

VILNIAUS UNIVERSITETAS

Gunda Petraitytė

Naujos kartos sekoskaitos metodais
nustatytų retų variantų patogeniškumo
vertinimas ir patogenezės mechanizmų
tyrimas atliekant retų variantų
molekulinį ir funkcinį charakterizavimą

DAKTARO DISERTACIJA

Medicinos ir sveikatos mokslai

Medicina (M 001)

VILNIUS 2025

Disertacija rengta 2019–2024 metais Vilniaus universiteto Medicinos fakulteto Biomedicinos mokslų institute.

Mokslinius tyrimus rėmė Lietuvos mokslo taryba (S-MIP-21-15 ir S-MIP-17-19), Vilniaus universiteto mokslo skatinimo fondas (MSF-JM-2/2020).

Mokslinė vadovė – prof. dr. Eglė Preikšaitienė (Vilniaus universitetas, medicinos ir sveikatos mokslai, medicina, M 001).

Mokslinė konsultantė – doc. dr. Aušra Sasnauskienė (Vilniaus universitetas, gamtos mokslai, biochemija, N 004).

Gynimo taryba:

Pirmininkas – prof. dr. Juozas Rimanats Lazutka (Vilniaus universitetas, gamtos mokslai, biologija N 010).

Nariai:

prof. dr. Audronė Jakaitienė (Vilniaus universitetas, gamtos mokslai, informatika N 009),

dr. Dalia Seaborne (Genomics England įmonė, Didžioji Britanija, medicinos ir sveikatos mokslai, medicina, M 001),

doc. dr. Sonata Šaulytė–Trakymienė (Vilniaus universitetas, medicinos ir sveikatos mokslai, medicina M 001),

prof. dr. Rasa Ugenskienė (Lietuvos sveikatos mokslų universitetas, medicinos ir sveikatos mokslai, medicina M 001).

Disertacija ginama viešame Gynimo tarybos posėdyje 2025 m. gruodžio mėn. 23 d. 11 val. Vilniaus universiteto Medicinos fakulteto Medicinos mokslo centre Oranžinėje auditorijoje. Adresas: Žaliųjų Ežerų g. 2, Vilnius, Lietuva, tel. +37052398700; el. paštas mf@mf.vu.lt.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: <https://www.vu.lt/naujienos/ivykiu-kalendorius>

VILNIUS UNIVERSITY

Gunda Petraitytė

Molecular and Functional Characterization of Rare Variants Identified Using Next-generation Sequencing to Determine the Pathogenicity of Rare Variants and Disease Mechanism

DOCTORAL DISSERTATION

Medical and Health Sciences
Medicine (M 001)

VILNIUS 2025

The dissertation was prepared between 2019 and 2024 at Vilnius University.

The research was supported by the Research Council of Lithuania (S-MIP-21-15 and S-MIP-17-19) and the Research Promotion Fund of Vilnius University (MSF-JM-2/2020).

Academic Supervisor – Prof. Dr. Eglė Preikšaitienė (Vilnius University, Medical and Health Sciences, Medicine, M 001).

Academic Consultant – Assoc. Prof. Dr. Aušra Sasnauskienė (Vilnius University, Natural Sciences, Biochemistry, N 004).

This doctoral dissertation will be defended in a public meeting of the Dissertation Defence Panel:

Chairman – Prof. Habil. Dr. Juozas Rimanats Lazutka (Vilnius University, Natural Sciences, Biology, N 010).

Members:

Prof. Dr. Audronė Jakaitienė (Vilnius University, Natural Sciences, Informatics, N 009),

Dr. Dalia Seaborne (Genomics England, Great Britain, Medical and Health Sciences, Medicine, M 001),

Assoc. Prof. Dr. Sonata Šaulytė–Trakymienė (Vilnius University, Medical and Health Sciences, Medicine, M 001),

Prof. Dr. Rasa Ugenskienė (Lithuanian University of Health Sciences, Medical and Health Sciences, Medicine, M 001).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 11 a.m. on the 23rd of December 2025 in the Orange auditorium of the Medical Science Center of Vilnius University.

Address: Žaliųjų Ežerų g. 2, Vilnius, Lithuania, tel. +37052398700; e-mail: mf@mf.vu.lt.

The text of this dissertation can be accessed at the libraries of Vilnius University, as well as on the website of Vilnius University: www.vu.lt/lt/naujienos/ivykiu-kalendorius

SANTRUMPOS

ACMG – Amerikos medicinos genetikos ir genomikos kolegija (angl. *American College of Medical Genetics and Genomics*)

AMP – Amerikos patologijos asociacija (angl. *Association of Molecular Pathology*)

ClinVar – NCBI duomenų bazė, kurioje kaupiami duomenys apie genomo variantų patogeniškumą

DAPI – (angl. *4',6-diamidino-2-phenylindole*) 4,6-diamino-2-fenilindolas

DMEM – (angl. *Dulbecco's Modified Eagle Medium*) ląstelių kultivavimo *Dulbecco* modifikuota erelio terpė

EDTA – (angl. *Ethyleneiaminetetraacetic Acid*) etilendiamintetraacto rūgštis

ENCODE – tarptautinis funkcinį genomo elementų analizės konsorciūmas (angl. *Encyclopedia of DNA Elements*)

FDR – (angl. *False Discovery Rate*) klaidingų atradimų dažnis

gnomAD – Genomo agregacijos duomenų bazė (angl. *Genome Aggregation database*)

HGMD – Žmogaus genų patogeninių variantų duomenų bazė (angl. *The Human Gene Mutation Database*)

inRNR – ilga nekoduojanti RNR (angl. *long non-coding RNA*)

iPSC – indukuotos daugiagalės kamieninės ląstelės (angl. *induced Pluripotent Stem Cells*)

IRDiRC – Tarptautinis retų ligų tyrimų konsorciūmas (angl. *The International Rare Diseases Research Consortium*)

JSA – jaučio serumo albuminas

kDNR – kopijinė DNR

kPGR – kiekybinė PGR

KSP – kopijų skaičiaus pokyčiai

lcDNR – laisvai cirkuliuojanti DNR (angl. *cell-free DNA*)

LGH – lyginamoji genomo hibridizacija

mtDNR – mitochondrinė DNR

NCBI – Nacionalinis biotechnologinės informacijos centras (angl. *National Center for Biotechnology Information*)

NDS – natrio dodecilsulfatas

NKS – naujos kartos sekoskaita

NMD – priešlaikinės baigmės kodono sąlygota iRNR degradacija (angl. *Nonsense-mediated mRNA decay*)

PBS – (angl. *Phosphate-Buffered Saline*) fosfatinis buferinis tirpalas

PGR – polimerazės grandininė reakcija

PI – propidžio jodidas

RIPA – (angl. *Radioimmunoprecipitation Assay buffer*) radio imunoprecipitacijos buferinis tirpalas

SV – struktūrinis variantas

VES – viso egzomo sekoskaita

VGS – viso genomo sekoskaita

vLGH – vektorinė lyginamoji genomo hibridizacija

VNV – vieno nukleotido variantai

VUS – neaiškios klinikinės reikšmės variantas (angl. *Variant of Uncertain Significance*)

X-gal – (angl. *5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside*) β galaktozidazės substratas

TURINYS

ĮVADAS.....	10
1. LITERATŪROS APŽVALGA	14
1.1 Retas paveldimas ligas lemiančių genų kandidatų atranka	14
1.2 Naujos kartos sekoskaitos metodai / strategijos retos ligos genetinei kilmei nustatyti, jų pranašumai ir trūkumai	17
1.3 DNR sekos variantų patogeniškumo nustatymo problematika	22
1.4 Funkciniai tyrimai DNR variantų patogeniškumui patvirtinti	26
1.4.1 Tyrimai molekulinio lygmeniu	27
1.4.2 Tyrimai ląstelių lygmeniu	29
2. METODAI.....	34
2.1 Tyrimo pagrindas ir tiriamieji.....	34
2.2 Tyrimo metodai	34
2.2.1 DNR gryninimas	36
2.2.2 Laisvai cirkuliuojančios DNR gryninimas.....	36
2.2.3 RNR gryninimas	36
2.2.4 Kopijinės DNR sintezė	36
2.2.5 Fibroblastų kultūros kultivavimas ir užšaldymas.....	37
2.2.6 PGR.....	37
2.2.7 PGR produkto gryninimas iš agarozės gelio	38
2.2.8 Sangerio sekoskaita.....	38
2.2.9 Kiekybinė PGR	38
2.2.10 Naujos kartos sekoskaita.....	39
2.2.11 Western blot analizė.....	39
2.2.12 Ląstelių transfekcija mažomis interferuojančiomis RNR ..	39
2.2.13 Ląstelių transfekcija plazmidėmis.....	40
2.2.14 Tėkmės citometrija.....	40
2.2.15 Fibroblastų senėjimo tyrimas	41
2.2.16 Imunofluorescencinė mikroskopija.....	41

2.2.17 Ląstelių elektrofiziologiniai matavimai „patch-clamp“ metodu	42
2.2.18 Bioinformatinė analizė	42
2.2.19 Statistinė duomenų analizė	43
3. REZULTATAI IR JŲ APTARIMAS	44
3.1 Tiriamųjų apibūdinimas ir atrinkti genai kandidatai tolesniems tyrimams	44
3.2 Genomo struktūrinių persitvarkymų apibūdinimas	56
3.2.1 KSP charakterizavimas	56
3.2.2 Chromotripsio charakterizavimas ir jo įtakos genų raiškai analizė	68
3.3 VNV įtakos baltymo lygmeniu nustatymas bioinformatiniais ir molekuliniiais tyrimais	76
3.3.1 <i>RLIM</i> gene nustatyto VNV įtakos baltymo–baltymo sąveikai ir X chromosomos inaktyvacijai analizė	76
3.3.2 <i>NAA15</i> gene nustatyto varianto nulemti koduojamo baltymo raiškos pokyčiai	79
3.4 VNV ir nedidelės iškritos poveikio ląstelės procesams įvertinimas	81
3.4.1 <i>WRN</i> gene nustatytos vieno nukleotido iškritos poveikio ląstelių senėjimui analizė	82
3.4.2 <i>DYNCH1</i> gene nustatyto sukirpimo vietos varianto įtaka KIFC1 ir KIFC3 baltymų raiškai bei Goldžio komplekso lokalizacijai	84
3.4.3 <i>FAS</i> gene nustatytos nedidelės iškritos įtaka apoptozės indukcijai	93
3.4.4 <i>DNMT3A</i> gene nustatyto varianto įtakos metilinimo sutrikimui ląstelėse analizė	98
3.4.5 <i>KCNQ1</i> gene nustatyto varianto lemiami kalio jonų judėjimo ląstelių membrana sutrikimų analizė	106
IV IŠVADOS	110
LITERATŪROS SĄRAŠAS	112
PRIEDAI	129
SUMMARY OF DOCTORAL DISSERTATION	139

PUBLIKACIJŲ SĄRAŠAS	200
PRANEŠIMŲ SĄRAŠAS	202
CURRICULUM VITAE	203
PADĖKA.....	205

IVADAS

Tiriamoji problema ir aktualumas

Retos ligos yra plati sveikatos sutrikimų grupė, kuriai būdingas didelis heterogeniškumas keliais aspektais: geografiniu paplitimu, ligos etiologija ir fenotipinių bruožų įvairove bei jų sunkumo laipsniu. Dažniausiai šios ligos pasireiškia ankstyvame amžiuje, sukelia lėtines sveikatos būkles, neretai paveikia nervų sistemą ir varijuoja priklausomai nuo individualių sergančiojo biologinių ypatybių [1, 2]. Visas retas ligas vienijantis bruožas yra mažas jų paplitimas. Europoje reta liga yra laikoma ta, kuria serga rečiau nei vienas iš 2000 asmenų bendrojoje populiacijoje (Orphanet duomenų bazė). Mažas šių ligų dažnis yra viena iš pirminių priežasčių, apsunkinusių jų pažinimo procesą, kadangi lėmė nevieningas pastangas pasaulyje, skirtas ištirti retų sutrikimų priežastis ir vystymąsi, kurti naujus gydymo būdus ir sveikatos priežiūros strategijas, taip siekiant palengvinti sergančiojo diagnostinį kelią, neretai trunkantį apie penkerius metus [2]. Vis dėlto skaičiuojama, kad šiuo metu 48-iose Europos šalyse reta būkle ar liga gali sirgti apie 30 mln. asmenų [2]. Taigi, retos ligos individualaus sergančiojo asmens ir visuomenės požiūriu yra problema ir didelis diagnostinis iššūkis tiek pasaulio, tiek Lietuvos sveikatos priežiūros sistemai, nes reikalauja gausių finansinių, laiko ir žmogiškųjų išteklių. Be to, ilgas kelias iki diagnozės patvirtinimo kelia emocinių sunkumų sergančiajam ir jo artimiesiems.

Per pastarąjį dvidešimtmetį vystytos ir toliau tobulėjančios molekulinės genetikos (pavyzdžiui, naujos kartos sekoskaita (NKS)) ir citogenetikos (pavyzdžiui, lyginamoji genomo hibridizacija (LGH)) technologijos iš esmės pakeitė biologijos, genetikos ir medicinos mokslus. Tai sukėlė neabejotinai reikšmingą proveržį ir retų ligų srityje: paskatino jų genetinio charakterizavimo tyrimus, pritraukė daugiau ne tik mokslininkų, bet ir visuomenės susidomėjimo. Remiantis disertacijos rašymo metu analizuota literatūra, 72–80 proc. retų ligų ar būklių yra nulemti genetinių (paveldimų) priežasčių [2, 3]. Didžioji dalis retų paveldimų ligų yra monogeninės, t. y. jas lemia konkretaus geno vieno nukleotido variantai (VNV), rečiau nustatomos nedidelės iškritos/intarpai ar struktūriniai variantai (SV), taip pat kopijų skaičiaus pokyčiai (KSP) [4, 5]. Remiantis Tarptautinio retų ligų tyrimų konsorciumo (*IRDiRC*) duomenimis, su retomis paveldimomis būklėmis jau susieta apie 4500 genų, tačiau didžiosios dalies šiuose genuose nustatytų variantų įtaka geno koduojamo baltymo funkcijai ir su juo susijusiems biologiniams procesams yra neištirta. Todėl dažnai susiduriama su problema, kai nustačius unikalų ir retą DNR sekos pokytį, jo klinikinė reikšmė arba

patogeniškumas yra neįrodytas. Šis procesas yra dinamiškas ir atliekamas vadovaujantis Amerikos medicininės genetikos ir genomikos kolegijos (ACMG) aprašytais gairėmis, kurios remiasi bioinformatine analize, sukauptomis žiniomis apie populiacijų ir ligų genetinę įvairovę, skirtingais biologinių procesų moksliniais tyrimais [6]. Nors NKS, kai taikiny yra egzomas ar genomas, leidžia pasiekti 24–50 proc. diagnostinį efektyvumą, maždaug 22 proc. nustatytų variantų priskiriami neaiškos klinikinės reikšmės variantų klasei [7-9]. Tai dažniausiai lemia žinių apie genetinės srities funkcionalumą, taip pat gene nustatyto varianto sukeliamą molekulinį patogenezę mechanizmą stoka. Verta pabrėžti, kad iš esmės pagrįsta molekulinė diagnozė gaunama tik nustčius DNR pokyčius, kurie priskirtini patogeninių variantų klasei. Todėl, net ir galimai patogeninių ar galimai nepatogeninių variantų, nustatytų žinomuose ligos genuose, atveju yra aktualūs ir vertingi papildomi funkciniai tyrimai, kuriais siekiama nuodugniau patikrinti DNR varianto poveikį molekuliniam ir biologiniam procesams.

Taigi, pažangiausios molekulinės genetikos ir citogenetikos technologijos sudarė sąlygas plačiu mastu tyrinėti retų paveldimų ligų genetines charakteristikas, taip pat patikslinti šių ligų klasifikaciją. Vis dėlto tik sekoskaita (egzomo, genomo) ir bioinformatinė rezultatų analizė neužtikrina galutinės diagnozės. Kombinuojant genetinę, bioinformatinę ir funkcinę analizę dažniau gaunama patvirtinta molekulinė diagnozė, kuri yra svarbi efektyviam genetiniam konsultavimui ir pacientui. Ligos patogenežės mechanizmo supratimas leidžia ne tik patvirtinti ar atmesti reto įgimto sutrikimo diagnozę, bet yra svarbus epistemiologiniu požiūriu. Retų genetinių ligų tyrimai gali suteikti žinių ir įžvalgų apie sudėtingus molekulinis procesus, kurie aktualūs ir labiau paplitusių ar daugiaveiksnių ligų bei įprastų požymių atžvilgiu. Platesnis ir gilesnis retų paveldimų sutrikimų vystymąsi lėmusių procesų pažinimas svarbus naujoms biologinėms terapijoms kurti šių patologijų gydymo srityje.

Tyrimo naujumas ir reikšmė

Šiame darbe, lyginant su panašaus pobūdžio moksliniais tyrimais Lietuvoje, genetinei įtariamai retos paveldimos patologijos kilmei nustatyti taikyta ne tik LGH, genų grupių ir egzomo sekoskaita ar egzomo duomenų reanalizė – kai kuriems tiriamiesiems atlikta genomo sekoskaita. Taikytos sekoskaitos strategijos leido nustatyti anksčiau literatūroje neaprašytą KSP, SV ir VNV žinomuose ligos genuose. Užsienio mokslininkų neaprašytiems, galimai patogeniniams ar neaiškos klinikinės reikšmės KSP, SV ir VNV,

nustatyti galimai įtaką kitų genų raiškai turinčiuose genuose naudotas ne tik kiekybinės PGR metodas, bet taikyta transkriptomo sekoskaita, kuri nėra plačiai naudojama retų genetinių ligų srityje Lietuvoje. Šie tyrimai leido platesniu mastu analizuoti molekulinius procesus ir ištirti mažai literatūroje aprašytus patogenezės mechanizmus, kuriuos lėmė šiame darbe nustatyti konkrečių genų DNR pokyčiai. DNR sekos pokyčių analizėje taikytos skirtingos funkcinų tyrimų strategijos, kuriose pasitelkiami metodai, leidžiantys analizuoti įvairius molekulinius procesus ląstelėse (konfokalinė mikroskopija, tėkmės citometrija, imunoblotingas, genų raiškos mažinimas naudojant siRNA) – tai, turimomis žiniomis, Lietuvoje anksčiau netaikyta retų genetinių ligų srityje.

Darbe atliktas retais paveldimais sutrikimais sergančiųjų genotipinis ir atrinktų genų sekos pokyčių molekulinis ir/ar funkcinis charakterizavimas yra indėlis į šių sutrikimų genetinės įvairovės ir molekulinės biologijos kompleksiskumo pažinimą. Aprašytas mokslinis tyrimas parodo retų paveldimų ligų funkcinų tyrimų galimybes mokslinės produkcijos atžvilgiu, kurios dar nėra visiškai išnaudojamos Lietuvoje. Fundamentinis retų paveldimą ligą lėmusio patogenezės mechanizmo ištyrimas yra tęstinis procesas, kuris gali tapti pagrindu naujiems moksliniams tyrimams, naujos kartos gydymo strategijoms ir vaistams kurti bei toks tyrimo modelis gali būti pritaikytas ir kitų retų ligų tyrimams.

Tyrimo tikslas – įvertinti su tiriamųjų retomis ligomis siejamuose genuose nustatytų DNR sekos pokyčių patogeniškumą, reikšmę RNR, baltymų ir ląstelėse vykstančių procesų lygmeniu atliekant bioinformatinius, molekulinius ir/ar funkcinus tyrimus.

Tyrimo uždaviniai:

1. Atlikti naujos kartos sekoskaitos tyrimais nustatytų DNR sekos pokyčių patogeniškumo vertinimą naudojant duomenų bazių, literatūros šaltinių, *in silico* analizę ir baltymų struktūros modeliavimą bei atrinkti genus kandidatus funkciniam tyrimams.
2. Išanalizuoti genomo struktūrinius persitvarkymus ir jų reikšmę genams kandidatams nustatyti.
3. Nustatyti atrinktų DNR sekos pokyčių reikšmę baltymo lygmeniu taikant bioinformatinių ir molekulinų tyrimų metodus.
4. Įvertinti atrinktų DNR sekos pokyčių įtaką ląstelėje vykstančių procesų lygmeniu atliekant molekulinius ir/ar funkcinus tyrimus.

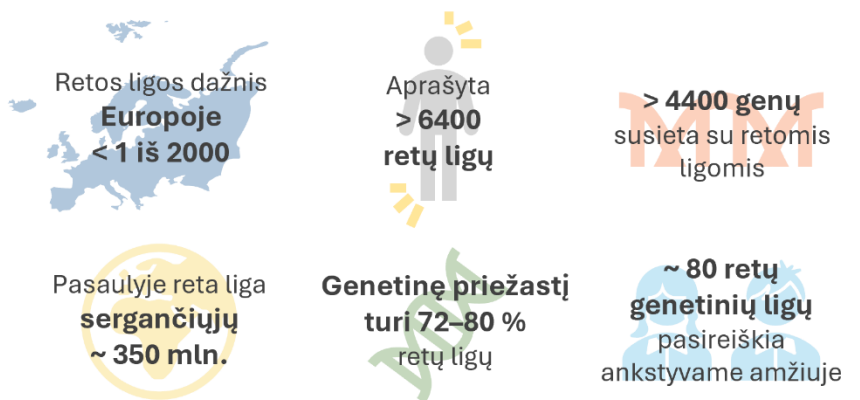
Ginamieji teiginiai:

1. Genų kandidatų atrankai funkciniam tyrimams didelę įtaką turi analizuotuose ligos genuose nustatytų variantų patogeniškumo įvertinimas, DNR pokyčio naujumas ir fundamentalios žinios bei aprašytų funkcinių tyrimų rezultatai apie geno koduojamą baltymą, jo vaidmenį ląstelėse.
2. Genomo struktūriniai persitvarkymai derinant su tiriamojo klinikiniu fenotipu reikšmingai susiaurina genų kandidatų paieškos lauką, taip pat potencialiai nusako reguliacinių genomo elementų įtaką paveldimos ligos patogenezei.
3. DNR sekos variantų, keičiančių aminorūgščių seką ar lemiančių jų praradimą, sąlygotą baltymo netaisyklingą susilankstymą ar degradaciją įvertinti tikslinga taikyti bioinformatinius metodus, tokius kaip baltymo struktūros modeliavimo įrankius ir konservatyvumo analizės algoritmus, bei molekulinis metodas baltymo kiekiui ir funkcijai nustatyti, kadangi jų kombinacija įgalina atskleisti DNR variantų biologinę reikšmę ir galimą patogenezės mechanizmą.
4. Molekuliniai ir funkciniai DNR variantų tyrimai, pavyzdžiui, raiškos analizė RNR ir baltymų lygmeniu bei molekulinis procesų, kuriuose dalyvauja pakitę baltymai, tyrimai *in vitro* sudaro galimybę įvertinti unikalius patogenezės mechanizmo aspektus ląstelėse.

1. LITERATŪROS APŽVALGA

1.1 Retas paveldimas ligas lemiančių genų kandidatų atranka

Remiantis Orphanet duomenų baze, kurios duomenis seka ir naudoja IRDiRC (angl. *The International Rare Diseases Research Consortium*), kasmet aprašoma keliasdešimt naujų retų ligų ir daugiau nei šimtas naujų genų yra susiejama su retomis paveldimomis ligomis (duomenys nuo 2016 m. iki 2022 m.). Ši statistika iliustruoja didelį retų įgimtų ligų genetinį heterogeniškumą. Be to, vieną ligą gali lemti keliuose skirtinguose genuose nustatyti variantai arba tame pačiame gene, tik kitose jo pozicijose nustatyti variantai lemia skirtingus fenotipus. Todėl supratimas ir gebėjimas taikyti įvairius būdus genams kandidatams atrinkti yra neabejotinai svarbus bei tampriai susijęs procesas su ligą lemiančių DNR pokyčių nustatymu retų paveldimų sutrikimų srityje.



1.1 paveikslas. Pagrindiniai retas ligas apibūdinantys statistiniai duomenys. Parengta remiantis Orphanet duomenų baze ir Lee et al. 2020 [1].

Taikant šiuolaikinius metodus ir tyrimo strategijas, ligos genų atradimo mastai ir greitis neabejotinai skiriasi nuo naudotų iki 2003-ųjų, kai buvo užbaigtas Žmogaus genomo projektas ir pirmą kartą nusekvenuotas žmogaus genomas. Genetinės sankibos analizė, kurios principas yra nepriklausomas požymių paveldėjimas ir rekombinacija, buvo vienas iš pagrindinių genetinio kartografavimo būdų. Paremtas statistiniais skaičiavimais ir sudėtinga laboratorinių tyrimų dalimi, šis metodas sudaro galimybę susieti tiriamojo fenotipą su specifine chromosomine vieta, kurioje, tikėtina, slypi ligos genas [10, 11]. Norint atlikti rezultatyvią ir statistiškai reikšmingą sankibos analizę,

reikalingas išankstinis ligos paveldėjimo ir geno regiono numatymas (išankstinė hipotezė), būtina surinkti ir ištirti bent kelias, patikimiau dideles šeimas, kuriose yra ne vienas specifinį fenotipą turintis šeimos narys – tai itin sudėtinga retų paveldimų ligų atžvilgiu. Net ir surinkus pakankamą tiriamųjų grupę, genetinei sankibai nustatyti turi būti parinkti genetiniai žymenys (pavyzdžiui, restrikcijos fragmentų ilgio polimorfizmai), kurių molekulinė ir statistinė analizė leistų susieti fenotipą su specifine vieta chromosomoje, toliau atlikus smulkesnį genotipavimą ir pozicinį klonavimą patikimai susiaurintį regioną [12]. Genetinės sankibos analizė yra istoriškai itin svarbus ir reikšmingas, bet kartu sudėtingas būdas genams kandidatams atrinkti, turintis nemažai trūkumų, pavyzdžiui, nėra tikslinga taikyti ne pagal Mendelio dėsnius ir ne monogeninėms ligoms tirti bei esant nepilnam penetrantiškumui. Taip pat šios analizės mastas geno atžvilgiu yra nedidelis ir atliekamas sergančiųjų šeimose, todėl nėra pilnai įvertinamas genetinis heterogeniškumas – nėra atskleidžiamas visas ligos etiologijos vaizdas [12].

NKS technologijos, kuriomis vienu metu galima analizuoti koduojančią geno dalį (egzomas), genų rinkinius, mitochondrinę DNR (mtDNR) arba visą genomą bei epigenomą ar net transkriptomą, yra populiarius ir plačiai taikomas metodas naujausiuose tyrimuose atliekant genų kandidatų paiešką bei išankstinių hipotezių apie ligos paveldėjimą ar genominę poziciją [13]. Vieno tiriamojo egzomo sekoskaitos duomenyse gali būti nustatyti apie 40–60 tūkst. DNR pokyčių, o geno sekoskaitos duomenyse sugeneruojama apie 4–5 mln. DNR variantų [14, 15]. Todėl būtina tikslingai pasirinkti ir atsakingai atlikti žingsnius bei strategijas, kuriomis galima būtų efektyviai ir patikimai atrinkti konkrečius geno regionus (genus) ir DNR pokyčius, sietinus su tiriamojo retu klinikiu fenotipu. Mažų (<1000bp) KSP analizė, taikant chromosominį mikrogardelių (angl. *chromosomal microarray*) metodą viso geno mastu, leidžia ne tik nustatyti genetinę retos ligos priežastį žinomuose genuose, bet ir prisideda nustatant naujus, su liga nesusietus genus [16]. Viena iš ligos geno nustatymą lengvinančių strategijų yra kartu tirti sergančiojo šeimos narius, pavyzdžiui, triadą, tetradą. Taip kartu atliekama ir segregacijos šeimoje analizė, leidžianti nustatyti *de novo* variantus, kurie yra dažna retų ligų kilmės priežastis, gaunama informacija apie *cis* ir *trans* varianto padėtį recesyvaus paveldėjimo atveju, taip pat sveiki šeimos nariai gali būti kaip atskira kontrolinių asmenų grupė [13]. Neretai taikoma strategija yra egzomo ir geno duomenų analizė pagal galimus paveldėjimo tipus.

Retas įgimtas patologijas lemiantys variantai daugeliu atveju bendrojoje populiacijoje yra reti (dažnis <1 proc.), todėl svarbus žingsnis yra populiacijų tyrimų duomenų analizė, siekiant patikrinti nustatyto varianto dažnį nesergančių asmenų grupėse ir, esant galimybei, dažnį nustatyti tiriamojo

kilmės populiacijoje. Tai svarbu, nes demografiniai, istoriniai ir evoliuciniai veiksniai turi įtakos variantų pasiskirstymui ir dažniui skirtingose populiacijose [1]. Genų kandidatų atrankoje susiaurinti paiešką padeda turimų žinių apie genominę sritį surinkimas ir bioinformatinių įrankių taikymas. Pastarieji sudaro galimybę vertinti varianto pozicijos evoliucinį konservatyvumą geno koduojamame produkte, poveikį produkto sekai ir struktūrai, konformacijos dinamikai, sąveikai su kitomis molekulėmis [17]. Yra sukurti algoritmai, kuriais analizuojant struktūriškai ir funkciškai panašius baltymus, sudarančius kompleksus, galima nustatyti komplekso baltyminių elementą (taigi ir jį koduojantį geną), kuris gali turėti įtakos fenotipo vystymuisi, taip nustatyti naują, sietiną su reta liga geną [18]. Nustatyti genus kandidatus pasitelkiama ir transkriptomo bei metilomo sekoskaita. Tai viso transkriptomo (ang. *Transcriptome-wide Association Studies*) ir geno lygmens metilomo asociacijų analizės (angl. *Gene-level Methylome-wide Association analysis*) tyrimai, kurie prisideda prie retus ir labiau paplitusius sutrikimus galimai lėmusių genų paieškos [19, 20]. Kadangi retų paveldimų ligų dažnis yra labai mažas, surinkti pakankamą tiriamųjų grupę, kurios pakaktų padaryti reikšmingą geno su tiriamojo fenotipiniais bruožais sąsają, yra sudėtinga. Todėl neabejotinai svarbu bendradarbiauti su kitų šalių tyrėjais ir dalintis informacija tam specialiai skirtose duomenų bazėse, siekiant efektyviai patvirtinti geno kandidato reikšmę tiriamojo retam sutrikimui [21].

Svarbu pabrėžti, kad siekiant patikimai nustatyti geno sąsają su tiriamojo fenotipu gali būti atliekami integraciniai tyrimai – kombinuojami skirtingi analizės metodai ir strategijos. Pavyzdžiui, šeimų analize pagrįsti genomo asociacijų tyrimai, kuriuose surenkama ne tik didelė tiriamųjų grupė, bet įtraukiamos šeimos, o analizė atliekama kombinuojant asociacijų analizę su genetinės sankibos įvertinimu. Tokiu būdu nustatomi ne tik dažni, bet ir reti variantai, turintys poveikį specifiniam tiriamojo fenotipui vystytis [22]. Egzomo ir genomo sekoskaita kartu su genetinės sankibos analize, suteikia galimybę egzomo arba genomo mastu efektyviai nustatyti konkretų geną kandidatą, kuris turi statistiškai stiprią genetinę sankibą esant didelei tiriamųjų grupei, taip pat išvengiama sekoskaitos procese susidariusių klaidų, kai analizuojami giminytės ryšiais nesusiję tiriamieji [22]. Tačiau bet koku būdu ar strategija nustačius geną kandidatą, jo reikšmė tiriamojo fenotipui patvirtinama tik atlikus funkcinis tyrimus [23]. Tai atskiras genų kandidatų atrankos etapas, kuriuo pereinama į *in vitro* ir *in vivo* tyrimų lygmenis. Šiame etape siekiama ištirti geno kandidato įtaką molekuliniams, biologiniams procesams ląstelėse, fiziologiniams, morfologiniams ir anatominiams pokyčiams organizme [23]. Verta atkreipti dėmesį, kad šiandien genų atranka yra pagrįsta DNR pokyčių nustatymu plačiu mastu (egzomo, genomo,

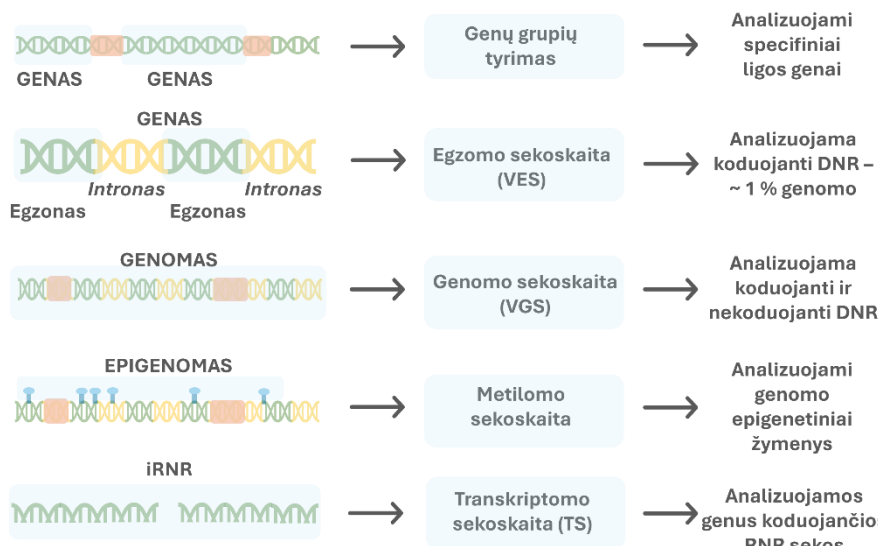
transkriptomo, metilomo), todėl tikslinga funkciškai ištirti ir konkretaus varianto reikšmę geno kandidato koduojamam produktui. Tai svarbu, nes reikšmė gali skirtis priklausomai nuo pokyčio dydžio, tipo, pozicijos, kuri gali lemti skirtingus (net priešingus) fenotipus. Be to, kaip minėta anksčiau, varianto fenotipiniam pasireiškimui įtakos turi ir paties individo genetiniai ypatumai. Todėl net ir su paveldimomis ligomis susietuose genuose nustatytų variantų išsamesnė funkcinė analizė yra itin vertinga ir reikšminga siekiant pilnai suprasti galimą geno vaidmenį retų paveldimų ligų patogenezės mechanizmuose (funkciniai tyrimai plačiau aprašyti tolesniuose disertacijos poskyriuose).

1.2 Naujos kartos sekoskaitos metodai / strategijos retos ligos genetinei kilmei nustatyti, jų pranašumai ir trūkumai

Iki naujos kartos sekoskaitos (NKS) išradimo tyrėjams reikėdavo nepalyginti daugiau laiko ir finansinių resursų nustatyti retos ligos įgimtą priežastį, kadangi pagrindinis įrankis buvo Sangerio sekoskaita. NKS technologijų, pagrįstų trumpų (keliasdešimties ar kelių šimtų) DNR fragmentų sekoskaita, spartus tobulėjimas, galimybė vienu metu nuskaityti didelį kiekį paveldimos informacijos ir palaipsniui mažėjanti šių metodų taikymo kaina moksliniuose tyrimuose ir diagnostikoje suteikė galimybę greitai ir efektyviai nustatyti retų sutrikimų genetinę kilmę.

Klinikinėje praktikoje neretai taikomas NKS būdas tirti retos ligos genetinę priežastį yra genų grupių sekoskaita (žr. 1.2 schemą). Šis metodas yra efektyvesnis laiko, kainos ir duomenų apdorojimo požiūriu, lyginant su VES ir VGS [24]. Genų grupės sudaromos iš žinomų ligos genų, sietinų su tam tikromis retų paveldimų ligų grupėmis ar atskiromis ligomis (pavyzdžiui, neurovystymosi sutrikimai, paveldimos širdies ir kraujagyslių ligos, paveldima spastinė paraplegija). Verta atkreipti dėmesį, kad genų grupių tyrimo atveju dažniausiai sekvenuojami tik genų egzonai, rečiau įtraukiami specifiniai reguliaciniai regionai, netransliuojamos sritys [25]. Dėl šios priežasties pasiekiamas itin didelis egzonų taikinių sekoskaitos padengimas, leidžiantis nustatyti ne tik VNV ar nedideles iškritas/intarpus, bet ir KSP bei mozaikinius variantus [25]. Didelis duomenų padengimas ir tik atrinktų, konkrečių genų analizė sudaro sąlygas rečiau nustatyti atsitiktinius radinius, nesusijusius su tiriamojo ligos fenotipu, taip pat sumažina tikimybę gauti klaidingai teigiamus/neigiamus rezultatus [26]. Diagnostinis genų grupių tyrimo efektyvumas gali būti panašus ar net pranokti VES efektyvumą, tačiau jam stiprią įtaką daro tikslinga tiriamųjų atranka ir išankstinė įžvalga apie genus, lemiančius tiriamojo fenotipą, t. y., stipri sąsaja tarp tiriamojo

fenotipinių ypatybių ir į grupę įtrauktų genų. Tai iliustruoja Jiman ir kolegų [27] atlikta sergančiųjų paveldimomis akių tinklainės ligomis analizė taikant genų grupių sekoskaitą. Bendrai genetinė diagnozė nustatyta 49 proc. tiriamųjų. Tačiau išskyrus dvi tiriamųjų grupes – pirmoji, kuriai priskirti tiriamieji su klinicine diagnoze, ir antroji, kuriai priklauso tiriamieji tik su preliminarine klinicine diagnoze, – diagnostinis efektyvumas pirmosios grupės atveju buvo 71 proc., antrajai nustatytas tik 25 proc. efektyvumas. Vis dėlto akivaizdus genų grupių sekoskaitos trūkumas – nauji ligos genai ir genomo sričių pokyčiai nėra nustatomi, nes analizuojama nedidelė (žymiai mažesnė nei 1 proc. genomo) koduojančios genomo srities dalis, susieta su retomis paveldimomis ligomis. Tačiau šiuo tyrimu yra nustatomi nauji neaprašyti žinomų ligos genų variantai, kurių klinikinė reikšmė neretai turi būti patvirtinama funkciniais tyrimais. Be to, kasmet nustatoma bent keliasdešimt naujų, retus įgimtus sutrikimus lemiančių genų, todėl genų grupės sudėtis turi būti reguliariai peržiūrima ir atnaujinama (Orphanet duomenų bazė). Tad atlikus genų grupių NKS ir nenustačius molekulinės ligos priežasties, turi būti atliekama pakartotinė ar kitų genų grupių sekoskaita arba pereinama prie VES ar VGS tyrimo.



1.2 schema. Dažniausiai pasitelkiami NKS metodai retų ligų kilmės priežastims nustatyti.

Vienas populiariausių NKS metodų yra viso egzomo sekoskaita (VES) (žr. 1.2 schemą), kuri yra pranašesnė už viso genomo sekoskaitą dėl mažesnių laiko ir finansinių resursų, paprastesnės bioinformatinės analizės ir mažesnio kiekio duomenų saugojimo. VES tyrimo objektas grindžiamas žiniomis apie

aprašytas genetines žmogaus ligas, kurių didžiąją dalį (apie 85 proc.) lemia baltymus koduojančiose geno sekose (egzonuose) nustatyti DNR pokyčiai [28]. Egzomas sudaro apie 1 proc. žmogaus geno (apie 30 Mb), todėl sekvenuojant šiuolaikinėmis technologijomis galima pasiekti didesnę sekoskaitos vertikalų padengimą būtent aminorūgštis koduojančiuose egzonuose ir kanoninėse sukirpimo vietose [29]. Kadangi viso egzomo sekoskaitos taikinyje yra koduojančios geno sritys, viso geno atžvilgiu (horizontalus) padengimas yra nedidelis, todėl šiuo tyrimu neanalizuojamos įvairios reguliacinės dalys, kurios gali turėti įtakos genų raiškai ir retos paveldimos ligos patofiziologijai. Taip pat VES pasiekiamas pakankamas padengimas, leidžiantis analizuoti KSP genų srityse ir ląstelių mozaikiškumo nulemtus variantus, taip pat sukurti algoritmai, sudarantys galimybę šiuo metodu įtarti didesnius nei 50 bazių porų struktūrinius variantus (SV) [30, 31]. Remiantis literatūra, egzomo sekoskaitos diagnostinis efektyvumas gali skirtis priklausomai nuo tiriamos ligų grupės, tiriamųjų populiacijos, pasirinktos tyrimo strategijos, detalaus fenotipinio aprašymo ir ištyrimo bei bioinformatinės analizės, tačiau daugeliu atveju jis yra 20–50 proc. ribose. Tiong Yang Tan ir kolegos [32] VES atliko vaikams (2–18 m. amžiaus), kuriems buvo įtarta monogeninė liga ir kurie nebuvo tirti kitais NKS metodais. Nustatyti genetinę diagnozę pavyko 52 proc. tiriamųjų. Tam įtakos galėjo turėti tiriamųjų atrankos kriterijai ir egzomo duomenų vertinimas tik žinomuose ligos genuose. Kitas tyrimas, atliktas Retterer ir kolegų [33], parodė, kad siekiant nustatyti ligos genetinę priežastį efektyviau yra tirti triadą nei vieną probandą, kadangi molekulinės diagnozės nustatymas skyrėsi 7 proc. (atitinkamai 31 proc. ir 23,6 proc.). Įdomus ir reikšmingas tyrimas publikuotas 2023 m., atliktas JAV mokslininkų [34]. Jame siekta nustatyti VES diagnostinio efektyvumo skirtumus tarp dviejų grupių, sudarytų iš vaikų ir prenatalinių atvejų: vieną sudaro tiriamieji, priskiriami populiacijoms, turinčioms bendrą Europos protėvį ir daugiausiai analizuotoms didelio masto sekoskaitos projektuose, kitą sudaro tiriamieji iš tautinių mažumų ar mažiau atstovaujama populiacijų (lotynų, amerikiečių, azijiečių, afroamerikiečių). Mokslininkai, tirdami šias dvi sergančiųjų grupes, nenustatė reikšmingų genetinės diagnozės skaičiaus skirtumų.

