LT-03101 Vilnius, tel. (8 5) 239 8701, (8 5) 239 7800, faks. (8 5) 239 8705,  
el. p. [mf@mf.vu.lt](mailto:mf@mf.vu.lt)

#### VILNIAUS REGIONINIS BIOMEDICININIŲ TYRIMŲ ETIKOS KOMITETAS

M.K. Čiurlionio g. 21/27, LT-03101 Vilnius, tel. (8 5) 268 6998, el. p. [rbtek@mf.vu.lt](mailto:rbtek@mf.vu.lt)

### LEIDIMAS ATLIKTI BIOMEDICININĮ TYRIMĄ

2017-12-05 Nr.158200-17-962-469

Tyrimo pavadinimas:

#### Intelektinės negalios ir įgimtų anomalijų genetinės architektūros Identifikavimas INGENES

Protokolo Nr.:	IN-17-01
Versija:	03
Data:	2017 11 17
Informuoto asmens sutikimo forma:	03 (12-17 metų vaikams) 2017 11 17 03 (suaugusiam) 2017 11 17 03 (tėvams) 2017 11 17
Pagrindinis tyrėjas:	<b>Eglė Preikšaitienė</b>
Įstaigos pavadinimas:	Vilniaus universiteto Medicinos fakulteto Biomedicinos mokslų instituto Žmogaus ir medicininės genetikos katedra
Adresas:	Santariškių g. 2
Leidimas galioja iki:	<b>2020 04 30</b>

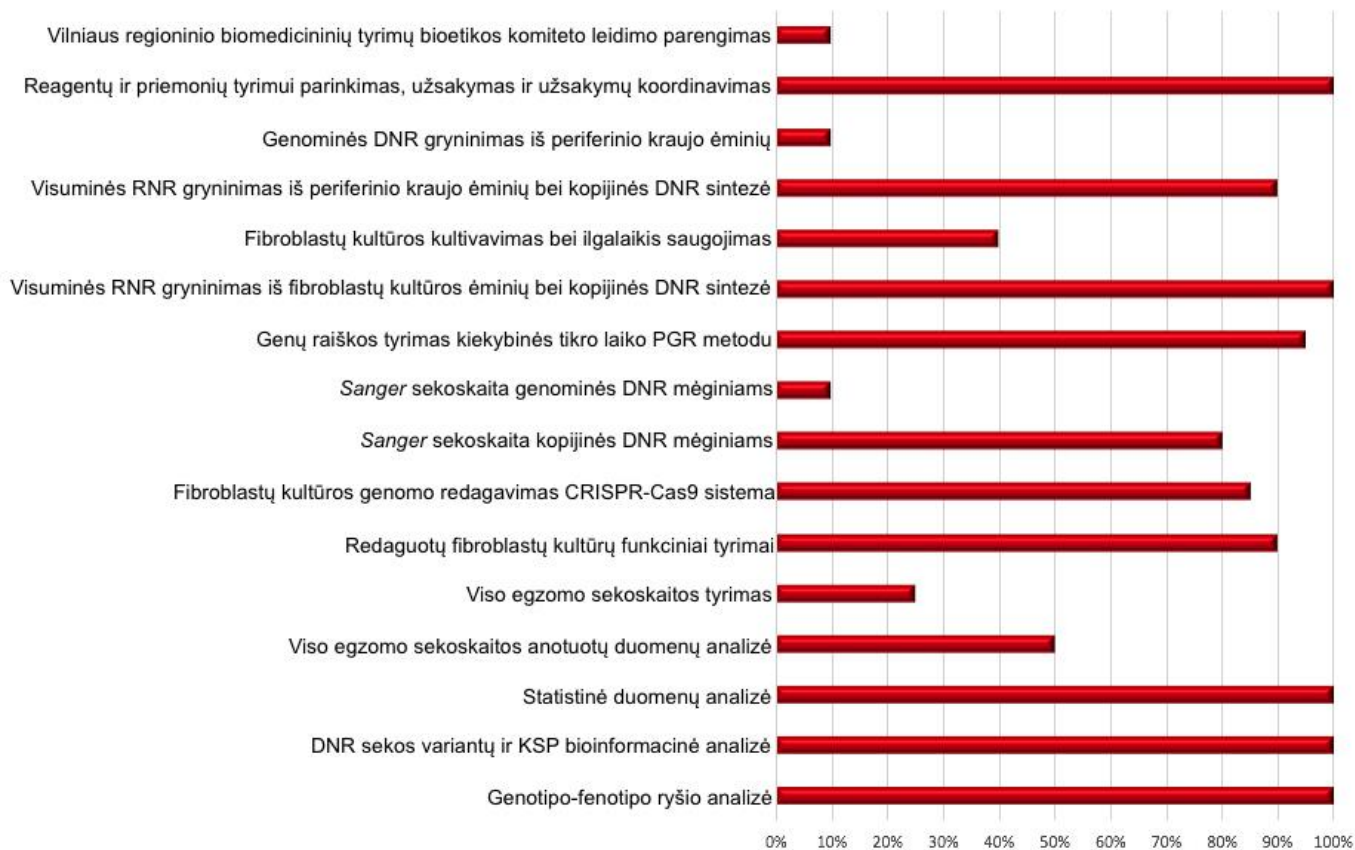
Leidimas išduotas Vilniaus regioninio biomedicininų tyrimų etikos komiteto posėdžio (protokolas Nr. 158200-2017/12), vykusio 2017 m. gruodžio 5 d. sprendimu.

Pirmininkas



 Saulius Vosylius

## II PRIEDAS. Disertacijos autorės atliktų darbų indėlis



### III PRIEDAS. Pradmenų sekos ir PGR reakcijos sąlygos

III.1 lentelė. gDNR ir kDNR pradmenų sekos

Tiriamasis	Genas (Referentinė seka)	Tiriamoji medžiaga	Pradmens seka (5' → 3')	Vieta genome	Pradmens lydimosi temp.	Produkto ilgis
#1	<i>ARID1B</i> (NM_020732.3)	gDNR	F: GAGAGTACGTGGGCTTTGGA	19 egzonas	59,4 °C	182 bp
			R: GACGGTCAGTGCTCATCAAC	19 intronas	58,9 °C	
		kDNR	F: TACCCACAGCAGCCGAATTA	17–18 egzonų jungtis	59,1 °C	1162 bp
			R: CGCTGTCTTCCTCCTCATCC	20 egzonas	59,9 °C	
#2	<i>GLI3</i> (NM_000168.6)	gDNR	F: CTGCAGTACCTCACAGAGCT	13 intronas	59,1 °C	374 bp
			R: TCTTCAAAGCCCAGTGGACA	14 intronas	59,1 °C	
			F: GCCAGCTCCATTCACCAGA	3 intronas	59,7 °C	
			R: TGGTTACAGCGTCATTTTAGGAC	4 intronas	59 °C	
		kDNR	F: CCACTACCACCCCTCCTCATCT	3–4 egzonų jungtis	60,3 °C	166 bp
			R: CCTAATGAAGGGCAGGTCCG	5 egzonas	60,2 °C	
			F: CCAGCACCCTTCTAATGAGGAT	2–3 egzonų jungtis	60,1 °C	
			R: CCTAATGAAGGGCAGGTCCG	5 egzonas	60,2 °C	
#3	<i>SLC9A6</i> (NM_001042537.1)	gDNR	F: CTTGCTATATTCCACGAGCTTCA	6 egzonas	58,6 °C	212 bp
			R: GCCTCATTCTTCCAAATGTTTCC	6 intronas	58,5 °C	
	<i>TAF9B</i> (NM_015975.4)	kDNR	F: ACGCTGATGAAGGTAACGGG	5 egzonas	60,1 °C	431 bp
			R: GCTTCAGCCAAGAGGAAGGT	8 egzonas	60 °C	
			F: ATAGTTTGCGTAAGCTGAGTG	5 intronas	58,7 °C	
			R: GCCACCTTCTGTTGGGTATT	6 intronas	60,2 °C	
<i>TENMI</i> (NM_001163278.1)	gDNR	F: GGGTTCGTGTTGAAAAATTGCC	21 intronas	59,5 °C	393 bp	
		R: GTGGATGGGCTAAAGTGTGG	22 intronas	58,5 °C		