Žmogaus geno sekoskaita NKS technologijomis (žr. 1.2 schemą) įgauna vis didesnę reikšmę ne tik moksliniuose tyrimuose, bet ir klinikinėje praktikoje. Tai didelis žingsnis personalizuotos medicinos link, kadangi nuskaičius tiriamojo genomą gaunama informacija ne tik apie retos ligos genetinę priežastį, bet galima analizuoti tiriamojo farmakogenetinius, metabolizmo ypatumus, nustatyti riziką sirgti tam tikromis ligomis [35]. Genomo NKS duomenų padengimas yra skirtingas, neretai gaunamas

mažesnis nei pasiekiamas atliekant VES ir genų grupių tyrimą, tačiau jis yra vientisesnis ir padengiantis visą genomą, todėl sudaro sąlygas nustatyti KSP, mtDNR, SV ir analizuoti baltymų nekoduojančią geno dalį [26]. ENCODE (angl. *Encyclopedia of DNA Elements*) projektui paskelbus pirmosios fazės rezultatus, daugelis tyrėjų buvo nustebinti, jog apie 80 proc. žmogaus geno gali turėti funkcinę reikšmę [36], todėl tiriamųjų analizė VGS metodu yra neabejotinai reikšminga retų paveldimų ligų srityje. Daugelį metų moksliniai ir diagnostiniai resursai buvo nukreipti į baltymus koduojančias geno sritis (genus), tačiau vis dažniau atkreipiamas dėmesys, jog retos paveldimos ligos gali turėti genetinę priežastį mažiau analizuotose geno srityse. Pavyzdžiui, Bagnall ir kolegos [37] keturioms šeimoms, kurių nariams diagnozuota hipertrofinė kardiomiopatija, VGS metodu nustatė tolimus introninius variantus, lėmusius tiriamųjų ligą. Šis tyrimas sufleruoja ir apie sudėtingą duomenų analizės procesą, kuris yra neatsiejamas nuo geno sekoskaitos. Ilgalais didelių duomenų laikymas bei didelių pajėgumų informacinės technologijos ir vis tobulėjančios bioinformatinės analizės sistemos yra svarbūs ir neretai sunkumų keliantys poreikiai norint atlikti VGS analizę. Gausus duomenų kiekis vienam tiriamajam lemia daugiau nustatomų genų kandidatų, kuriuos tyrėjas turi atmesti arba patvirtinti, taip pat atsitiktinių radinių, sukeliančių papildomų etinių klausimų [38]. Vis dėlto diagnostinio efektyvumo atžvilgiu geno sekoskaita yra įrankis, neretai pranašesnis už genų grupių ir egzomo sekoskaitą, taip pat apimantis ne tik šių, bet ir kai kurių citogenetinių tyrimų galimybes (KSP, SV nustatymas). Tiriant paveldimas akių ligas VGS metodu, mokslininkams pavyko nustatyti genetinę sutrikimo kilmę daugiau nei 57 proc. tiriamųjų [39]. Verta atkreipti dėmesį, kad šiame tyrime beveik 13 proc. sudarė SV ir DNR pokyčiai nekoduojančiose geno srityse. Genomo NKS neretai taikoma tiriamiesiems, kuriems VES metodu genetinė diagnozė nenustatyta. Tokiais atvejais patogeninius pokyčius pavyksta nustatyti 34 proc. tiriamųjų, sergančių įtariama Mendelinio paveldėjimo liga [40]. Kitas svarbus geno sekoskaitos aspektas yra finansinės išlaidos. Ewans ir kolegos [40] palygino, kiek skiriasi sąnaudos taikant tik geno sekoskaitą, VES ir VGS, arba VES ir egzomo duomenų peranalizavimą po kelerių metų. Tyrimas atliktas Mendelinio paveldėjimo ligomis sergančiųjų grupėje ir nustatyta, kad geno NKS tyrimas, taikant jį kaip pirmąjį genetinį testą, yra optimalus tyrimas kainos atžvilgiu, siekiant didžiausio diagnostinio efektyvumo. Vis dėlto, jei tinkamai įvertinama genetinės diagnozės nenustatymo rizika, optimaliausias kainos atžvilgiu yra VES tyrimas su egzomo duomenų peranalizavimu po kelerių metų [40].

Ištyrus visą sergančiojo genomą, ne visada pavyksta nustatyti, kas lėmė įtartą retą paveldimą ligą ar pagrįsti, kad literatūroje neaprašytas DNR pokytis

yra patogeninis. Tai viena iš pagrindinių priežasčių, skatinančių tyrėjus ieškoti ligos priežasties kituose lygmenyse, turinčiuose įtakos genetinės informacijos nurašymui ir pasireiškimui. Transkriptomo NKS (žr. 1.2 schemą) leidžia tirti visą organizmo informaciją RNR lygmeniu, t. y. nustatyti genų raiškos pokyčius, transkriptų įvairovę ir jų sekų pokyčius, taip pat analizuoti mažas ir ilgas nekoduojančias RNR. Be to, tiriant genomą ir transkriptomą kartu atsiranda galimybė ne tik nustatyti pokyčius RNR molekulių atžvilgiu, bet ir sieti nustatytus genomo pokyčius su funkcinė reikšme [41]. Taip gali būti patvirtinamas neaprašytų DNR pokyčių, nustatytų įvairiose genomo srityse, patogeniškumas, todėl pagrįstų molekulinų diagnozių skaičius gali išaugti. Lee ir kolegos [41] publikavo tyrimą, kuriame nustatė, kad atliekant genomo ir transkriptomo sekoskaitą pasiekiamas 38 proc. diagnostinis efektyvumas, kuris yra 7 proc. mažesnis atliekant tik VES arba VGS retomis paveldimomis ligomis sergančiųjų grupėje. Nors diagnostinio efektyvumo atžvilgiu indėlis nėra didelis, transkriptomo sekoskaitos reikšmė yra ypač svarbi siekiant įveiklinti funkcinį tyrimų dalį tiek diagnostikoje, tiek mokslinėje veikloje ir prisidėti prie patogenezės mechanizmo išaiškinimo.

Epigenetinis genomo lygmuo yra kitas, labai svarbus organizmo mechanizmas, reguliuojantis genų raišką, todėl neabejotinai darantis įtaką retiems paveldimams sutrikimams vystytis. Pokyčių DNR metilinimo lygmeniu nustatymui naudojama NKS technologijomis pagrįsta viso metilomo arba jo dalies sekoskaita [42] (žr. 1.2 schemą). Retas ligas lemiantys epigenetiniai pokyčiai dažniausiai yra somatiniai variantai, paveldimi mitozės būdu ir neperduodami iš kartos į kartą, todėl jų poveikis fenotipui priklauso nuo organizmo vystymosi laikotarpio, kuriame jie susidarė [43]. Epigenome susidarę pakitimai gali prisidėti prie retos paveldimos ligos vystymosi arba tiesiogiai ją lemti [43, 44]. Literatūroje aprašyta retų atvejų, kai tiriamojo fenotipas sufleruoja apie recesyvaus paveldėjimo ligą, tačiau nustatomas tik vienas pakitęs alelis, – tokiu atveju galima įtarti epigenetinį pokytį kitame alelyje [45, 46]. Epigenetiniai pokyčiai gali lemti fenokopiją, t. y. fenotipą, kurį įprastai lemtų konkrečiame gene nustatomi DNR pokyčiai, lemia DNR metilinimo pokytis šio geno reguliacinėje, dažniausiai promotoriaus pozicijoje [47, 48]. Be to, metilomo sekoskaita yra svarbi siekiant nustatyti ir pagrįsti variantų patogeniškumą genuose, kurie koduoja DNR metilinimui svarbius baltymus, ir lemia su jais siejamas retas įgimtas ligas [43]. NKS metilomo tyrimas nėra dažnai taikomas klinikinėje praktikoje, tačiau jis turi įtakos diagnostiniam efektyvumui. Sadikovic ir kolegos [49] aprašė tyrimą, parodantį, kad beveik 11 proc. tiriamųjų, kuriems kitais genetiniais metodais nebuvo nustatyta molekulinė diagnozė, metilomo sekoskaitos metodu nustatyti epigenetiniai pokyčiai, sietini su tiriamųjų fenotipu.

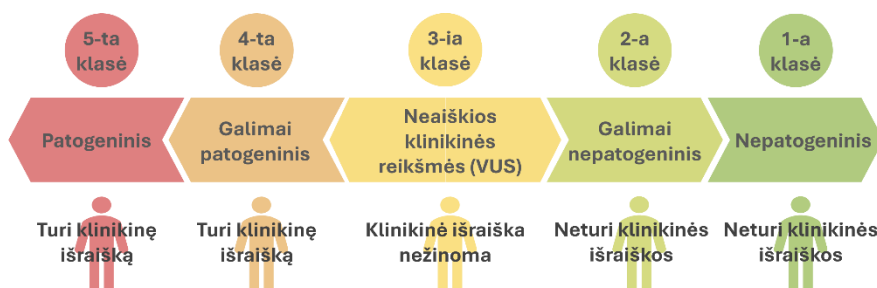
Apibendrinant, NKS technologijos pakeitė retų paveldimų ligų mokslinių tyrimų ir diagnostikos sritis: sutrumpino laiką nuo įtarimo iki genetinės diagnozės, praplėtė galimybes vienu metodu analizuoti didelį kiekį paveldimos informacijos ir nustatyti įvairius DNR sekos pokyčius tiek koduojančiose, tiek nekoduojančiose sekose bei suteikė galimybę genomo mastu analizuoti galimą pokyčių funkcinę reikšmę. Šios technologijos paneigė įsisenėjusią prielaidą, kad vienas genas lemia vieną fenotipą. NKS tyrimai atskleidė, kad tą pačią retą genetinę ligą gali lemti skirtinguose genuose nustatyti variantai, kelis fenotipus gali lemti tame pačiame gene nustatyti skirtingo tipo variantai. Taip pat nustatyti atvejai, kai tiriamajam patvirtintos kelios diagnozės, nes identifikuoti DNR variantai keliuose skirtinguose genuose, lemiančiuose skirtingas ligas. Be to, epigenomo ir transkriptomo sekoskaita parodė, kokios kompleksiškos gali būti ne genominės priežastys, lemiančios vieno fenotipo pasireiškimą. Iš esmės NKS leidžia tiksliau klasifikuoti ir kategorizuoti ligas, remiantis jų genetinė, transkriptomine ir epigenetine architektūra. Vis dėlto NKS technologijos dar nėra itin tikslios sekvenuojant GC turtingas sritis, taip pat homologinių genomo regionų ir sekų analizė yra komplikauta, nes atliekant šių sričių bioinformatinę analizę šios sekos neretai pašalinamos dėl labai mažos kokybės, kurią lemia jų prilygiavimas daugelyje genomo vietų [50, 51]. Todėl populiarėjančios trečios kartos sekoskaitos technologijos, kuriomis nuskaitomi ilgi (kelių ar keliolikos tūkstančių bazių porų) DNR fragmentai be polimerizacijos žingsnio ir neretai realiu laiku, yra būdas, leidžiantis išvengti NKS spragų ir gerinantis tyrėjų galimybes nustatyti retas genetines ligas lemiančius genomo, transkriptomo ir metilomo pokyčius [52].

1.3 DNR sekos variantų patogeniškumo nustatymo problematika

Nustatyti retus DNR variantus taikant NKS yra pirmas ligos priežasties įvertinimo etapas. Kitas kritiškai svarbus žingsnis yra nustatytų variantų interpretacija arba jų patogeniškumo vertinimas, kai siekiama įrodyti, kad variantas lemia tiriamojo fenotipą. Šis etapas tyrėjams kelia iššūkių, reikalauja nemažai laiko resursų ir patirties, nes tikimasi iš tiriamojo egzome arba genome nustatytų atitinkamai keliasdešimties tūkstančių ar milijonų variantų atrinkti tuos DNR pokyčius (dažniausiai vieną), kurie galimai lėmė tiriamojo ligos pasireiškimą.

Niehaus ir kolegos 2019 m. [53] publikavo apklausos rezultatus, kurioje 97 proc. dalyvavusių molekulinis tyrimus atliekančių laboratorijų JAV ir kitose pasaulio šalyse nurodė, kad vadovaujasi Amerikos medicininės genetikos ir genomikos kolegijos (angl. *American College of Medical*

Genetics and Genomics (ACMG)) bei Molekulinės patologijos asociacijos (angl. *Association of Molecular Pathology* (AMP)) sukurtais gairėmis variantams interpretuoti ir klasifikuoti [6]. Šios gairės sukurtos siekiant palengvinti interpretavimą ir suvienodinti genetinių tyrimų rezultatus, todėl apibrėžia variantų patogeniškumo vertinimo procesą ir sudaro sąlygas tarptautiniu mastu standartizuotai vertinti molekulines genetikoje ir citogenetikoje rezultatus. Gairėse pateikiami konkretūs kriterijai ir jų taikymo būdai, kurių laikydamiesi tyrėjai visame pasaulyje gali užtikrinti sklandų informacijos dalijimąsi ir rezultatų atkuriamumą. Remiantis ACMG/AMP rekomendacijomis, DNR variantai pagal klinikinę reikšmę klasifikuojami į 5-ias pagrindines klases: nepatogeniniai, galimai nepatogeniniai, neaiškios klinikinės reikšmės (angl. *variant of uncertain significance*, (VUS)), galimai patogeniniai, patogeniniai (žr. 1.3 schemą). Variantas priskiriamas vienai iš klasių tuo atveju, jei atitinka kriterijus, sudarytus iš įrodymų, kurie patogeninių ir galimai patogeninių variantų atžvilgiu skirstomi į labai stiprius, stiprius, vidutinio stiprumo ir palaikančius, o nepatogeninių ir galimai patogeninių variantų atžvilgiu – į savarankiškus, stiprius ir palaikančius [6].



1.3 schema. DNR variantų klasifikavimas atsižvelgiant į ACMG rekomendacijas. Parengta remiantis Richards et al. 2015 [6].

Taigi, interpretuojant variantus yra svarbūs įrodymai, kurie surenkami analizuojant literatūrą, naudojant *in silico* įrankius ir įvairias duomenų bazines. Vienas pirmųjų filtrų, kuris leidžia tyrėjams sumažinti gausų genomo variantų skaičių pirminiuose duomenyse, yra jų dažnio nustatymas pasitelkiant duomenų rinkinius viešai prieinamose populiacijų tyrimų duomenų bazėse. Pirmoji duomenų bazė, kurioje talpinta informacija apie bendrą populiacijų genetinę variaciją, buvo 1000 Genomų projekto, tačiau šiuo metu pagrindinė ir didžiausia duomenų bazė, talpinanti DNR variantų informaciją iš daugiau nei 195 tūkst. tiriamųjų egzomo ar genomo sekoskaitos duomenų, yra Genomo agregacijos duomenų bazė (angl. *Genome Aggregation Database* (gnomAD)) [54]. Duomenų bazės kaip gnomAD padeda atskleisti kliniškai

neriekšmingus genetinius skirtumus tarp populiacijų ir atskirų individų bei parodyti DNR pokyčių įvairovę, kuri, tikėtina, neturi įtakos ligų vystymuisi. Išanalizavus turimą gnomAD duomenų rinkinį, parodyta, kad varianto retumas yra ne visada tinkamas jo patogeniškumo įrodymas, kadangi vieno individo genomo koduojančioje sekoje vidutiniškai nustatoma apie , variantų, kurių dažnis itin mažas ($<0,1$ proc.), tačiau pabrėžtinai ne visi turi reikšmę fenotipui [55]. Kita vertus, nustatoma dažnų variantų (≥ 5 proc.), kurie turi įtakos retos ligos pasireiškimui, tačiau dėl jų žemo penetrantiškumo dažnis gali būti klaidinantis [56]. Svarbu įvertinti, kad gnomAD duomenų bazėje egzomo ir genomo sekoskaitos duomenys yra sveikų tiriamųjų, tačiau ne visiems buvo atlikta išsami sveikatos apžiūra, todėl sunku užtikrinti, jog visi tiriamieji yra išskirtinai sveiki [55]. Nors gnomAD talpina apie 200 tūkst. individų genomo duomenis, jie vis dar neatskleidžia visų galimų, kliniškai įtakos neturinčių DNR pokyčių, todėl tikslinga toliau plėsti nesergančių tiriamųjų genetinių duomenų rinkinius, kartu didinant tiriamųjų demografinę ir geografinę įvairovę.

Variantams interpretuoti taip pat reikalingos duomenų bazės, talpinančios informaciją apie kliniškai reikšmingus variantus, t. y. tos, kuriose skirtingos tyrėjų grupės DNR pokyčiui priskyrė patogeniškumo vertinimo kriterijus ir susiejo su konkrečia liga ar fenotipiniais bruožais, pavyzdžiui, ClinVar, *Human Gene Mutation Database* (HGMD), LOVD duomenų bazės. Remiantis disertacijos rašymo metu ClinVar duomenų bazėje pateikta statistika, joje talpinama apie 2,5 mln. unikalių DNR sekos pokyčių (HGMD apie 0,5 mln.) ir šis skaičius kasmet didėja. Tačiau svarbu tai, kad apie 36 proc. pateiktų skirtingo tipo variantų yra priskiriama VUS klasei [57]. Į ją patenka variantai, kurių klinicinei reikšmei patvirtinti ar atmesti neužtenka įrodymų, todėl tokie DNR pokyčiai ypač apsunkina diagnostinį darbą ir mažina diagnostinį efektyvumą. Taip pat 5 proc. ClinVar variantų yra vertinami prieštaringai, nėra priskirti vienai iš ACMG/AMP pasiūlytų variantų klasių [57]. Tai tik paryškina, kad variantų patogeniškumui įvertinti tyrėjui reikalingi papildomi įrankiai ir būdai.

Duomenų bazės pabrėžia informacijos dalijimosi, tiriamųjų skaičiaus didinimo ir informacijos atnaujinimo svarbą. GeneMatcher yra unikalus internetinis įrankis, skirtas dalintis informacija apie retus ligas lemiančius genetinius pokyčius (genus ar variantus), kurie domina tyrėjus ir apie juos nepakanka žinių bei publikuotų tyrimų užtvirtinti geno ar varianto sąsają su tiriamojo fenotipu. Todėl kitų tyrėjų pasaulyje patirtis ir jų nustatyti panašūs atvejai (nepublikuoti literatūroje) reikalingi siekiant nustatyti genetinę diagnozę [58]. Šiuo įrankiu taip pat gali naudotis modelinėse organizmų

sistemose tyrimus atliekantys mokslininkai, todėl sudaroma galimybė prisijungti prie DNR pokyčių funkcinės reikšmės tyrimų [58].

Variantams interpretuoti ir vertinti jų galimą įtaką molekulinio lygmeniu padeda bioinformatiniai įrankiai. Tai matematiniai skaičiavimo metodai, kompiuteriniai algoritmai, kurie gali įvertinti, pavyzdžiui, varianto, nustatyto koduojančioje genomo sekoje, įtaką baltymui remiantis jo pozicijos evoliuciniu konservatyvumu ir baltymo struktūrinėmis savybėmis (SIFT, PolyPhen-2, MutationTaster įrankiai) [59]. Kadangi reguliaciniai ir nekoduojantys regionai gali būti susiję su paveldimų ligų vystymusi, kuriami pažangūs įrankiai juose nustatytiems variantams įvertinti. Sukirpimo vietose nustatytų pokyčių įtaka informacinės RNR (iRNR) sukirpimo procesui vertinama pasinaudojant giliu mokymusi (dirbtinio intelekto poaibis) pagrįstu įrankiu SpliceAI [60]. Parodyta, kad šio įrankio specifiškumo ir jautrumo rodikliai yra geresni, lyginant su kitais neretai naudojamais įrankiais sukirpimo variantų įtakai modeliuoti [61]. Šis pavyzdys pabrėžia poreikį naudoti naujausius kompiuterinius modelius, tobulinant ir kuriant naujus *in silico* tyrimų įrankius genetinių variantų patogeniškumo vertinimui. Vis dėlto bioinformatiniai įrankiai nėra visiškai patikimi, jų rezultatai apie varianto reikšmę turi būti vertinami tikslingai ir kritiškai, todėl rekomenduojama išvadas daryti pasitelkiant ne vieną *in silico* įrankį, skirtą tam tikro tipo (aminorūgštį keičiančiam ar sukirpimo vietos) variantui vertinti. Be to, jog yra sukurti paieškos varikliai, įrankiai, tokie kaip *Varsome*, kurie apjungia daugelį anksčiau paminėtų duomenų bazių ir bioinformatinių įrankių bei literatūros šaltinių analizę, yra patogūs ir parankūs tyrėjo atžvilgiu, kadangi patogeniškumo vertinimui būtina informacija pateikiama vienoje vietoje [62]. *Varsome* įrankyje integruota ir automatinė variantų klasifikacijos sistema, kuri, remdamasi turimais įrodymais apie DNR pokytį, pateikia klasę, kuriai variantas galėtų būti priskirtas [62]. Tačiau net ir tokiu atveju tyrėjas priima galutinį sprendimą apie genetinio varianto funkcinę reikšmę remdamasis ne tik surinktomis žiniomis, publikuotais tyrimais ir įrankių rezultatais, bet ir savo profesine patirtimi.

Apibendrinant ankstesniuose ir šiame poskyryje pateiktą informaciją galima teigti, kad skirtingomis NKS technologijomis nustatytų DNR pokyčių reikšmės fenotipui vertinimas yra sudėtingas ir dinamiškas procesas, reikalaujantis tarpdisciplininių žinių ir patirties. Nė vienas NKS metodas nepasiekia šimtaprocentinio diagnostinio efektyvumo, kuriam didelę įtaką turi ir patogeniškumo vertinimo etapas. Nors populiacijų tyrimų duomenų bazės plečiasi, integruojami naujausi skaičiavimo modeliai kuriant bioinformatinius įrankius ir publikuota naujų genų ir juose nustatytų variantų funkcinę reikšmę analizuojančių tyrimų, daugiau nei trečdalis lieka VUS klasėje (remiantis

ClinVar duomenų bazė). Pavyzdžiui, *DYNC1H1* gene (MIM# 600112), kuris siejamas su neurovystymosi ir neuroraumeniniais sutrikimais (MIM# 614228, MIM# 614563, MIM# 158600), ClinVar duomenų bazėje daugiau nei 36 proc. (1425/3919) variantų yra VUS. Svarbu atkreipti dėmesį, kad tiek galimai patogeniniai, tiek galimai nepatogeniniai variantai žinomuose ligos genuose taip pat reikalauja papildomų įrodymų, siekiant juos priskirti konkrečiai patogeninių arba nepatogeninių variantų klasei. Taigi, neabejotinai išryškėja funkcinių tyrimų, kaip galutinio variantų patogeniškumo interpretacijos etapo, svarba, siekiant patvirtinti varianto vaidmenį patogenezės mechanizme ar suteikti naujų, fundamentinių žinių apie tam tikrus biologinius procesus. Kita svarbi funkcinių tyrimų misija personalizuotos medicinos atžvilgiu yra nustatyti varianto įtaką ne tik molekuliniu, ląstelių lygmeniu, bet ir individo lygmeniu, kadangi ne visada konkretus variantas turės tą patį poveikį skirtingiems individams. Visa tai yra ir didžiulė paskata kurti šiuolaikines biologinės terapijos metodus ir įrankius, kurie suteiktų galimybę sergantiejiems retomis paveldimomis ligomis pagerinti gydymo būdus ir tobulinti sveikatos priežiūros strategijas.

1.4 Funkciniai tyrimai DNR variantų patogeniškumui patvirtinti

Empiriniai įrodymai yra reikšmingi, tačiau nelengvai surenkami siekiant patvirtinti varianto patogeniškumą – tai pabrėžiama ir ACMG/AMP rekomendacijose. Šie įrodymai gaunami tik eksperimentiškai ištyrus koduojančiose ar nekoduojančiose sekose nustatytus DNR pokyčius ar kandidatinius genus ir jiems priskyrus (arba atmetus) funkcinę reikšmę. Todėl funkciniai tyrimai ir platesne apimtimi funkcinė genomika yra pagrindinis įrankis siekiant atskleisti DNR pokyčių sąsają su molekuliniais pokyčiais, signaliniais keliais ar procesais ląstelėse, kurių sutrikimas galimai lemia tiriamojo reto sutrikimo fenotipinių bruožų vystymąsi [13]. Pasaulyje ryškėja tendencija funkcinę genomo variantų analizę, kuri dažniausiai atliekama mokslinių tyrimų srityje, taikyti klinikinėje praktikoje, taip skatinti mokslinių ir diagnostinių padalinių glaudesnę bendradarbiavimą ir didinti diagnostinį efektyvumą. Australų mokslininkai [63] publikavo nacionaliniu lygiu atliktą tyrimą, kuriuo siekė parodyti integracinių tyrimų naudą, siekiant efektyviau diagnozuoti retomis ligomis sergančius naujagimius ir vaikus. Atlikus genomo sekoskaitą, pasiektas 47 proc. diagnostinis efektyvumas, tačiau pritaikius papildomus transkriptomo, proteomo ir mažesnio masto, labiau specializuotus funkcinius tyrimus šis rodiklis padidėjo iki 54 proc. [63]. Šis tyrimas dar kartą parodo funkcinių tyrimų vertę tiek epistemiologiniu, tiek pritaikomumo aspektais, kurie yra reikšmingi tobulinant mokslinę ir

sveikatos priežiūros sistemoje taikomus retų paveldimų ligų tyrimus ir diagnostikos strategijas.

1.4.1 Tyrimai molekulinio lygmeniu

Molekuliųjų tyrimų mastas ir našumas, analizuojant retomis paveldimomis ligomis sergančiųjų DNR, RNR ar baltymus, gali skirtis priklausomai ne tik nuo finansinių ir laiko resursų, bet ir paties ištyrimo tikslo. Tuo atveju, kai nustatomas variantas RNR sukirpimui svarbioje geno vietoje, tikslas yra įvertinti šio varianto įtaką iRNR sekos struktūrai, kurios pokyčiai sutrikdytų normalų sukirpimo procesą, toliau geno raiškai ar net galimai priešlaikinio baigmės kodono sukeltai iRNR degradacijai (angl. *Nonsense-mediated mRNA decay* (NMD)). Šis procesas aptinkamas visuose eukariotiniuose organizmuose ir atlieka tiek nepakitusių transkriptų kiekio reguliaciją arba adaptaciją esant aplinkos pokyčiams, tiek priešlaikinės baigmės kodoną turinčių transkriptų pašalinimą [64, 65]. Kopijinės DNR (kDNR), kuri susintetinama pagal tiriamojo RNR, Sangerio sekoskaita, naudojant sukirpimo vietos pokytį apimančius pradmenis, panaudota siekiant nustatyti, kokią poveikį *CAPN3* iRNR sekai turi heterozigotinis akceptorinės sukirpimo vietos variantas c.2440-1G>A [66]. Šia tyrimo strategija mokslininkai nustatė *CAPN3* 24-to egzono iškritą, todėl padaryta išvada, kad geno koduojamas baltymas sutrumpėja ir, tikėtina, sutrinka jo funkcija. Atlikę kiekybinę PGR (kPGR), nustatė dvigubai mažesnę *CAPN3* geno iRNR raišką lyginant su kontroliniais asmenimis, tad šiuo atveju tikėtina, kad aktyvuojamas NMD kelias ir sutrumpėjusi iRNR yra suskaidoma [66]. Šie rezultatai įrodo varianto patogeniškumą, bet neatskleidžia viso patogenezės mechanizmo vaizdo. Panaši strategija pasirinkta analizuojant *PHF21A* gene nustatytą heterozigotinį donorinės sukirpimo vietos variantą c.153+1G>C [67]. Tik šiuo atveju tiriamajam ir jo tėvams pirmiausia buvo atlikta transkriptomo sekoskaita, kuria nustatytas pakitęs *PHF21A* transkriptas be 6-to egzono. Šis radinys patvirtintas atliekant kDNR Sangerio sekoskaitą ir *in silico* nustatyta, kad ji lemia vidugeninę deleciją, dėl kurios geno koduojamas baltymas praranda aminorūgštis, koduojančias funkcinį leucino užtrauktuko domeną [67]. Įdomu tai, kad 6-to egzono iškrita nelemia rėmelio poslinkio ir priešlaikinės baigmės kodono susidarymo, todėl tikėtina, kad NMD mechanizmas neaktyvuojamas, tačiau kPGR tyrimo rezultatai parodė du kartus sumažėjusią 6-to egzono raišką tiriamojo RNR mėginyje [67]. Taigi, tyrėjai, taikydami pasirinktą tyrimo strategiją, pademonstravo sutrikusį sukirpimo procesą, jo įtaką RNR raiškai ir galimus baltymo funkciją sutrikdančius pokyčius. Beje, sukirpimo vietos variantams analizuoti neretai

pasirenkamas ir *in vitro* minigenų sukirpimo tyrimas (angl. *In vitro Minigene splicing assay*), kuris leidžia modeliuoti sukirpimo vietos variantus pasirinkto tipo ląstelėse, todėl neretai taikomas, kai tiriamo geno raiška nevyksta prieinamuose mėginiuose, pavyzdžiui, kraujyje. Taikant šį metodą, nustatyta, kad *NAA15* geno heterozigotinis donorinės sukirpimo vietos variantas c.1410+5G>C lemia 12-o introno dalies (224 bp) tarpą [68]. Tai leido tyrėjams įrodyti, kad variantas lemia sutrikusį iRNR sukirpimo procesą ir, tikėtina, pakitusią baltymo funkciją, todėl yra susijęs su reto sutrikimo vystymusi [68].

Verta pabrėžti, kad viena pakitusi kanoninė sukirpimo vieta gali lemti kelių pakitusių transkriptų raišką. Sanoguera-Miralles ir kolegos [69] pritaikė minigenų sukirpimo analizę, siekiant funkciškai įvertinti *CHEK2* gene nustatytiems sukirpimo vietos variantams. Tyrėjams pavyko nustatyti, jog c.319 + 5G > T lėmė galimai 11-kos skirtingų transkriptų susidarymą, kuris iš dalies parodo dar neatskleistas genomo koduojamos informacijos reikšmes, kurios yra svarbios sukirpimo procese dalyvaujantiems molekuliniams kompleksams [69]. Taip pat svarbu paminėti, kad aminorūgštį keičiantys vieno nukleotido pokyčiai gali turėti įtakos sukirpimo procesui. Naudojant minigenų tyrimą, Dionnet ir kolegos [70] ištyrė 21-ą aminorūgšties pokytį sukeliantį DNR variantą, esantį toliau nuo 5' ir 3' egzono galų, ir nustatė, kad 8-i variantai lemia įvairius iRNR sukirpimo pokyčius: egzono iškritą, introno dalies tarpą, alternatyvios splaisingo vietos susidarymą. Šių variantų galima įtaka sukirpimo procesui nebuvo nustatyta naujausiais bioinformatiniais įrankiais. Visi anksčiau paminėti molekuliniai tyrimai pabrėžia naudą klinikinėje praktikoje ir įrodo svarbą fundamentinių žinių aspektu, kadangi atskleidžia sudėtingus molekulinis procesus, darančius įtaką pre-iRNR sukirpimui ir tolesniam apdorotos iRNR įveiklinimui – transkripcijai ir transliacijai.

Skirtingi molekuliniai metodai taikomi siekiant įvertinti DNR variantų įtaką ir baltymo lygmeniui. Neretai šiuo būdu pasirenkama analizuoti aminorūgštį keičiančius, rėmelio poslinkį ar priešlaikinės baigmės kodono susidarymą lemiančius genomo sekos pokyčius. *CHM* gene nustatytą aminorūgštį keičiantį variantą p.(Leu457Pro) analizuojant Western blot metodu, sergančiojo fibroblastuose tyrėjai nustatė pakitusio baltymo raiškos sumažėjimą 86 proc., lyginant su sveikų asmenų nepakitusio baltymo raiška fibroblastų kultūrose [71]. Šis kiekio pasikeitimas įrodo neigiamą varianto įtaką geno koduojamam baltymui, kuris, tikėtina, yra suardomas [71]. Chen ir kolegos [72] siekė nustatyti vienos aminorūgšties iškritą p.(Thr312del) lemiančio varianto reikšmę *KCNQ1* koduojamo kalio jonų kanalo funkcijai, tačiau pritaikė Western blot analizę, transfekuotose žmogaus embriono inkstų

ląstelėse šio baltymo raiškos skirtumų nenustatė, lyginant nepakitusio baltymo ir pakitusio baltymo raišką turinčiose ląstelėse. Tai leido įrodyti, jog variantas neturi didelės įtakos kalio kanalo geno raiškai arba transliacijai [72]. Šiuo atveju pritaikyta ir koimunoprecipitacija, siekiant nustatyti, ar KCNQ1 baltymo vienos aminorūgšties iškirta neturi įtakos jo sąveikai su partneriu KCNE1 baltymu – skirtumų taip pat nenustatyta [72]. Šis tyrimas pabrėžia baltymo kiekybinės analizės svarbą net ir nustačius iškirtą lemiantį genomo pokytį. Tokio tipo pokyčio įtaka molekuliniais procesams ne visada pasireiškia neigiamai raiškos ar molekulių tarpusavio sąveikos atžvilgiu – tai dar kartą parodo patogenezės mechanizmų sudėtingumą ir skirtingumą kiekvieno varianto atveju. Hemizigotinė duplikacija (3,8 Mb), apimanti *RLIM* geną, taip pat analizuota taikant Western blot metodą, kuriuo įrodyta, jog tiek RNR, tiek baltymo lygmeniu *RLIM* geno koduojamo produkto kiekis reikšmingai išauga lyginant sergančiųjų ir kontrolinių asmenų ląsteles [73]. Taip patvirtinta nustatyto genomo pokyčio įtaka kitu genomo koduojamos informacijos pasireiškimo lygmeniu, todėl gaunamas papildomas funkcinės reikšmės įrodymas. Anksčiau aprašytų tyrimų atvejais analizuojamas vienas arba keli baltymai, kuriuos koduoja DNR pokytį turintis genas ar geno koduojamo produkto sąveikos partneriai. Tačiau retų paveldimų ligų srityje taikomi ir plataus masto proteomo ar metabolomo tyrimai (kaip skysčių chromatografija – masių spektrometrija (LC-MS)), kurie sukuria galimybę analizuoti baltymų ar metabolitų pokyčius neturint išankstinės fenotipo kilmės hipotezės tam tikrų ligų grupėse, pavyzdžiui, paveldimų metabolizmo sutrikimų ar mitochondrinių ligų, siekiant pirmiausia nustatyti fiziologinius tiriamojo individo pokyčius, toliau genetinius variantus, galimai lėmusius įgimtą tiriamojo fenotipą [74-76]. Tai strategija, kuri gali būti taikoma siekiant nepraleisti genomo variantų, kurių analizė molekuliniiais ar funkciniais tyrimais nepagrindžia jų klinikinės reikšmės, tačiau proteomo ar metabolomo lygmenimis jų funkcinė įtaka gali būti atskleista [75, 76]. Vis dėlto, analizuojant RNR, baltymus ir metabolitus molekuliniiais metodais, neretai gaunamos tik įžvalgos ir gairės apie galimas funkcines pasekmes ir tam tikriems baltymams, kompleksams, ar mechanizmus, sukeliančius pokyčius organizmo ląstelėse vykstantiems procesams, todėl susidaro poreikis detaliau paanalizuoti atrinktų taikinių lemiamą patogenezės kelią, taikant eksperimentinius tyrimus ląstelėse.

1.4.2 Tyrimai ląstelių lygmeniu

Ląstelės yra vienas iš nedaugelio prieinamų, funkcinei analizei ir eksperimentinėms manipuliacijoms naudojamų objektų ligoms modeliuoti,

kuris leidžia *in vitro* arba *in vivo* tirti retų genetinių ligų patogenezės procesus, nulemtus pokyčių genome. Tiriamojo pirminės fibroblastų kultūros, kultivuojamos iš odos dermos biopstatų, neretai pasirenkamos tyrimams dėl palyginti su kitomis žmogaus ląstelėmis ne itin sudėtingos priežiūros, taip pat tada, kai žinoma, jog juose vyksta analizuojamo geno raiška [77-80]. Be to, tiriamųjų fibroblastai sudaro galimybę analizuoti DNR pokyčius unikalioje individo genetinėje ir epigenetinėje aplinkoje, kuri gali turėti įtakos fenotipiniams bruožams pasireikšti. Goldžio komplekso mikroskopinė atsikūrimo analizė po poveikio nokodazolu tiriamųjų pirminiuose fibroblastuose leido įrodyti, kad *DYNC1H1* gene nustatyti aminorūgščių keičiantys heterozigotiniai variantai (c.3581A>G p.(Gln1194Arg) ir c.9142G>A p.(Glu3048Lys)) lėmė lėtesnį Goldžio komplekso atsikūrimą lyginant su kontrolinių asmenų ląstelėmis [81]. Šiuo eksperimentu mokslininkai parodė, kad atliekant nesudėtingas manipuliacijas ląstelėse galima patvirtinti analizuojamų variantų patogeniškumą ir atskleisti tam tikrus patogenezės mechanizmo aspektus – sutrikusį Goldžio komplekso membranų judėjimą, už kurį ląstelėse atsakingas citoplazminis dineinas, o pastarojo sunkioji grandinė koduojama *DYNC1H1* geno [81]. Analizuojant Vernerio sindromu sergančiųjų fibroblastus, tyrėjai nustatė, kad *WRN* geno VNV, lemiantys priešlaikinės baigmės kodono susidarymą ir vienos aminorūgšties pokytį (atitinkamai p.(Gln748Ter) ir p.(Phe1074Leu)), lėmė reikšmingai sulėtėjusį ląstelių dalijimąsi, padidėjusį oksidacinį stresą dėl šiose ląstelėse sumažėjusio aktyvių deguonies formų nukenksminimo [82]. Molekulinės ir ląstelės biologijos metodais analizuojant tiriamųjų fibroblastus, pavyko identifikuoti molekulinis kelius, kurie sukelia patofiziologinius pokyčius sergančiajam.

Esama skirtingų priežasčių, dėl kurių tiriamojo ląstelių kultūros gali būti neprieinamos: asmeninės, medicininės, techninės, etinės, logistinės. Taip pat ne visada analizuojamo geno raiška nustatoma prieinamuose ląstelių ėminiuose (kaip fibroblastuose). Tokiais atvejais eksperimentams naudojamos išvestos ląstelių linijos, kurios neabejotinai pranašesnės už pirminių ląstelių kultūras nemirtingumu – gebėjimu neribotai dalintis nepereinant į senėjimo būseną [83]. Tačiau net kai analizuojamo geno raiška ląstelėse nenustatoma arba reikalinga pakitusio geno koduojamo baltymo raiška tam tikrose ląstelių linijose, pasitelkiama transfekcija ar transdukcija, kurių pagalba galima vykdyti norimo produkto raišką. Atliekant pakitusio kalio kanalo (turinčio p.Thr312del pokytį), koduojamo *KCNQ1* geno, funkcinę analizę, į HEK293 (žmogaus embrioninės inkstų) ir CHO (kinų žiurkėno kiaušidžių) ląsteles cheminės transfekcijos būdu buvo įvestos plazmidės, koduojančios nepakitusį ir pakitusį KCNQ1 baltymus [72]. Po

manipuliacijos ląstelėms atlikti elektrofiziologiniai tyrimai „*patch-clamp*“ metodu leido nustatyti, jog vienos aminorūgšties delecija lemia kalio kanalo disfunkciją – kalio jonų srovė nustatyta reikšmingai maža arba nenustatyta iš viso [72]. Taip pat pastebėta, kad šis pokytis, tikėtina, turi dominuojantį neigiamą efektą, kadangi ląstelėse, kuriose vyko abiejų plazmidžių raiška, kalio jonų srovė sumažėjo daugiau nei 50 proc., – vadinasi, pakinta kalio kanalo laidumas [72]. Taikant šią tyrimo strategiją, mokslininkai taip pat gebėjo ląstelėse sukelti ir KCNQ1 subvieneto KCNE1, reikalingo kanalo funkcijai, raišką – taip nustatė, jog sąveika tarp jų sutrinka [72]. Šis tyrimas pabrėžia ir tai, kad šiuolaikiniais transfekcijos metodais įmanoma tirti itin specifiniams audiniams (pavyzdžiui, širdies audiniui, jį sudarantiems kardiomiocitams) būdingus baltymus kaip jonų kanalus, kurių analizė izoliuotomis ir kontroliuojamomis sąlygomis pasirinktose ląstelių linijose sudaro galimybę specifiskai tirti vieno baltymo ar kelių jo partnerių veiklą, funkciją, molekulinis patogeneizės kelius, sukeltus tiriamo varianto.

Specifiniams audiniams (širdies, nerviniam) būdingų baltymų analizė jiems įprastoje ląstelinėje aplinkoje gali būti pasiekta taikant indukuotų pluriotentinių kamieninių ląstelių (angl. *induced Pluripotent Stem Cells* (iPSC)) technologijas. Perprogramavus tiriamojo ląstelės į iPSC, skirtingų transkripcijos veiksnių, cheminių reagentų ir kitų veiksnių bei aplinkos manipuliacijų pagalba galima diferencijuoti ląsteles į skirtingus tipus [84]. Tačiau perprogramavimo procese gali būti prarandamos unikalios epigenetinės žymės, būdingos tiriamojo genomui. Kadangi retos ligos neretai paveikia ne vieną organizmo sistemą, naudojant tiriamojo iPSC galima specifinį retą sutrikimą ir jį lėmusį geno pokytį tirti kelių audinių atžvilgiu. Liu ir kolegos [85] panaudojo sergančiųjų Hutchinson–Gilford progerija pirminių fibroblastų kultūras, siekiant nustatyti *LMNA* geno VNV p.(Gly608Gly) sukeltą ligos patogeneizės mechanizmą. Pritaikę retrovirusinę transdukciją, tiriamųjų fibroblastai perprogramuoti į iPSC ir toliau specifiniais transkripcijos veiksniais diferencijuoti į lygiųjų raumenų ląsteles [85]. Šiame tyrime mokslininkams pavyko nustatyti, kad genetinis pokytis lemia greitesnį lygiųjų raumenų ląstelių senėjimą, kuris pasireiškia per baltymą progeriną [85]. Kartu tirdami iPSC bei pirmines ląsteles, nustatė tam tikrus neaprašytus molekulinis senėjimo mechanizmus, būdingus kamieninėms ir diferencijuotoms sveikų asmenų ląstelėms [85]. Verta atkreipti dėmesį, kad anksčiau aprašytuose tyrimuose tiriamas genetinis pokytis, esantis sergančiojo ląstelėse, arba pokytis įvedamas mutagenezės metodu į plazmidės, o pastarosios transfekuojamos (ar transdukuojamos) į ląsteles, taip pat laikinai tiriamo geno produkto raiškai sumažinti taikomos ir mažos interferuojančios RNR (angl. *small interfering RNA* (siRNR)). Tačiau

tobulėjant genų inžinerijos technologijoms, norimą genų raiškos pokyčio efektą galima gauti įvedant DNR pokytį į pasirinktų ląstelių genomą naujausiais genų redagavimo metodais [86]. Viena iš itin populiarių ir dažniausiai taikomų technologijų yra CRISPR-Cas9 sistema [87-89]. Ši technologija leidžia tiriamų ląstelių genome sukurti VNV, intarpus, iškritas bei didesnius struktūrinius variantus naudojant Cas9 baltymą, atliekantį DNR grandinės kirpimą, atitinkamą gidinę RNR (gRNR), kuri Cas9 baltymui nurodo kirpimo vietą, ir esant poreikiui donorinę DNR, reikalingą siekiant įterpti naują ar pakeisti buvusią DNR seką genome [87, 88]. Ward ir kolegės [90] CRISPR-Cas9 technologija į izogeninių iPSC ląstelių genomą įterpė *NAA15* geno heterozigotinį variantą, lemiantį aminorūgšties pokytį (p.Arg276Trp). Ištyrę tiek redaguotas iPSC, tiek iš šių ląstelių diferencijuotus kardiomiocitus, nustatė, kad sutrinka NAA15 baltymo sąveika su ribosominiu kompleksu, todėl sutrinka tam tikrų baltymų translacija ir raiška, taip pat sutrinka specifinių baltymų N-galo acetilinimas, kuriame dalyvauja NAA15, ir kai kurie iš pakitusios raiškos baltymų susieti su paveldimomis širdies ligomis [90]. Siekiant įrodyti keturių aminorūgšties pokyčių (p.Pro587Arg, p.Arg599Cys, p.Tyr356Cys, p.Arg387Cys), nustatytų su X susieta intelektine negalia sergančiųjų *RLIM* genuose, šiuos pokyčius lėmusius DNR variantus įvedė į pelės embrioninių kamieninių ląstelių genomus (angl. *mouse Embryonic Stem Cells* (mESC)) [91]. Mokslininkai nustatė sumažėjusį *RLIM* (E3 ubikvitino ligazė) baltymo katalizinį aktyvumą, atsiradusį dėl skirtinguose domenuose nustatytų variantų, taip pat atskleidė šio baltymo svarbą mESC ląstelių diferenciacijai į nervines ląsteles – neuronų vystymuisi [91].

Taigi, pasitelkiant įvairias ląstelių manipuliacijos technikas tyrėjai turi galimybę atkurti reto paveldimo sutrikimo modelį ląstelėse arba iširti tiriamojo sergančių ląstelių linijas ir skirtingų DNR pokyčių tipų sukeltus patofiziologinius procesus. Išaiškintas patogenezės mechanizmas yra svarbus biologinėms terapijoms kurti, naujiems vaistų taikiniams nustatyti, taip pat fundamentinėms žinioms pagilinti ar papildyti. Be to, mokslinis genų redagavimo taikymas kuriant reto paveldimo sutrikimo modelį žinduolių ląstelėse prisideda ir kaip šios sistemos terapinis taikymas, kadangi CRISPR-Cas9 technologija galima ne tik įvesti, bet ir ištaisyti genomo pokytį sutrikusiose ląstelėse [89]. Tai itin perspektyvi sistema tokių retų ligų atvejais, kuriose pažeidžiamos specifinių audinių ląstelės arba yra įmanomos *ex vivo* sergančiojo ląstelių manipuliacijos, po kurių atliekama autologinė ląstelių transplantacija [92]. Nors tyrimai ląstelėse neabejotinai reikšmingi siekiant atskleisti molekulinį procesų ypatybes ir įrodyti funkcinę genomo pokyčių reikšmę, tačiau 2D ląstelių kultūros neleidžia susidaryti pilno patogenezės mechanizmo vaizdo audinių, organų ar viso organizmo atžvilgiu. Dėl to

funkciniai genų kandidatų ar DNR pokyčių žinomuose ligos genuose tyrimai atliekami ir yra reikalingi 3D ląstelių kultūrose, audiniuose ar modeliniuose organizmuose. Šis poreikis yra svarbi paskata tolesniems moksliniams tyrimams, kurie padėtų užtikrintai įvertinti genetinių variantų įtaką įvairiais gyvo organizmo genetinės informacijos įgyvendinimo lygmenimis ir turėtų reikšmingą bei teigiamą įtaką retais įgimtais sutrikimais sergančiųjų diagnostinėje odisėjoje ir efektyvaus gydymo perspektyvoje.

2. METODAI

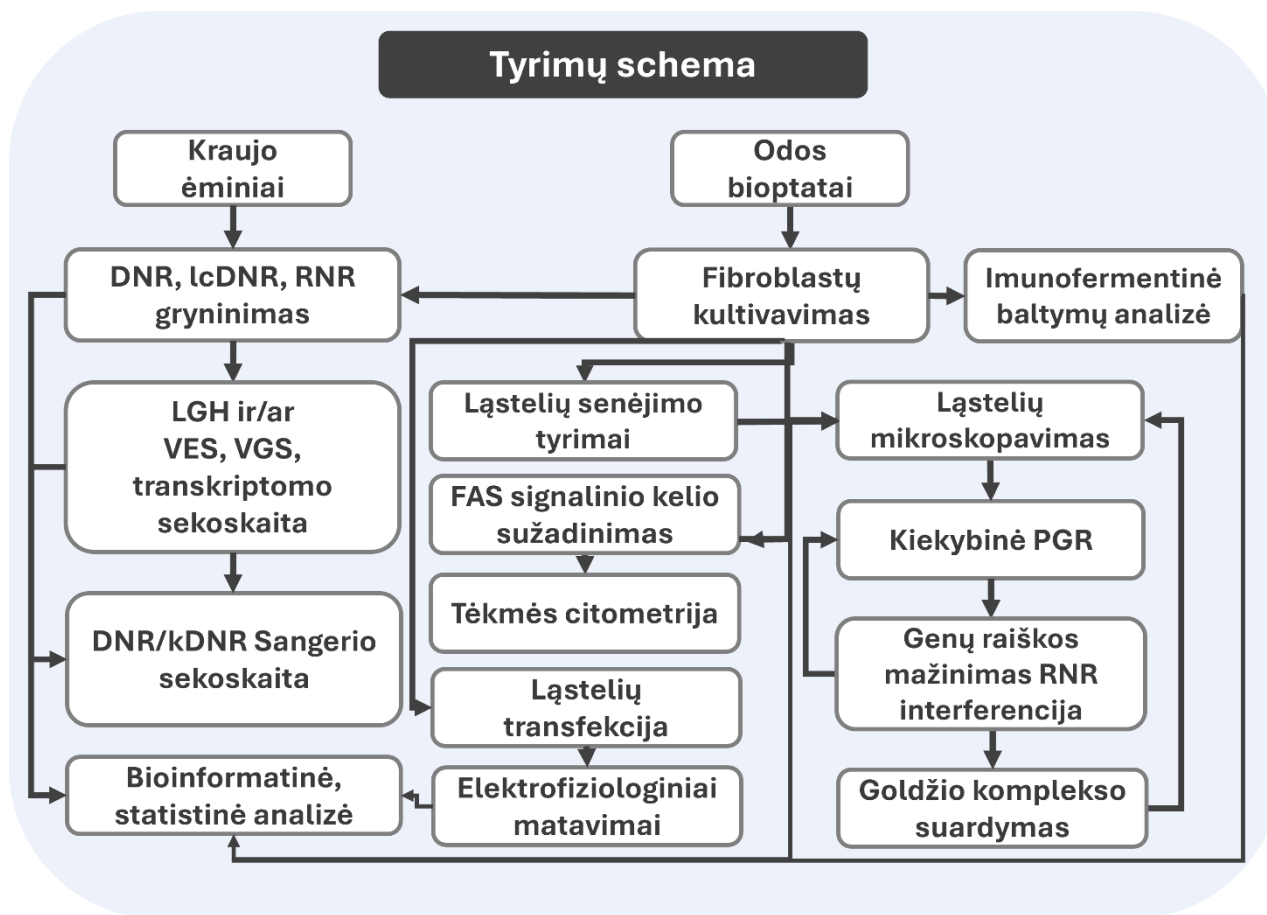
2.1 Tyrimo pagrindas ir tiriamieji

Pagrindinis disertacijos tyrimas atliktas Lietuvos mokslų tarybos finansuoto mokslinio projekto „Retų paveldimų ligų genetinių priežasčių integraciniai tyrimai“ pagrindu (akronimas ATGC; sutartis Nr. S-MIP-21-15; projekto biudžetas 149994 EUR; projekto vad. prof. Eglė Preikšaitienė). Dalis šio darbo tyrimo atlikta naudojant duomenis iš mokslinio projekto „Intelektinės negalios ir įgimtų anomalijų genetinės architektūros identifikavimas“ (akronimas INGENES; sutarties Nr. S-MIP-17-19/LSS-150000-1179; projekto biudžetas 99918 EUR; projekto vad. prof. Eglė Preikšaitienė). Kai kurie INGENES projekto tiriamieji pakviesti dalyvauti ATGC moksliniame projekte papildomiems arba tolesniems molekuliniams ir/ar funkciniais tyrimams. Taip pat, nedidelė disertacinio darbo tyrimo dalis finansuota VU Mokslo skatinimo fondo, laimėjus konkursą, skirtą jaunųjų mokslininkų idėjomis finansuoti (projektas „Genomo ir transkriptomo analizė retų paveldimų ligų patogenezės tyrimuose“; Nr. MSF-JM-2/2020; projekto biudžetas 9970,24 EUR; projekto vad. dokt. Gunda Petraitytė). Tyrimas vykdytas VU MF BMI Žmogaus ir medicininės genetikos katedroje bei VU GMC BI Biochemijos ir molekulinės biologijos katedroje. Tyrimams atlikti buvo gauti Vilniaus regioninio biomedicininio tyrimų etikos komiteto leidimai (INGENES Nr. 158200-17-962-469 ir ATGC Nr. 2021/9-1373-849). Į tyrimą įtraukti tiriamieji yra pasirašę Informuoto asmens sutikimo formas dalyvauti INGENES ir ATGC biomedicininuose moksliniuose tyrimuose.

Į mokslinį tyrimą įtraukti tiriamieji, konsultuoti gydytojų genetikų VUL Santaros klinikų Medicininės genetikos centre ir atitinkantys vieną pagrindinį kriterijų: pagrindiniam tiriamajam – probandui – įtariama reta genetinė patologija. Vadovaujantis šiuo kriterijumi, į tyrimą įtraukta 70 probandų ir 125 turintys arba neturintys ligos požymių jų šeimos nariai, t. y., jei yra galimybė, abu tėvai ir broliai/seserys (sibai), taigi iš viso 195 tiriamieji. Molekuliniams ir funkciniais tyrimams atlikti į tyrimą įtraukti negiminingi, ligos požymių neturintys sveiki asmenys (n=22).

2.2 Tyrimo metodai

Disertacinio darbo tyrimo tikslui įgyvendinti atlikti molekuliniai ir/ar funkciniai eksperimentiniai tyrimai taikant įvairius metodus (išsamiau aprašyti šiame poskyryje). Tyrimų taikymo strategijos (eiga) nurodytos 2.2 paveiksle. Disertacijos autorės indėlis atliktuose darbuose nurodytas 1-ame priede.



2.2 paveikslas. Tyrimų taikymo strategijų schema.

2.2.1 DNR gryninimas

Tiriamųjų genominė DNR gryninta iš periferinio kraujo leukocitų taikant standartinį fenolio–chloroformo–izoamilio alkoholio metodą [93]. DNR koncentracijai ir švarumui nustatyti naudotas spektrofotometras *NanoDrop® 2000* (ThermoFisher Scientific, JAV).

2.2.2 Laisvai cirkuliuojančios DNR gryninimas

Iš *DNMT3A* geno variantą turinčių tiriamųjų kraujo plazmos gryninta lasvai cirkuliuojanti DNR (lcDNR). Gryninimui naudotas *MagMAX™ Cell-Free DNA Isolation Kit* (AppliedBiosystems, JAV) rinkinys pagal pateiktą gamintojo protokolą. DNR koncentracijai ir švarumui nustatyti naudotas spektrofotometras *NanoDrop® 2000* (ThermoFisher Scientific, JAV).

2.2.3 RNR gryninimas

Didžiosios dalies tiriamųjų suminės RNR gryninta iš periferinio kraujo ėminių, surinktų į RNR stabilizuojantį reagentą turinčius *Tempus™ Blood RNA Tube* (ThermoFisher Scientific, JAV) mėgintuvėlius. Gryninimas atliktas naudojant *Tempus™ Spin RNA Isolation Kit* (ThermoFisher Scientific, JAV) rinkinį pagal pateiktą gamintojo protokolą. RNR koncentracija ir švarumas nustatyti naudojant spektrofotometrą *NanoDrop® 2000* (ThermoFisher Scientific, JAV).

Dalies tiriamųjų suminė RNR gryninta iš pirminių fibroblastų kultūrų. RNR gryninimui surenkama $1\text{--}3 \times 10^6$ ląstelių. Ląstelės surenkamos jų monosluoksniį praplaunant 0,02% EDTA tirpalu, toliau ląstelės atkeliamos 0,05% tripsino tirpalu, surinktos ląstelės centrifuguojamos (7 min, $1000 \times g$ 4°C). Nusėdusios ląstelės kartą praplaunamos PBS tirpalu ir dar kartą centrifuguojamos (7 min, $1000 \times g$ 4°C). Toliau RNR gryninimas atliekamas naudojant *RNeasy® Mini Kit* (Qiagen, Vokietija) rinkinį pagal pateiktą gamintojo protokolą. RNR koncentracija ir švarumas nustatyti naudojant spektrofotometrą *NanoDrop® 2000* (ThermoFisher Scientific, JAV).

2.2.4 Kopijinės DNR sintezė

Kopijinė DNR (kDNR) sintetinta iš suminės RNR mėginių naudojant *High Capacity RNA-to-cDNA™ Kit* (ThermoFisher Scientific, JAV) rinkinį pagal gamintojo pateiktą protokolą ir taikant jame nurodytas atvirkštinės transkripcijos PGR (angl. *Reverse transcription PCR* (AT-PGR)) sąlygas. Kopijinės DNR sintezei naudota 500 ng RNR. AT-PGR vykdyti naudotas

ProFlex PCR System (ThermoFisher Scientific, JAV) arba *Mastercycler (Eppendorf, Vokietija)* termociklerio kaitinimo blokai.

2.2.5 Fibroblastų kultūros kultivavimas ir užšaldymas

Tiriamųjų pirminiai fibroblastai gauti iš odos biopatų. Šios ląstelės kultivuojamos sudaro monosluoksnį. Fibroblastai kultivuojami DMEM tipo terpėje su 10% FBS (angl. *Fetal Bovine Serum*; *Gibco*, JK), pridedant 0,4% amfotericino B (*Gibco*, JK), arba *AmnioMAX C-100* terpėje su 14,2% *AmnioMAX C-100* terpės priedu (*Gibco*, JK), pridedant 0,4% amfotericino B. Kultivuojama 37°C temperatūroje, 5% CO₂ ir 95% drėgmės atmosferoje. Ląstelės persėjamos du arba tris kartus per savaitę, kai kultūra pasiekia 90–100% konfluenciją. Persėjimui ląstelės pirmiausia praplaunamos 0,02% EDTA tirpalu, toliau atkeliamos 0,05% tripsino tirpalu ir praskiedžiamos 1:2 santykiu.

Ilgalaikiam fibroblastų saugojimui skysto azoto diuaruose ruošta šaldymo terpė, sudaryta iš 90% FBS ir 10% DMSO (angl. *Dimethyl Sulfoxide*; *OriGen Biomedical*, JAV). Ląstelės šaldomos 1 ml šaldymo terpės, saugomos kriomėgintuvėliuose (*ThermoFisher Scientific*, JAV).

2.2.6 PGR

Pradmenys PGR sukurti naudojant viešai prieinamą *Primer-BLAST* įrankį [94]. Sukurtų pradmenų poros, naudotos DNR ir kDNR fragmentams pagausinti, nurodytos 2-ame priede.

DNR ir kDNR fragmentai gausinti PGR metodu naudojant polimerazės mišinį *2× Phusion Flash High-Fidelity PCR MasterMix (ThermoFisher Scientific, JAV)*. Į reakcijos mišinio sudėtį taip pat įeina dejonizuotas vanduo, pradmenų pora ir DNR/kDNR mėginys. PGR mišinys daugeliu atveju atitiko gamintojo protokolą, kai kuriais atvejais mišinio sudėtis optimizuota (pakeitimai nurodyti 3-iame priede). PGR sąlygos nustatytos remiantis polimerazės gamintojo rekomendacijomis, kai kuriais atvejais sąlygos optimizuotos (pakeitimai nurodyti 4-ame priede). PGR vykdyti naudotas *ProFlex PCR System* termociklerio kaitinimo blokas (*ThermoFisher Scientific*, JAV).

PGR produkto ilgiui, kiekiui ir specifiškumui įvertinti taikyta elektroforezė. Produktams išfrakcionuoti naudotas 1,5% agarozės *Top Vision Agarose (ThermoFisher Scientific, JAV)* gelis su 10mg/ml etidžio bromidu ((EtBr) *Roth*, Vokietija; 2–3 µl EtBr 100 ml agarozės gelio), 3× užnešimo dažas *DNA LoadingDye (ThermoFisher Scientific, JAV)* ir DNR molekulinės masės ir dydžio standartas *100 bp GeneRuler (ThermoFisher Scientific, JAV)*.

Elektroforezė atlikta *EC 105* horizontalios elektroforezės sistemoje (*E-C Apparatus Corporation*, JAV). Gelis vizualizuotas UV transiliumatoriuje (*E.A.S.Y. 442K*, *Herolab*, Vokietija), fotografuotas *EasyWin 32* (*Herolab*, Vokietija) gelių analizės programa.

2.2.7 PGR produkto gryninimas iš agarozės gelio

PGR produktai iš agarozės gelio gryninti *DYNC1H1* geno tyrimo atveju naudojant *GeneJET Gel Extraction Kit* (*ThermoFisher Scientific*, JAV) rinkinį pagal gamintojo pateiktą protokolą ir laikantis nurodytų saugumo rekomendacijų.

2.2.8 Sangerio sekoskaita

PGR produktų paruošimas sekoskaitai išsamiai aprašytas 5-ame priede. Sangerio sekoskaita atlikta naudojant *BigDye® Terminator v.3.1 Cycle Sequencing Kit* (*ThermoFisher Scientific*, JAV) rinkinį pagal gamintojo pateiktą protokolą. Kapiliarinė elektroforezė atlikta genetiniu analizatoriumi *ABI PRISM® 3130xl Genetic Analyzer* (*Hitachi*, JAV). Elektroforegramos analizuotos programa *Sequence Analysis v.5.1* (*Thermo Fisher Scientific*, JAV), toliau gautos elektroforegramos vizualizuotos ir analizuotos naudojant *Chromas v.2.4.4* (*Technelysium Pty Ltd*) programą. *ExPASy LALIGN* [95] bioinformatinis įrankis naudotas nuskaitytoms sekoms palyginti su referentinėmis genų sekomis.

2.2.9 Kiekybinė PGR

Kiekybinės PGR (kPGR) metodui atlikti naudoti komerciniai oligonukleotidiniai hidrolizės žymenys *TaqMan™* (*ThermoFisher Scientific*, JAV). Specifiniams ir namų ruošos genams naudotų žymenų sąrašas pateikiamas 6-ame priede. kPGR mišiniui naudotas universalus reakcijos mišinys *TaqMan™ Universal PCR Master Mix* (*ThermoFisher Scientific*, JAV). Reakcijos mišinio sudėtis paruošta ir reakcijos sąlygos nustatytos pagal gamintojo pateiktą protokolą. kPGR atlikti naudotas instrumentas *StepOne™ Real-Time PCR System* (*Applied Biosystems™*, JAV). To paties instrumento programinė įranga naudota kPGR duomenų analizei. Genų raiškos skirtumai nustatyti taikant standartinį $\Delta\Delta C_t$ skaičiavimo metodą. kPGR rezultatai pateikiami kaip genų raiškos pokyčiai kartais (angl. *Fold Change*), normalizuojant pagal namų ruošos geną *GAPDH* (arba *ACTB*) ir palyginus su tiriamo geno raiška sveikų kontrolinių asmenų ląstelėse.

CFLIP/CFLAR geno raiškos sumažinimo efektyvumas tikrintas kPGR metodu, naudojant *Maxima SYBR Green/ROX qPCR MasterMix (2X)* (*Thermo Fisher Scientific*) reagentą pagal gamintojo nurodytas rekomendacijas ir protokolą. Šioje reakcijoje naudoti tiesioginiai ir atvirkštiniai pradmenys pateikiami 2-ame priede. kPGR rezultatai normalizuoti pagal namų ruošos geną *HPRT*.

2.2.10 Naujos kartos sekoskaita

Tiriamųjų užkoduoti DNR mėginiai viso žmogaus egzomo ir/ar genomo NKS siūsti sekoskaitos paslaugų tiekėjui *Novogene*. Egzomo sekoskaitai naudota platforma *Illumina NovaSeq 6000*, kuria atlikta poruotų galų, 150 bp ilgio fragmentų sekoskaita su 50× sekoskaitos padengimu. Genomo sekoskaitai taip pat naudota *Illumina NovaSeq 6000*, kuria atlikta poruotų galų, 150 bp ilgio fragmentų sekoskaita su 30× sekoskaitos padengimu. Tiriamųjų užkoduoti RNR mėginiai transkriptomo sekoskaitai (taikiny s iRNR) siūsti tam pačiam sekoskaitos paslaugų tiekėjui. Transkriptomo sekoskaitai taip pat naudota *Illumina NovaSeq 6000*, kuria atlikta poruotų galų, 150 bp ilgio fragmentų sekoskaita sugeneruojant ≥ 20 milijonų poruotų fragmentų vienam RNR mėginiui.

2.2.11 Western blot analizė

Baltymų kiekybinė analizė atlikta Western blot metodu pagal Ausubel et al. (2003) [96]. Tikslus tyrime taikytas protokoląs pateikiamas 7-ame priede. Analizėje naudotų pirminių antikūnų sąrašas pateikiamas 8-ame priede. Baltymai perneštoje membranoje nustatomi naudojant su krienų peroksidaze konjuguotą antrinį antikūną prieš triušio IgG (*ThermoFisher Scientific*, JAV), arba prieš pelės IgG (*ThermoFisher Scientific*, JAV).

2.2.12 Ląstelių transfekcija mažomis interferuojančiomis RNR

Genų raiškai laikinai nuslopinti naudotos mažos interferuojančios RNR (siRNR). Šių molekulių transfekcijos efektyvumui vertinti kPGR metodu ląstelės sėtos į 6-ių šulinėlių plokštelę 1×10^6 ląstelių/ml tankiu, o stebėjimams mikroskopu sėtos į 24-ių šulinėlių plokštelės vieno šulinėlio tūrį atitinkančią 8-ių šulinėlių mikroskopijai skirtą plokštelę 4×10^4 ląstelių/ml tankiu. Ląstelės eksperimentui kultivuojamos mitybinėje terpėje be amfotericino B.

Transfekcijai atlikti naudotas transfekcijos reagentas *Lipofectamine™ RNAiMAX Transfection Reagent* (*ThermoFisher Scientific*, JAV). Transfekcija atlikta pagal gamintojo pateiktą protokolą. Siekiant laikinai

nuslopinti KIFC1 baltymo raišką ląstelėje, naudota *KIFC1* transkriptui savita siRNR (s7907; *ThermoFisher Scientific*, JAV). Siekiant laikinai nuslopinti KIFC3 baltymo raišką ląstelėje, naudota *KIFC3* transkriptui savita siRNR (s7843; *ThermoFisher Scientific*, JAV). Abiem atvejais palyginimui naudota siRNR (#4404021; *ThermoFisher Scientific*), kuri nėra savita jokiam žmogaus transkriptui.

2.2.13 Ląstelių transfekcija plazmidėmis

Ląstelių transfekcijai plazmidėmis HEK293T/17 (ATCC) ląstelės sėtos į 24-ių šulinėlių plokšteles (*Biolite 24-well Multidish*, *Thermo Scientific*) 80 tūkst. ląstelių/ml tankiu dieną prieš transfekciją. Transfekcijai naudotas reagentas *Lipofectamine® 3000* (*Thermo Fisher Scientific*), remiantis gamintojo pateiktomis rekomendacijomis ir protokolu. KCNE1 baltymą koduojanti plazmidė pIRES_EGFP_GA_KCNE1 (Addgene, #173160) ir baltymą KCNQ1 koduojanti plazmidė pIRES(A7Mut)_CyOFP1_KCNQ1 (Addgene, #173162), dovanotos Al George, transfektuotos skirtingomis kombinacijomis (KCNQ1_WT; KCNQ1_MUT; KCNQ1_WT+KCNE1; KCNQ1_MUT+KCNE1) išlaikant plazmidžių DNR kiekį, rekomenduotą transfekcijos reagento gamintojo. Po transfekcijos ląstelės inkubuotos (37°C, 5% CO₂) 24–48 val. Toliau atliekami elektrofiziologiniai matavimai.

2.2.14 Tėkmės citometrija

Citometrijai ląstelės kultivuotos 12-os šulinėlių plokštelę 40 tūkst. ląstelių/ml tankiu. Analizei citometru terpė kartu su ląstelėmis surenkamos į mėgintuvėlį. Ląstelės kartą praplaunamos PBS buferiniu tirpalu. Ląstelių suspensija centrifuguojama 1000 aps./min 7 min 4 °C. Supernatantas nusiurbiamas, o likusios ląstelės dažomos *Tali™ Apoptosis Kit - Annexin V Alexa Fluor™ 488 & Propidium Iodide* (*Thermo Fisher Scientific*) reagentu pagal gamintojo pateiktą protokolą ir rekomendacijas. Fluorescencija matuojama 1000 ląstelių naudojant *BDA FACSymphony™ AI Cell Analyzer* (*BD Biosciences*) tėkmės citometrą ir *BD FACSDiva™* programinę įrangą. Duomenys analizuoti *Flowing Software 2.5.1* programinę įrangą. Ląstelės klasifikuotos į šias grupes: aneksino V ir propidžio jodido (PI) teigiamos, aneksino V teigiamos ir PI neigiamos, aneksino V neigiamos ir PI teigiamos, aneksino V ir PI neigiamos. Tarp mėginių vertintos žūstančios ląstelės, kurios patenka į aneksino V teigiamų ir PI neigiamų ląstelių grupę.

2.2.15 Fibroblastų senėjimo tyrimas

Fibroblastų senėjimui nustatyti ląstelės sėtos į 60 mm skersmens lėkšteles 5×10^5 ląstelių/ml tankiu. Senėjimui nustatyti naudotas rinkinys *Senescence Cells Histochemical Staining Kit (Sigma-Aldrich, JAV)*. Tyrimas atliktas pagal gamintojo pateiktą protokolą ir rekomendacijas. Po nurodytos inkubacijos, ląstelės mikroskopuotos invertuotu mikroskopu (*Olympus CKX41*, naudojant $10\times$ objektyvą, NA 0.25), atsitiktine tvarka nufotografuojant ne mažiau kaip 10 vieno auginimo indo vietų.

2.2.16 Imunofluorescencinė mikroskopija

Goldžio kompleksas buvo vaizdinamas naudojant konfokalinę mikroskopiją po fibroblastų paveikio nokodazolu. Ląstelės užsėtos į mikroskopijai skirtas 8-ių šulinėlių plokšteles (*C8-I-N, Cellvis, Vokietija*) 3×10^5 ląstelių/ml tankiu. Po užsėjimo praėjus 24-ioms val., ląstelės paveiktos $5 \mu\text{g/ml}$ koncentracijos nokodazolo tirpalu (*Fisher Scientific, Vokietija*). Ląstelių kultūros skirtinguose poveikiu nokodazolu laiko taškuose 20 min fiksuotos 4% formaldehido tirpalu, po to plaautos 3 kartus po 5 min PBS tirpalu. Ląstelių membranos pralaidintos 15 min inkubuojant su 0,1% Tritono X-100 tirpalu ir toliau 3 kartus po 5 min praplautos PBS tirpalu. Ant fiksuotų ir permeabilizuotų ląstelių užpiltas blokavimo tirpalas (2% BSA tirpalo PBS'e), laikyta 60 min. Po to ląstelės per naktį 4°C inkubuotos su blokavimo tirpale praskiestu (santykiu 1:200) pirminiu antikūnu prieš DYNC1H1 arba GM130 baltymą. Nusiurbus pirminio antikūno tirpalą, ląstelės 3 kartus po 5 min plaautos PBS tirpalu. Po plovimo ląstelės inkubuotos 1 val. tamsoje blokavimo tirpale, kuriame yra 1:200 santykiu praskiestas su *Alexa Fluor® 647* konjuguotas antrinis antikūnas prieš triušio IgG (A21244; *Invitrogen, JAV*). Nusiurbus antrinio antikūno tirpalą ląstelės 3 kartus po 5 min praplautos PBS tirpalu ir 15 min inkubuota su DAPI dažais, praskiestais 1:1000 santykiu PBS tirpale. Po inkubacijos ląstelės 3 kartus po 5 min praplautos PBS tirpalu. DYNC1H1 ir GM130 imunofluorescencija registruota bei GM130 lokalizacija vertinta konfokaliniu mikroskopu *Leica SP8*, naudojant $63\times$ objektyvą, skaitinė apertūra (NA) 1.4, esant $\lambda_{\text{suž.}} = 604 - 644 \text{ nm}$, $\lambda_{\text{em.}} = 672 - 712 \text{ nm}$. FAS receptoriaus sužadinimo atveju apoptozė vertinta invertuotu fluorescenciniu mikroskopu *Olympus IX-81*, naudojant $40\times$ objektyvą, NA 0,9, esant $\lambda_{\text{suž.}} = 480 \text{ nm}$, $\lambda_{\text{em.}} = 530 \text{ nm}$ ir turinčiu *Hamamatsu ORCA-FLASH 4.0LT* kamera. Mikroskopijos vaizdai analizuoti ir vertinti *Fiji 2.15.0* programa [97].

2.2.17 Ląstelių elektrofiziologiniai matavimai „patch-clamp“ metodu

Elektrofiziologiniai matavimai „patch-clamp“ metodu atliekami 24–48 val. po transfekcijos. Ant stikliukų kultivuotos ląstelės perkeltos į *Nikon FN-S2N* mikroskopo registravimo kamerą (naudojant 40× objektyvą, NA 0.80), kuria parenkamos fluorescuojančios ląstelės.

Elektros srovės registruotos kambario temperatūroje naudojant *Multiclamp 700B* (*Molecular Devices*) stiprintuvą, valdomą programine įranga *pCLAMP 10.5* (*Molecular Devices*). Registravimo mikroelektrodai („patch“) ištraukti iš borosilikatinio stiklo kapiliarų, naudojant *Flaming/Brown* mikropipetę (*P-1000*; *Sutter Instrument Co.*), turinčią 2–4 MΩ varžą ir užpildytą specialiu ląsteliniu tirpalu (mM): 140 KCl, 1 Na₂-ATP, 2 EGTA, 10 HEPES, 0,1 CaCl₂, 1 MgCl₂ ir 10 gliukozės, pH 7,4. Užląstelinį tirpalą sudarė (mM): 140 NaCl, 4 KCl, 2 CaCl₂, 1 MgCl₂, 10 HEPES ir 10 gliukozės, pH 7,4.

Ląstelių membranų srovė fiksuota pradedant nuo –80 mV membranų laikančiu potencialu iki depoliarizacijos potencialo ties +30 mV, toliau registruota naudojant 2 s trukmės kas 20 mV (iki + 60 mV) didėjančias potencialo komandas. Gautos vertės buvo apskaičiuotos normalizuojant užregistruotas sroves pagal ląstelės talpą (pA/pF).

2.2.18 Bioinformatinė analizė

Bioinformatinė analizė apėmė pirminių egzomo, genomo ir transkriptomo duomenų paruošimą interpretacijai. Ši analizė taip pat apėmė ir pirminį tiriamo varianto patogeniškumo ir įtakos RNR bei baltymo lygmeniu įvertinimą taikant bioinformatinius įrankius, atliekant duomenų bazių, literatūros analizę. Disertacinio darbo metu naudotos duomenų bazės ir bioinformatiniai įrankiai pateikti 9-ame priede.

Pirminiai egzomo ir genomo duomenys sulygiuoti pagal žmogaus genomo referentinę seką, pateikiamą *UCSC* (hg38) duomenų bazėje, naudojant *BWA* (*Burrows-Wheeler Aligner*) programą [98]. NKS duomenų analizės algoritmui naudotas įrankis *GATK* [99]. Nustatytų variantų anotacijai naudoti *ANNOVAR* v.2018Apr16 [100] ir *VarAFT* v.2.16 [101] įrankiai.

Pirminiai transkriptominiai duomenys buvo įvertinti naudojant *FastQC* v0.11.9 [102], o žemos kokybės bazės ir adapterių sekos buvo apkarpytos naudojant *Trimmomatic* v0.39 [103]. Išvalytos sekos sulygiuotos pagal žmogaus genomą (GRCh38, Ensembl 111 versija) naudojant *STAR* v2.7.11a [104]. Genų lygmeniu sekoskaitos fragmentų skaičiavimai atlikti naudojant *featureCounts* v2.0.6 [105]. Diferencinė genų raiškos analizė atlikta naudojant *DESeq2* v1.42.0 [106]. Šioje analizėje probandas 1 (P1) ir probandas 2 (P2)

palyginti su amžių atitinkančiais vyr. lyties kontroliniais asmenimis (K), o motinos (M) palyginimas atliktas su amžių atitinkančiais mot. lyties kontroliniais asmenimis (K). Log2 raiškos pokyčių korekcija atlikta naudojant *lfcShrink* funkciją su *apeglm* parinktimi *DESeq2* įrankiu. Pakitusi genų raiška nustatyta taikant ribą: FDR pakoreguota p reikšmė $< 0,1$ ir raiškos pokytis $> 1,5$. Genų ontologijos (GO) terminų praturtinimo analizė buvo atlikta naudojant *clusterProfiler* v4.10.1 [107]. Į GO analizę įtraukti visi pakitusios raiškos genai. GO terminų praturtinimo analizė atlikta naudojant *enrichGO* funkciją *clusterProfiler* programoje. Parametrai buvo sukonfigūruoti įtraukiant visas tris pagrindines GO ontologijas: biologinį procesą, ląstelių komponentą ir molekulinę funkciją, nurodant „ont = 'ALL'“. „pAdjustMethod“ metodas pasirinktas Benjamini–Hochberg. „OrgDb“ parametras nustatytas kaip „org.Hs.eg.db“, kad atitiktų žmogaus genų ID ontologijos žodyną. Kiti parametrai buvo numatytieji. GO terminai buvo laikomi reikšmingais esant „pvalueCutoff“ mažiau nei 0,05. Visos statistinės analizės buvo atliktos naudojant R versiją 4.3.2.

2.2.19 Statistinė duomenų analizė

Eksperimentų duomenys apdoroti ir dalis grafikų braižyta naudojant *SigmaPlot 14.0* programą (*Systat Software*), kita dalis – *MS Office Excel* programą. Grafikuose pateikiami duomenys gauti iš 2–3 nepriklausomų biologinių eksperimentų. Vertinant santykinę geno raišką kiekviename eksperimente naudoti ≥ 2 techniniai pakartojimai. Duomenų pasiskirstymas pagal normalųjį skirstinį tikrintas taikant *Shapiro-Wilk* normalumo kriterijų. Jei duomenys pasiskirstę pagal normalųjį skirstinį, taikoma parametrinė statistika – t kriterijus arba *Kruskal-Wallis* kriterijus skirtumams palyginti. Jei duomenys nepasiskirstę pagal normalųjį skirstinį, taikoma neparametrinė statistika – *Mann-Whitney U* kriterijus. Genų raiškos, mikroskopijos vaizdų, tėkmės citometrijos ir elektrofiziologinių matavimų duomenys laikyti statistiškai reikšmingais, kai $p \leq 0,05$.

3. REZULTATAI IR JŲ APITARIMAS

Siekiant nustatyti genomo pokytį ir, esant galimybei, įvertinti nustatyto pokyčio pasireiškimą molekuliniu ir/ar ląstelėse vykstančių procesų funkcinio lygmeniu, svarbu tinkamai suplanuoti tyrimo strategiją, pasirinkti metodus, kurie turimomis sąlygomis būtų efektyviausi, bei įvertinti galimas tyrimo rizikas. Tai pirminės šiame darbe atliktų eksperimentinių tyrimų užduotys, kurias atliekant reikia įvertinti, ar keliama funkcinių tyrimų tikslai gali būti įgyvendinti tiriamojo fibroblastų kultūroje, kadangi ne visų genų (ar baltymų) raiška aptinkama šiose ląstelėse. Taip pat reikia atsižvelgti, jog analizuojamas ląstelinis procesas ir jo reguliacija gali skirtis tarp ląstelių tipų (pvz., fibroblastuose ir nervinėse ląstelėse). Itin svarbu atsižvelgti į tai, kad individualios genominės, epigenominės ir metabolominės tiriamųjų savybės turi įtakos biologinei variacijai, kuri išryškėja atliekant kiekybinius (ar kokybinius) tyrimus sergančiųjų ir kontrolinių asmenų ląstelėse, todėl kontroliniai asmenys atrenkami pagal tiriamojo lytį ir amžių. Tai funkciniais tyrimams specifinės sąlygos, kurios apibrėžia ne tik šių tyrimų pasirinkimo, bet ir jų vykdymo sudėtingumą bei kompleksiskumą.