### III.1 lentelė (tęsinys). gDNR ir kDNR pradmenų sekos

Tiriamasis	Genas (Referentinė seka)	Tiriamoji medžiaga	Pradmens seka (5' → 3')	Vieta genome	Pradmens lydimosi temp.	Produkto ilgis
#4	<i>CHD7</i> (NM_017780.4)	gDNR	F: TCGGTATGCCTGATGCCAAG	26 egzonas	60,2 °C	713 bp
			R: CCTTGAAAGCAAAGCAAGAAACAG	27 intronas	59,7 °C	
		kDNR	F: TCGGTATGCCTGATGCCAAG	26 egzonas	60,2 °C	444 bp
			R: TTCTTGTCCACTTTTGCCGC	29–30 egzonų jungtis	59,6 °C	
#5	<i>TGFBR2</i> (NM_001024847.2)	kDNR	F: TAACAGTGGGCAGGTGGGA	5–6 egzonų jungtis	60,5 °C	449 bp
			R: TTAGGGAGCCGTCTTCAGGA	8 egzonas	60 °C	
		kDNR	F: TAACAGTGGGCAGGTGGGA	5–6 egzonų jungtis	60,5 °C	638 bp
			R: CTATGTCACCCACTCCCTGC	3'UTR	60 °C	
#6	<i>BLM</i> (NM_000057.2)	kDNR	F: CCTGGGGTCACTGTTGTCAT	9 egzonas	59,6 °C	452 bp
			R: AAGCTCATGCTAAACACCTGAG	12–13 egzonų jungtis	58,7 °C	
#7	<i>CAPN3</i> (NM_000070.2)	kDNR	F: GCTTCGCCATCTACGAGGTT	11–12 egzonų jungtis	60,2 °C	380 bp
			R: GCTTGTTTTGCCCTTGCCCT	16 egzonas	60,2 °C	

III.1 lentelė (tęsinys). gDNR ir kDNR pradmenų sekos

Tiriamasis	Genas (Referentinė seka)	Tiriamoji medžiaga	Pradmens seka (5' → 3')	Vieta genome	Pradmens lydymosi temp.	Produkto ilgis	
#10	<i>MED13L</i> (NM_015335.4)	kDNR	<b>F:</b> CGGGAATCAAATGGCGTAGG	2 egzonas	59 °C	953 bp	
			<b>R:</b> AGCACTCTGCATACCTCCAC	8 egzonas	59,5 °C		
			<b>F:</b> CGGGAATCAAATGGCGTAGG	2 egzonas	59 °C	1262 bp	
			<b>R:</b> CCCTGGTTGAGATACTGTGGG	10 egzonas	59,8 °C		
			<b>F:</b> CGGGAATCAAATGGCGTAGG	2 egzonas	59 °C	2524 bp	
			<b>R:</b> GGTGGAGTGGGAAACATCCT	15 egzonas	59,3 °C		
		kDNR (Genų raiškos tyrimas)	<b>F:</b> CGAGCCTGGAGGATTGTCA	1 egzonas	59,1 °C	118 bp	
			<b>R:</b> GCTGGGGCTGAAATTATGGG	2 egzonas	59 °C		
		gDNR (CRISPR-Cas9 tyrimas)	<b>TaqMan™ žymuo:</b> ACCTCTTTTCGCTGGCTGAACTCAC		1–2 egzonų jungtis	61,2 °C	559 bp
			<b>F:</b> CTCCACCTTAAATTCGAGAAGA	1 intronas	58,5		
				<b>R:</b> GGGGTGAGAATGTTCTTTTGAG	2 intronas	59,5	



**III.2 lentelė.** Optimizuoto PGR mišinio sudėtis

Tiriamasis	Genas	Tiriamoji medžiaga	PGR mišinio sudėtis vienai reakcijai					Bendras mišinio tūris
			Dejonizuotas vanduo	2x Phusion Flash PCR MasterMix	F (10 pmol/μl)	R (10 pmol/μl)	kDNR	
#1	<i>ARID1B</i> (MIM #614556)	kDNR	4 μl	10 μl	0,5 μl	0,5 μl	5 μ	20 μl
#3	<i>SLC9A6</i> (MIM #300231)	kDNR	4 μl	10 μl	0,5 μl	0,5 μl	5 μ	
#4	<i>CHD7</i> (MIM #608892)	kDNR	4 μl	10 μl	0,5 μl	0,5 μl	5 μ	
#6	<i>BLM</i> (MIM #210900)	kDNR	5 μl	10 μl	0,5 μl	0,5 μl	4 μ	
#7	<i>CAPN3</i> (MIM #114240)	kDNR	4 μl	10 μl	0,5 μl	0,5 μl	5 μl	