Šio tyrimo tikslas – įvertinti su tiriamųjų retomis ligomis siejamuose genuose nustatytų DNR sekos pokyčių patogeniskumą, reikšmę RNR, baltymų ir ląstelėse vykstančių procesų lygmeniu, atliekant bioinformatinius, molekulinius ir/ar funkcinius tyrimus. Rezultatai šiame skyriuje yra pateikiami pagal keturis išsikeltus uždavinius. Pirmiausia pateikiami NKS ir Sanger sekoskaitos tyrimais nustatyti bei išsamia duomenų bazių, literatūros šaltinių ir *in silico* analize atrinkti galimai patogeniniai ar patogeniniai DNR sekos pokyčiai tolesniems molekuliniams ir/ar funkciniais tyrimams. Toliau apibūdinami genomo struktūriniai persitvarkymai ir genai kandidatai, kurie galėtų būti susiję su klinikinių požymių vystymusi. Siekiant gilesnio supratimo apie atrinktų VNV poveikį baltymams ir jų funkcijoms, aprašyta šių pokyčių įtaka baltymo struktūrai ir funkcijai, taikant bioinformatinius ir molekulinės biologijos metodus. Galiausiai, atlikus molekulinius ir/ar funkcinius tyrimus, įvertinama atrinktų VNV ar nedidelės iškritos reikšmė ląsteliniams procesams ir galimas patogenezės mechanizmas.

3.1 Tiriamųjų apibūdinimas ir atrinkti genai kandidatai tolesniems tyrimams

Disertacinio darbo vykdymo metu į tyrimą buvo įtraukta 70 tiriamųjų (probandų). Iš jų 36-iems tiriamiesiems (ir kai kuriais atvejais jų šeimos nariams) atlikta molekulinių, citogenetinių ir/ar NKS tyrimų duomenų analizė, siekiant nustatyti įtariamą retą paveldimą patologijos genetinę

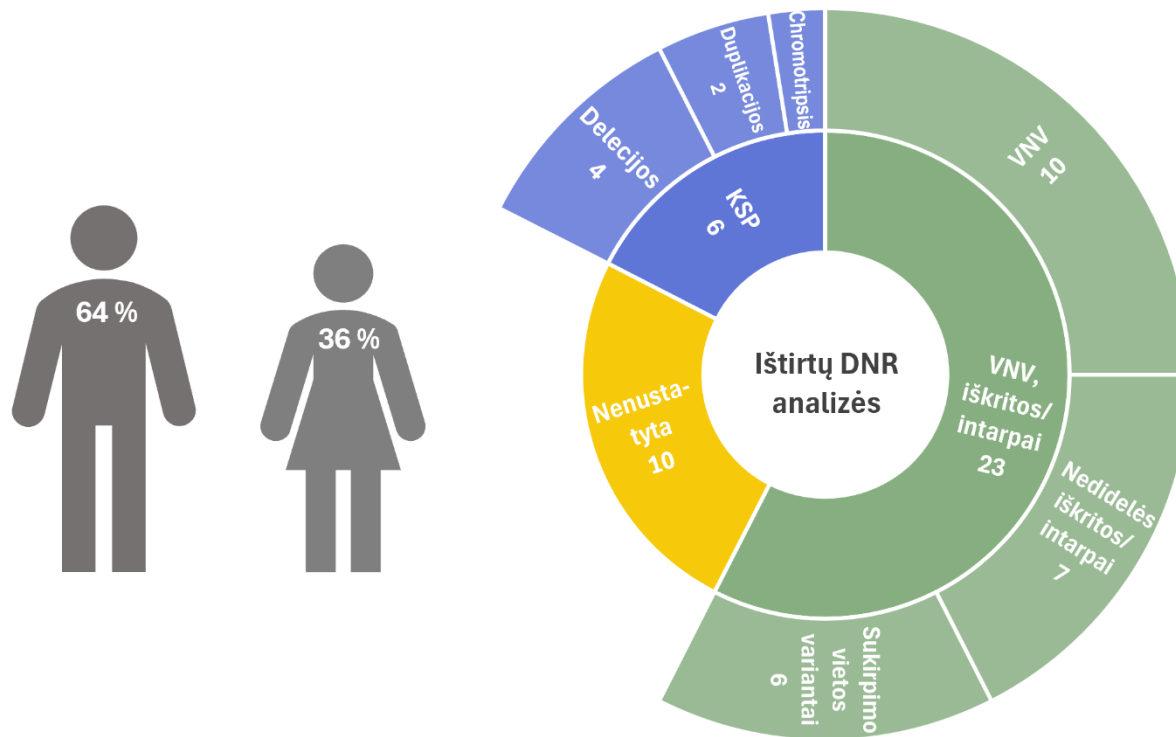
priežastį. Projektas, kurio pagrindu atlikti šio darbo tyrimai ir duomenų analizė, yra tęsiamas, todėl disertacijos rašymo metu ir po jo bus atliekama kitų į tyrimą įtrauktų 34-ių tiriamųjų (jų duomenys pateikti 10-ame priede) NKS duomenų analizė.

Iš 36-ių anksčiau minėtų tiriamųjų, kurių 64 proc. (23/36) sudaro vyriškos lyties asmenys ir 36 proc. (13/36) sudaro moteriškos lyties asmenys (žr. 3.1 grafikas), 20-čiai (55 proc. tiriamųjų) nustatyti neaišrios klinikinės reikšmės, galimai patogeniniai, patogeniniai VNV arba nedidelės iškritos/intarpai žinomuose ligos genuose (žr. 3.1 lentelė ir 3.3 lentelė). Sangerio sekoskaitos būdu DNR pokyčiai buvo nustatyti dviems tiriamiesiems (2/20), genų grupių NKS metodu – dviems tiriamiesiems (2/20), o egzomo sekoskaita – likusiems tiriamiesiems (16/20). DNR pokyčiai buvo nustatyti heterozigotinėje, homozigotinėje ir sudėtinėje heterozigotinėje būklėse. VNV sudaro 43,5 proc. visų nustatytų variantų (10/23), sukirpimo vietos variantai – 26 proc. (6/23), o nedidelės iškritos/intarpai – 30,5 proc. (7/23) (žr. 3.1 grafikas).

Šešiems tiriamiesiems (17 proc.) nustatyti KSP (žr. 3.2 lentelė ir 3.3 lentelė), lėmę probandų klinikinio fenotipo pasireišimą. VNP-LGH metodu KSP nustatyti keturiems tiriamiesiems (4/6), vLGH metodu vienam tiriamajam (1/6) ir genomo sekoskaita taip pat vienam tiriamajam (1/6). Iš viso nustatyti septyni skirtingi KSP: du duplikacijų atvejai (2/7), keturi delecijų atvejai (4/7) ir vienas chromotripsio atvejis (1/7) (žr. 3.1 grafikas). Šie tiriamieji ir jiems nustatyti KSP plačiau charakterizuojami 3.2.1 poskyryje.

Atlikus egzomo sekoskaitos duomenų analizę, 10-čiai tiriamųjų kliniškai reikšmingi DNR pokyčiai, kurie galėtų būti reto genetinio fenotipo vystymosi priežastimi nebuvo nustatyti (žr. 3.4 lentelę bei 3.1 grafiką).

Iš 36-ių tiriamųjų papildomi molekuliniai ir/ar funkciniai tyrimai disertacinio darbo vykdymo metu atlikti septyniems tiriamiesiems, kuriems nustatyti literatūroje neaprašyti DNR pokyčiai žinomuose ligos genuose (žr. 3.3 lentelė). Šie tiriamieji, jiems nustatyti DNR pokyčiai ir atlikti tyrimai išsamiau aprašyti 3.3 ir 3.4 poskyriuose.



3.1 paveikslas. Disertaciniame darbe analizuotiems tiriamiesiems nustatyti DNR pokyčiai. Iš viso disertacinio mokslinio darbo metu išanalizuoti 36-i tiriamųjų DNR mėginiai. Nustatyti 23-ys VNV ar nedidelės iškritos/intarpai, šeši KSP. Likusiems tiriamiesiems taikytais molekuliniais genetinės analizės metodais DNR pokyčiai nenustatyti.

3.1 lentelė. Tiriamieji, kuriems nustatyti DNR variantai, lemiantys retą paveldimą ligą.

Tiriamąjo (probando) kodas, lytis	Ištirti šeimos nariai	Varianto nustatymo metodas	Genas	Genotipas, nustatytas variantas	Atlikti papildomi molekuliniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–39–P, vyr.	Motina, tėvas	Sangerio sekoskaita	<i>PDE4D</i> MIM#600129	Heterozigotinis NM_001104631.2:c.581G > C	-	Akrodizostozė (MIM#614613)
IN17–03–P1, mot.	Motina, tėvas	Egzomo sekoskaita	<i>SLC27A4</i> MIM#604194	Homozigotinis NM_005094.3:c.1528C>T	Sangerio sekoskaita	Ichtiozė (MIM#608649), prikurtimas, regos sutrikimas
IN17–09–P1, vyr.	Motina, tėvas	Genų grupių NKS, egzomo sekoskaita	<i>MAST1</i> MIM#612256	Heterozigotinis NM_014975:c.1549G>A	Sangerio sekoskaita	Psichomotorinės raidos atsilikimas, ventrikulomegalija, <i>Dandy-Walker</i> anomalija
IN17–19–P1, vyr.	Brolis	Sangerio sekoskaita	<i>SLC6A8</i> MIM#300036	Hemizigotinis NM_:c.1248C>A	-	Intelektinė negalia
IN17–15–P1, mot.	Motina, tėvas	Egzomo sekoskaita	<i>CHD7</i> MIM#608892	Heterozigotinis NC_000008.11(NM_017780.4):c.5535- 1G>A	Sangerio sekoskaita (DNR, kDNR)	<i>CHARGE</i> sindromas (MIM#214800)

3.1 lentelė (tęsinys). Tiriamieji, kuriems nustatyti DNR variantai, lemiantys retą paveldimą ligą.

Tiriamąjo (probando) kodas, lytis	Ištirti šeimos nariai	Varianto nustatymo metodas	Genas	Genotipas, nustatytas variantas	Atlikti papildomi molekuliniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–22–P1, mot.	Motina, tėvas, močiutė	Egzomo sekoskaita	<i>GLI3</i> MIM#165240	Heterozigotinis NC_000007.14(NM_000168.6):c.473+3A>T	Sangerio sekoskaita (DNR, kDNR)	<i>Greig</i> cefalopolisindaktilijos sindromas (MIM#175700)
IN17–37–P, mot.	Motina, tėvas	Egzomo sekoskaita	<i>ARID1B</i> MIM#614556	Heterozigotinis NC_000006.12(NM_020732.3):c.5025+2T>C	Sangerio sekoskaita (DNR, kDNR)	<i>Coffin-Siris</i> sindromas (1-o tipo; MIM#135900)
IN17–47–P, vyr.	Motina, tėvas	Genų grupių NKS	<i>CAPN3</i> MIM#114240	Sudėtinis heterozigotinis NM_000070.2:c.598_612delTTCTGG AGTGCTCTG NC_000015.10(NM_000070.2):c.1746-20C>G	Sangerio sekoskaita	Galūnių–juosmens raumenų distrofija (1-o tipo; MIM#253600)
IN17–48–P, vyr.	Motina, tėvas	Genų grupių NKS	<i>TGFBR2</i> MIM#190182	Heterozigotinis NC_000003.12(NM_001024847.2):c.1600-2A>G	Sangerio sekoskaita, kPGR	<i>Loeys-Dietz</i> sindromas (2-o tipo; MIM#610168)
IN17–54–P, mot.	Motina, tėvas	Egzomo sekoskaita	<i>PCCA</i> MIM#232000	Sudėtinis heterozigotinis NG_008768.1(NM_000282.4):c.300+2dup NG_008768.1(NM_000282.4):c.(1746+1_1747-1)_(1845+1_1846-1)del	Sangerio sekoskaita (DNR, kDNR)	Propiono acidemija (MIM# 606054)

3.1 lentelė (tęsinys). Tiriamieji, kuriems nustatytas DNR variantai, lemiantys retą paveldimą ligą.

Tiriamąjo (probando) kodas, lytis	Ištirti šėimos nariai	Varianto nustatymo metodas	Genas	Genotipas, nustatytas variantas	Atlikti papildomi molekuliniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–44–P, vyr.	Motina, tėvas	Egzomo sekoskaita	<i>POMK</i> MIM#615247	Homozigotinis NM_032237.5:c.136C>T	Sangerio sekoskaita	Distroglikanopatija su smegenų anomalijomis (MIM# 615249)
IN17–55–P, vyr.	-	Egzomo sekoskaita	<i>PHF21A</i> MIM#608325	Heterozigotinis NM_016621.3:c.959_967delinsTT ACA	Sangerio sekoskaita	Intelektinė negalia, elgesio sutrikimai (MIM# 618725)
ATGC21–1–P, vyr.	Motina, tėvas	Egzomo sekoskaita	<i>NAA15</i> MIM#608000	Heterozigotinis NM_057175.5:c.382C>T	Sangerio sekoskaita, Western blot analizė	Intelektinė negalia, skeleto anomalijos (MIM#617787)
ATGC21–35–P, vyr.	Motina, tėvas, sesuo	Egzomo sekoskaita	<i>PIGN</i> MIM#606097	Sudėtinis heterozigotinis NM_176787.5:c.1247_1251del NM_176787.5:c.1942G>T	Sangerio sekoskaita	Igimtos anomalijos, hipotonija (MIM# 614080)

3.2 lentelė. Tiriamieji, kuriems nustatyti DNR variantai – KSP, lemiantys retą paveldimą sutrikimą.

Tiriamąjo (probando) kodas, lytis	Ištirti šėimos nariai	KSP nustatymo metodas	KSP	Atlikti papildomi molekuliniai ir/ar funkciniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–05–P1, mot.	-	Vektorinė LGH	arr[hg19] 13q31.3(91,166,748- 92,010,901)×3	Egzomo sekoskaita, kPGR	Intelektinė negalia, aukštas ūgis, skeleto anomalijos, įgimta širdies yda
IN17–20–P1, mot.	Motina, tėvas, dvi seserys	VNP-LGH	arr[hg19] 4q13.3(72,647,749– 74,208,199)x1	FISH	Intelektinė negalia, skeleto anomalijos, įgimta širdies yda
IN17–21–P1, vyr.	Motina, tėvas, brolis	VNP-LGH	arr[hg19] 1q21.1q21.2(146,476,526- 147,820,342)x3,4q12(53,531,672- 56,238,263)x1	-	Psichomotorinės raidos atsilikimas, kriptorchizmas, piebaldizmas
IN17–26–P1, mot.	-	VNP-LGH	arr[hg19] 16p13.11p12.3 (15,493,046- 18,069,465)×1	-	Psichomotorinės raidos atsilikimas, autizmas, kraniosinostozė, smegenėlių kirmino hipoplazija

3.2 lentelė (tęsinys). Tiriamieji, kuriems nustatyti DNR variantai – KSP, lemiantys retą paveldimą sutrikimą.

Tiriamąojo (probando) kodas, lytis	Ištirti šeimos nariai	KSP nustatymo metodas	KSP	Atlikti papildomi molekuliniai ir/ar funkciniai tyrimai	Pagrindiniai tiriamąojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–28–P1, vyr.	-	VNP-LGH	arr[hg19] 12q24.21(116,523,305- 116,621,185)x1	Sangerio sekoskaita (kDNR), kPGR, redagavimas, taikant CRISPR-Cas9 technologiją, Western blot analizę, transkriptomo sekoskaita	<i>MED13L</i> haplonepakankamumo sindromas (MIM#616789)

3.3 lentelė. Tiriamieji ir jiems nustatyti DNR pokyčiai, atrinkti papildomiems molekuliniams ir/ar funkciniais tyrimams.

Tiriamojo (probando) kodas, lytis	Ištirti šeimoniariai	DNR pokyčio nustatymo metodas	Genas (-ai)	Genotipas, nustatytas variantas/KSP	Atlikti papildomi molekuliniai ir/ar funkciniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
ATGC21–34–P, vyr.	-	Genomo sekoskaita	-	Chromotripsis	Kariotipo nustatymas, FISH, transkriptomo sekoskaita	Intelektinė negalia, epilepsija, regos sutrikimas, aukštas ūgis, skeleto anomalijos
ATGC21–3–P, vyr.	Motina, tėvas, brolis	Egzomo sekoskaita	<i>DYNC1H1</i> MIM# 600112	Heterozigotinis NC_000014.9(NM_001376.5): c.6405+1G>C	Sangerio sekoskaita (DNR, kDNR), kPGR, raiškos slopinimas siRNR, fluorescencinė mikroskopija, Western blot analizė	Intelektinė negalia ir smegenų žievės pakitimai
ATGC21–4–P, vyr.	Motina, tėvas	Egzomo sekoskaita	<i>RLIM</i> MIM#300379	Hemizigotinis NM_016120.4:c.1721T>A	Sangerio sekoskaita, baltymo struktūros modeliavimas, X inaktyvacijos tyrimas	<i>Tonne-Kalscheuer</i> sindromas (MIM#300978)

3.3 lentelė (tęsinys). Tiriamieji ir jiems nustatyti DNR pokyčiai, atrinkti papildomiems molekuliniams ir/ar funkciniais tyrimams.

Tiriamojo (probando) kodas, lytis	Ištirti šeimos nariai	DNR pokyčio nustatymo metodas	Genas (-ai)	Genotipas, nustatytas variantas/KSP	Atlikti papildomi molekuliniai ir/ar funkciniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
ATGC21–12–P1, vyr.	Motina, tėvas, brolis	Egzomo sekoskaita	<i>DNMT3A</i> MIM#602769	Heterozigotinis NM_022552.4:c.2324C>A	Sangerio sekoskaita (DNR, kDNR), baltymo struktūros modeliavimas, transkriptomo sekoskaita	<i>Tatton-Brown-Rahmane</i> sindromas (MIM#615879)
ATGC21–15–P, vyr.	-	Egzomo sekoskaita	<i>WRN</i> MIM#604611	Homozigotinis NM_000553.6:c.1578del	Senėjimo tyrimas, Western blot analizė	<i>Werner</i> sindromas (MIM#277700)
ATGC21–18–P, mot.	Motina, tėvas	Egzomo sekoskaita	<i>FAS</i> MIM#134637	Heterozigotinis NC_000010.11(NM_000043.6): c.16_30+1del	Sangerio sekoskaita (DNR, kDNR)	Autoimuninis limfoproliferacinis sindromas (1A tipo; MIM#601859)
ATGC21–19–P, vyr.	Motina, tėvas	Egzomo sekoskaita	<i>KCNQ1</i> MIM#607542	Heterozigotinis NM_000218.3:c.1111G>C	Ląstelių transfekcija, elektrofiziologiniai matavimai	Ilgo QT sindromas (MIM#192500)

3.4 lentelė. Tiriamieji, kuriems nenustatyti DNR pokyčiai, galimai lemiantys įtariamą retą paveldimą sutrikimą, atlikus citogenetinius, NKS ir/ar kitus molekulinis tyrimus.

Tiriamąo (probando) kodas, lytis	Ištirti šeimos nariai	Atlikti citogenetiniai, molekuliniai tyrimai	Genai kandidatai	Molekulinė diagnozė	Pagrindiniai tiriamąo klinikiniai požymiai/retas paveldimas sutrikimas
IN17-01-P1, mot.	-	VNP-LGH, Sangerio sekoskaita	<i>HYLS1</i> <i>POMK</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Intrauterinis augimo atsilikimas, galvos smegenų centrinių struktūrų įgimtos anomalijos
IN17-04-P1, vyr.	-	VNP-LGH, genų grupių NKS, egzomo sekoskaita	-	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Hipotonija, vėluojanti motorinė raida, skilvelių pertvaros defektas, atviras arterinis latakas
IN17-06-P1, vyr.	Motina, tėvas	VNP-LGH, genų grupių NKS, egzomo sekoskaita, Sangerio sekoskaita	<i>NRCAM</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Autizmas
IN17-07-P1, vyr.	Motina, tėvas	Vektorinė LGH, FISH, egzomo sekoskaita, Sangerio sekoskaita	<i>RAC1</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Psichomotorinės raidos atsilikimas, astigmatizmas
IN17-08-P1, vyr.	Motina, tėvas	VNP-LGH, egzomo sekoskaita	<i>HERC2</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Kairės akies mikroftalmija, koloboma, rankos 3-4 pirštų sindaktilija
IN17-11-P, mot.	Motina, tėvas	VNP-LGH, egzomo sekoskaita	-	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Specifinis mišrus raidos sutrikimas, įgimtos galvos smegenų cistos, toliaregystė

3.4 lentelė (tęsinys). Tiriamieji, kuriems nenustatyti DNR pokyčiai, galimai lemiantys įtariamą retą paveldimą sutrikimą, atlikus citogenetinius, NKS ir/ar kitus molekulinis tyrimus.

Tiriamąjo (probando) kodas, lytis	Ištirti šeimos nariai	Atlikti citogenetiniai, molekuliniai tyrimai	Genai kandidatai	Molekulinė diagnozė	Pagrindiniai tiriamojo klinikiniai požymiai/retas paveldimas sutrikimas
IN17–14–P1, vyr.	Motina, tėvas	Egzomo sekoskaita, Sangerio sekoskaita	<i>ANKRD11</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Psichomotorinės raidos vėlavimas, kraujagyslinių rezginių cistos, saikinga ventrikulomegalija, ptozė
IN17–30–P, vyr.	Motina, tėvas	VNP-LGH, genų grupių NKS, egzomo sekoskaita, Sangerio sekoskaita, kPGR	<i>RIPK4</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Įgimta širdies yda, išangės ir tiesiosios žarnos atrezija, pasagos formos inkstas, regos sutrikimas, mikrocefalija, nykščių aplazija
IN17–40–P, vyr.	Motina, tėvas, sesuo	VNP-LGH, egzomo sekoskaita, Sangerio sekoskaita	<i>TENM1</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Sunki intelektinė negalia
IN17–42–P, mot.	Motina, tėvas	VNP-LGH, egzomo sekoskaita, Sangerio sekoskaita	<i>BLM, FANCD2</i>	Kliniškai reikšmingi DNR pokyčiai nenustatyti	Aplastinė anemija

3.2 Genomo struktūrinių persitvarkymų apibūdinimas

KSP (delecijos ar duplikacijos) identifikavimas molekuliniiais genetiniais ar citogenetiniais metodais yra reikšmingas žingsnis siekiant nustatyti, ar sergančiojo klinikinių požymių priežastis glūdi jo genome. Šiame skyriuje plačiau charakterizuojami moksliniame tyrime nustatyti KSP ir jų įtaka tiriamojo fenotipinių požymių pasireiškimui. KSP nustatymas yra tik vienas iš genomino pokyčio charakterizavimo būdų. Kitas reikšmingas ir tolesniems nustatyto KSP vertinimo žingsniams svarbus etapas, atliktas šiame darbe, yra geno kandidato (–ų) nustatymas, kai tiriamajam būdingi fenotipiniai požymiai siejami su žiniomis apie konkretų geną, jo funkciją, sąsaja su konkrečiais žmogaus ir jo vystymosi etapais bei fenotipiniais bruožais.

3.2.1 KSP charakterizavimas

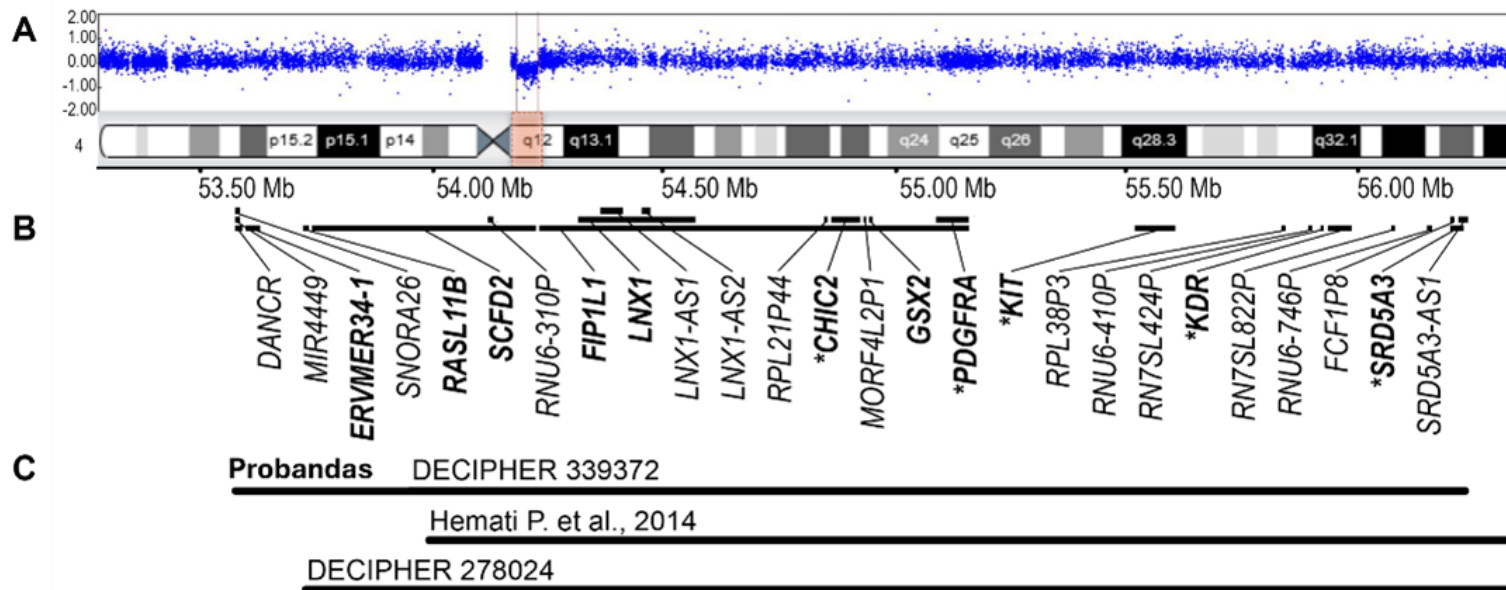
Šiame moksliniame darbe analizuota 4q12 (arr[hg19] 4q12(53,531,672-56,238,263)x1) srities 2,7 Mb dydžio šeiminė mikrodelecija, VNP oligonukleotidų lustų metodu nustatyta vienerių metų berniukui (IN17–21–P1, žr. 3.2 lentelę) su įtariamu piebaldizmu (pagrindiniai požymiai yra baltų plaukų sruoga kaktos srityje, baltos odos dėmės liemens srityje). Mikrodeleciją probandas paveldėjo iš piebaldizmo požymių turinčio tėvo. Sesers ir mamos genome ši delecija nenustatyta. Tiek mama, tiek sesuo neturi probandai būdingų klinikinių požymių [108].

Remiantis literatūra, 4q12 srities delecijos, apimančios *KIT* geną (MIM#164920), arba tik šio geno patogeniniai variantai, siejami su piebaldizmu, kuris būdingas ir mūsų tiriamajam [109-112]. Piebaldizmas (OMIM#172800) yra autosominio dominantinio paveldėjimo retas genetinis sutrikimas, kai tam tikrose odos ir plaukų srityse nėra melanino pigmentą gaminančių ląstelių melanocitų.

Orphanet duomenų bazėje pažymima, kad su piebaldizmu siejami du genai: *KIT* ir *SNAI2*. Mūsų tiriamajam ir jo tėvui nustatyta mikrodelecija iš viso apima 38-is genus (žr. 3.2.1.1 paveikslą), iš kurių keturi yra ilgų nekoduojančių RNR (inRNR) genai ir 15-a koduoja baltymus: *LNXI*, *CHIC2*, *GSX2*, *PDGFRA*, *KIT*, *KDR*, *SRD5A3*, *TMEM165*, *CLOCK*, *PDCL2*, *NMU*, *EXOC1L*, *EXOC1*, *CEP135*, *CRACD*. Disertacijos rašymo metu (2024 m. kovas) DECIPHER duomenų bazėje aštuoni iš šių genų siejami su ligomis (angl. *morbid genes*): *CHIC2*, *PDGFRA*, *GSX2*, *KIT*, *KDR*, *SRD5A3*, *CEP135*, *TMEM165*. Žemas geno haplonepakankamumo rodiklis ($pHaplo \leq 0,55$) reiškia, kad geno kopijų skaičiaus sumažėjimas, tikėtina, neturi įtakos fenotipui. Šiems genams priskirtini *CHIC2* ($pHaplo$ 0,19), *SRD5A3* ($pHaplo$ 0,09), *TMEM165* ($pHaplo$ 0,32) ir

CEP135 (pHaplo 0,27). Aukštesnę haplonepakankamumo rodiklio vertę turintys genai, kurie OMIM duomenų bazėje susieti su klinikiniais požymiais, yra *GSX2* (MIM#616253), *KDR* (MIM#191306) ir *PDGFRA* (MIM#173490). Tačiau, remiantis ta pačia duomenų baze, su pastaraisiais genais siejami klinikiniai požymiai nėra būdingi tiriamajam ir jo tėvui, todėl, tikėtina, įtakos fenotipui neturi. Heterozigotinė *KIT* geno iškrita (*pHaplo 0,87*) OMIM duomenų bazėje ir anksčiau minėtuose literatūros šaltiniuose siejama su piebaldizmu ir, tikėtina, lemia probando ir jo tėvo klinikinį fenotipą.

KIT genas koduoja transmembraninį ląstelės paviršiaus tirozino kinazės receptorių, dar neretai vadinamą protoonkogenu c-Kit, kurio ligandas yra kamieninių ląstelių augimo veiksnys (angl. *Stem Cell Growth Factor*, SCF). Geno reguliuojamas melanogenezės signalinis kelias ir jo veikla yra svarbi melanoblastų, vėliau ir melanocitų gyvenimo ciklui (migracijai, homeostazei, išgyvenamumui) tiek vystantis organizmui, tiek ir suaugusiame organizme [113]. Įdomu tai, kad piebaldizmo klinikinė išraiška priklauso nuo varianto pozicijos baltymo grandinėje. KIT baltymas sudarytas iš trijų pagrindinių domenų: N-galo užląstelinio regiono, sudaryto iš 5-ių į imunoglobuliną panašių domenų, viduryje esančio transmembraninio domeno ir C-galo viduląstelinio tirozino kinazės katalitinio domeno [114]. Variantai, nustatomi tirozino kinazės regione, lemia sunkias piebaldizmo formas, o transmembraniniame ir užląstelinuose baltymo regionuose nustatyti variantai ar visa vienos *KIT* geno kopijos delecija lemia švelnesnius piebaldizmo požymius [115]. Taigi, atlikta duomenų bazių ir literatūros analizė leidžia daryti prielaidą, kad *KIT* genas, patenkantis į nustatytos 4q12 srities mikrolelecijos sritį, lemia sutrikusį melanogenezės procesą dėl sumažėjusio geno kopijų skaičiaus ir yra pagrindinė tiriamojo klinikinio fenotipo vystymosi priežastis.

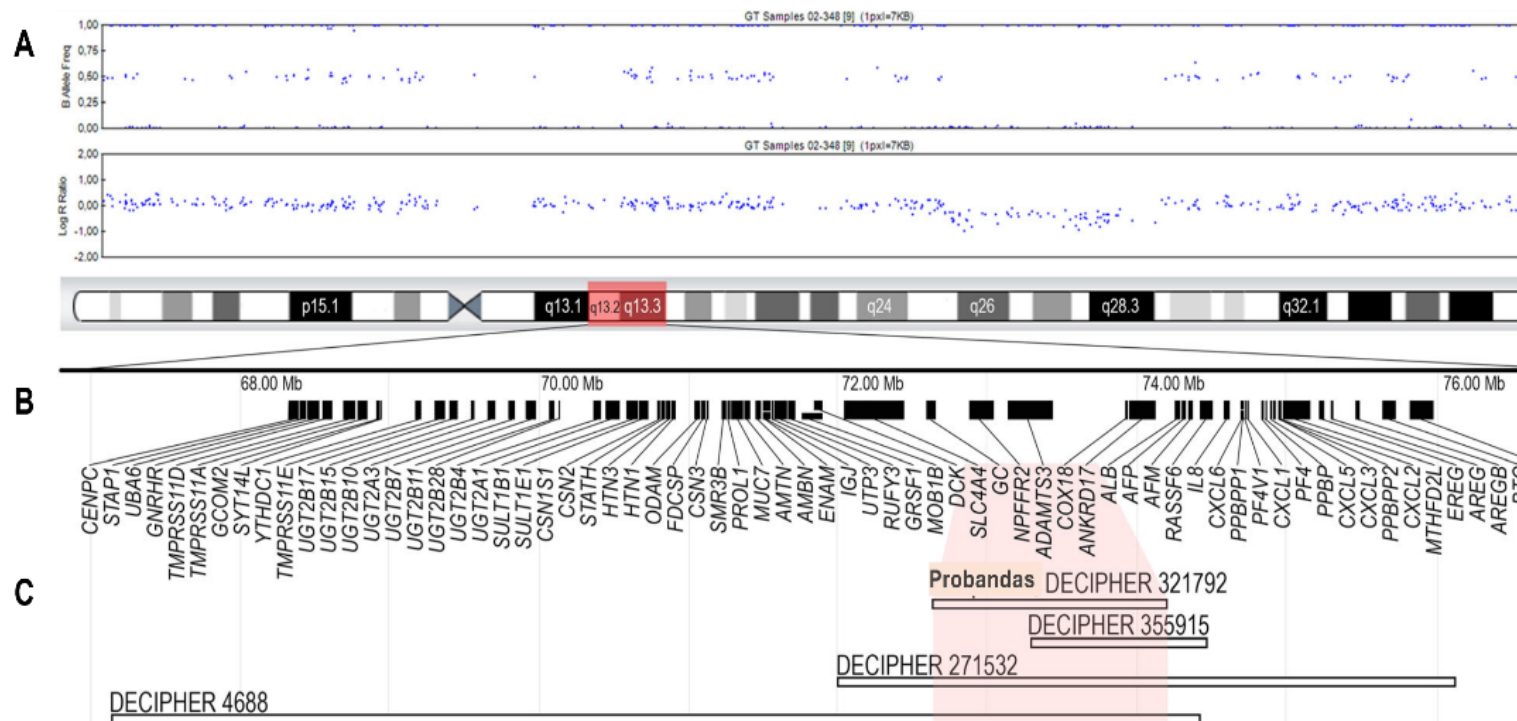


3.2.1.1 paveikslas. Mikrodelecijos schema. A – mikrodelecija. B – genai, kuriuos apima nustatytas KSP. Baltymus koduojantys genai yra paryškinti, žvaigždutės žymi fenotipo vystymuisi įtakos turinčius genus. C – DECIPHER duomenų bazėje aprašytos delecijos, persidengiančios su šiame tyrime nustatyto KSP (Probandas). Paveikslas pagal Arasimavičius et al. 2020 [108].

Kita šiame darbe analizuota 1,56 Mb dydžio 4q13.3 srities (arr [hg19] 4q13.3(72,647,749_74,208,199)×1) šeiminė mikrolelecija VNP oligonukleotidų lustų metodu nustatyta devyniolikos metų moteriai (IN17–20–P1, žr. 3.2 lentelę), kuriai nustatyta mikrocefalija, mikrognatija, taip pat būdingi grubūs veido bruožai, trumpas kaklas, stuburo deformacijos, vėluojantis brendimas, lengvas neurosensorinis klausos sutrikimas, širdies struktūros pokyčiai ir lengva intelektinė negalia. FISH analizė patvirtino tą patį KSP seseriai ir mamai. Tirta ir sveika sesuo – ši mikrolelecija nebuvo nustatyta [116].

Sergančiosios sesers pagrindiniai klinikiniai požymiai: psichomotorinės raidos atsilikimas, intelektinė negalia, netaisyklinga laikysena, trumpas kaklas, grubūs veido bruožai, mikrognatija, pirmo ir antro kojos pirštų klinodaktilija, dažnos viršutinių kvėpavimo takų infekcijos, anemija, taip pat nustatyta aortos vožtuvo anomalija, osteochondrozė, *cavum septi pellucidi*, dešiniojo inksto hipoplazija, abiejų akių astigmatizmas ir žvairumas bei lengvas neurosensorinis klausos sutrikimas. Probandės motinai nuo vaikystės nustatyta stuburo deformacija ir trumpesnė dešinė koja, intelektinė negalia, stebėti grubūs veido bruožai, hipotelorizmas, netaisyklinga laikysena, plati krūtinė, arterinė hipertenzija, hipertenzinė kardiopatija, hemodinamiškai nereikšmingas prieširdžių pertvaros defektas, neurosensorinis klausos sutrikimas, šlapimo pūslės akmenys, osteochondrozė ir stenozę [116].

Remiantis mokslinės literatūros ir DECIPHER duomenimis, aprašyti kiti keturi tiriamieji su persidengiančia 4q13.3 srities monosomija (iki 10 Mb; [117]). Quintela ir kolegų (2015) [117] nustatyta 6,85 Mb delecija tiriamajam, kuriam būdingas psichomotorinės raidos atsilikimas, lengva intelektinė negalia, elgesio sutrikimai, žemas ūgis, veido anomalijos. DECIPHER duomenų bazėje įrašytiems tiriamiesiems su 4q13.3 srities delecija taip pat būdinga intelektinė negalia, vėluojantis kalbos vystymasis arba autizmas, taip pat vienam iš aprašytų šioje duomenų bazėje (žr. paveikslė 3.2.1.2) nustatytas žvairumas ir skeleto anomalijos. Taigi, kai kurie klinikiniai požymiai atitinka mūsų tiriamos šeimos narių fenotipą. Sutampanti mikrolelecijos sritis iš viso apima 7-is genus (žr. 3.2.1.2 paveikslą), iš kurių trys koduoja baltymus: *ADAMTS3*, *ANKRD17*, *COX18*. Žemiausias haplonepakankamumo įvertis (*pHaplo* 0,19) yra *COX18*, todėl tikėtina, kad šio geno kopijų skaičius neturi reikšmingos įtakos tiriamojo fenotipui.



3.2.1.2 paveikslas. Mikrodelecijos schema. A – mikrodelecija. B – genai, kuriuos apima nustatytas KSP ir aplinkinės sritys. Baltymus koduojantys genai, patenkantys į mūsų tiriamojo KSP sirtį, pažymėti raudoname fone. C – DECIPHER duomenų bazėje aprašytos delecijos, persidengiančios su šiuo tyrime nustatytu KSP (Probandas). Paveikslas pagal Maldžienė et al. 2020 [116].

ADAMTS3 genas (*pHaplo* 0,49; MIM#605011) koduoja ADAM metalopeptidazę, turinčią 1-o tipo trombospondino 3-ią motyvą. Šis baltymas svarbus I–III tipo prokolagenų apdorojimui įvairiose ląstelėse ir audiniuose [118]. Šio geno raiška intensyvi kremzliniuose audiniuose tiek vystymosi metu, tiek po gimimo, o pelių tyrimuose nustatyta, kad geno raiška vyksta ir jungiamajame audinyje bei smegenų žievėje [119]. Mokslinėje literatūroje aprašyta, kad *ADAMTS3* geno variantai lemia kaukolės ir veido pakitimus bei sulėtėjusį augimą, todėl galima daryti prielaidą, jog šio geno haplonepakankamumas turi įtakos tokių klinikinių požymių, kaip skeleto anomalijos ir žemas ūgis bei intelektinė negalia, būdingų tiriamiesiems, vystymuisi.

ANKRD17 genas (*pHaplo* 0,99; MIM#615929) koduoja baltymą, priklausančią ankirino pasikartojimo motyvus turinčių baltymų šeimai, svarbų kraujagyslių vientisumui palaikyti embriogenezės metu [120]. Aprašytas mokslinis tyrimas, kuriame tyrėjai nustatė, jog nutildžius *Ankrd17* geną pelėms išsivysto įvairūs širdies defektai, hemoragijos ir sumažėja kraujagyslių lygiųjų raumenų ląstelių skaičius, galiausiai šios pelės nugaišta [120]. Šio geno koduojamas baltymas svarbus DNR replikacijos procese, nes ANKRD17 baltymo praradimas arba sumažėjimas lemia šio proceso sustojimą ir ląstelės ciklo progresavimo slopinimą [121]. Remiantis Chopra ir kolegų (2021) [122] straipsniu, kuriame aprašyti 34-i tiriamieji su *ANKRD17* geno koduojamo baltymo funkcijos praradimą lemiančiais variantais, taip pat šį geną apimančia 1,16 Mb dydžio mikroleccija, teigiama, kad šis genas yra pagrindinė su 4q13.3 sritimi siejamų anksčiau aprašytų fenotipų (intelektinė negalia, augimo sutrikimai, epilepsija, pasikartojančios infekcijos, oftalmologiniai sutrikimai, sąnarių hiper mobilumas, dismorfiniai veido bruožai) vystymosi priežastis. Taigi, šio geno kopijų skaičiaus sumažėjimas, tikėtina, lemia intelektinės negalios, dismorfinių veido bruožų, oftalmologinių sutrikimų bei širdies defektų, būdingų mūsų tiriamiesiems, pasireiškimą. Apibendrinant, atlikta duomenų bazių ir literatūros analizė leidžia įvardinti *ADAMTS3* ir *ANKRD17* genus, patenkančius į 4q13.3 srities mikrolecciją, kaip kandidatus, lėmusius tiriamosios klinikinio fenotipo vystymąsi. Vis dėlto, remiantis naujausia literatūra, *ANKRD17* baltymo praradimas, galimai lemiantis DNR replikacijos sutrikimą, dėl kurio sustoja ląstelės ciklas, yra pagrindinis patogeneizės mechanizmo dalyvis.

Šiame darbe taip pat analizuota 97,88 kb dydžio mikroleccija 12q24.21 srityje, nustatyta VNP–LGH metodu penkerių metų berniukui (IN17–28–P1, žr. 3.2 lentelė) su psichomotorinės raidos atsilikimu, kalbinės raidos vėlavimu, plagiocefalija, regos sutrikimu ir žvairumu. Nustatytas KSP yra vidugininė *MED13L* (MIM#608771) geno mikroleccija, kurios dydis, taikyto metodo duomenimis, nebuvo tikslus – gali varijuoti nuo 97,880 bp iki 280,666 bp. Šio

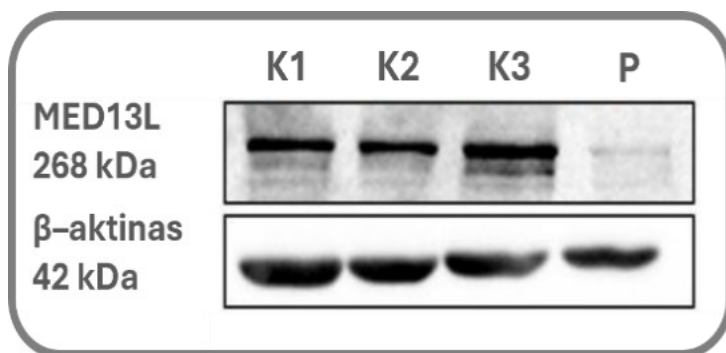
KSP sritis, t. y. iškritę egzoni, patikslinti INGENES projekto vykdytojų [123]. Kiekybinės PGR ir Sangerio sekoskaitos metodais nustatyta, kad iškreinta 3–9 egzoni *MED13L* gene (NG_023366.1(NM_015335.5):c.(310+1_311-1)(1280+1_1281-1)del). Nustatyta *in silico*, kad delecija lemia skaitymo rėmelio poslinkį ir priešlaikinės baigmės kodono susidarymą, dėl kurio baltymas, tikėtina, reikšmingai sutrumpėja (p.(Val104Glyfs*5)) ir praranda savo funkciją [123]. Nustatytas genomo pokytis yra *de novo*, tiriamojo tėvai bei brolis ir sesuo yra sveiki.

MED13L genas OMIM duomenų bazėje siejamas su retu, autosominio dominantinio paveldėjimo *MED13L* haplonepakankamumo sindromu (sutrikęs intelekto vystymasis su specifiniais veido bruožais su ar be širdies defektų; MIM#616789), kurio fenotipiniai požymiai būdingi ir tiriamajam. Remiantis literatūra ir DECIPHER duomenų baze, *MED13L* geno delecijas turinčiųjų ir šiuo sindromu sergančiųjų fenotipai gali varijuoti, tačiau visi (taip pat ir mūsų tiriamasis) turi intelektualinę negalią, psichomotorinės raidos atsilikimą ir kalbos vystymosi vėlavimą [124, 125].

MED13L genas koduoja mediatoriaus komplekso (angl. *Mediator complex*) 13L subvienetą. Mediatoriaus kompleksas yra evoliuciškai konservatyvus, masyvus baltymų kompleksas, kuris žmogaus ląstelėse veikia kaip II-osios RNR polimerazės transkribuojamų genų transkripcijos koaktyvatorius [126]. Šio geno raiška vyksta daugelyje žmogaus audinių ir yra nustatyta, kad jo koduojamas baltymas svarbus ankstyvam širdies, smegenų ir centrinės nervų sistemos vystymuisi [127, 128]. Moksliniais tyrimais parodyta, kad *MED13L* subvienetas svarbus Rb/E2F signaliniam keliui, kuris reguliuoja ląstelės ciklo progresavimą iš G1 į S fazę. Aktyvi E2F veiksnio forma gali paskatinti ląstelės ciklui svarbių genų raišką [129].

Siekdami patvirtinti nustatytos vidugeninės mikrodelecijos funkcinę reikšmę fibroblastuose, INGENES projekto vykdytojai pritaikė CRISPR-Cas9 technologiją ir sveiko asmens fibroblastų genome sėkmingai sukėlė heterozigotinę *MED13L* geno 2-o egzono iškritą, lemiančią skaitymo rėmelio poslinkį su priešlaikinės baigmės kodonu ir baltymo sutrumpėjimą, būdingą ir tiriamam probandui (p.(Ala25Cysfs*13)) [123]. Kontrolinių asmenų fibroblastų kDNR mėginiai po genomo redagavimo ir sergančiojo asmens fibroblastų kDNR mėginiai buvo tiriami kiekybinės PGR metodu, analizuojant pokyčio įtaką ląstelės cikle dalyvaujančių genų *RB1*, *E2F1* ir *CCNC* raiškai. Nustatyta, kad šių genų raiška yra keliolika kartų (kartais atitinkamai 22,3, 17,6, ir 11,6) sumažėjusi sergančiojo ląstelėse ir keliolika kartų (kartais atitinkamai 12,9, 11,7 ir 19,8) sumažėja redaguotose ląstelėse, lyginant su kontrolinių asmenų neredaguoto genomo ląstelių kultūrų tyrimo rezultatais [123]. Atlikto tyrimo duomenys rodo, kad tripano mėliu nusidažiusių žuvusių ląstelių yra du kartus daugiau ląstelių,

redaguotu genomu, kultūroje, lyginant su to paties asmens neredaguotomis ląstelėmis, o senėjimo tyrimas atskleidė, kad tiriamojo probando fibroblastuose yra trigubai daugiau senstančių ląstelių, lyginant su kontrolinio asmens fibroblastais [123]. Šio disertacinio darbo mokslinio tyrimo metu užbaigti funkciniai *MED13L* gene nustatyto pokyčio tyrimai atlikus imunofermentinę *MED13L* baltymo analizę Western blot metodu, kuriuo nustatyta, kad sergančiojo probando fibroblastuose *MED13L* baltymo kiekis sumažėja beveik penkis kartus, lyginant su sveikų asmenų ląstelėmis (žr. 3.2.1.3 paveikslą). Taigi, atlikta literatūros analize bei molekuliniais ir funkciniais tyrimais parodyta, kad *MED13L* geno 3–9 egzonų delecija lemia *MED13L* baltymo kiekio sumažėjimą. Šis kiekio pasikeitimas tiek RNR, tiek baltymo lygmeniu, tikėtina, turi neigiamos įtakos baltymo funkcijai ląstelėje, dėl kurios sutrinka Mediatoriaus komplekso, kaip transkripcijos koaktyvatoriaus, veikla, lemianti tiriamojo klinikinio fenotipo vystymąsi [130].

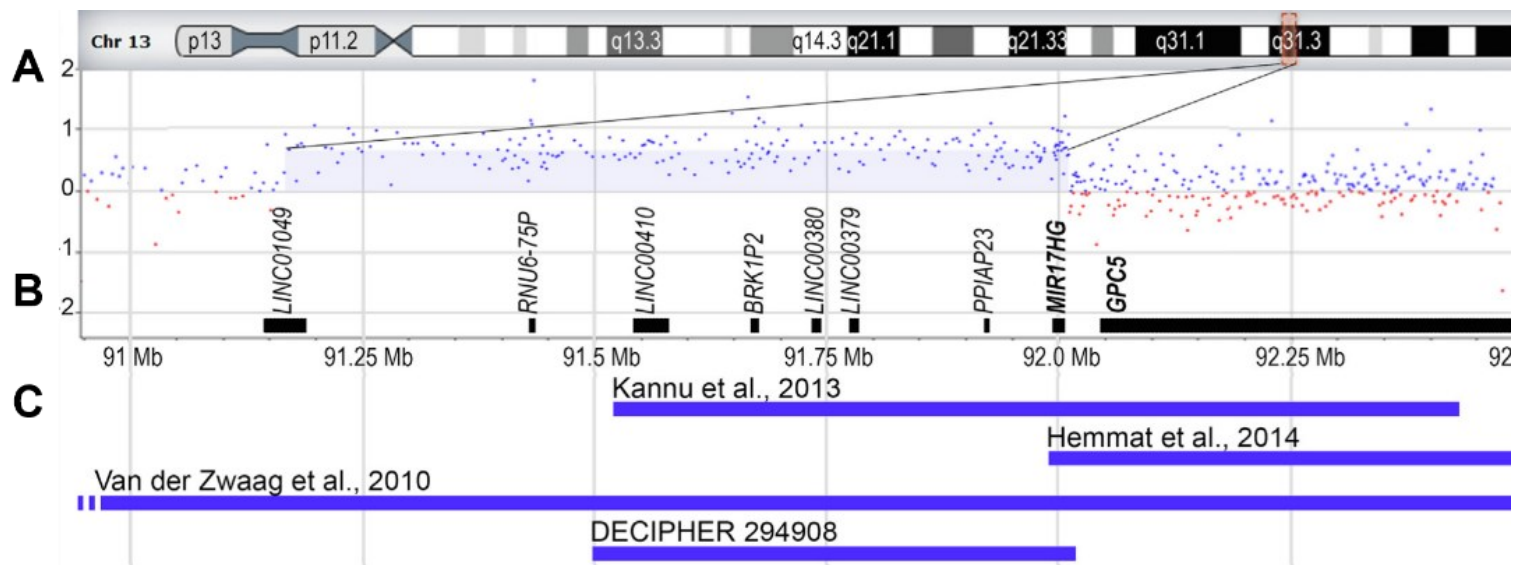


3.2.1.3 paveikslas. *MED13L* baltymo kiekis sergančiojo fibroblastuose. Imunofermentinio metodo Western blot membranos vaizdas, kuriame K1, K2, K3 – kontrolinių sveikų asmenų mėginiai, P – probando mėginys.

Moksliniame darbe analizuota vLGH metodu nustatyta 844 kb dydžio mikroduplikacija 13q31.3 srityje septyniolikos metų mergaitei (IN17–05–P1, žr.

3.2 lentelė), kurios pagrindiniai klinikiniai požymiai yra hipotonija naujagimystėje, hidrocefalija, Arnold–Chiari malformacija, psichomotorinės raidos atsilikimas, kalbinės raidos vėlavimas, regos nervo atrofija, mitralinio vožtuvo prolapsas, aukštas ūgis, *pectus excavatum*, skoliozė, hipermobilūs sąnariai, ilgi pirštai, šiurkštūs plaukai, apatinio žandikaulio asimetrija, netaisyklingas dantų sąkandis. FISH tyrimu patvirtinta, jog šis pokytis yra *de novo* [131].

Nustatyta mikroduplikacija apima keturis inRNR genus (*LINC01049*, *LINC00410*, *LINC00380*, *LINC00379*), *BRK1P2*, *PPIAP23* ir *RNU6-75P* pseudogenus bei *MIR17HG* geną, koduojantį mikro RNR (miRNR, miR) klasterį miR-17 ~ 92, kuriame yra šešios miRNR: *MIR17*, *MIR18A*, *MIR19A*, *MIR20A*, *MIR19B1*, *MIR92A1* (žr. 3.2.1.4 paveikslą). *MIR17HG* genas (MIM#609415) OMIM duomenų bazėje nėra susietas su liga. Tačiau literatūroje pažymima, kad šio geno koduojamas miRNR klasteris gali turėti įtakos augliams formotis, taip pat šios srities hemizigotinė delecija susieta su monogeniniu antro tipo Feingold sindromu (FGLDS2; MIM#614326; [132, 133]). Šio klasterio miRNR *MIR96* (MIM#611606) heterozigotiniai variantai susieti su autosominiu dominantiniu būdu paveldimu kurtumu 50 (DFNA50; MIM#613074), taip pat *MIR184* (MIM#613146) heterozigotiniai variantai siejami su endoteline distrofija, rainelės hipoplazija, įgimta katarakta ir ragenos stromos išplonėjimu (EDICT; MIM#614303). Verta atkreipti dėmesį, kad į nustatytą mikroduplikaciją nepatenka telomerinėje pokyčio srityje esantis genas *GPC5* (MIM#602446). Siekdami atmesti *GPC5* geno įtaką tiriamosios fenotipo vystymuisi, INGENES projekto vykdytojai atliko šio geno raiškos analizę kPGR metodu ir nustatė, jog raiška yra nepakitusi, lyginant su kontrolinių sveikų asmenų RNR mėginiais [131]. Taip pat, probandai atlikta VES, kuria nenustatyti kliniškai reikšmingi DNR pokyčiai [131].



3.2.1.4 paveikslas. Mikroduplikacijos schema. A – 844 kb dydžio mikroduplikacijos vieta. B – genai, kuriuos apima KSP ir aplinkinės sritys. C – DECIPHER duomenų bazėje aprašytos delecijos, persidengiančios su šiame tyrime nustatytu KSP. Paveikslas pagal Siavrienė et al. 2020 [131].

Idomu tai, kad FGLDS2 sindromo kritine laikoma hemizigotinės delecijos sritis apima tiek *MIR17HG* geną, tiek pirmuosius *GPC5* geno egzonus, kurie, manoma, lemia pagrindinius šio sindromo klinikinius požymius: mikrocefaliją, žemą ūgį, pirštų anomalijas, širdies anomalijas ir lengvą intelektinę negalią [133]. Vienos *MIR17HG* geno kopijos išjungimas pelių modelyje sukelia FGLDS2 sindromui būdingus požymius [133]. Priešingai, aprašyti keli mikroduplikacijų, apimančių *MIR17HG* geną ir *GPC5* geno pirmuosius egzonus, atvejai, iš kurių vienam tiriamajam būdingas vystymosi vėlavimas, autizmo požymiai, augimo pagreitėjimas, makrocefalija ir postaksialinė A tipo polidaktilija, o kitu atveju pasireiškė žemas ūgis ir brachidaktilija [134, 135]. Kadangi mūsų tiriamajai būdingas aukštas ūgis, makrocefalija, vystymosi atsilikimas, skeleto ir pirštų anomalijos, tikėtina, kad padidėjusi *MIR17HG* srities raiška gali lemti minėtų fenotipinių požymių vystymąsi. Remiantis analizuota literatūra ir 13q31.3 srities mikroduplikacijos nustatymo faktu, teigiama, kad miR-17 ~ 92 klasterio duplikacija lemia naują sindromą, kurio klinikiniai požymiai yra priešingi 2-o tipo Feingold sindromui [131]. Tačiau verta atkreipti dėmesį, kad Hemmat ir kolegų (2014) [135] aprašytu, šį klasterį apimančios duplikacijos atveju tiriamajam pasireiškia žemas ūgis, todėl tikėtina, kad veikia dar neištirtas molekulinis mechanizmas, dėl kurio *MIR17HG* geno kopijų skaičiaus padidėjimas gali lemti tiek aukštą, tiek žemą ūgį. Taigi, atlikta duomenų bazių, literatūros analizė bei molekuliniai tyrimai leidžia daryti prielaidą, kad tiriamosios fenotipo vystymuisi didžiausią įtaką turi į 13q31.3 srities mikroduplikaciją patenkantis *MIR17HG* genas, kurio padidėjęs kopijų skaičius lemia sutrikusį, jo koduojamų miRNR genų raiškos veiklą.

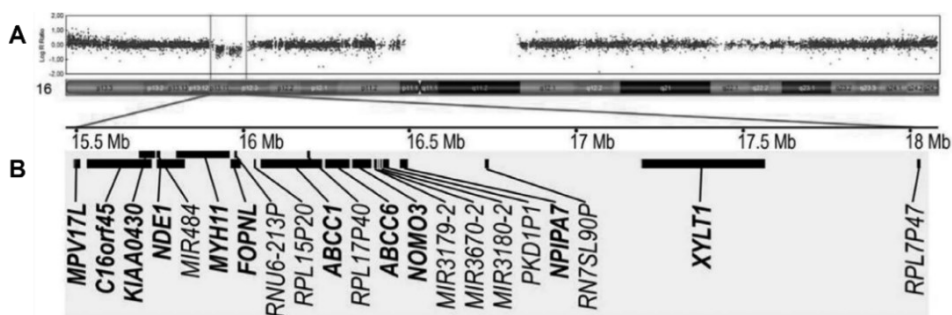
Šiame darbe analizuota 2,6 Mb delecija 16p13.11-p12.3 genomineje srityje, nustatyta VNP–LGH metodu ketverių metų mergaitei (IN17–26–P1, žr. 3.2 lentelę), kuriai būdingas psichomotorinės ir kalbinės raidos atsilikimas, dolichocefalija, sagitalinė kraniosintozė, smegenėlių kirmino hipoplazija, pusiausvyros sutrikimai, mažosios veido anomalijos. Ši mikrodelecija VNP–LGH metodu nenustatyta sveikų tėvų DNR mėginiuose [136].

Remiantis DECIPHER duomenų baze, įvairaus dydžio mikrodelecijų, apimančių 16p13.11 sritį, nustatyta daugiau nei 100-ai atvejų. Paskelbtuose klinikiniuose atvejuose atsikartojantys fenotipiniai požymiai yra psichomotorinės raidos atsilikimas, intelektinė negalia, epilepsija, autizmas, mažosios veido ir įgimtos anomalijos, taip pat įvairios skeleto anomalijos (žemas ūgis, sindaktilija, kamptodaktilija, skoliozė, krūtinkaulio anomalijos). Priešingai, tiriamajai būdinga kraniosintozė aprašyta asmeniui su 16p13.11 mikroduplikacija [137].

Mikrodelecija iš viso apima 21-ą geną (žr. 3.2.1.5 paveikslą), iš kurių keturi koduoja miRNR (*MIR484*, *MIR3179-2*, *MIR3180-2*, *MIR6770-2*), šeši yra

pseudogenai (*PKDIP1*, *RN7SL90P*, *RNU6-213P*, *RPL15P20*, *RPL17P40*, *RPL7P47*) ir 11-ka koduoja baltymus: *ABCC1*, *ABCC6*, *C16orf45*, *FOPNL*, *KIAA0430*, *MPV17L*, *MYH11*, *NDE1*, *NOMO3*, *NPIPA7*, *XYLT1*. Remiantis DECIPHER duomenų baze, *ABCC6*, *MYH11*, *NDE1* ir *XYLT1* genai yra išskiriami kaip, tikėtina, turintys įtaką klinikinio fenotipo vystymuisi. Tačiau OMIM duomenų bazėje su *ABCC6* (MIM#603234), *MYH11* (MIM#160745) ir *NDE1* (MIM#609449) siejami tiek autosominiu dominantiniu būdu, tiek autosominiu recesyviu būdu paveldimi klinikiniai fenotipai neatspindi mūsų tiriamajai būdingų klinikinių požymių.

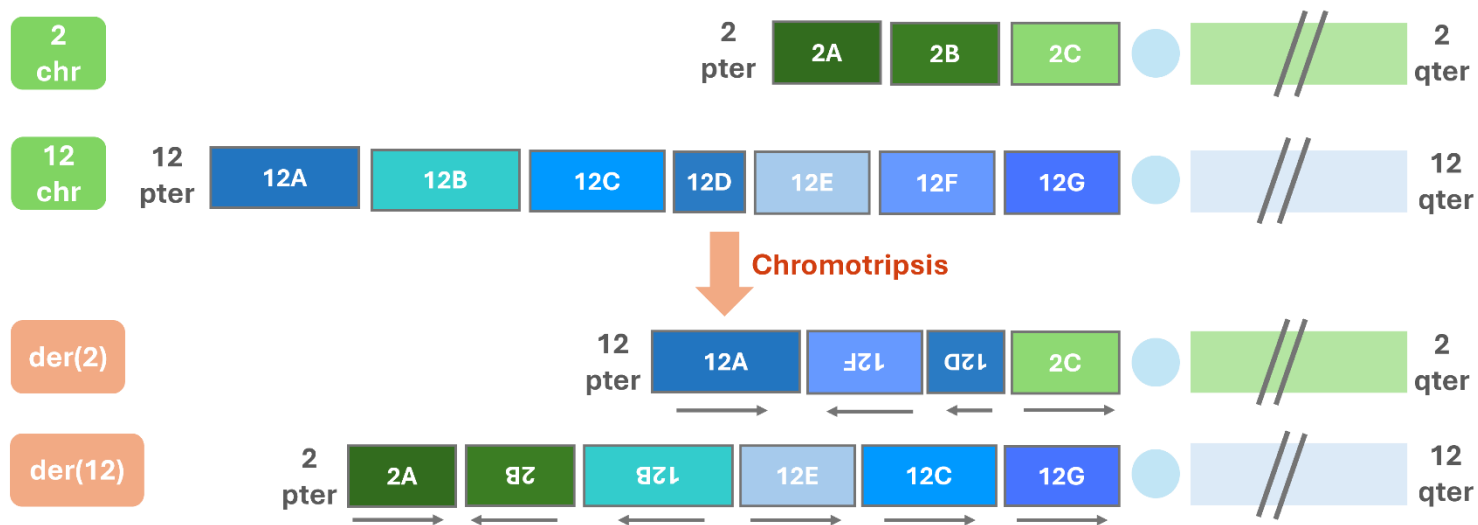
XYLT1 genas (MIM#608124; *pHaplo* 0,93) koduoja fermentą ksiloziltransferazę 1, kuris yra svarbus glikozaminoglikano grandinių biosintzei [138]. OMIM duomenų bazėje šio geno homozigotiniai variantai siejami su *Desbuquois* displazija 2, kuri pasireiškia skeleto anomalijomis. Aprašytas atvejis, kai berniuko, kurio pagrindiniai požymiai yra trumpų galūnių skeleto displazija, vLGH metodu viename alelyje nustatyta delecija, apimanti *XYLT1* geną, o kitame alelyje Sangerio sekoskaita nustatyta keliolikos nukleotidų delecija šio geno 7-o introno ir 8-o egzono jungties srityje [139]. Tikėtina, kad *XYLT1* genas gali turėti įtakos ir mūsų tiriamosios skeleto anomalijų vystymuisi. Vis dėlto tvirtų genų kandidatų, paaiškinančių kitus tiriamosios klinikinius požymius, įvardinti šiuo metu nėra galimybės. Tai tik signalizuoja apie galimus reguliacinius elementus nustatyto KSP srityje, kurie galimai daro įtaką specifinių genų raiškai. Taigi, atlikta duomenų bazių, literatūros analizė bei molekuliniai tyrimai leidžia daryti prielaidą, kad *XYLT1* genas, patenkantis į 16p13.11-p12.3 srities mikrodeleciją, lemia bent dalies tiriamajai būdingų klinikinių požymių vystymąsi dėl sutrikusios geno koduojamo baltymo fermentinės reakcijos, svarbios biosintezėje glikozaminoglikanų, būtinų įvairių audinių struktūrai ir funkcijai palaikyti.



3.2.1.5 paveikslas. Mikrodelecijos schema. A – nustatyta mikrodelecija. B – genai, kuriuos apima nustatytas KSP. Paryškinti baltymus koduojantys genai. Paveikslas pagal Maldžienė et al. 2019 [136].

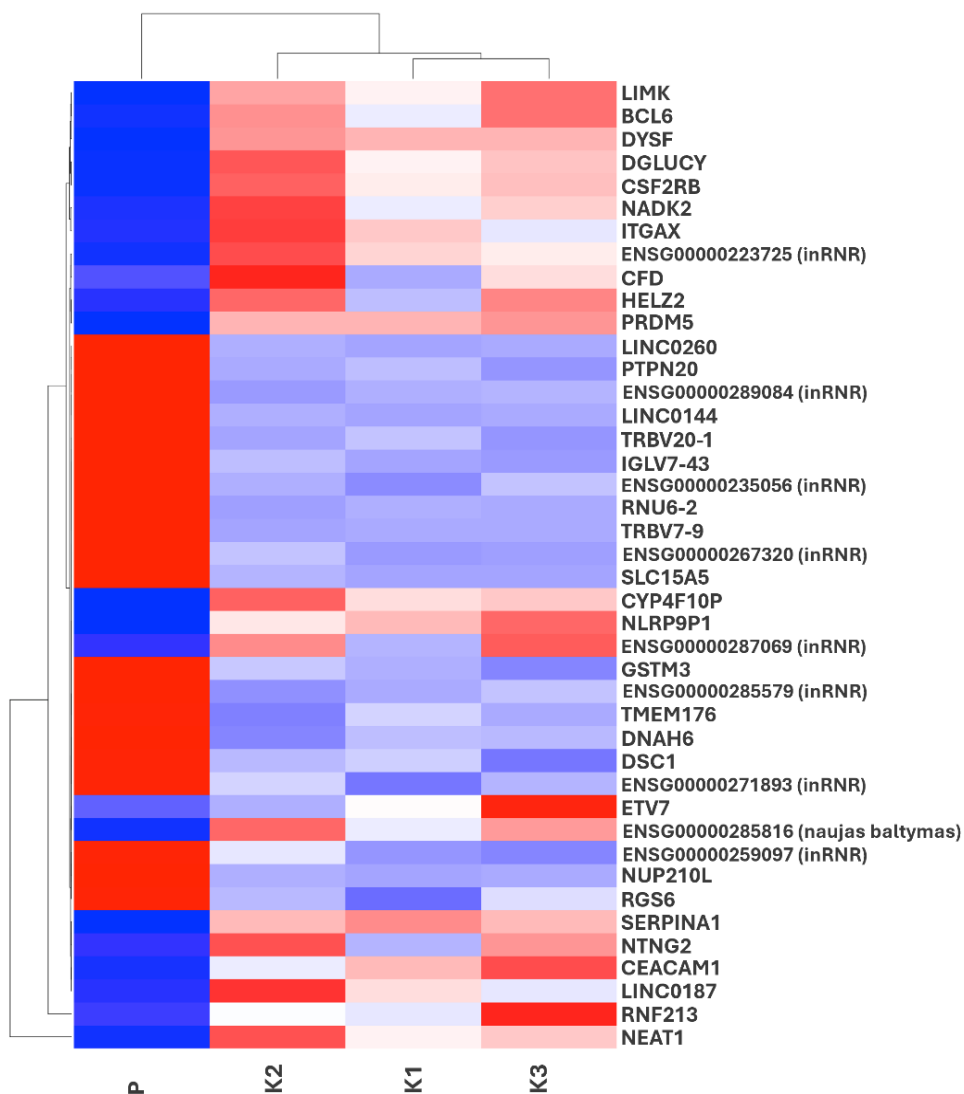
3.2.2 Chromotripsio charakterizavimas ir jo įtakos genų raiškai analizė

Šiame moksliniame darbe analizuoti išskirtinio tiriamojo (ATGC21–34–P, žr. 3.3. lentelę), kuriam diagnozuota intelektinė negalia, epilepsija, ataksija, tremoras, skoliozė, arachnodaktilija, žvairumas, toliaregystė ir aukštas ūgis, tyrimų duomenys. Įtariant retą paveldimą sutrikimą, atliktas kariotipo tyrimas, kuriuo nustatyta *de novo* subalansuota translokacija tarp antros ir dvyliktos chromosomų trumpųjų pečių ($t(2;12)(p16.2;p12.2)$). Atlikus vLGH tyrimą, patvirtinta, kad translokacija yra subalansuota. Siekiant nustatyti, ar translokacijos srityje nėra pertrauktų genų, atliktas tyrimas FISH metodu – antroje chromosomoje, netoli citogenetiniu tyrimu nustatytų trūkio taškų esantis *NRXN1* (MIM#600565) genas nepakitęs. Nuspręsta patikslinti chromosominio pokyčio trūkio taškus atliekant tiriamojo DNR mėginio, gryninto iš periferinio kraujo ėminio, viso genomo sekoskaitą. Šiuo metodu buvo nustatytas sudėtingas ir retas chromosominis persitvarkymas, vadinamas chromotripsiu (ang. *chromothripsis*). Antros chromosomos trumpajame petyje nustatyti du trūkio taškai, tarp kurių yra ~ 276 kb atstumas. Dvyliktos chromosomos trumpojo peties ~ 4,11 Mb ilgio fragmente nustatyti šeši trūkio taškai. Chromotripsio procese susidariusių fragmentų persitvarkymo, kuris apima translokaciją, inversijas ir insercijas, schema nurodyta 3.2.2.1 paveiksle. Persitvarkymų metu nebuvo prarasta genetinė informacija – pokytis yra subalansuotas. Antros chromosomos trumpojo peties fragmentas ir artimiausios sekos aplink trūkio taškus genų neapima. Tačiau dvyliktos chromosomos trumpojo peties persitvarkymo regione yra šių genų sekos: *LMO3*, *SKPIP2*, *LINC02378*, *RERGL*, *PIK3C2G*, *PLCZ1*, *PLEKHA5*, *AEBP2*, *LINC02398*. Šioje chromosomoje trūkio taškai pertraukia du genus: pirmasis trūkio taškas yra *MGST1* gene, o paskutinis yra *PDE3A* (MIM#123805) gene. Vienas genas (*PDE3A*) OMIM duomenų bazėje yra žinomas ligos genas. Kiti septyni koduoja baltymus, tačiau nėra susieti su žinomais sutrikimais, o dar du (*LINC02378*, *LINC02398*) yra inRNR genai, kurių įtaka sveikatos sutrikimams nėra žinoma.

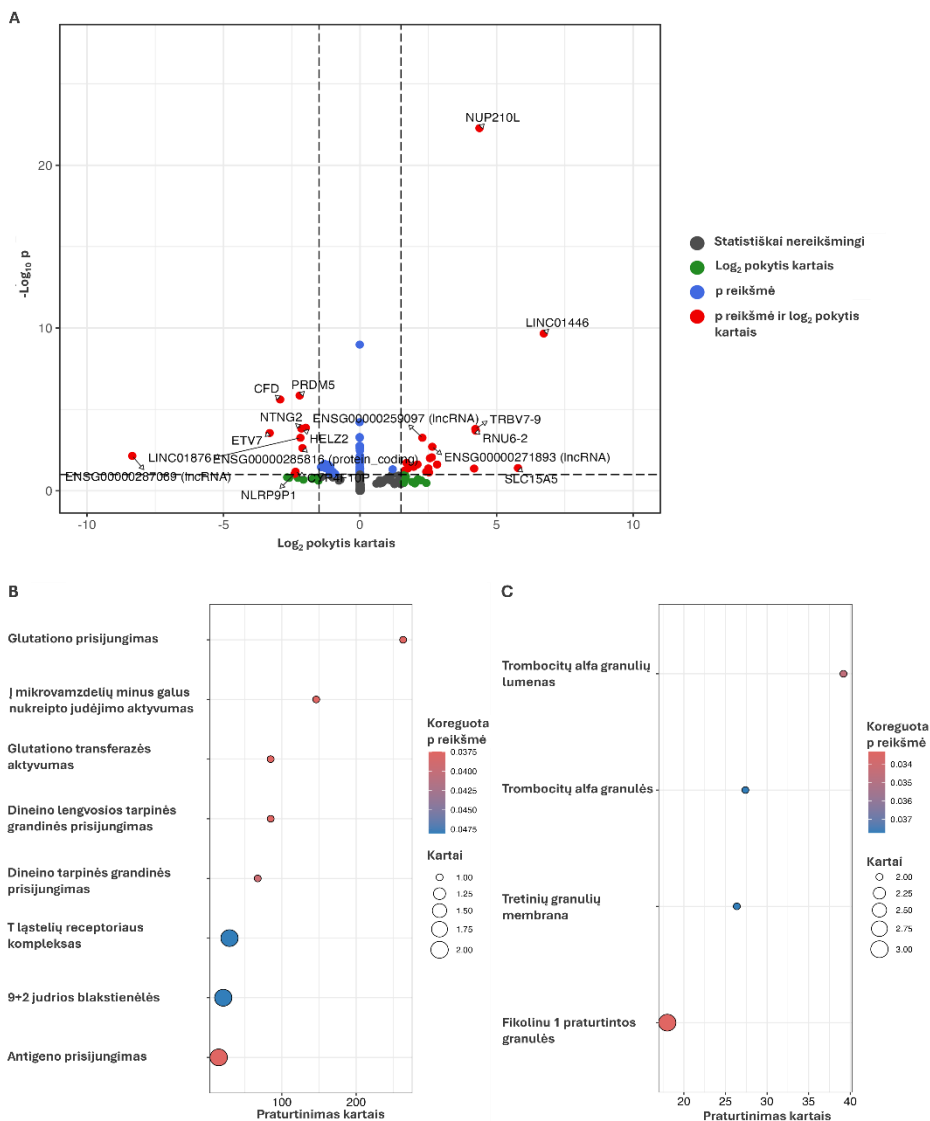


3.2.2.1 Chromotripsio schema. Chromosominiai persitavrkymai (translokacijos, insercijos, inversijos), įvykę tarp 2-os (2 chr) ir 12-os (12 chr) chromosomų. Antros chromosomos trumpajame petyje (2 pter) susidarė trys skirtingo ilgio fragmentai (2A–2C), o 12-os chromosomos trumpajame petyje (12 pter) – septyni skirtingo ilgio fragmentai (12A–12G). Chromosomų ilguosiuose pečiuose (2 qter, 12 qter) persitavrkymų nenustatyta. Pilkos rodyklės žemiau fragmentų žymi 5'→3' kryptį. Eilutėse der(2) ir der(12) pavaizduotos derivatinės (pakitusios) 2-a ir 12-ta chromosomos.

Siekiant nustatyti chromotripsį lėmusių persitvarkymų metu susidariusių trūkio taškų ir fragmentų įtaką į regioną patenkančių ir nepatenkančių genų raiškai, kadangi regione, tikėtina, lokalizuoti ir reguliaciniai genomo elementai, atlikta tiriamojo RNR mėginio (gryninto iš periferinio kraujo ėminio) transkriptomų sekoskaita. Nustatyta, kad tiriamojo mėginyje statistiškai reikšmingai pakito 42-jų genų raiška, lyginant su trimis panašaus amžiaus kontrolinių sveikų asmenų mėginiais (žr. paveikslą 3.2.2.2). Sumažėjusios raiškos genų grupei priklauso 22 genai, o padidėjusios – 20 genų. Iš 42-jų genų trisdešimt koduoja baltymus, 11-a yra su ligomis nesusieti inRNR genai ir vienas transkriptas įvardinamas naujo baltymo genu (žr. paveikslą 3.2.2.2, ENSG00000285816). Iš baltymus koduojančių genų 9-ni OMIM duomenų bazėje yra susieti su ligomis: *NUP210L*, *PRDM5*, *NTNG2*, *RNF213*, *CFD*, *DYSF*, *CSF2RB*, *NADK2*, *SERPINA1*. *NUP210L* geno raiška yra padidėjusi, o visų kitų išvardintų genų raiška probando RNR mėginyje – sumažėjusi (žr. 3.2.2.3 paveikslą). Pastarųjų genų lemiami fenotipai yra autosominio recesyvaus paveldėjimo. Chromotripsio procese dalyvavusių 2-os ir 12-os chromosomų trumpuosiuose pečiuose lokalizuojasi tik trys genai (*DNAH6*, *SLC15A5*, *DYSF*), iš jų *DYSF* (MIM#603009) yra žinomas ligos genas. Su *NTNG2* (MIM#618689) genu susietam neurovystymosi sutrikimui (MIM#618718) būdingi fenotipiniai bruožai tik iš dalies persidengia su mūsų tiriamojo ir galimai galėtų būti intelektinės negalios, tremoro, ataksijos priežastis. Tačiau *NTNG2* gene homozigotiniai variantai lemia bendrą augimo atsilikimą ir žemą ūgį, kai tiriamasis pasižymi aukštu ūgiu. Kitų genų lemiami fenotipai nepersidengia tiriamojo.



3.2.2.2 paveikslas. Genų raiškos pokyčiai tiriamojo, kuriam nustatytas chromotripsis, kraujo RNR mėginyje. Genų raiškos profilyje pažymėti statistiškai reikšmingą raiškos pokytį turintys genai. Kiekviena eilutė atitinka geną, o stulpelis reprezentuoja tiriamąjį. P – probandas, K1–K3 – kontroliniai sveiki asmenys. Raudona – raiškos padidėjimas, mėlyna – raiškos sumažėjimas, balta – vidutinė/nepakitusi raiška. Grafikas sukurtas dr. A. Konovalovo.



3.2.2.3 paveikslas. Chromotripsio poveikio tiriamojo genų raiškai analizė. A – statistiškai reikšmingai pakitusios raiškos genų tarp tiriamojo su chromotripsiu ir sveikų kontrolinių asmenų grafikas. X ašis reprezentuoja genų raiškos pokytį kartais $\log_2$ skalėje, o Y ašis rodo FDR neigiamą dešimtainį logaritmą. Reikšmingai pakitusios raiškos genai ($FDR < 0,1$) ir daugiau nei 1,5 karto didesniu raiškos pokyčiu vaizduojami virš horizontalios brūkšninės linijos ir už vertikalių brūkšnių. B – statistiškai reikšmingai padidėjusios raiškos genų GO terminų praturtinimo taškinis grafikas. C – statistiškai reikšmingai sumažėjusios raiškos genų GO terminų praturtinimo taškinis grafikas. Grafikai sukurti dr. A. Konovalovo.

Atlikus GO terminų praturtinimo analizę, padidėjusios raiškos grupei priklausantys genai daugiausiai susieti su ląstelės ir joje vykstančiu judėjimu mikrovamzdeliais, su imunine sistema susijusia veikla ir keliais glutationo apykaitos aspektais (žr. 3.2.2.3 B paveikslą), o lyginant su kontroliniais sveikais asmenimis, mažesnę raišką turintys genai praturtino su imunine sistema susijusius GO terminus (žr. 3.2.2.3 C paveikslą). Atsižvelgiant į tiriamojo klinikinius bruožus, imuninės sistemos veiklos sutrikimai tiesiogiai nesisieja su sergančiojo fenotipu, tačiau ląstelėje vykstantis judėjimas mikrovamzdeliais, kuris yra svarbus įvairių ląstelės krovinių pernašai, todėl turi įtakos įvairiems ląstelės procesams, yra potenciali molekulinė sąsaja su tiriamajam būdingais klinikiniais fenotipo požymiais.

Šiame darbe probandai nustatytas chromotripsis yra retas chromosominis persitvarkymas, kadangi dažniausiai jis pasireiškia vėžinėmis ligomis sergantiesiems kaip somatinis pokytis ląstelėse, o konstitucinio chromotripsio atvejai nustatomi retai [140, 141]. Chromotripsiu apibūdinamas procesas, kai daugybiniai trūkio taškai suformuoja įvairaus ilgio fragmentus, kurių persitvarkymas apima skirtingus chromosominius pokyčius: inversijas, insercijas, translokacijas, delecijas [142]. Literatūroje yra aprašyti šio molekulinio fenomeno, apimančio skirtingas chromosomas, įgimti atvejai tiriamiesiems su įvairiomis įgimtomis anomalijomis ir/ar intelektine negalia, neurologiniais sutrikimais ir pakartotiniais persileidimais [140, 141, 143-145].