**III.3 lentelė.** PGR termociklerio programos etapai ir optimizuotos sąlygos

Tiriamasis	Genas	Tiriamoji medžiaga	PGR etapai					Pabaiga
			Pradinė denatūracija	30 ciklų		Galutinė sintezė		
				Denatūracija	Pradmenų prijungimas		Sintezė	
#1	<i>ARID1B</i> (MIM #614556)	kDNR	98 °C, 30 s	98 °C, 2 s	60 °C, 10 s	72 °C, 20 s	72 °C, 60 s	4°C, ∞
#3	<i>SLC9A6</i> (MIM #300231)	gDNR	98 °C, 30 s	98 °C, 2 s	62 °C, 10 s	72 °C, 7 s	72 °C, 60 s	4°C, ∞
		kDNR	98 °C, 30 s	98 °C, 2 s	64 °C, 10 s	72 °C, 15 s	72 °C, 60 s	4°C, ∞
#4	<i>CHD7</i> (MIM #608892)	gDNR	98 °C, 30 s	98 °C, 2 s	62 °C, 10 s	72 °C, 20 s	72 °C, 60 s	4°C, ∞
#7	<i>CAPN3</i> (MIM #114240)	kDNR	98 °C, 30 s	98 °C, 2 s	62 °C, 10 s	72 °C, 20 s	72 °C, 60 s	4°C, ∞

#### IV PRIEDAS. Tyrimų metu naudotos duomenų bazės ir bioinformaciniai įrankiai

Pavadinimas	Nuoroda internetiniame puslapyje	Trumpas apibūdinimas	Cita- vimas
<b>Mokslinės literatūros duomenų bazė</b>			
<i>NCBI PubMed</i>	<a href="https://www.ncbi.nlm.nih.gov/pubmed/">https://www.ncbi.nlm.nih.gov/pubmed/</a>	Duomenų bazė, turinti daugiau nei 30 mln. medicinos mokslo publikacijų iš MEDLINE, gyvybės mokslų žurnalų bei elektroninių knygų.	[90]
<b>Pagrindinės genomo naršyklės</b>			
<i>ENSEMBL</i>	<a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a>	Genomo naršyklė, kurioje kaupiama informacija apie genus, jų referentines sekas, gali būti atliekama genų anotacija, sekų sulygiavimas, nuspėjama genų reguliacinė funkcija ir kt.	[95]
<i>UCSC Genome Browser</i>	<a href="https://genome.ucsc.edu/">https://genome.ucsc.edu/</a>	Genomo naršyklė, skirta interaktyviai vizualizuoti genomo duomenims, sulygiuoti sekoms, nuspėti variantų funkciniam poveikiui ir kt.	[194]
<b>Genų ir / ar genomo duomenų bazės</b>			
<i>GeneCards</i>	<a href="https://www.genecards.org/">https://www.genecards.org/</a>	Duomenų bazė, kaupianti informaciją apie visus anotuotus žmogaus genus. Ši duomenų bazė integruoja genomine, transkriptomine, proteomine, genetinė, klinikinę ir funkcinę informaciją iš ~150 skirtingų šaltinių.	[261]
<i>NCBI Gene</i>	<a href="https://www.ncbi.nlm.nih.gov/gene/">https://www.ncbi.nlm.nih.gov/gene/</a>	Duomenų bazė, kurioje pateikta informacija apie genų nomenklatūrą, genų raišką, referentines sekas, variantus, fenotipus ir kt.	[288]
<i>NCBI Genome</i>	<a href="https://www.ncbi.nlm.nih.gov/genome/">https://www.ncbi.nlm.nih.gov/genome/</a>	Duomenų bazė, kurioje pateikta informacija apie genomus ir genų referentines sekas.	[289]
<i>GTEx portal</i>	<a href="https://www.gtexportal.org/home/">https://www.gtexportal.org/home/</a>	Kaupiami Genotipo-audinių raiškos (angl. <i>Genotype-Tissue Expression</i> ) projekto duomenys apie audiniams būdingą genų raišką ir jos reguliaciją. Projekto metu tirti ~1 000 asmenų mėginiai, kurie buvo išskirti iš 54 pažeistų audinių.	[118]
<i>The Expression Atlas</i>	<a href="https://www.ebi.ac.uk/gxa/home/">https://www.ebi.ac.uk/gxa/home/</a>	Skirtingų rūšių genų raiškos atlasas.	[290]