Chromotripsio klinikinė reikšmė dažniausiai priklauso ir yra vertinama atsižvelgiant į persitvarkymo regionuose esančius genus. Probando atveju vienas iš trūkio taškų nustatytas *PDE3A* geno srityje, kurio dalis iš dvyliktos chromosomos translokuojama į antros chromosomos trumpąją petį. OMIM duomenų bazėje šis genas yra siejamas su autosominio dominantinio paveldėjimo hipertenzijos ir brachidaktilijos sindromu (MIM#112410). Šiam sindromui būdingi pagrindiniai fenotipiniai bruožai yra žemas ūgis, anksti pasireiškianti hipertenzija ir trumpi bei stori plaštakų ir pėdų kaulai (E tipo brachidaktilija). Šie klinikiniai požymiai neatitinka mūsų probando fenotipo, taip pat kai kurie iš paminėtų požymių (žemas ūgis, trumpi pėdų ir plaštakų pirštai) yra priešingi mūsų tiriamajam. Likę chromotripsio regionų genai OMIM duomenų bazėje nėra susieti su specifiniais sutrikimais. Į neaiškią klinikinę reikšmę turinčių DNR sekų grupę patenka ir dvi inRNR, kurios dar nėra funkciškai ištirtos. Todėl galima daryti prielaidą, kad trūkio taškai ir įvykę chromosominiai persitvarkymai pažeidžia ar paveikia reguliacines genomo sekas (pvz., genų promotorius, slopiklius, stipriklius ar nutolusius reguliacinius elementus), kurios gali daryti įtaką specifiniams fenotipiniams požymiams vystytis.

Middelkamp ir kolegos (2017) [141] atliko tiriamojo su įvairiomis įgimtomis anomalijomis (kraniosintoze, veido dismorfizmais, augimo atsilikimu) ir intelektine negalia, kuriam nustatytas paveldimas chromotripsis, indukuotų nervinių ląstelių transkriptomo sekoskaitą, siekiant įvertinti funkcinę chromosominio sutrikimo reikšmę genų raiškai. Tyrėjai sėkmingai nustatė pakitusią chromotripsio persitvarkymo regionuose esančių genų ir netolimose aplinkinėse sekose esančių genų (iš viso 67-ių baltymus koduojančių genų) raišką [141]. Atlikus mūsų probando iš kraujo grynintos RNR transkriptomo sekoskaitą, taip pat nustatyta, kad pakinta 42-jų genų raiška, iš kurių trisdešimt yra baltymus koduojantys genai, tačiau tik devyni iš jų yra žinomi ligos genai OMIM duomenų bazėje. Įdomu tai, kad tik trys iš OMIM genų patenka į tiriamojo chromotripsio sritį. Fenotipo atžvilgiu tiriamojo klinikiniai bruožai iš dalies persidengia su keliais *NTNG2* geno siejamam neurovystymosi sutrikimui būdingais požymiais: intelektine negalia, tremoru, ataksija. Tačiau pastarajam sutrikimui būdingas bendras augimo atsilikimas ir žemas ūgis, o tiriamasis yra aukšto ūgio. Taip pat svarbu, kad šio geno raiška mūsų sergančiojo RNR mėginyje nustatyta mažesnė lyginant su sveikais asmenimis, o *NTNG2* lemiamas sutrikimas yra autosominio recesyvaus paveldėjimo tipo, todėl tikėtina, kad abejuose geno aleliuose koduojamo baltymo funkcija turėtų būti sutrikusi tam, kad pasireikštų daugelis su *NTNG2* susietų klinikinių požymių, ypač išreikšta intelektinė negalia ir kalbos vystymosi sutrikimas. Genas koduoja netriną G2, kuris nurodo gaires aksonams, yra svarbus sinapsių formavimuisi ir jų plastiškumo susidarymui centrinės nervų sistemos vystymosi metu. Dias ir kolegų (2019) aprašytame moksliniame tyrime, kuriame analizuoti *NTNG2* homozigotiniai variantai, pabrėžiama, kad pakankama geno raiška yra reikalinga neurotipiniam vystymuisi [146]. Kai kuriems mokslininkų aprašytiems sergantiesiems, be anksčiau minėtų požymių, būdinga ir epilepsija bei žvairumas – šiuos požymius turi ir tiriamasis. Siekdami išsiaiškinti patogenezės mechanizmą, Dias ir kolegos siRNR molekulėmis sumažino *NTNG2* raišką (pasiekta du kartus mažesnė raiška, lyginant su nepaveiktomis ląstelėmis) pelių neuroblastomos N2A ląstelėse ir nustatė, kad mažesnę geno raišką turinčiose kultūrose neuritų (dendritų ir/ar aksonų pirmtakų) kiekis, ilgis ir dendritų vystymosi lauko plotas yra sumažėję [146]. Įdomu tai, kad Pan ir kolegų (2010) aprašytame tyrime *NTNG2* padidėjusi raiška nustatyta epilepsija sergančių žmonių temporalinės neožievės biopstatuose ir žiurkių hipokampo bei greta esančios smegenų žievės audinių ląstelėse [147], o tiriamajam būdinga epilepsija ir su mažesne šio geno raiška. Aprašyti tyrimai leidžia daryti prielaidą, kad tiriamojo neurologiniai požymiai galėtų būti nulemti *NTNG2* koduojamo baltymo funkcijos sutrikimo. Tačiau prielaidai pagrįsti geno raiškos pokytis turėtų būti patvirtintas kitu metodu (kPGR) ir neturėtų būti

atmesta kitų pakitusios raiškos genų, kurie nėra susieti su ligomis ar koduoja inRNR, įtaka sergančiojo klinikiškai požymiams vystytis.

Idomu tai, kad GO terminų praturtinimo analizė sumažėjusios raiškos genams statistiškai reikšmingai priskyrė imuninės sistemos veiklas, o padidėjusios raiškos genams – net kelis su ląstelės išoriniu judėjimu ir plazmoje vykstančiu krovinių pernešimu susijusius procesus bei glutatono transsferazės, svarbios ląstelės dalijimuisi, apoptozės ir nekrozės procesuose, veikla. Vis dėlto, atsižvelgiant į tiriamojo fenotipą, anksčiau aprašytą galimą kandidatinių genų, išvegti tiesioginę sąsają tarp pakitusių genų raiškos profilio ir tiriamojo klinikiškai požymių yra sudėtinga, todėl ją patvirtinti reikalingi papildomi bioinformatiniai tyrimai, kuriais būtų analizuojamos mažesnės padidėjusios ar sumažėjusios raiškos genų grupės, siekiant tikslingai atsekti specifinius genus, potencialiai lemiančius patogenezės mechanizmą.

Atlikti bioinformatiniai ir molekuliniai tyrimai bei literatūros analizė pagrindžia probandui nustatyto chromotripsio klinikinę ir funkcinę reikšmę. Nustatyti genų raiškos pokyčiai leidžia įtarti 42-jų baltymų ir inRNR koduojančių genų įtaką tiriamojo fenotipo vystymuisi, pabrėžiant *NTNG2* geno sumažėjusią raišką ir jo koduojamo neutrino G2 sumažėjimą kaip galimą geną kandidatą ir jo vaidmenį patogenezėje.

Šiame tyrimo rezultatų ir jų aptarimo poskyryje pateikta KSP ir genomo struktūrinio persitvarkymo chromotripsio charakterizavimas atliekant duomenų bazių ir literatūros analizę leido įvardinti genus kandidatus, kaip patogenezės mechanizmo dalyvius, lėmusius tiriamųjų klinikinio fenotipo pasireiškimą. Kai kuriais atvejais tikėtina ir reguliacinių genomo elementų įtaka, kadangi vieno geno kandidato nepakanka pilnam fenotipui paaiškinti. Nustatytos 2,6 Mb delecijos 16p13.11-p12.3 genomineje srityje, *MED13L* mikrodelecijos ir chromotripsio atvejais atlikti papildomi molekuliniai ir/ar eksperimentiniai tyrimai pagrindė genomo struktūrinių pokyčių klinikinę reikšmę bei padėjo iš dalies atskleisti sudėtingus nustatytų pokyčių lemiamus molekulinis procesus, dėl kurių sutrinka ne vieno geno koduojamo produkto veikla, kadangi į patogenezės mechanizmą įtraukiami genų raiškos reguliacijai svarbūs molekuliniai veiksniai.

Vis dėlto kandidatų atranka nėra galutinis KSP charakterizavimo žingsnis, kadangi toliau kandidatų įvardintas genas turėtų būti analizuojamas siekiant patvirtinti arba atmesti jo funkcinę reikšmę fenotipui vystytis – tai įvairių resursų reikalaujanti, bet galimai naujų įžvalgų ar net atradimų žadanti sritis, kuri galėtų būti paskata ateityje atlikti įvairius funkcinis tyrimus tiek ląstelėse, tiek ir audinių modeliuose ar modeliniuose organizmuose.

3.3 VNV įtakos baltymo lygmeniu nustatymas bioinformatiniais ir molekuliniais tyrimais

Atrinktus vieno nukleotido DNR pokyčius genų koduojamose sekose pirminiam biologinės reikšmės įvertinimui bioinformatiniais įrankiais atlikta baltymo struktūros bei sąveikos analizė ir taip numatytas potencialus jų poveikis geno koduojamo produkto funkcionalumui. Siekiant pagrįsti šių pokyčių biologinę reikšmę, šiame darbe dviem tiriamiesiems taikyti papildomi molekuliniai tyrimai, siekiant atskleisti konkretaus DNR varianto sukeltus ląstelių procesų sutrikimus ir detalizuoti galimus molekulinius mechanizmus.

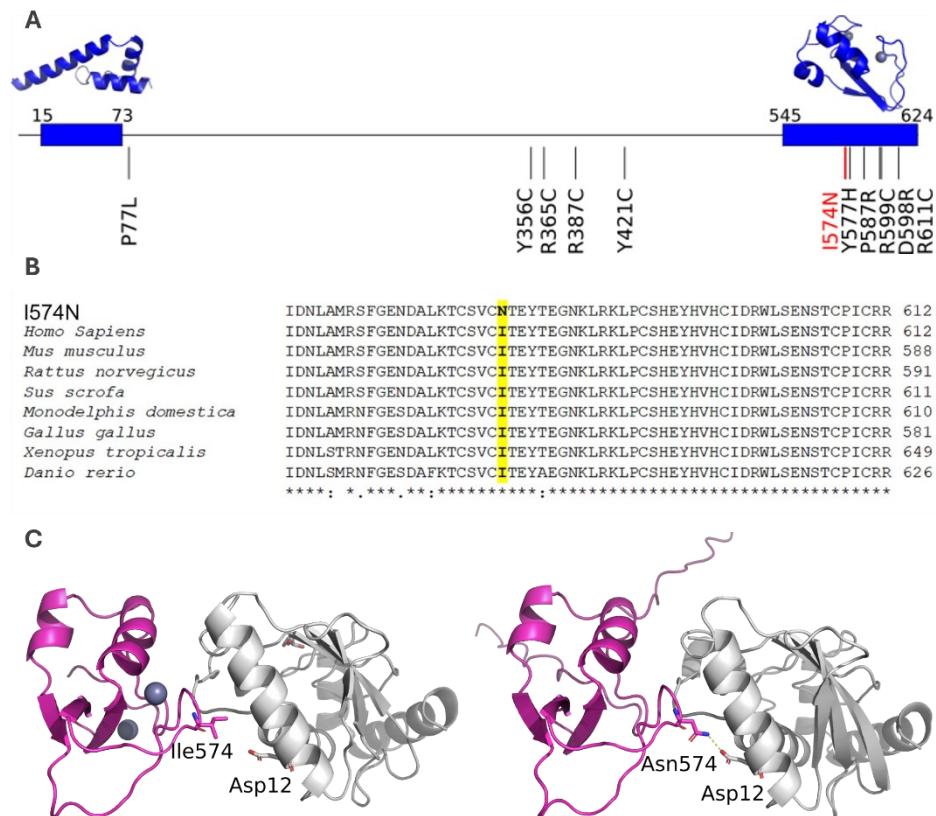
3.3.1 *RLIM* gene nustatyto VNV įtakos baltymo–baltymo sąveikai ir X chromosomos inaktyvacijai analizė

Moksliniame darbe analizuotas egzomo sekoskaita nustatytas hemizigotinis variantas NM_016120.4:c.1721T>A žinomame ligos gene *RLIM* (MIM#300379) tiriamajam (ATGC21–4–P, žr. 3.3 lentelę) su intelektine negalia, žvairumu ir kriptorchizmu. Tiriamojo mama ir jaunesnis brolis yra sveiki. Sangerio sekoskaita patvirtinta mamos nešiotojos būklė. Kitas brolis mirė kūdikystėje dėl dauginių raidos defektų. OMIM duomenų bazėje *RLIM* genas susietas su Tonne–Kalscheuer sindromu (MIM#300978), su X chromosoma susijusio paveldėjimo. Šio sindromo fenotipiniai bruožai sutampa su mūsų tiriamojo.

Nustatytas VNV yra literatūroje neaprašytas, retas variantas, neaptiktas gnomAD duomenų bazėje. Remiantis *PolyPhen-2*, *Sift*, *MutationTaster* ir *AlphaMissense* bioinformatinių įrankių įverčiais, šis variantas prognozuotas galimai patogeniniu. Nustatytas variantas baltymo lygmeniu lemia vienos aminorūgšties pokytį (NP_057204.2:p.(Ile574Asn)). Atlikus evoliuciškai nutolusių gyvūnų rūšių *RLIM* baltymo sekos palyginį, patvirtintas nustatyto pokyčio pozicijos evoliucinis konservatyvumas (žr. 3.3.1.1 paveikslą). Segregacijos ir *in silico* tyrimai leidžia teigti nustatyto varianto *RLIM* gene įtaką tiriamojo klinikinio fenotipo pasireiškimui.

Remiantis *GTEX* duomenų baze, *RLIM* geno raiška vyksta beveik visuose žmogaus audiniuose. *RLIM* (sinonimas *RNF12*) genas koduoja RING–H2 cinko pirštelių domeną turintį baltymą (angl. *Ring Finger Protein*, *LIM Domain Interacting*). Tai didžiausios baltymų ubikvitino ligazių E3 šeimos narys, kuris jungiasi prie LIM besijungiančio domeno 1 (*LIMBD1*), sudaro dimerą ir toliau reguliuoja LIM homeodomeną turinčių transkripcijos veiksnių veiklą, į kurią patenka nervinių struktūrų ir galūnių vystymasis, ląstelių diferenciacija [148]. Taip pat žinoma, kad *RLIM* dalyvauja TGF-β signalinio kelio, svarbaus ląstelių

proliferacijai, diferenciacijai, migracijai ir išgyvenamumui, reguliacijoje, nes šiame kelyje dalyvaujantis inhibitorius SMAD7 (angl. *Small Mothers Against Decapentaplegic 7*) yra ubiquitilinamas RLIM ligazės, o tai nulemia SMAD7 suirimą proteosomose [149]. Taip pat svarbu, kad ši ubiquitino ligazė RLIM yra reikšminga X chromosomos inaktyvacijos procese. Pelių modelyje parodyta, kad RLIM gali ubiquitilinti transkripcijos veiksnį REX1 (angl. *Reduced Expression 1*), nuo kurio priklauso X chromosomos inaktyvacijos proceso inicijavimas [150].



3.3.1.1 RLIM gene nustatyto VNV poveikio baltymo struktūrai analizė.

A – žmogaus RLIM baltymo schema. Paryškintos sritys, kurios AlphaFold duomenų bazės modelyje prognozuojamos labai patikimai (pLDDT > 70): trys α -spiralės N gale ir RING-H2 cinko pirštų domenas C gale. Žemiau pateikti žinomi patogeniniai RLIM variantai. B – daugybiniis RLIM baltymo sekų palyginys su aštuoniomis evoliuciškai nutolusiomis rūšimis. Tirtas baltymo pokytis pažymėtas geltonai. Žemiau aminorūgščių sekų specialiais simboliais pažymėtas konservatyvumas: konservatyvi seka (*), konservatyvus variantas (:), pusiau konservatyvus variantas (.) ir nekonservatyvus variantas žymimas tarpu. C – žmogaus E3 ubiquitino ligazės RLIM RING-H2 domeno (purpurinė spalva) komplekso su E2 ubiquitiną prisijungiančiu fermentu UBE2D2 (pilka spalva) eksperimentinė struktūra (kairėje) ir RLIM_I574N-UBE2D2 sąveikos AlphaFold modelis (dešinėje). Paveikslo dalys sukurtos dr. J. Dapkūno.

Atsižvelgiant į RLIM baltymo reikšmę X chromosomos inaktyvacijai, atliktas *AR* geno, esančio X chromosomoje, metilinimo tyrimas, kuriuo siekta įvertinti X inaktyvaciją sergančiojo probando, jo mamos ir dviejų sveikų, moteriškos lyties asmenų DNR mėginiuose. Pagausintų CAG kartotinę seką turinčių fragmentų ilgis probando ir motinos mėginiuose sutapo (24 CAG pasikartojimai, 277 nt ilgio amplikonas). Tačiau panaudojus restrikcijos endonukleazę HpaII, probando mėginyje pagausintas fragmentas nebuvo kirptas, todėl DNR jo mėginyje nemetilita. Motinos mėginyje pagausintų fragmentų sukarpymas ir analizė kapiliarine elektroforeze atskleidė ryškų X chromosomos inaktyvacijos nuokrypį (>95% X chromosomos, kurią probandas paveldėjo iš motinos, inaktyvuota), tuo tarpu kontrolinių asmenų mėginiuose nustatyta atsitiktinė X inaktyvacija (santykis būdingas sveikiems asmenims 53:47) [151]. Šis tyrimas patvirtina prielaidą apie tikėtiną patogeninio varianto buvimą sveikos mamos inaktyvuotoje X chromosomoje, kurią paveldėjo probandas. Atsižvelgiant į tai, kad probando X chromosomoje kitų variantų, sietinų su patologija, nebuvo nustatyta, didėja tikimybė, kad mamos neatsitiktinė X inaktyvacija yra susijusi su neigiamu *RLIM* gene nustatyto varianto c.1721T>A poveikiu koduojamam baltymui.

Siekiant įvertinti galimą RLIM baltyme pakitusios aminorūgšties įtaką šio baltymo erdvinei struktūrai ir sąveikai su domenais/kitais baltymais, atlikta kompiuterinė baltymo struktūrų analizė naudojant AlphaFold ir DISOPRED3 įrankius (žr. 3.3.1.1 paveikslą). Pagrindiniai ir žinomi RLIM baltymo struktūros domenai yra N galo trys α -spiralės ir C galo RING–H2 cinko pirštų domenas, būtinas sąveikai su ubikvitina konjuguojančiais E2 baltymais [152]. Nustatytas pokytis patenka į šių baltymų sąveikos sritį, kuri yra evoliuciškai konservatyvi, todėl tikėtinas neigiamas poveikis baltymo–baltymo sąveikai.

Iki šiol mokslinėje literatūroje Tonne–Kalscheuer sindromu sergantiems tiriamiesiems buvo aprašyti tik *RLIM* geno aminorūgštį keičiantys variantai ir jie yra žinomuose RLIM baltymo N galo arba C galo domenuose [91]. Remiantis šia informacija, tikėtina, kad baltymo funkcijos praradimą lemiantys DNR pokyčiai yra nesuderinami su gyvybe. Pagrindiniai Tonne–Kalscheuer sindromui būdingi klinikiniai požymiai, kurie pasireiškė ir tiriamajam, yra intelektinė negalia (ypač kalbinės raidos atsilikimas), veido anomalijos, augimo sutrikimas, elgesio problemos. Pastebėta, kad RING–H2 domene nustatyti variantai lemia sunkesnę sindromo formą dėl dažnesnio diafragmos išvaržos pasireiškimo [91]. Tai atliepia probando brolio, mirusio kūdikystėje, klinikinius požymius, kurių priežastis galėjo būti RLIM baltymo funkcijos sutrikimas, tačiau jo DNR mėginys nebuvo gautas molekuliniais tyrimams atlikti. RING–H2 domene nustatyti ir žinomi aminorūgščių variantai *AlphaMissense* bioinformatinio įrankio yra prognozuojami kaip patogeniniai. Tai neabejotinai pabrėžia šio domeno, kuriame yra ir mūsų nustatyto varianto pozicija, kritinę reikšmę baltymų

tarpusavio sąveikai. Galima daryti prielaidą, kad dėl pakitusios baltymų tarpusavio sąveikos sutrinka nuo ubikvitino priklausomas baltymų degradavimas, kuris gali būti kaip tikėtinas patogenezės mechanizmas tiriamojo klinikinio fenotipo vystymuisi. Vis dėlto reikalingi eksperimentiniai tyrimai, kuriuose būtų analizuojamas nuo šio baltymo priklausomas baltymų ubikvitilinis, siekiant pilnai atskleisti patogenezės mechanizmą ir galimą unikalų *RLIM* gene nustatyto varianto poveikį ląstelėse vykstantiems procesams.

Atlikti bioinformatiniai ir molekuliniai tyrimai parodo *RLIM* gene nustatyto varianto NM_016120.4:c.1721T>A funkcinę ir klinikinę reikšmę. Bioinformatiniais įrankiais prognozuoti *RLIM* baltymo sąveikos pokyčiai ir nustatyta sutrikusi X chromosomos inaktyvacija leidžia daryti prielaidą, jog sutrikęs specifinių baltymų ubikvitilinis ir suirimas yra kaip galimas patogenezės mechanizmas.

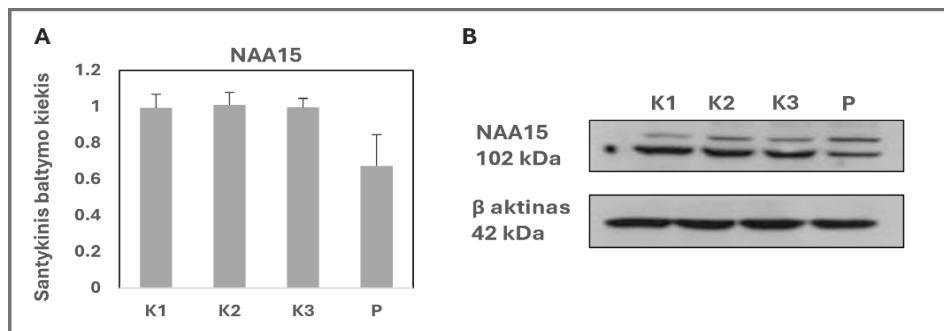
3.3.2 *NAA15* gene nustatyto varianto nulemti koduojamo baltymo raiškos pokyčiai

Moksliniame darbe tirtas *NAA15* gene (MIM#608000) egzomo sekoskaita nustatytas ir literatūroje neaprašytas, heterozigotinis NM_057175.5:c.382C>T variantas tiriamajam (ATGC21–1–P, žr. 3.3 lentelę) su įtariamu intelektinės negalios sindromu. Tiriamajam stebėtas ir psichomotorinės raidos vėlavimas, labilios emocijos, sutrikusi judesių koordinacija. Segregacijos tyrimu patvirtinta, kad tiriamasis variantą paveldėjo iš lengvą intelektualinę negalią turinčio tėvo. Genealogijos duomenimis, šeimoje iš tėvo pusės yra ir daugiau intelektualinę negalią turinčių asmenų. OMIM duomenų bazėje šis genas yra žinomas ligos genas, kurio DNR pokyčiai lemia autosominio dominantinio paveldėjimo raidos vėlavimą su elgesio sutrikimais (MIM#617787). Šiam sutrikimui būdingi fenotipiniai bruožai atitinka mūsų tiriamojo klininius požymius.

Nustatytas *NAA15* geno variantas *gnomAD* duomenų bazėje yra labai retas (dažnis 4×10^{-6}), nerastas 1000G duomenų bazėje. Baltymo lygmeniu variantas lemia priešlaikinės baigmės kodono susidarymą (NP_476516.1:p.(Arg128Ter)). Šio kodono pozicija yra transkripto pradžioje (NM_057175.5 transkriptą sudaro 6174 bazių poros, priešlaikinės baigmės kodonas prasideda 382 pozicijoje), todėl tikėtina, kad inicijuojamas NMD mechanizmas, kuris suardo sutrumpėjusį transkriptą. Tai reikštų, jog yra prarandamas ir pakitęs *NAA15* baltymas.

Kadangi nustatytas *NAA15* geno variantas lemia priešlaikinės baigmės kodono susidarymą (p.(Arg128Ter)) baltymo pradžioje, nuspręsta įvertinti baltymo kiekį tiriamojo ir sveikų kontrolinių asmenų fibroblastuose taikant imunofermentinį metodą Western blot ir naudojant antikūną prieš *NAA15* baltymo C galą (epitopas ties 730–750 a.r.). Nustatytas statistiškai reikšmingas

($p < 0,036$) baltymo kiekio sumažėjimas beveik 2 kartus tiriamojo fibroblastuose, palyginus su kontrolinėmis ląstelėmis (žr. 3.3.2.1 paveikslą). Nors tikėtina, kad sutrumpėjusi baltymo forma ląstelėse netransliuojama dėl NMD, šiame tyrime naudotas antikūnas jungiasi tik prie nepakitusios NAA15 baltymo formos.



3.3.2.1 paveikslas. NAA15 baltymo raiškos pokyčiai tiriamojo, kuriam nustatytas *NAA15* baltymą trumpinantis variantas, fibroblastuose. **A** – baltymo kiekis nustatytas ImageJ programa. **B** – Western blot membranų vaizdas. P – tiriamo asmens baltymų mėginys; K1, K2, K3 – sveikų asmenų baltymų mėginiai.

NAA15 genas koduoja pagalbinį N galo acetiltransferazės A (NatA) subvienetą (angl. *Auxillary subunit of the N-terminal acetyltransferase A (NatA) 15 complex*, NAA15). NatA yra heterodimerinis baltymas, kurio struktūrą sudaro evoliuciškai konservatyvus NAA15 ir acetiltransferazės aktyvumą turintis NAA10 subvienetas [153]. NAA15 prisijungimas prie NAA10 lemia šio katalizinio subvieneto konformacinius pokyčius, kurie apsprendžia NatA baltymo selektyvumą substratams [153]. Nors žmogaus organizme aptinkamos 7-ios N galo acetiltransferazės, NatA yra išskirtinė tuo, kad ji gali acetilinti ~40 proc. visų žmogaus baltymų [154]. Acetilinimas yra svarbi baltymų potransliacinė modifikacija, kuri gali paveikti apie 90 proc. baltymų [155]. Žinoma, kad ši modifikacija svarbi ir gali turėti įtakos baltymų lokalizacijai ląstelėse, erdvinei struktūrai, baltymų kompleksams sudaryti ir jiems suardyti [155]. Ląstelėse baltymų modifikacijos sutrikimas gali pasireikšti įvairių ir svarbių procesų, pavyzdžiui, apoptozės, ląstelės ciklo reguliavimo, genų raiškos reguliavimo, sutrikimais [155]. Remiantis literatūra, galima daryti prielaidą, kad NAA15 subvieneto kiekio sumažėjimas gali turėti neigiamą įtaką NatA funkcijai ir sukelti ryškių pokyčių ląstelės fiziologijai.

Literatūroje šio baltymo haplonepakankamumas dėl *NAA15* gene nustatytų baltymą trumpinančių variantų yra įvardijamas kaip galimas patogenezės mechanizmas, dėl kurio pasireiškia įvairūs nervinės sistemos vystymosi sutrikimai [156]. Vis dėlto neatmetama galimybė, kad kai kurie variantai

(akcentuojami aminorūgšties pokytį sukeliantys DNR variantai) gali lemti baltymo pokyčius, suteikiančius dominuojantį neigiamą pakitusio baltymo (angl. *dominant-negative*) efektą arba papildomų funkcijų įgavimo (angl. *gain-of-function*) efektą, dėl kurių pasireikštų specifiniai ląstelės procesų sutrikimai [156]. Ward ir kolegų (2021) [90] atliktas eksperimentas, kuriame į žmogaus iPSC CRISPR-Cas9 technologija įvedus pokyčius į *NAA15* geną, lemiančius haplonepakankamumą, sutrinka tam tikrų baltymų N galo acetilinimas. Redaguotose ląstelėse taip pat nustatyta, kad pasikeičia NatA sąveika su ribosoma, dėl kurios sutrinka kai kurių baltymų transliacija [90]. Vis dėlto, norint susieti nustatytus NAA15 baltymo pokyčius su specifiniais fenotipiniais bruožais, būtų tikslinga ištirti, kurių NatA substratų (baltymų) acetilinimas sutrinka ir kokį poveikį tai turi jų atliekamai funkcijai.

Atlikti bioinformatiniai ir molekuliniai tyrimai bei literatūros analizė pagrindžia *NAA15* gene nustatyto varianto NM_057175.5:c.382C>T klinikinę ir funkcinę reikšmę. Šiais tyrimais nustatyti NAA15 baltymo kiekio pokyčiai tiriamojo ląstelėse leidžia daryti prielaidą apie šio baltymo haplonepakankamumą ir, tikėtina, sutrikusį specifinių baltymų N galo acetilinimą kaip galimą patogenezės mechanizmą.

Šiame rezultatų ir jų aptarimo poskyryje aprašyti vieno nukleotido DNR sekos pokyčiai remiantis, literatūros, duomenų bazių ir *in silico* analize, įvardinti kaip tiriamųjų paveldimus sutrikimus lemiantys genetiniai pokyčiai. Molekuliniiais eksperimentiniais tyrimais atskleisti RLIM ir NAA15 baltymų funkcijos ir/ar kiekio pokyčiai, pagrindžiantys molekulinį procesų, kuriuose dalyvauja minėti genų produktai, trikdžius, dėl kurių pakinta ląstelės būklė, galiausiai, tikėtina, ir organizmo biologiniai ir fiziologiniai procesai. Tačiau papildomi didesnio masto molekuliniai ir funkciniai tyrimai leistų išsamiau apibūdinti klinikinių požymių vystymosi patogenezėje dalyvaujančių skirtingų molekulinį veiksmų įtaką.

3.4 VNV ir nedidelės iškritos poveikio ląstelės procesams įvertinimas

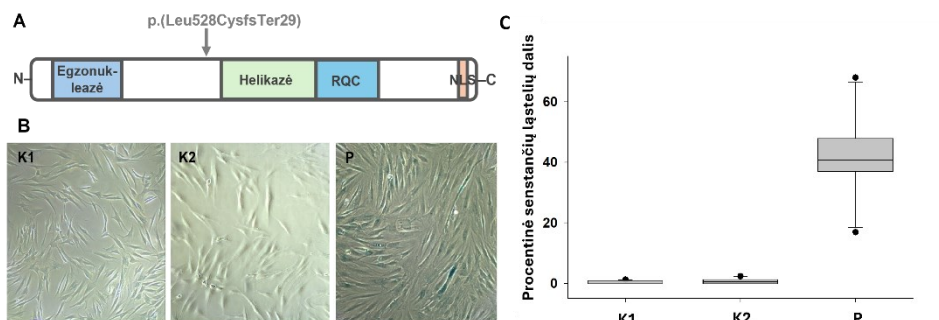
NKS metodais nustatyti vieno ar kelių nukleotidų DNR pokyčiai genų sekose ir specifinėse jų reguliacinėse srityse leidžia numatyti galimus iRNR ir/ar baltymo kokybinius bei kiekybinius pokyčius. Remiantis koduojamo baltymo funkcija, galimos išankstinės įžvalgos ir apie viduląstelių procesų sutrikimus. Siekiant išsiaiškinti šiuos sutrikimus, šiame poskyryje išsamiau aprašyti penki tiriamieji, kuriems klinikinė ir funkcinė nustatytų DNR variantų reikšmė charakterizuota įgyvendinus kiekvienam atvejui pritaikytą molekulinį ir/ar funkcinį tyrimų planą.

3.4.1 *WRN* gene nustatytos vieno nukleotido iškritos poveikio ląstelių senėjimui analizė

Šiame darbe analizuotas egzomo sekoskaita nustatyta homozigotinė vieno nukleotido iškrita NM_000553.6:c.1578del žinomame ligos gene *WRN* (MIM#604611) vyriškos lyties tiriamajam (ATGC21–15–P, žr. 3.3 lentelę) su itin retu ankstyvo senėjimo sutrikimu – progerija. *WRN* genas OMIM duomenų bazėje siejamas su autosominio recesyvaus paveldėjimo ankstyvo senėjimo Werner sindromu (MIM#277700). Tiriamojo genealogija nebuvo informatyvi ir segregacijos tyrimams atlikti nebuvo galimybės ištirti tėvų DNR mėginius.

Nustatyta nedidelė iškrita *in silico* lemia rėmelio poslinkį ir priešlaikinės baigmės kodono susidarymą (NP_000544.2:p.(Leu528CysfsTer29)), taip pat *MutationTaster* įrankiu vertinama kaip patogeninis DNR pokytis. Baltymo pokytis susidaro kelių aminorūgščių atstumu iki helikazės domeno, todėl, tikėtina, sutrumpėjęs baltymas praranda savo funkciją ir/ar yra degraduojamas. Remiantis literatūra, baltymą trumpinantys ir baltymo funkcijos praradimo variantai, nustatyti įvairiuose Vernerio helikazės regionuose, yra žinoma sindromo vystymosi priežastis [157, 158]. Šiame tyrime nustatytas p.(Leu528CysfsTer29) pokytis (žr. 3.4.1.1 A paveikslą) aprašytas Werner sindromu sergančiam asmeniui sudėtinėje heterozigotinėje būklėje su kitu patogeniniu, baltymą trumpinančiu *WRN* variantu [157]. Tai leidžia pagrįsti mūsų tiriamojo genetinę suaugusiųjų progerijos sindromo priežastį.

Remiantis duomenimis, kad *WRN* baltymą trumpinantys ir jo praradimą lemiantys variantai skatina ląstelių senėjimą [159, 160], atliktas β -galaktozidazės aktyvumo tyrimas mūsų tiriamojo ir dviejų sveikų kontrolinių asmenų fibroblastų kultūrose. Po inkubacijos su β -galaktozidazės substratu X-gal didelė dalis tiriamojo ląstelių nusidažė žalsvai mėlyna spalva, kuri parodo, kad tiriamojo fibroblastuose yra prasidėjęs ankstyvas senėjimas, kai to paties pasažo (ketvirtojo pasažo) kontrolinių asmenų fibroblastuose žalsvai mėlyna spalva nusidažė tik pavienės ląstelės (žr. 3.4.1.1 B paveikslą). Išreiškus senstančių ląstelių skaičių procentais nuo bendro ląstelių skaičiaus, kontrolinių, sveikų asmenų ląstelių kultūroje nusidažė vidutiniškai 0,59 proc. ląstelių, o p.(Leu528CysfsTer29) pokytį turinčio tiriamojo ląstelių kultūroje senstančių ląstelių nustatyta 42 proc. – keliasdešimt kartų daugiau ($p < 0,001$; žr. 3.4.1.1 C paveikslą). Tai pagrindžia *WRN* helikazės funkcijos praradimo nulemtą ankstyvą ląstelių replikacinį senėjimą, kuris yra vienas iš šiame gene nustatomų DNR pokyčių nulemtų patogenezės mechanizmų.



3.4.1.1 paveikslas. Ląstelių senėjimo tyrimas *WRN* gene vieno nukleotido iškritą turinčiuose fibroblastuose. A – WRN baltymo domenų schema su nustatytu DNR pokyčiu baltymo lygmeniu. B – fibroblastų kultūrų nuotraukos po inkubacijos su X-gal substratu. K1, K2 yra kontrolinių asmenų, o P yra tiriamo asmens ląstelės. C – procentinė ląstelių dalis, pasižyminti senėjančioms ląstelėms būdingu β -galaktozidazės aktyvumu.

Remiantis *GTEX* duomenų baze, Vernerio helikazę koduojančio geno raiška vyksta daugelyje žmogaus audinių, o tai pabrėžia šio baltymo reikšmę organizmui. *WRN* genas koduoja RecQ pošeimio DNR helikazę (dar vadinama Vernerio helikaze; angl. *Werner Syndrome RecQ Like Helicase*). Tai branduolyje veikiantis baltymas, kuris išvynioja įvairias DNR struktūras, todėl yra labai svarbus DNR replikacijos, transkripcijos, DNR pažeidų taisymo procesuose bei telomerų palaikyme [158]. WRN baltymas turi kelis pagrindinius domenus: N galo 3'→5' egzonukleazinį domeną, centriniame baltymo regione nuo ATF (adenozino trifosfato) priklausomą 3'→5' helikazinį domeną ir RecQ konservatyvų regioną bei į branduolį nukreipiantį signalą C gale [161].

Werner sindromą turintiems asmenims pasireiškia ankstyvo senėjimo požymiai (sulėtėjęs augimas, žemas ūgis, anksti pražilę plaukai ar jų slinkimas, paodžio atrofija, katarakta, vėžiniai susirgimai, pasireiškia kitos su senėjimu siejamos ligos). Ankstyvas senėjimas pastebimas ir WRN baltymo disfunkciją turinčiųjų asmenų ląstelėse. Nustatyta, kad molekuliniais metodais slopinama *WRN* geno raiška lemia greitą ląstelių senėjimą [159]. Vernerio helikazės disfunkciją turinčiose ląstelėse atliktais tyrimais parodyta sulėtėjusi ląstelių proliferacija, taip pat jose sumažėja su ląstelės ciklu siejamų genų ir su DNR pažeidų taisymu susijusių genų raiška [162]. Tai vieni iš požymių, nurodančių ankstyvą ląstelių senėjimą, šiuo atveju susijusį su WRN baltymo funkcijos praradimo lemiamu genomo nestabilumu ir telomerų trumpėjimu [159, 162]. Senstančios ląstelės nesidalija, tačiau yra metaboliškai aktyvios ir gyvybingos [163]. Nuo ramybės būsenos arba galutinai diferencijuavusių ir nebesidalijančių ląstelių senstančios ląstelės skiriasi morfologiniais pokyčiais, nes tampa

didesnės, suplokštėja, o citoplazmoje padaugėja membraninių pūslelių ir atsiranda su senėjimu susijusios β -galaktozidazės (SA- β -gal) aktyvumas [164], kurio padidėjimas nustatytas ir tiriamojo fibroblastuose. Tai lizosominė hidrolazė, kurios substratai yra β -d-galaktozidai (kaip laktozė, keratino sulfatai, sfingolipidai), o jos aktyvumo ląstelėse didėjimas koreliuoja su ląstelių senėjimu [164]. Vis dėlto, atsižvelgiant į nustatytą genetinį variantą ir homozigotinį genotipą, būtų naudinga atlikti *WRN* geno ir jo koduojamo baltymo raiškos nustatymą, taip pat iRNR sekoskaitą tiriamojo RNR, siekiant detaliau atskleisti nustatytos nedidelės iškritos lemiamą unikalų patogenezės mechanizmą. Šiam tiriamajam pastarųjų tyrimų atlikti nebuvo galimybės dėl riboto gautų fibroblastų skaičiaus bei dėl ypač lėtos šių ląstelių proliferacijos.

Atlikti bioinformatiniai, molekuliniai ir funkciniai tyrimai bei literatūros analizė pagrindžia *WRN* gene nustatytos vieno nukleotido iškritos NM_000553.6:c.1578del klinikinę ir funkcinę reikšmę. Nustatytas pagreitėjęs tiriamojo fibroblastų senėjimas leidžia daryti prielaidą, kad ankstyvas replikacinis ląstelių senėjimas dėl sutrikusios Vernerio helikazės veiklos yra pagrindinis patogenezės mechanizmas.

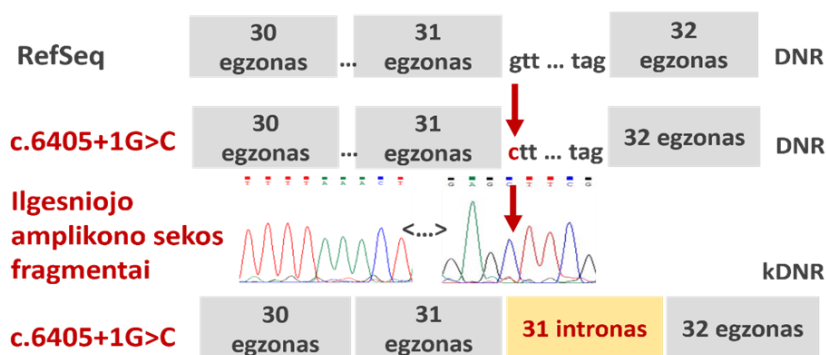
3.4.2 *DYNCH1* gene nustatyto sukirpimo vietos varianto įtaka KIFC1 ir KIFC3 baltymų raiškai bei Goldžio komplekso lokalizacijai

Moksliniame darbe analizuotas *DYNCH1* gene (MIM#600112) nustatytas ir literatūroje neaprašytas heterozigotinis variantas NC_000014.9(NM_001376.5):c.6405+1G>C iRNR sukirpimui svarbioje pozicijoje. DNR pokytis nustatytas tiriamajam (ATGC21–3–P, žr. 3.3 lentelę) su psichomotorinės raidos, ypač kalbinės, atsilikimu. *De novo* varianto kilmė patvirtinta ištyrus sergančiojo probando sveikų tėvų mėginius.

OMIM duomenų bazėje *DYNCH1* genas yra siejamas su keliais neuroraumeniniais (MIM#614228, MIM#158600) ir vienu neurovystymosi (MIM#614563) sutrikimu. Pastarajam priskiriami fenotipiniai bruožai (kaip psichomotorinės raidos atsilikimas) būdingi ir tiriamajam. Taip pat pabrėžiama, kad šį sutrikimą lemiantys variantai dažniausiai yra *de novo* kilmės. Įdomu tai, kad disertacijos rašymo metu (2024 m. gruodis) literatūroje aprašyti tik trys baltymų trumpinantys (p.Arg728Ter; c.11941+2T>A; c.7473+105C>T) ir, tikėtina, tiriamųjų klinikinio fenotipo vystymąsi lemiantys *DYNCH1* geno variantai [165-167]. Visi kiti, tiriamųjų neurovystymosi arba neuroraumeninius sutrikimus lemiantys DNR variantai, *DYNCH1* baltymo lygmeniu sukelia vienos aminorūgšties pokytį.

Nustatyto sukirpimo vietos varianto nėra *gnomAD* duomenų bazėje, o *MutationTaster*, *HumanSplicingFinder* ir *SpliceAI* bioinformatiniai įrankiai šį variantą vertina kaip galimai sutrikdantį kanoninį iRNR sukirpimo procesą. Šis

poveikis įvertintas probando kDNR mėginyje gausinant specifinius fragmentus PGR metodu ir toliau juos sekvenuojant Sangerio metodu. Naudojant tiesioginį pradmenį, kuris hibridizuojasi ant 29–30 egzono jungties, ir atvirkštinį, kuris jungiasi prie 33-io egzono sekos, gautas fragmentas, kurio sekoskaitos chromatogramoje išryškėjo papildomos smailės, žyminčios 31-o introno dalį. Siekiant patvirtinti papildomo fragmento seką, PGR produktai agarozės gelyje frakcionuoti ilgesnį laiką. Gelyje atsiskyręs papildomas fragmentas iškirptas iš agarozės gelio, o naudojant tą patį tiesioginį pradmenį atlikus Sangerio sekoskaitą nustatytas iRNR fragmentas su 31-o introno sekos dalimi. Sukurti papildomi pradmenys, siekiant specifiskai nusekvenuoti tik pakitusį fragmentą – patvirtinta, kad *DYNCH1* geno sukirpimo vietos variantas lemia 31-o introno intarpą (žr. 3.4.2.1 paveikslą). Kanoninėse sukirpimo vietose nustatomi DNR variantai dažniau lemia egzono iškritą ar alternatyvios sukirpimo vietos susidarymą, rečiau yra nustatomos kelių egzono iškritos ar viso introno intarpai [168, 169].

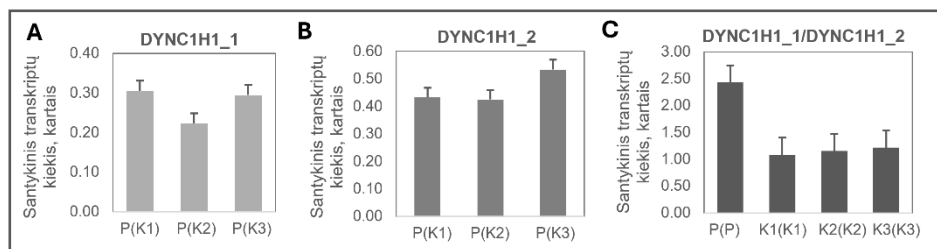


3.4.2.1 paveikslas. *DYNCH1* geno sukirpimo vietos varianto poveikio iRNR sukirpimui schema. Nustatytas variantas lemia papildomo fragmento, kuriame yra 31-o introno intarpas, susidarymą.

DYNCH1 genas koduoja didelio (~1,5 MDa) baltymų komplekso, vadinamo citoplazminiu dineinu 1 (arba dineinu), subvienetą – sunkiąją grandinę 1 (angl. *cytoplasmic dynein 1 heavy chain 1*, DYNCH1). Šis baltymų kompleksas yra variklinis baltymas, priklausantis dineinų superšeimai. *In silico*, 31-o introno intarpas lemia skaitymo rėmelio poslinkį ir priešlaikinės baigmės kodono susidarymą (NP_001367.2:p.(Ile2136LeufsTer20)). Kadangi naujo baigmės kodono pozicija iRNR sekoje yra arčiau transkripto vidurio, tikėtina, kad yra aktyvinamas priešlaikinės baigmės kodono inicijuotas iRNR suardymas (NMD). Svarbu tai, kad šiame procese suardomi pakitę transkriptai, kurių translacijos produktai gali būti sutrikusios funkcijos baltymai, galimai turintys neigiamą poveikį (angl. *dominant-negative effect*) organizmo fenotipui [65]. Taigi, tai natūralus, ląstelėse veikiantis mechanizmas, kurio iniciaciją lėmę DNR variantai

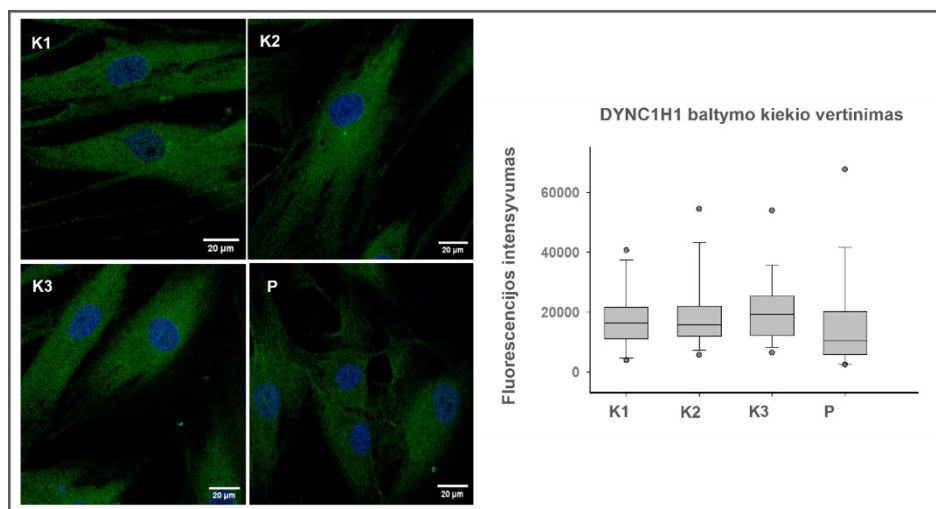
gali koreguoti paveldimam sutrikimui būdingą fenotipinių bruožų sunkumo ir įvairovės pasireišimą.

Siekiant įvertinti galimą pakitusio iRNR transkripto suskaldymą, atliktas *DYNC1H1* transkriptų kiekio vertinimas kPGR metodu, lyginant probando (P) fibroblastuose nustatytą genų raišką su trijų vyriškos lyties sveikų asmenų kontrolinėmis ląstelėmis (K1, K2, K3). Kadangi *DYNC1H1* geną sudaro 78-i egzonai (transkriptas NM_001376.5), o nustatytas pokytis yra 31-o egzono gale, analizei pasirinktos dvi sritys – pirmasis *TaqMan*TM zondas (*DYNC1H1_1*), kurio taikinytis yra transkripto pradžioje (15–16 egzonų jungtis), ir antrasis zondas (*DYNC1H1_2*), kurio taikinytis yra transkripto gale (51–52 egzonų jungtis). Tokiu būdu, pirmuoju zonu galima identifikuojami tiek pakitę, tiek nepakitę *DYNC1H1* geno transkriptai, tuo tarpu antruoju – tik nepakitę. Nustatyta, jog pirmąjį taikinį turinčių *DYNC1H1* transkriptų kiekis (suminį *DYNC1H1* transkriptų kiekį), probando ląstelėse sumažėjo apie 3 kartus, lyginant su kontrolinėmis ląstelėmis (žr. 3.4.2.2 A paveikslą). Antrąjį taikinį turinčių *DYNC1H1* transkriptų (nepakitę *DYNC1H1* transkriptai) kiekis probando fibroblastuose, lyginant su kontrolinių asmenų ląstelėmis, sumažėjo nežymiai – 1,3 karto (žr. 3.4.2.2 B paveikslą). Šie geno raiškos rezultatai leidžia daryti prielaidą, jog nustatyto sukirpimo vietos varianto lemiamas priešlaikinės baigmės kodono susidarymas inicijavo NMD mechanizmą, dėl kurio ląstelėse suskaldytas pakitęs *DYNC1H1* transkriptas. Verta atkreipti dėmesį, kad probando ląstelėse bendras transkriptų kiekis buvo 2,4 karto didesnis nei nepakitusių transkriptų kiekis (žr. 3.4.2.2 C paveikslą). Taigi, probando fibroblastuose yra daugiau *DYNC1H1* transkriptų, turinčių pirmųjų egzonų sekas nei galinių egzonų sekas. Kontroliniuose mėginiuose, pirmąjį ir antrąjį taikinį turinčių transkriptų kiekis nesiskyrė.



3.4.2.2 *DYNC1H1* geno transkriptų raiškos pokyčiai. A – santykinis nepakitusių ir pakitusių *DYNC1H1* transkriptų kiekis. B – santykinis nepakitusių *DYNC1H1* transkriptų kiekis. C – *DYNC1H1* geno transkriptų pradžios ir pabaigos santykinis kiekis. P(K1) reiškia probando transkriptų kiekio palyginimą su kontroliniu asmeniu K1, P(K2) ir P(K3) atitinkamai taip pat. *DYNC1H1_1* – pirmasis *TaqMan*TM zondas, kurio taikinytis yra 15–16 egzonų jungtis. *DYNC1H1_2* – antrasis *TaqMan*TM zondas, kurio taikinytis yra 51–52 egzonų jungtis.

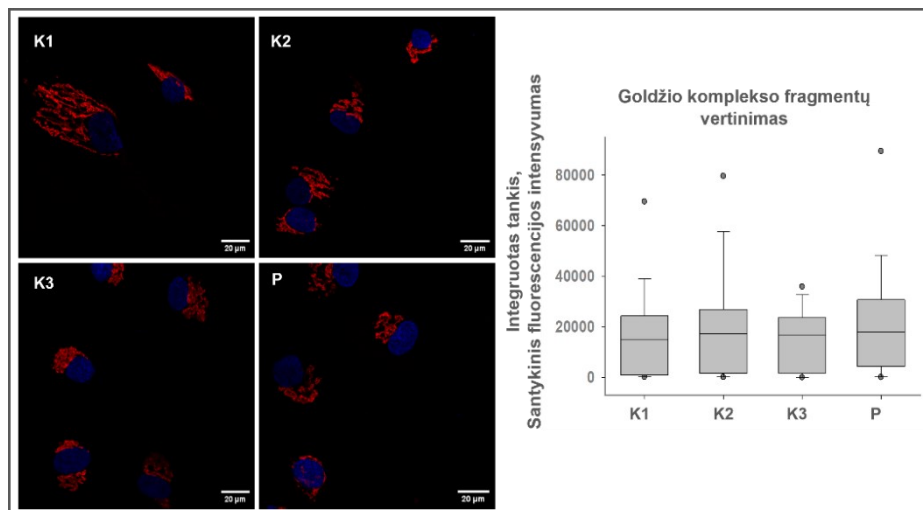
Bioinformatiniais įrankiais prognozuojama priešlaikinio baigmės kodono pozicija (p.(Ile2136LeufsTer20)) yra motoriniame variklinio baltymo domene, kuris yra pirmoje baltymo dalyje (baltymą sudaro 4646 aminorūgštys), todėl, tikėtina, baltymas praranda savo funkciją. Be to, nustatytas *DYNC1H1* geno raiškos pokytis probando fibroblastuose gali turėti įtakos baltymo raiškai. Siekiant įvertinti *DYNC1H1* baltymo kiekį, taikyta imunofluorescencinė mikroskopija. Nustatyta, kad probando fibroblastuose *DYNC1H1* kiekis sumažėja 1,7 karto ($p=0,014$), lyginant su K1 ląstelėmis, 1,4 karto ($p=0,001$), lyginant su K2 ląstelėmis ir 2 kartus ($p<0,001$), lyginant su K3 ląstelėmis (žr. 3.4.2.3 paveikslą). Šis tyrimas parodo, kad translacija nuo *DYNC1H1* pakitusio transkripto tikėtina nevyksta dėl aktyvuoto NMD mechanizmo, todėl sutrumpėjęs *DYNC1H1* baltymas ląstelėse neturėtų funkcionuoti.



3.4.2.3 paveikslas. *DYNC1H1* baltymo kiekio pokyčiai fibroblastuose. Mikroskopijos vaizduose žaliai žymėtas *DYNC1H1*, mėlynai – ląstelės branduoliai. P – tiriamojo asmens fibroblastai; K1, K2, K3 – kontrolinių asmenų fibroblastai. Grafike pavaizduotas santykinis baltymo kiekis.

Ląstelės gyvybiniais procesams yra svarbi baltymų ir lipidų modifikacija bei rūšiavimas, kurį atlieka Goldžio kompleksas. Tai iš cisternų ir pūslelių sudaryta dinamiška organelė, kurios judėjimą, kaitą ir lokalizaciją ląstelėje palaiko citoplazminis dineinas 1 [170, 171]. Atsižvelgiant į šią dineino funkciją, vertinta Goldžio komplekso lokalizacija ir membranų kiekis fibroblastuose taikant imunofluorescencinę mikroskopiją ir naudojant Goldžio komplekso žymenį GM130. GM130 žymens fluorescencijos intensyvumas, lyginant probando su kontrolinių asmenų ląstelėmis, reikšmingai nesiskyrė (žr. 3.4.2.4 paveikslas).

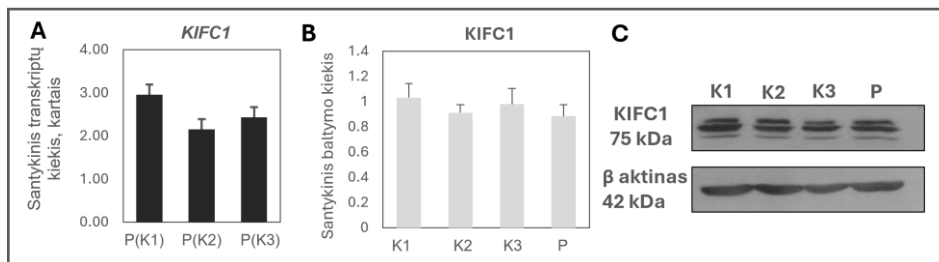
Skirtumai nenustatyti ir vertinant Goldžio komplekso membranų pasiskirstymą fibroblastuose.



3.4.2.4 paveikslas. Goldžio komplekso fragmentų vertinimas fibroblastuose.

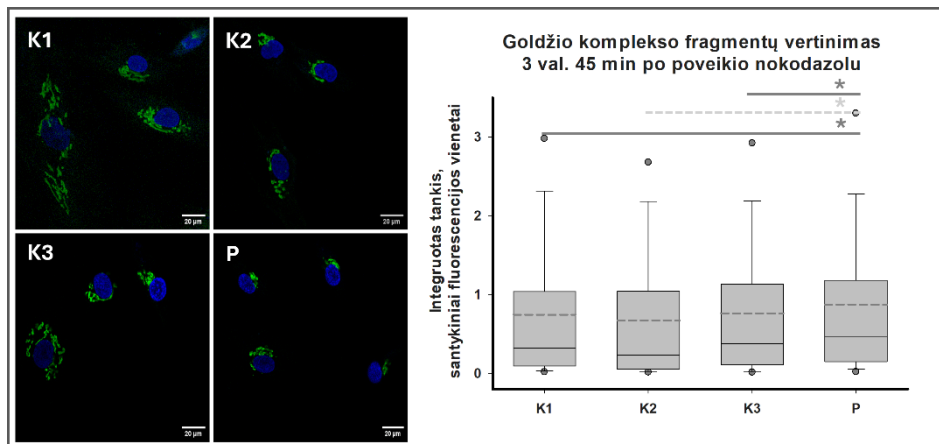
Mikroskopijos vaizduose raudonai žymėtas Goldžio komplekso baltymas GM130, mėlynai – ląstelės branduoliai. P – tiriamojo asmens fibroblastų kultūros; K1, K2, K3 – kontrolinių asmenų fibroblastai. Grafike pavaizduotas integruotas tankis apskaičiuojamas objekto plotą padauginus iš vidutinės pikselio fluorescencijos.

Kadangi tiriamojo fibroblastuose nustatytas mažesnis DYNC1H1 baltymo kiekis ir žinoma, kad dineinas yra svarbus Goldžio komplekso dinamikai palaikyti, tikėtasi užfiksuoti pakitusią Goldžio komplekso struktūrą, vietą ir/ar tūrį. Iškelta hipotezė, kad ląstelėse gali egzistuoti kompensacinis mechanizmas, inicijuotas kitų variklinių baltymų, dėl kurių suaktyvėjusios veiklos nėra matomi ryškūs organelės pokyčiai. Literatūroje pažymima, kad varikliniai baltymai, priklausantys kinezinų superšeimai (KIF), KIFC1 ir KIFC3 ląstelės krovinius taip pat transportuoja link mikrovamzdelių minus galų [172]. Įvertinti KIFC1 ir KIFC3 baltymus koduojančių genų raiškai taikyta kPGR. Nustatyta, kad *KIFC1* transkriptų kiekis probando ląstelėse padidėja 3 kartus, palyginus su K1 kontrole, 2 kartus palyginus su K2 kontrole ir 2,5 karto palyginus su K3 kontrole (žr. 3.4.2.5 A paveikslą). Lyginant kontrolinių asmenų mėginius tarpusavyje, KIFC1 transkriptų kiekis nesiskyrė. Nuspręsta įvertinti *KIFC1* raišką ir baltymo lygmeniu, tačiau atlikus imunofermentinę analizę Western blot metodu, nustatyta, kad KIFC1 baltymo kiekis probando ląstelėse nesiskiria nuo kontrolinių asmenų (žr. 3.4.2.5 B ir C paveikslą).



3.4.2.5 *KIFC1* geno ir jo koduojamo variklinio baltymo raiškos pokyčiai tiriamojo fibroblastuose. A – santykinis *KIFC1* transkriptų kiekis. B – santykinis KIFC1H1 baltymo kiekis. P(K1) reiškia probando transkriptų kiekio palyginimą su kontroliniu asmeniu K1, P(K2) ir P(K3) atitinkamai taip pat. C – KIFC1 baltymo imunofermentinė analizė Western blot metodu.

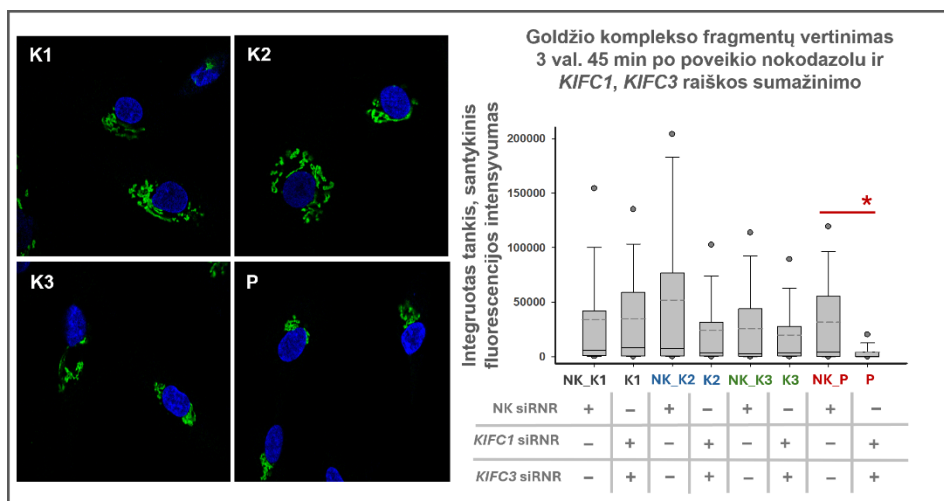
Remiantis analizuota literatūra, įvertinti dineino reikšmę Goldžio komplekso lokalizacijai palaikyti gali būti naudojamas nokodazolas [81, 173]. Tai junginys, kuris depolimerizuoja mikrovamzdelius jungdamasis prie mikrovamzdelių struktūrinio subvieneto β -tubulino, todėl Goldžio kompleksas suyra. Tad po poveikio Goldžio komplekso atsistatymas priklauso nuo dineino funkinių savybių – gebėjimo atstatyti Goldžio kompleksui būdingą struktūrą ir lokalizaciją prie ląstelės branduolio. [81]. Tiriamojo atveju, probando fibroblastus lyginant su kontrolinių asmenų fibroblastais, pastebėti Goldžio komplekso atsikūrimo pokyčiai (vertinta pagal Goldžio komplekso žymens GM130 integruotą tankį, kuris apskaičiuojamas objekto plotą padauginant iš vidutinio pikselio fluorescencijos intensyvumo) praėjus 3 val. ir 45 min po nokadozolo nuplovimo. Palyginus probando ir kontrolinių asmenų ląsteles, probando ląstelėse nustatytas didesnis GM130 fluorescencijos integruotas intensyvumas – siekė 1,4 karto, palyginus su K1 ($p=0,03$), 2 kartus, palyginus su K2 ($p<0,001$) bei 1,3 karto, palyginus su K3 ($p=0,047$) kontrolinių asmenų ląstelėmis (žr. 3.4.2.6 paveikslą). Tačiau lyginant K2 su kitais kontroliniais mėginiais, nustatytas statistiškai reikšmingas skirtumas, todėl išvados daryti naudojami K1 ir K3 mėginiai. Gali būti, kad tiriamojo fibroblastuose Goldžio komplekso struktūra minėtame laiko taške mažiau kompaktiška, plačiau išsidėsčiusi prie ląstelės branduolio.



3.4.2.6 paveikslas. Goldžio komplekso atsikūrimas po poveikio nokodazolu.

Grafike pavaizduotas GM130 fluorescencijos intensyvumas tirtuose mėginiuose 3 val. 45 min po poveikio nokodazolu. Vertintas integruotas tankis, apskaičiuojamas objekto plotą padauginant iš pikselio vidutinio fluorescencijos intensyvumo. Punktyrinė pilka linija dėžutėse žymi duomenų vidurkį, vientisa – medianą. Punktyrinė linija virš duomenų rodo, kad K2 kontrolinio asmens duomenys išvados nenaudoti, nes lyginant K2 duomenis su kitais kontroliniais mėginiais nustatytas statistiškai reikšmingas skirtumas. Mikroskopijos vaizduose žaliai žymėtas Goldžio komplekso baltymas GM130, mėlynai – ląstelės branduoliai. P – tiriamojo asmens fibroblastų kultūros; K1, K2, K3 – sveikų kontrolinių asmenų fibroblastai. Statistinis reikšmingumas $*p \leq 0,05$.

Vis dėlto KIFC1 ir KIFC3 variklinių baltymų veikla tiriamojo ląstelėse gali iš dalies maskuoti DYNC1H1 sutrikusios funkcijos (ar sumažėjusios raiškos) padarinius Goldžio komplekso atsikūrimo procesui. Todėl nuspręsta įvertinti DYNC1H1 gebėjimą atstatyti Goldžio kompleksą po nokodazolo poveikio ir nutildžius neigiamo mikrovamzdelių galo kryptimi krovinius nešančių KIFC1 ir KIFC3 baltymus koduojančių genų raišką naudojant siRNR. Tildymo efektyvumas tikrintas taikant kPGR metodą – 24 val. po ląstelių transfekcijos siRNR specifinių genų raiška sumažėja 3 kartus. Nutildžius *KIFC1* ir *KIFC3* kontrolinių asmenų ląstelėse, Goldžio fluorescencija 3 val. 45 min po nokodazolo nuplovimo nesiskyrė nuo neigiamos kontrolės (nespecifinės siRNR). Tačiau probando ląstelėse šių genų raiškos nutildymas 13 kartų mažino GM130 žymens fluorescenciją, palyginus su neigiamos kontrolės poveikiu ($p < 0,001$) (žr. 3.4.2.7 paveikslą). Šis tyrimas parodo, kad tiriamojo fibroblastuose citoplazminio dineinio 1, kurio pagrindinę ir motorinę dalį sudaro DYNC1H1, funkcija yra sutrikusi, todėl Goldžio kompleksas, neveikiant KIFC1 ir KIFC3 baltymams, įgauna kompaktiškesnę struktūrą ir lokalizuojasi arčiau ląstelės branduolio.



3.4.2.7 paveikslas. Goldžio komplekso atsikūrimo vertinimas po genų raiškos tildymo ir poveikio nokodazolu tiriamojo fibroblastuose. Mikroskopijos vaizduose žaliai žymėtas Goldžio komplekso baltymas GM130, mėlynai – ląstelės branduoliai. P – tiriamojo asmens fibroblastai (grafike raudonu šriftu); K1 (grafike tamsiai pilku šriftu), K2 (grafike mėlynu šriftu), K3 (grafike žaliu šriftu) – kontrolinių asmenų fibroblastai. Vertintas integruotas tankis, apskaičiuojamas objekto plotą padauginant iš pikselio vidutinio fluorescencijos intensyvumo. Punktyrinė pilka linija dėžutėse žymi duomenų vidurkį, vientisa – medianą. Grafike pavaizduotas integruotas tankis tirtuose mėginiuose po *KIFC1* ir *KIFC3* genų raiškos tildymo ir 3 val. 45 min po nokodazolo nuplovimo. Statistinis reikšmingumas $*p \leq 0,05$.

Tiriamojo DNR sukirpimo vietos variantas nustatytas DYNCH1 baltyme, kuris ląstelių viduje atlieka įvairių krovinių, tokių kaip ląstelės organelių, dalijimosi verpstės komponentų, virusų, nukleorūgščių ar kitų baltymų judėjimą citoskeletą sudarančiais mikrovamzdeliais link jų minus galo [170]. Judesiui atlikti naudojamas adenozinotrifosfatas, kadangi šis molekulinis variklis turi ATPazės aktyvumą [170]. ATPazinį aktyvumą turintį domeną (motorinis domenas), esantį baltymo C galo globulinėje galvutėje, suformuoja dvi vienodos sunkiosios grandinės, kurios yra įprasta dineinų struktūrinė dalis [174]. Sunkiosios grandinės suformuoja dar du funkciškai skirtingus domenų: krovinį surišančią uodegą (baltymo N gale) ir mikrovamzdelius surišantį kotelį (baltymo C gale) [174]. Varijuojanti dineinų dalis gali būti sudaryta iš kelių lengvųjų ir kelių tarpinių grandinių [170, 174]. Nustatytas sukirpimo vietos variantas, lemiantis *DYNCH1* geno raiškos sumažėjimą, baltymo variklinio domeno dalies praradimą, mūsų tiriamojo ląstelėje lėmė ir paties DYNCH1

baltymo raiškos efektyvumo pokyčius, atsispindėjusius konfokalinės mikroskopijos vaizduose.

Dineinas yra pagrindinis variklinis baltymas palaikant Goldžio komplekso dinamiką ląstelėje, todėl mikroskopuojant Goldžio kompleksą tikėtasi nustatyti vizualius šios organelės pokyčius, nulemtus pakitusios DYNC1H1 raiškos. Tačiau literatūroje aprašytais eksperimentiniais tyrimais nustatyta, kad kinezinų superšeimai priklausantys KIFC1 ir KIFC3, taip pat judantys į mikrovamzdelių minus galus, dalyvauja Goldžio komplekso padėties palaikyme [175, 176]. Vis dėlto minėtų baltymų RNR ir baltymo raiškos tyrimai mūsų tiriamojo fibroblastuose atskleidė prieštaringsus rezultatus – *KIFC1* geno transkriptų kiekis padidėjęs, o KIFC1 baltymo kiekis ląstelėse nepasikeitęs. Įdomu tai, kad transkriptų kiekio padidėjimas ląstelėse ne visada koreliuoja su baltymo raiškos pokyčiais, kaip ir šiuo atveju *KIFC1* raiška, o tam įtakos gali turėti įvairūs ląstelėse vykstantys procesai: transliacijos efektyvumas, kai transkriptai tarpusavyje konkuruoja dėl vietos transliacijos mechanizme; baltymo stabilumas, dėl kurio kiekio skirtumai nenustatomi, nes baltymo gyvavimo ląstelėje trukmė yra trumpesnė; sutrikusi baltymo lokalizacija ir pernaša dėl padidėjusio specifinio baltymo kiekio, kuris ląstelėse gali kauptis ne jam būdingose vietose, todėl gali būti nukreiptas suardymui [177-179]. Taigi, galima daryti prielaidą, kad DYNC1H baltymo sutrikusi veikla ląstelėje iš dalies kompensuojama KIFC1 variklinio baltymo. Tuo tarpu *KIFC3* transkriptų kiekis probando RNR mėginyje nesiskyrė nuo kontrolinių asmenų. Be to, siekiant įvertinti Goldžio komplekso tūrio pokyčius, pasirinkta transkriptų lygiu nustatyti Goldžio komplekso baltymo GOLGB1 raišką – raiška nekito, palyginus probando mėginį su kontrolinių asmenų mėginiais.

Kadangi Goldžio kompleksas yra nepastovi struktūra, neretai atliekami dinamikos ir padėties stebėjimo tyrimai. Tokie eksperimentai su nokodazolu, depolimerizuojančiu ląstelių citoskeletą sudarančius mikrovamzdelius, taikyti ir siekiant įvertinti *DYNC1H1* gene identifikuotų aminorūgšties pokytį lemiančių variantų (p.Gln1194Arg ir p.Glu3048Lys) poveikiui Goldžio komplekso dinamikai. Šiuo atveju nustatyta, kad tiriamojo ląstelėse Goldžio struktūra atsikuria lėčiau [81]. Tai gali lemti sutrikusi DYNC1H1 baltymo funkcija, dėl kurios suirusio Goldžio komplekso pūslelės link branduolio (mikrovamzdelių minus galų) juda lėčiau. Goldžio komplekso pūslelių judėjimo ir padėties prie branduolio pasikeitimai stebėti ir mūsų tiriamojo fibroblastuose. Tai lėmė tiek baltymo DYNC1H1 raiškos pokyčiai, tiek nustatyta pakitusi *KIFC1* raiška, kadangi šiame darbe tirtose ląstelėse Goldžio komplekso pūslelės prie branduolio išsidėstė kompaktiškiau, o judėjimo greičio skirtumų nepastebėta.

Taigi, *DYNCH1* sukirpimo vietos variantą turinčio probando fibroblastuose nustatyti šio geno raiškos pokyčiai RNR ir baltymo lygiu, aptiktas išaugęs *KIFC1* transkriptų kiekis bei nustatyti Goldžio komplekso atsikūrimo pokyčiai patvirtina *DYNCH1* gene nustatyto varianto klinikinę ir funkcinę reikšmę. Tačiau būtų naudinga atlikti papildomus tyrimus siekiant ištirti daugiau kontrolinių asmenų *KIFC1* geno raišką, pratęsti imunofluorescencijos tyrimus, siekiant nustatyti *KIFC1* baltymo kiekį ir stebėti citoplazminio dineino bei mikrovamzdelių dinamiką po poveikio nokodazolu dažnesniais laiko taškais, taip nustatyti tikslų Goldžio komplekso atsikūrimo pradžios tašką, kuriame labiausiai išsiskirtų tiriamojo ir sveikų asmenų fibroblastai. Atsižvelgiant į tiriamojo neurovystymosi fenotipą, tikslinga šiuos eksperimentus atlikti nervinėse ląstelėse, siekiant pagrįsti šio baltymo funkcijos sutrikimą ir poveikį nervinių ląstelių veiklai.

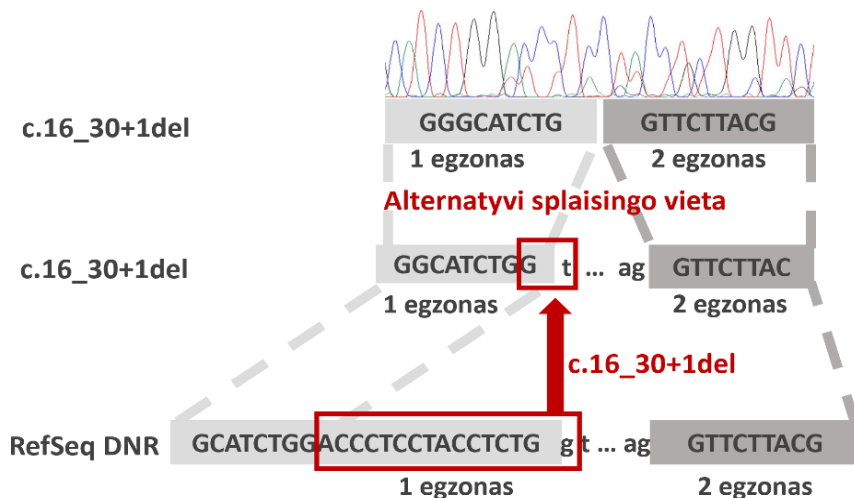
Atlikti bioinformatiniai, molekuliniai ir funkciniai tyrimai bei literatūros analizė pagrindžia *DYNCH1* gene nustatyto sukirpimo vietos varianto NC_000014.9(NM_001376.5):c.6405+1G>C patogeniškumą. Šiais tyrimais nustatyti *DYNCH1*, *KIFC1* kiekio pokyčiai ir sutrikusi Goldžio komplekso dinamika ląstelėje leidžia daryti prielaidą apie kitų ląstelės krovinių judėjimo sutrikimą kaip galimą patogeneizės mechanizmą.

3.4.3 *FAS* gene nustatytos nedidelės iškritos įtaka apoptozės indukcijai

Analizuotas egzomo sekoskaita nustatytas ir literatūroje neaprašytas, heterozigotinis pokytis NC_000010.11(NM_000043.6):c.16_30+1del *FAS* gene (MIM#134637) tiriamajai (ATGC21–18–P, žr. 3.3 lentelę), kuriai įtartas autoimuninis limfoproliferacinis sindromas (ALPS; MIM#601859). OMIM duomenų bazėje šis genas yra žinomas ligos genas, kuriame nustatyti DNR pokyčiai sukelia minėtą imuninės sistemos sutrikimą. Įdomu tai, kad ta pati nedidelė iškrita nustatyta ir tiriamosios mamai, kuri neturi ligos požymių. Remiantis prieinama literatūra, skaičiuojama, kad ALPS klinikinis ligos penetrantiškumas siekia ~ 60 proc., todėl dalis *FAS* geno sekos pokyčius turinčių asmenų yra besimptomiai ligos nešiotojai [180]. Vis dėlto kai kuriais atvejais neužtenka *FAS* geno pokyčio pasireikšti ALPS – šis fenomenas reikalauja išsamesnių eksperimentinių tyrimų.

Kadangi nustatyta nedidelė iškrita apima kanoninę sukirpimo vietą, atlikta kDNR, susintetintos naudojant tiriamosios ir jos mamos RNR mėginius, grynintus iš kraujo ėminių, sekoskaita Sangerio metodu. Fragmentams pagausinti naudotas tiesioginis pradmuo, kuris hibridizuojasi ant geno 5' netransliuojamos srities, ir atvirkštinis pradmuo, kuris jungiasi prie 3-io ir 4-o egzono jungties. Nustatyta, kad dėl nedidelės iškritos susidaro alternatyvi sukirpimo vieta (žr. 3.4.3.1 paviekslą). *In silico* tai lemia skaitymo rėmelio poslinkį ir priešlaikinės

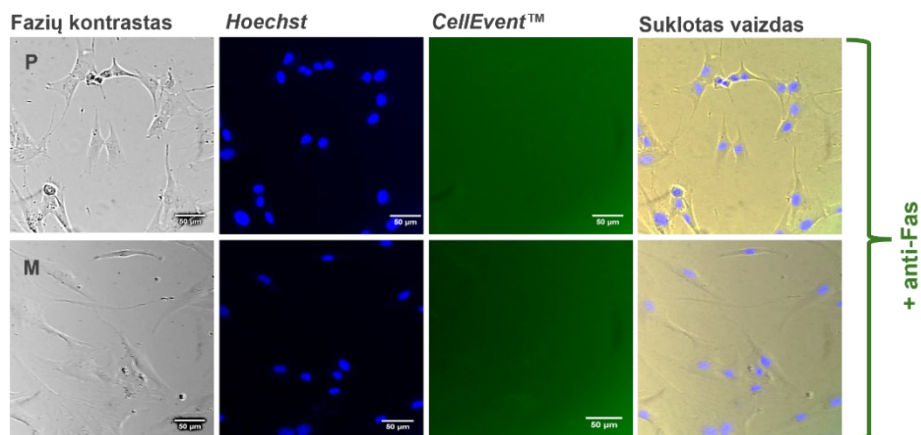
baigmės kodono susidarymą (NP_000034.1:p.(Thr6Phefs17Ter)). Šis kodonas susidaro arti *FAS* transkripto pradžios (5' galo), todėl tikėtina, kad yra inicijuojamas NMD mechanizmas ir pakitęs transkriptas ląstelėse yra suardomas.



3.4.3.1 paveikslas. *FAS* geno nedidelės iškritos, apimančios donorinę sukirpimo vietą, poveikio iRNR sukirpimui schema. Nustatytas variantas lemia alternatyvios sukirpimo vietos susidarymą ir skaitymo rėmelio poslinkį.

FAS genas koduoja TNF (angl. *Tumor Necrosis Factor*) receptorių superšeimos glikoproteiną, vadinamą ląstelės paviršiaus žūties receptoriumi Fas (angl. *Fas Cell Surface Death Receptor*). Pagrindinė šio receptoriaus funkcija daugelyje ląstelių yra programuotos ląstelės žūtis – apoptozė – reguliacija, kuri yra aktyvuojama prie *FAS* receptoriaus prisijungus jo ligandui FasL [181]. Remiantis *GTE*x domenų baze, *FAS* geno raiška vyksta įvairiuose audiniuose ir ląstelėse, didelė raiška vyksta ir fibroblastuose. Atsižvelgiant į galimą *FAS* baltymo kiekio sumažėjimą ir atitinkamą poveikį ląstelės programuotai žūčiai, nuspręsta tiriamašios ir jos mamos bei sveikų kontrolinių asmenų fibroblastuose įvertinti apoptozę. Kadangi tiriamajai diagnozuotas ALPS, kurio vieni iš klinikinių simptomų yra sutrikusi limfocitų apoptozė, padidėjęs skirtingus žymenis turinčių T ląstelių kiekis, iškelta hipotezė, kad fibroblastuose apoptozė taip pat sutrinka, tikėtina, yra mažiau efektyvi sergančiosios ląstelėse. Eksperimentui atlikti fibroblastai buvo 28 val. inkubuoti su Fas receptorių aktyvuojančiu antikūnu anti-Fas (Sigma-Aldrich) ir apoptozės indukcija buvo įvertinta komerciniu dažu *CellEvent™ Caspase-3/7 Green ReadyProbes™ Reagent* (Invitrogen). Ląstelių branduoliai žymėti *Hoechst* dažu. Po poveikio anti-Fas antikūnu apoptozei aktyvuoti, fibroblastų kultūros stebėtos ir vertintos fluorescencine mikroskopija. Atliekant mikroskopijos vaizdų analizę, lygintos

tiriamosios ir jos mamos ląstelės su atitinkamo amžiaus kontrolinių asmenų ląstelėmis – statistiškai reikšmingų skirtumų nenustatyta (žr. 3.4.3.2 paveikslą). Skirtumai taip pat neišryškėjo lyginant tiriamosios ir jos mamos fibroblastus. Apoptozės efektyvumas taip pat analizuotas tėkmės citometrijos metodu naudojant komercinį rinkinį *Dead Cell Apoptosis Kit with Annexin V for Flow Cytometry* (Invitrogen), tačiau reikšmingi skirtumai taip pat nenustatyti.



3.4.3.2 paveikslas. Apoptozės tyrimas fluorescencine mikroskopija po poveikio FAS aktyvuojančiu antikūnu (anti-FAS) tiriamųjų fibroblastuose. P – tiriamosios, kuriai nustatyta nedidelė iškrita *FAS* gene, fibroblastų kultūra. M – tiriamosios mamos, neturinčios ligos požymių, bet turinčios tą patį *FAS* DNR pokytį, fibroblastų kultūra. Šiuo tyrimu statistiškai reikšmingų skirtumų tarp *FAS* geno pokytį turinčių ir sveikų asmenų ląstelių nenustatyta.