#### IV PRIEDAS (tęsinys). Tyrimų metu naudotos duomenų bazės ir bioinformaciniai įrankiai

Pavadinimas	Nuoroda internetiniame puslapyje	Trumpas apibūdinimas	Cita- vimas
<b>Populiacinės duomenų bazės</b>			
<i>1 000 genome project</i>	<a href="https://www.internationalgenome.org/">https://www.internationalgenome.org/</a>	Internetiniame puslapyje pateikta informacija apie tarptautinio 1 000 genomų projekto tyrimų rezultatus, pateikiamas žmogaus genomo variantų katalogas bei genotipų duomenys.	[6]
<i>dbSNP</i>	<a href="https://www.ncbi.nlm.nih.gov/snp/">https://www.ncbi.nlm.nih.gov/snp/</a>	Vieno nukleotido polimorfizmų (VNP), smulkių delecijų / duplikacijų, mikrosatelitų bei nepolimorfinių variantų duomenų bazė.	[92]
<i>dbVAR</i>	<a href="https://www.ncbi.nlm.nih.gov/dbvar/">https://www.ncbi.nlm.nih.gov/dbvar/</a>	Duomenų bazė, apimanti struktūrinius genomo pokyčius: delecijas, duplikacijas, inversijas, insercijas, judriuosius genomo elementus, translokacijas.	[93]
<i>DGV</i>	<a href="http://dgv.tcag.ca/dgv/app/home">http://dgv.tcag.ca/dgv/app/home</a>	Struktūrinių genominių variantų duomenų bazė.	[94]
<i>gnomAD v.2.1.1 browser</i>	<a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a>	Pateikiami įvairių tarptautinių projektų vykdytų egzomo ir genomo tyrimų apibendrinti duomenys. Anksčiau egzomo duomenys buvo kaupiami atskirai vadinamojoje ExAC (angl. <i>Exome Aggregation Consortium</i> ) duomenų bazėje.	[96]
<b>Bioinformaciniai įrankiai sukirpimo varianto patogeniškumui įvertinti</b>			
<i>MutationTaster2</i>	<a href="http://www.mutationtaster.org/">http://www.mutationtaster.org/</a>	DNR sekos variantų patogeniškumo vertinimo įrankis.	[99]
<i>Human Splicing Finder v.3.1</i>	<a href="http://www.umd.be/HSF3/">http://www.umd.be/HSF3/</a>	Sukirpimo variantų ir jų poveikio iRNR <i>in silico</i> analizė.	[101]

#### IV PRIEDAS (tęsinys). Tyrimų metu naudotos duomenų bazės ir bioinformaciniai įrankiai

Pavadinimas	Nuoroda internetiniame puslapyje	Trumpas apibūdinimas	Cita- vimas
<b>PGR ar kPGR tyrimo metu naudojami pagalbiniai bioinformaciniai įrankiai</b>			
<i>NCBI Primer-blast</i>	<a href="https://www.ncbi.nlm.nih.gov/tools/primer-blast/">https://www.ncbi.nlm.nih.gov/tools/primer-blast/</a>	Bioinformacinis įrankis, skirtas sukurti pradmenims.	[188]
<i>Custom TaqMan Probes tool</i>	<a href="https://www.thermofisher.com/order/custom-oligo/custom-taqman-probes">https://www.thermofisher.com/order/custom-oligo/custom-taqman-probes</a>	Bioinformacinis įrankis, skirtas sukurti <i>TaqMan</i> <sup>TM</sup> žymenims.	[189]
<b>Bioinformaciniai įrankiai, skirti transkriptomo ar proteomo analizei atlikti</b>			
<i>ExpASy</i>	<a href="https://web.expasy.org">https://web.expasy.org</a>	<i>ExpASy LALIGN</i> : Įrankis, skirtas palyginti tiriamajai genominei sekai su referentine seka. <i>ExpASy Translate</i> : Įrankis, skirtas nukleotidų (DNR / RNR) sekai paversti baltymų seka.	[102]
<i>Clustal Omega</i>	<a href="https://www.ebi.ac.uk/Tools/msa/clustalo/">https://www.ebi.ac.uk/Tools/msa/clustalo/</a>	Įrankis, skirtas konservatyvumo analizei atlikti, lyginant RNR, DNR ar aminorūgščių sekas.	[252]
<i>Pfam v 32.0</i>	<a href="https://pfam.xfam.org/">https://pfam.xfam.org/</a>	Baltymų, jų funkcinių domenų, baltymų šeimų katalogas.	[104]
<i>UniProt</i>	<a href="https://www.uniprot.org/">https://www.uniprot.org/</a>	Kaupiama informacija apie baltymų sekas ir jų funkcijas.	[103]
<i>LiteMol</i>	<a href="https://www.litemol.org/">https://www.litemol.org/</a>	Įrankis, skirtas generuoti trimatės baltymų struktūros modeliams.	[239]
<i>RegNetwork</i>	<a href="http://regnetworkweb.org/about.jsp">http://regnetworkweb.org/about.jsp</a>	Žmogaus ir pelių reguliacinių elementų, veikiančių transkripcijos ir <i>posttranskripciniu</i> lygmeniu, duomenų bazė.	[276]

#### IV PRIEDAS (tęsinys). Tyrimų metu naudotos duomenų bazės ir bioinformaciniai įrankiai

Pavadinimas	Nuoroda internetiniame puslapyje	Trumpas apibūdinimas	Cita- vimas
<b>Duomenų bazės genotipo ir fenotipo ryšio analizei atlikti</b>			
<i>NCBI ClinVar</i>	<a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a>	Informacija apie ryšį tarp žmogaus variantų ir fenotipo.	[8]
<i>NCBI OMIM</i>	<a href="https://www.ncbi.nlm.nih.gov/omim/">https://www.ncbi.nlm.nih.gov/omim/</a>	Žmogaus genų ir fenotipų santraukų duomenų bazė.	[91]
<i>ORPHANET</i>	<a href="https://www.orpha.net/consor/cgi-bin/index.php?lng=EN">https://www.orpha.net/consor/cgi-bin/index.php?lng=EN</a>	Retų ligų, susijusių klinikinių požymių ir vaistų portalas.	[291]
<i>HGMD</i>	<a href="http://www.hgmd.cf.ac.uk/ac/index.php">http://www.hgmd.cf.ac.uk/ac/index.php</a>	Žmogaus genų variantų duomenų bazėje pateikiama informacija apie visus žinomus variantus, kurie yra siejami su žmogaus paveldimomis ligomis.	[34]
<i>DECIPHER</i>	<a href="https://decipher.sanger.ac.uk/">https://decipher.sanger.ac.uk/</a>	Kaupiami daugiau nei 35 tūkst. tiriamų asmenų genotipo ir fenotipo duomenys.	[108]
<i>Phenolyzer</i>	<a href="http://phenolyzer.wglab.org/">http://phenolyzer.wglab.org/</a>	Bioinformacinis įrankis, skirtas nustatyti ryšiui tarp genotipo ir fenotipo.	[110]
<i>The Human Phenotype Ontology</i>	<a href="http://human-phenotype-ontology.github.io/">http://human-phenotype-ontology.github.io/</a>	Duomenų bazė, skirta klinicinei diagnostikai bei bioinformacinei fenotipo ir ląstelinių / biocheminių mechanizmų sąsajų analizei pagal fenotipinius požymius.	[111]
<i>CHD7 database</i>	<a href="https://molgenis51.gcc.rug.nl/">https://molgenis51.gcc.rug.nl/</a>	Kaupiami duomenys apie <i>CHD7</i> variantus ir su šiuo genu siejamus fenotipus.	[225]
<i>LOVD v.3.0</i>	<a href="https://www.lovd.nl/">https://www.lovd.nl/</a>	<i>LOVD</i> (angl. <i>Leiden Open Variation Database</i> ) duomenų bazėje kaupiama informacija apie ryšį tarp DNR sekos variantų ir fenotipo.	[249]

## PADĖKA

Šv. Jonas Paulius II (1920–2005) yra pasakęs: „*Žmogus yra didis ne tuo, kiek turi turtų, bet tuo, kuo jis yra; ne tuo, ką turi, bet tuo, kuo dalinasi su kitais.*“ Todėl nuoširdžiai dėkoju visiems, kurie dalinosi savo žiniomis ir patirtimi man einant genetikos mokslo keliu, o ypač esu dėkinga:

doktorantūros studijų moksliniam vadovui akad. prof. habil. dr. Vaidučiui Kučinskui, kuris sugalvojo mokslinio darbo idėją ir suteikė galimybę ją įgyvendinti. Taip pat dėkoju gerbiamam Profesoriumi, kad nukreipė mane reikiama mokslo linkme, visada geranoriškai atsakydavo į iškilusius klausimus bei skatino dalyvauti mokslinėse stažuotėse ir konferencijose;

mokslinio darbo konsultantei ir INGENES projekto vadovei doc. dr. Eglei Preikšaitienei, kuri suteikė galimybę tapti šio projekto komandos dalimi. Gerbiamai Docentei dėkoju ne tik už pasitikėjimą, padaršinimą bei vertingus patarimus rengiant disertaciją, bet ir už turimas žmogiškas vertybes: draugiškumą, nuoširdumą ir supratingumą;

visiems INGENES projekto vykdytojams, kurių dėka gavau reikšmingos patirties visam gyvenimui. Esu nuoširdžiai dėkinga dr. Violetai Mikštienei, kuri išmokė naujų darbo metodų bei visada nuoširdžiai, rūpestingai ir geranoriškai konsultavo įvairiais iškilusiais techniniais klausimais; doc. dr. Živilei Maldžienei už nuoširdumą, draugiškumą ir bendrą darbą laboratorijoje; dr. Tautvydui Rančeliui už pagalbą sprendžiant klausimus, susijusius su bioinformacine analize. Taip pat širdingai dėkoju bendramintei draugei dokt. Gundai Petraitytei už draugiškumą, nuoširdumą, patarimus, palaikymą bei produktyvų bendrą darbą laboratorijoje;

disertacijos Gynimo tarybos pirmininkei prof. (HP) dr. Loretai Cimbalistienei bei nariams prof. dr. Vytautui Kasiulevičiui, prof. dr. Arvydui Kaminskui, prof. habil. dr. Limui Kupčinskui, prof. dr. Jāniui Klovinš'ui už įdomius klausimus diskusijai;

recenzentėms prof. dr. Rūtai Mameniškienei, dr. Ingridai Domarkienei ir dr. Birutei Burnytei už pastabas ir vertingus pasiūlymus, kurie padėjo patobulinti disertaciją;

visiems tiriamiesiems, kurie sutiko dalyvauti tyrime ir leido analizuoti savo genetinius ir klinikinius duomenis;

VU MF BMI ŽMGK vedėjui prof. (HP) dr. Algirdui Utkui už suteiktą galimybę įgyvendinti mokslinio darbo idėją;

visam VU MF BMI ŽMGK ir VšĮ VUL SK MGC kolektyvui už suteiktą pagalbą ir kompetentingus patarimus. Išskirtinė padėka yra dr. Valentinai Ginevičienei, dr. Alinai Urnikytei ir dr. Laurai Pranckėnienei už

palaikymą ir gerą nuotaiką darbe. Taip pat esu dėkinga Daliai Kuzminskienei už kantrybę ir vertingą pagalbą sprendžiant doktorantūros studijų metu iškilusius klausimus;

doc. dr. Aušrai Sasnauskienei ir dokt. Vilmantei Žitkutei už galimybę atlikti dalį šio darbo eksperimentų VU GMC bei suteiktą pagalbą vykdant mokslinius tyrimus;

prof. Alexandre'ui Reymond'ui, dr. Lucie Gueneau ir dr. Norine Voisin už suteiktą galimybę stažuotis Lozanos universiteto Integratyvosios genomikos centre. Šioje stažuotėje ne tik įgijau unikalių žinių ir praktinio darbo įgūdžių, bet ir praplėčiau savo draugų ratą;

pirmiesiems genetikos mokytojams – prof. habil. dr. Juozui Rimantui Lazutkai, prof. (HP) dr. Gražinai Slapšytei, dr. Gabrieliui Kundrotui, doc. dr. Veronikai Dedonytei, kurie ne tik užkrėtė meile genetikos mokslui, bet ir visada padrąsindavo šiltu bendravimu. Taip pat doc. dr. Veronikai Dedonytei dėkoju už atliktą seserinių chromatidžių mainų tyrimą;

savo mylimiausiai šeimai ir artimiausiems draugams dėkoju už palaikymą ir supratimą. Už kantrybę, įkvėpimą bei skatinimą siekti gyvenimo tikslų ir svajonių ypač dėkoju vyrui Markui, dukrai Olivijai ir mamai Leokadijai – Jūsų dėka nesustoju net ir sunkiausią akimirką.

Vilniaus universiteto leidykla  
Saulėtekio al. 9, LT-10222 Vilnius  
El. p. [info@leidykla.vu.lt](mailto:info@leidykla.vu.lt), [www.leidykla.vu.lt](http://www.leidykla.vu.lt)  
Tiražas 15 egz.