Literatūros duomenimis, tyrimai, kuriais analizuota nuo FAS receptoriaus priklausoma apoptozė įvairių tipų ląstelėse, atskleidė, kad ši programuota ląstelių žūtis gali būti slopinama specifinio baltymo FLIP (angl. *Cellular FLICE-like inhibitory protein*; akronimai FLIP, cFLIP, CFLAR) [182]. Šis baltymas slopina FAS apoptozės signaliniame kelyje veikiančios kaspazės–8 proteolitinę aktyvumą, dėl kurio apoptozė nėra aktyvinama [182]. Įdomu tai, kad FLIP gali veikti tiek kaip nuo FAS priklausomos apoptozės slopiklis, tiek kaip stipriklis – poveikis priklauso nuo FLIP raiškos intensyvumo. Tose ląstelėse, kuriose vyksta nuolatinė šio baltymo raiška, FAS apoptozinis kelias yra slopinamas [182]. Žmogaus odos fibroblastuose atliktais eksperimentiniais tyrimais parodyta, kad slopinant FLIP raišką specifiniais priešprasminiais oligonukleotidais arba naudojant baltymų sintezę sutrikdantį cikloheksamidą FAS receptoriaus sukeliama apoptozė aktyvuojama – fibroblastai tampa jautrūs FAS ligandui ir ilgainiui žūsta [183].

Remiantis anksčiau aprašytu tyrimu, *CFLIP* geno raiška sumažinta mūsų tiriamosios ir jos mamos fibroblastuose naudojant specifinę siRNR (slopinimo efektyvumas tikrintas kPGR, raiška sumažėja nuo 4 iki 16 kartų) siekiant įvertinti, ar sukirpimo pokyčiai iRNR lygmeniu keičia nuo FAS priklausomos apoptozės indukcijos efektyvumą. Žūstančioms ląstelėms atpažinti fibroblastų kultūrose naudotas *Tali™ Apoptosis Kit - Annexin V Alexa Fluor™ 488 & Propidium Iodide* rinkinys. Antikūnu ir siRNR nepaveikuose arba tik antikūnų paveiktuose probando ir jo mamos mėginiuose žūstančių ląstelių, kurios atpažįstamos kaip aneksiną V prisijungusios ląstelės, aptikta apie 10 proc. Po poveikio FAS aktyvuojančiu antikūnu ir *CFLIP* raiškos sumažinimo probando mėginiuose apoptozinių ląstelių procentinė dalis padidėja apytiksliai 3 kartus, lyginant su taikinio neturinčia siRNR ir anti-FAS antikūnu paveiktomis probando ląstelėmis ($p=0,045$; žr. 3.4.3.3 paveikslą). Mamos ląstelių atitinkamuose mėginiuose stebimas panašus efektas – apoptozinių ląstelių skaičius išauga daugiau nei 20 proc. po *CFLIP* raiškos sumažinimo ir FAS receptoriaus apoptozinio kelio indukavimo (žr. 3.4.3.3 paveikslą).

			Procentinė žūstančių ląstelių dalis			
anti-FAS ak.	NK siRNR	CFLIP siRNR		I	II	III
-	✓	-	P	5.0	8.4	7.8
-	-	✓		4.3	15.7	11.0
✓	✓	-		5.2	13.1	10.2
✓	-	✓		16.7	38.5	34.7
-	✓	-	M	NA	30.0	0.0
-	-	✓		NA	20.6	19.1
✓	✓	-		NA	33.6	13.8
✓	-	✓		NA	52.2	47.9

3.4.3.3 paveikslas. Apoptozinių ląstelių dalis po poveikio FAS aktyvuojančiu antikūnu ir (ar) *CFLIP* raiškos sumažinimo sergančiosios ir jos mamos fibroblastuose. P – probando fibroblastų mėginiai, M – probando mamos fibroblastų mėginiai. Romėniški skaičiai (I, II, III) žymi tris atskirus eksperimentus. NA (angl. *Not Available*) pažymėti duomenys, kurių dėl techninių priežasčių nebuvo galima surinkti. Grafiko kiekviena eilutė reprezentuoja atskirą ląstelių mėginį. Kiekviename mėginyje tėkmės citometru surinkti duomenys bent 800 ląstelių. Ląstelių reguliuojamai žūčiai nustatyti naudotas *Tali™ Apoptosis Kit - Annexin V Alexa Fluor™ 488 & Propidium Iodide* (Thermo Fisher Scientific) rinkinys.

Įdomu ir tai, kad po poveikio FAS antikūnu ir *CFLIP* raiškos slopinimo žūstančių mamos fibroblastų nustatyta apytiksliai 20 proc. daugiau, lyginant su atitinkamu probando mėginiu, nors statistinis reikšmingumas nebuvo pasiektas ($p=0,086$). Tačiau, biologine prasme, gautas rezultatas nurodo galimą tendenciją, kad probando fibroblastuose apoptozė yra mažiau efektyvi. Vis dėlto dėl techninių priežasčių ne visuose eksperimentuose pavyko surinkti mėginių duomenis, todėl statistiniam reikšmingumui pasiekti ir apoptozės efektyvumo tendencijai pagrįsti tikslingas eksperimentų pakartojimas ir atitinkamų eksperimentų atilikimas sveikų kontrolinių asmenų ląstelėse. Taip pat būtų naudinga patyrinėti FAS signaliniame kelyje dalyvaujančių ir su šiuo keliu susijusių baltymų (juos koduojančių genų) raišką probando ir mamos ląstelėse genomo ir transkriptomo sekoskaitos metodais, siekiant išsiaiškinti galimas ligos (ne)pasireiškimo priežastis signalinių kelių lygmeniu.

Kadangi FAS receptorius yra ląstelės paviršiuje, jo struktūra skirstoma į užląstelinę ir viduląstelinę dalis. Taigi, N galas yra ligandą prijungiantis užląstelinis domenas (angl. *Prie-ligand Assembly Domain*, PLAD), toliau seka transmembraninis domenas ir receptoriaus C gale yra citoplazminis žūties domenas (angl. *Death Domain*, DD) [184]. Paprastai FAS receptorius ląstelėse yra monomerinėje būsenoje, tačiau po FasL prisijungimo sudaro homotrimerą, kuris reikalingas inicijuoti tolesnius viduląstelinius signalinius procesus (DISC komplekso susidarymą, kaspazių kaskadą), lemiančius apoptozę [185]. Mūsų tiriamajai ir jos požymių neturinčiai mamai nustatytai iškritai, lėmusiai alternatyvų sukirpimą, prognozuota priešlaikinės baigmės kodono pozicija (p.(Thr6Phefs17Ter)) leidžia daryti prielaidą, jog kelių amino rūgščių baltymo liekana ląstelėje turėtų nefunkcionuoti ir apskritai neturėtų būti transliuojama. *FAS* gene dažniau nustatomi variantai, kurie lemia DD domeno pokyčius, dėl kurių ląstelėse susidaro pakitęs baltymas, galintis užląsteliniu domenu sąveikauti su nepakitusiais FAS baltymais ir sukelti „dominuojančią interferenciją“ (angl. *dominant-interference*), kai susidaręs mutavęs FAS baltymas dėl sąveikos konkuruoja su nepakitusiais FAS monomerais ir sudaro neveiksnius heterotrimerus [180]. Rečiau nustatomi *FAS* pokyčiai, dėl kurių sumažėja šio geno koduojamo receptoriaus kiekis ląstelėje ir pasireiškia šio baltymo haplonepakankamumas [180]. Šiame darbe atlikti eksperimentai ne imuninės sistemos ląstelėse, bet sergančiosios ir jos mamos fibroblastuose, tačiau stebėti apoptozės efektyvumo skirtumai tarp šių dviejų tiriamųjų leidžia daryti prielaidą, kad nustatytas baltymą trumpinantis variantas turi mažesnę poveikį ląstelių žūčiai probando ląstelėse, lyginant su mamos ląstelėmis. Tai sufleruoja apie galimą kompensacinį mechanizmą, lemiantį nevisišką *FAS* geno lemiamos ligos penetrantiškumą.

Atlikti bioinformatiniai, molekuliniai ir funkciniai tyrimai bei literatūros analizė parodo *FAS* gene nustatytos nedidelės išskritos NC_000010.11(NM_000043.6):c.16_30+1del galimą patogeniškumą. Šiais tyrimais nustatyti *FAS* iRNR sukirpimo proceso pokyčiai leidžia daryti prielaidą, kad *FAS* indukuojamos apoptozės sumažėjęs efektyvumas tiriamosios ląstelės yra galimas ligos patogenezės mechanizmas. Vis dėlto būtų tikslingi papildomi genomo, transkriptomų tyrimai siekiant nustatyti signalinių kelių skirtumus tarp ligos požymius turinčios tiriamosios ir nesergančios, *FAS* geno iškritą nešiojančios mamos ląstelių.

3.4.4 *DNMT3A* gene nustatyto varianto įtakos metilinimo sutrikimui ląstelėse analizė

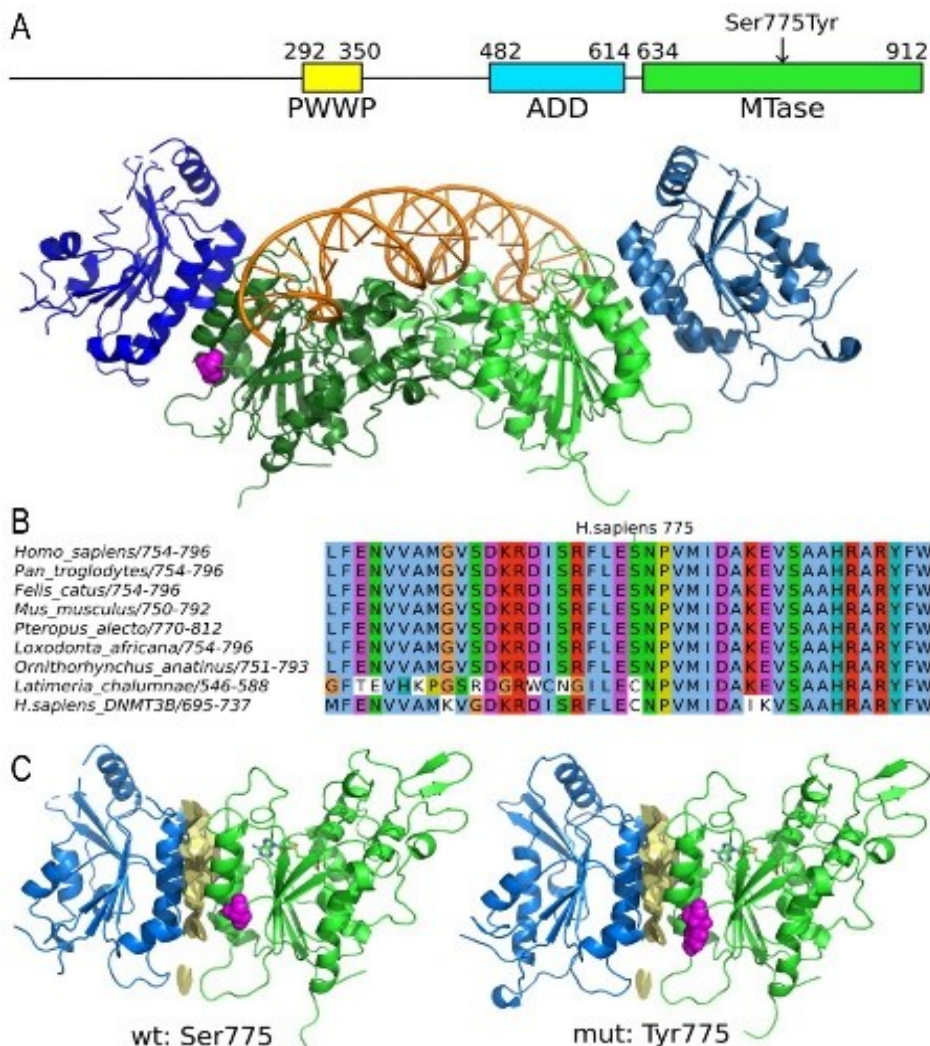
Moksliniame darbe analizuotas *DNMT3A* gene (MIM#602769) viso egzomo sekoskaita nustatytas, literatūroje neaprašytas heterozigotinis variantas NM_022552.4:c.2324C>A tiriamajam (ATGC21–12–P1, žr. 3.3 lentelę) su kardiologine patologija (aortos šaknies išsiplėtimas, kairiojo skilvelio nekompaktinė kardiomiopatija, aritmija), intelektine negalia, aukštu ūgiu, skeleto anomalijomis ir nutukimu. Atlikus segregacijos analizę Sangerio metodu, nustatyta, kad šį variantą probandas paveldėjo iš mamos, turinčios kardiologinę patologiją ir gretutines ligas. Variantas patvirtintas ir skeleto anomalijų, aukštą ūgį bei kardiologinių sutrikimų turinčiam broliui. OMIM duomenų bazėje *DNMT3A* yra žinomas ligos genas, kuriame nustatomi somatiniai DNR variantai lemia vėžinius sutrikimus (MIM#601626), o germinaciniai variantai lemia autosominio dominantinio paveldėjimo mikrocefalinį žemo ūgio sindromą (MIM#618724) arba aukšto ūgio Tatton-Brown-Rahman sindromą (MIM#615879). Pastarajam priskiriami fenotipiniai požymiai būdingi mūsų tiriamiesiems.

Nustatytas variantas neaptiktas *gnomAD* duomenų bazėje. Keliais bioinformatiniais įrankiais (*SIFT*, *PolyPhen*) variantas prognozuotas kaip galimai patogeninis. Nors, remiantis *SpliceAI* įrankiu, varianto poveikis sukirpimui labai mažai tikėtinas, *MutationTaster* įrankiu šis variantas vertintas kaip galimai turintis įtakos iRNR sukirpimo procesui. Labiausiai tikėtina dėl to, kad yra 20-o egzono 5' galo 2-oje pozicijoje (netoli kanoninės akceptorinės sukirpimo vietos). Nuspręsta atlikti tiriamojo kDNR, susintetintos naudojant tiriamojo RNR mėginį, grynintą iš periferinio kraujo ėminio, sekoskaitą Sangerio metodu. Fragmentams pagausinti naudotas tiesioginis pradmuo, kuris hibridizuojasi ant 18-o egzono ir atvirkštinis pradmuo, kuris hibridizuojasi ant

21-o ir 22-o egzono jungties. Vis dėlto, pakitusios kDNR sekos fragmentai nebuvo nustatyti.

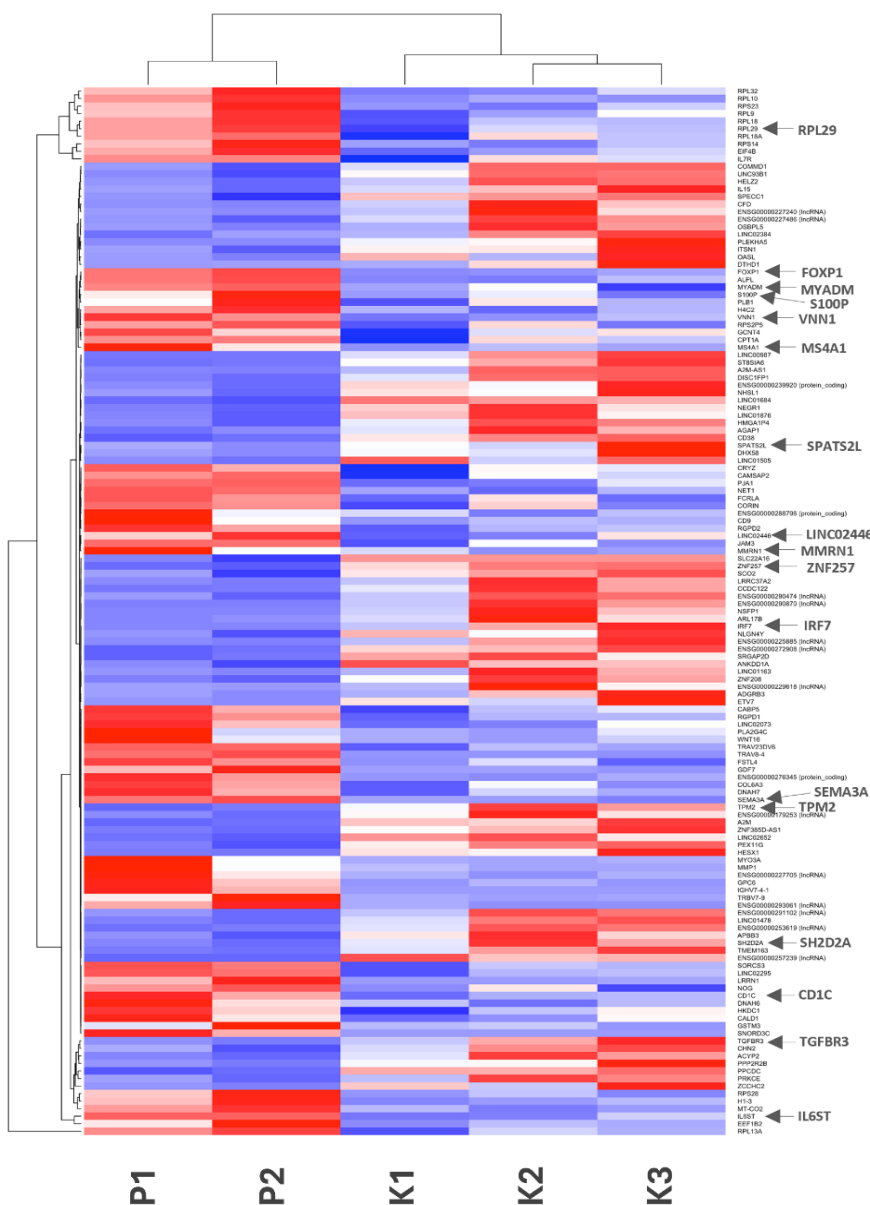
DNMT3A genas koduoja *de novo* DNR metilinimą atliekantį baltymą DNR metiltransferazę 3 α (angl. *DNA Methyltransferase 3 Alpha*). Jo raiška aptinkama beveik visuose organizmo audiniuose ir intensyviausia nepilnai diferencijuotose ląstelėse, kadangi sukuria metilinimo žymes nemetilintose DNR pozicijose [186]. DNMT3A atliekama epigenetinė DNR modifikacija yra vienas iš būdų ląstelei reguliuoti genų raišką, kadangi neretai metilinamos genų promotorių sekos, o dėl pasikeitusios chromatinio būsenos specifinio geno raiška yra slopinama. Todėl ši metiltransferazė yra itin svarbi embrioniniame etape dar tik vystantis organizmui, taip pat X chromosomos inaktyvacijai ir formuojant imprintingo žymes [186].

DNMT3A baltymas sudarytas iš dviejų pagrindinių domenų: N galo reguliacinio domeno (sudaro PWWP ir ADD domenai), reikalingo sąveikai su chromatinu, ir C galo katalizinio domeno, atliekančio metiltransferazės funkciją [186]. Baltymo lygmeniu nustatytas c.2324C>A variantas lemia aminorūgšties pokytį NP_072046.2:p.(Ser775Tyr) metiltransferazės domene, kuris yra evoliuciškai konservatyvus. Siekiant *in silico* įvertinti nustatytos aminorūgšties pokyčio įtaką baltymo struktūrai ir funkcijai, atlikti kompiuterinio modeliavimo tyrimai. Nustatyta, kad pakitusi aminorūgštis patenka į sritį, svarbią sąveikai su DNMT3L baltymu. Pastarasis yra katalizinio aktyvumo neturintis baltymas, kuris veikia kaip DNMT3A reguliacinis veiksnys, nes skatina katalizinį aktyvumą [187]. Teisinga DNMT3A sąveika su DNMT3L sukelia konformacinius DNR metiltransferazės pokyčius, dėl kurių atsiveria katalitinė DNMT3A dalis ir paskatinamas DNR prisijungimas [187]. Todėl tikėtina, kad esant sąveikos pokyčiams, galimas DNR metilinimo efektyvumo sutrikimas. Svarbu ir tai, kad Ser775 yra arti katalizinio domeno ir ADD domeno sąveikai svarbios srities (jungiklio srities). *In silico*, dėl serino pokyčio į didesnę aminorūgštį tiroziną, gali pakisti baltymo termodinaminės savybės ir sutrikti šių domenų sąveika, kuri, tikėtina, turi neigiamos įtakos DNMT3A atliekamai funkcijai (žr. 3.4.4.1 paveikslą).

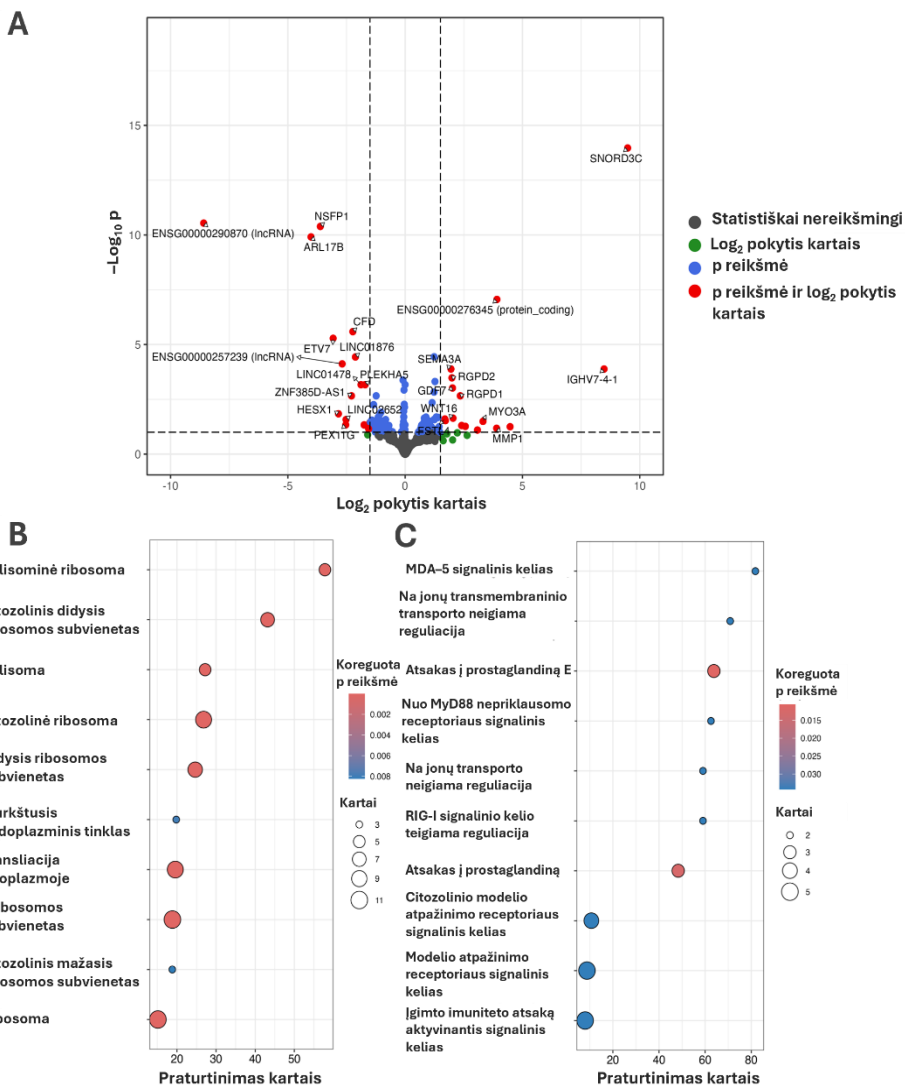


3.4.4.1 paveikslas. DNMT3A baltymo struktūriniai pokyčiai dėl aminorūgšties pakaitos Ser775Tyr. A – DNMT3A domenų organizacija ir struktūra komplekse su DNMT3L ir DNR (PDB: 5YX2). DNMT3A – žalia, DNMT3L – mėlyna, DNR – oranžinė, o Ser775 aminorūgštis pavaizduota magenta spalvos sferomis. B – daugybinis, žinduoliuose nustatytų DNMT3A baltymų sekų palyginys vaizduoja konservatyvų regioną, kuriame yra Ser775 aminorūgštis. C – pakitusi aminorūgštis 775 pozicijoje yra netoli DNMT3A ir DNMT3L baltymų sąveikos (PDB: 5YX2). Sąveika pavaizduota geltonai. Paveikslas sukurtas dr. J. Dapkūno.

Atsižvelgiant į gautus baltymo kompiuterinio modeliavimo rezultatus ir į DNMT3A atliekamą funkciją, atlikta tiriamojo ir jo brolio bei panašios amžiaus grupės, vyriškos lyties, sveikų asmenų (n=3) RNR mėginių transkriptomo sekoskaita. Nustatyta, kad, lyginant probando ir jo brolio mėginius su sveikais kontroliniais asmenimis, pakinta 139-ių genų raiška: 69-ių genų raiška padidėja, 70-ies genų raiška sumažėja (žr. 3.4.4.2 paveikslą ir 3.4.4.3). Daugiau nei trys ketvirtadaliai pakitusios raiškos genų koduoja baltymus, taip pat yra ilgų nekoduojančių RNR (inRNR) bei su konkrečia funkcija nesusietų transkriptų. Verta pabrėžti, kad *DNMT3A* geno raiška tirtuose mėginiuose nėra reikšmingai pakitusi. Kai kurie iš nustatytų genų analizuotoje literatūroje yra sietini su kardiovaskulinės sistemos biologija ir fiziologija. Vienas iš padidėjusios raiškos genų yra *FOXP1* (angl. *Forkhead Box P1*). Šis genas koduoja transkripcijos veiksnių, svarbų įvairiems audiniams vystytis, taip pat ir širdies audiniui [188]. Įdomu tai, kad *Foxp1* siejamas su širdies anomalijomis pelių embrioninių ląstelių tyrimuose ir yra aprašytas literatūroje kaip DNMT3A baltymo metilinas genas [188, 189]. Kiti su širdies ir kraujagyslių sistema susiję genai, kurių raiška reikšmingai padidėjusi probando ir jo brolio mėginyje, yra baltymų sintezėje dalyvaujantys (pvz. *RPL29*), imuninei sistemai svarbūs baltymai (pvz. *IL6ST*, *MS4A1*, *CD1C*) ir signaliniuose keliuose dalyvaujantys veiksniai (pvz. *SEMA3A*, *WNT16*) [190-193]. Reikšmingai sumažėjusi raiška genų, koduojančių su transkripcijos reguliavimu siejamus baltymus (pvz. *ZNF257*) ir signaliniais keliais (pvz. *SPATS2L*, *SI00P*), kurių reikšmė kardiovaskulinei sistemai analizuojama įvairiais tyrimais [194]. Be to, dalis pakitusios raiškos transkriptų (pvz. *MYADM*, *MMRN1*, *IRF7*, *LINC02446*, *CD9*, *VNN1*, *SH2D2A*), nustatytų mūsų tiriamiesiems, identifikuoti Smith ir kolegų (2021) [195] atliktame tyrime, kuriame analizuotas kitas DNMT3A baltymo pokytis (Arg882His) tiriamajam su Tatton-Brown-Rahman sindromo diagnoze. Nors tiesioginė sąsaja nustatytiems genams su tiriamųjų fenotipu kol kas eksperimentiškai neiširta, užuomazgų apie jų reikšmę kardiovaskulinės sistemos biologiniams ir fiziologiniams procesams yra, todėl neabejotinai svarbūs išsamesni funkciniai tyrimai sąsajai patvirtinti bei patogenezės mechanizmams atskleisti.

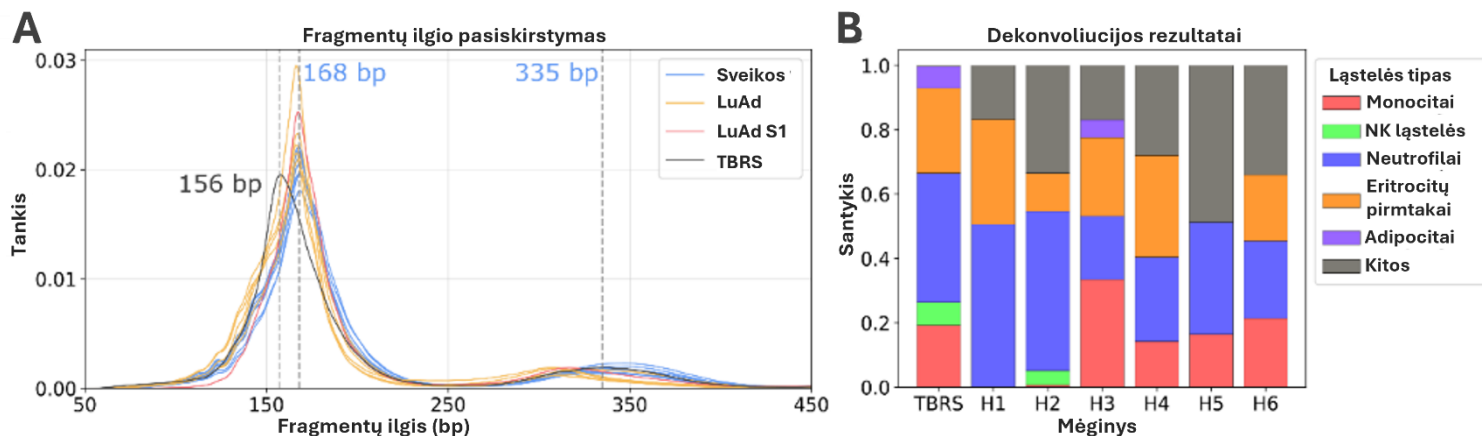


3.4.4.2 paveikslas. Pakitusios raiškos genai, nustatyti periferinio kraujo mėginiuose tiriamiesiems, kuriems nustatytas variantas *DNMT3A* gene. Pakitusios raiškos genų profilis probandų RNR mėginiuose, kuriems nustatytas variantas *DNMT3A* gene, lyginant su kontroliniais asmenimis. Genų raiškos profilis atspindi statistškai reikšmingą pakitusią genų raišką pacientų (P1 ir P2), ir kontrolinių asmenų (K1, K2, K3) grupėse. Kiekviena eilutė atitinka konkretų geną, o kiekvienas stulpelis reprezentuoja atskirą tiriamąjį. Raudona – padidėjusi raiška, mėlyna – sumažėjusi raiška, balta – vidutinė/nepakitusi raiška. Grafikas sukurtas dr. A. Konovalovo.



3.4.4.3 paveikslas. Sumažėjusios ir padidėjusios raiškos genai, nustatyti periferinio kraujo mėginiuose tiriamiesiems, kuriems nustatytas variantas *DNMT3A* gene. A – statistiškai reikšmingai pakitusios raiškos genų tarp tiriamojo su *DNMT3A* variantu ir sveikų kontrolinių asmenų Volcano grafikas. X ašis reprezentuoja genų raiškos pokytį kartais log2 skalėje, o Y ašis rodo FDR neigiamą dešimtainį logaritmą. Genai su reikšmingai pakitusia raiška ($FDR < 0,1$) ir daugiau nei 1,5 karto didesniu raiškos pokyčiu vaizduojami virš horizontalios brūkšninės linijos ir už vertikalių brūkšnių. B – statistiškai reikšmingai padidėjusios raiškos genų GO terminų praturtinimo taškinis grafikas. C – statistiškai reikšmingai sumažėjusios raiškos genų GO terminų praturtinimo taškinis grafikas. Grafikai sukurti dr. A. Konovalovo.

Siekiant įvertinti galimai pakitusio DNMT3A baltymo funkciją, atlikta laisvai cirkuliuojančios DNR (lcDNR) analizė trečios kartos sekoskaita (*Oxford Nanopore* technologija) tiriamojo broliui (tyrimui atlikti buvo galima gauti tik šio asmens mėginį). Šiuo tyrimu analizuojami nukleosomą apšiviję DNR fragmentai, kurių ilgis gali varijuoti priklausomai nuo DNR metilinimo [196]. Nustatytas trumpesnis mononukleosomos fragmento ilgis (156 bp) tiriamajam, lyginant su laisvai prieinamų duomenų bazėse esančių kontrolinių asmenų duomenimis (168 bp; žr. 3.4.4.4 paveikslą). Remiantis literatūra, trumpesni fragmentai būdingi esant hipometilinimui ir didesniame nukleazijų aktyvumui vėžiu sergančiųjų mėginiuose [197, 198]. Taip pat trečios kartos sekoskaita sudaro galimybę nustatyti lcDNR būdingą metilinimą (5mC) ir hidroksimetilinimą (hm5C). Šiuo atveju tiriamojo lcDNR mėginyje nustatytas bendras CpG salų modifikavimas, kurį lemia DNMT3A: 75,8% 5mC ir 2,4% 5hmC, iš viso 78,2% (žr. 3.4.4.4 paveikslą). Nustatytas kiekis yra panašus į Smith ir kolegų (2021) [195] publikuotame tyrime nustatytą metilinimą periferinio kraujo mėginiuose sergantiesiems tuo pačiu sindromu ir turintiems kitą aminorūgščių keičiantį variantą *DNMT3A* gene. Taigi, tiriamiesiems nustatytas *DNMT3A* geno variantas lemia jo koduojamo produkto funkcijos sutrikimus, kurie turi įtakos ne tik metilinimo ir hidroksimetilinimo procesams, bet ir kitų genų raiškai. Šie tyrimai patvirtina *DNMT3A* geno varianto patogeniškumą ir potencialiai atveria sudėtingą biologinių procesų valdymą, lemiantį tiriamojo Tatton-Brown-Rahman sindromo fenotipą.



3.4.4.4 paveikslas. Tiriamojo brolio kraujo lcDNR nanoporų sekoskaitos analizė. A – lcDNR fragmentų tankio grafikas. Fragmentų dydžio pasiskirstymas palygintas tarp *DNMT3A* geno variantą turinčių sergančiųjų (TBRS), septynių sveikų asmenų ir šešių plaučių adenokarcinomos pacientų (LuAd). Mėginys S1 yra plaučių adenokarcinomos paciento mėginys, kurio lcDNR biblioteka paruošta artimiausiu būdu mūsų tirtam asmeniui. B – lcDNR fragmentų ląstelinės kilmės nustatymas naudojant Moss et al. metodą [198]. Mūsų tiriamojo brolio (P2) rezultatai palyginti su šešiais sveikais (H1–H6) asmenimis iš Katsman et al. 2022 tyrimo [199]. Paveikslas sukurtas ir tyrimai atlikti dr. M. Tomkuvienės ir J. Balčiūno.

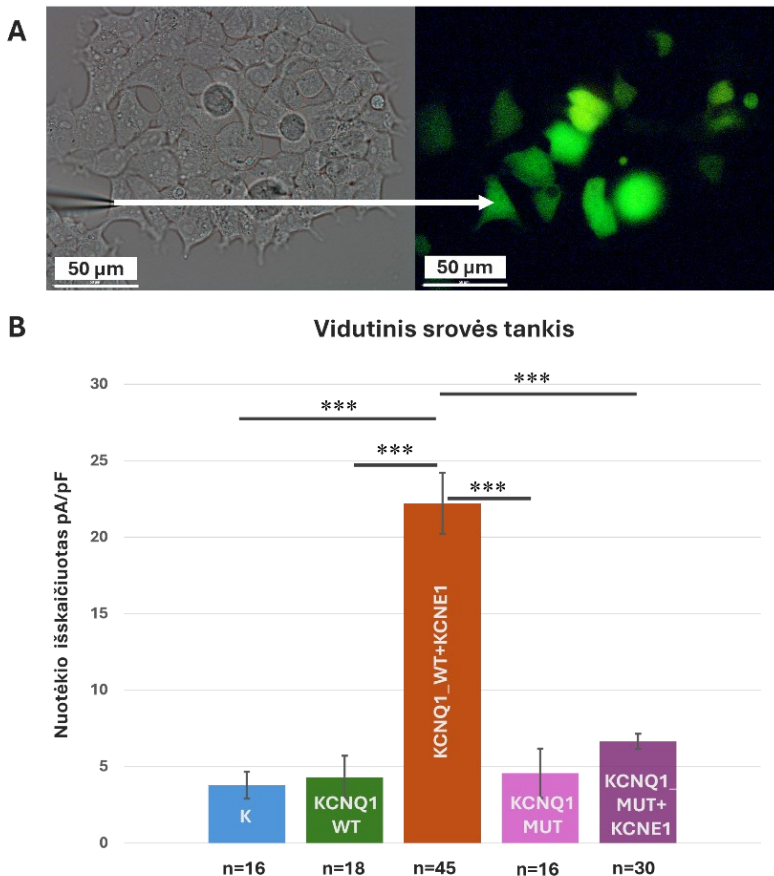
Atlikti bioinformatiniai ir molekuliniai tyrimai bei literatūros analizė pagrindžia *DNMT3A* gene nustatyto varianto NM_022552.4:c.2324C>A klinikinę ir funkcinę reikšmę. Nustatyti DNMT3A baltymo funkcijos pokyčiai tiriamųjų mėginiuose atliekant RNR ir lcdNR sekoskaitos tyrimus leidžia daryti prielaidą apie mažiau efektyvų metilinimo procesą ir sutrikusią specifinių genų raišką kaip galimą patogeneizės mechanizmą.

3.4.5 *KCNQ1* gene nustatyto varianto lemiami kalio jonų judėjimo ląstelių membrana sutrikimų analizė

Moksliniame darbe analizuotas *KCNQ1* gene (MIM#607542) egzomo sekoskaita nustatytas, literatūroje neaprašytas heterozigotinis variantas NM_000218.3:c.1111G>C tiriamajam (ATGC21–19–P, žr. 3.2 lentelę), kuriam diagnozuotas ilgo QT sindromas. OMIM duomenų bazėje šis genas yra žinomas ligos genas, kurio DNR variantai siejami su autosominio recesyvaus paveldėjimo Jervell ir Lange-Nielsen sindromu (MIM#220400) bei autosominio dominantinio paveldėjimo šeiminiu prieširdžių virpėjimu (3-io tipo; MIM#607554), trumpo QT sindromu (2-o tipo; MIM#609621) ir ilgo QT sindromu (1-o tipo; MIM#192500). Pastarajam sutrikimui būdingi fenotipiniai bruožai atitinka tiriamojo kardiopatologiją.

Nustatytas variantas baltymo lygmeniu (NP_000209.2:p.(Ala371Pro)) patenka į KCNQ1 C galo A spiralę, svarbią kalio kanalo sąveikai su kalmodulinu, kuris reguliuoja tiek kanalo atsidarymą, tiek tetramerinės struktūros susidarymą [199]. Todėl tikėtina, kad nustatytas variantas gali turėti neigiamos įtakos kalio kanalo funkcijai ar net visai tetramerinei struktūrai. Siekiant įvertinti galimą pakitusio KCNQ1 disfunkciją, nuspręsta stabilioje ląstelių linijoje sukurti pakitusio ir nepakitusio baltymų raiškos sistemą bei atlikti elektrofiziologinius matavimus. Tyrimams pasirinkta literatūroje aprašoma HEK293T ląstelių linija, kurių genome yra *KCNQ1* genas, tačiau reikšminga šio geno raiška ir koduojamo baltymo procesai šiose ląstelėse nevyksta [200]. Naudojant cheminę transfekciją, į ląsteles įvestos nepakitusi (KCNQ1_WT) ir pakitusi (KCNQ1_MUT) baltymą koduojančios plazmidės bei KCNE1 baltymą koduojanti plazmidė, kurių raiška tikrinta fluorescenciniu mikroskopu. Praėjus 24–48 val. po transfekcijos, specialiose plokštelėse ant stikliuko augančioms ląstelėms atlikti elektrofiziologiniai matavimai „patch-clamp“ metodu (žr. 3.4.5.1 A paveikslą). Sėkmingai užregistruoti keli atsakų pavyzdžiai. Netransfekuotose HEK293T ląstelėse užfiksuota minimali išeinanti srovė ($3,78 \pm 0,88$ pA/pF; žr. 3.4.5.1 B paveikslą), kuri aprašyta ir kituose tyrimuose [200]. Statistiškai reikšmingi skirtumai nenustatyti lyginant netransfekuotas su KCNQ1_WT ($4,31 \pm 1,39$ pA/pF) arba KCNQ1_MUT ($4,56 \pm 1,62$ pA/pF) plazmidė transfekuotomis ląstelėmis (žr. 3.4.5.1 B paveikslą). Tačiau transfekavus ląsteles KCNQ1_WT ir

KCNE1 plazmidžių kombinacija, išeinanti srovė statistiškai reikšmingai padidėjo ($22,25 \pm 1,91$ pA/pF), o KCNQ1_MUT ir KCNE1 plazmidžių kombinaciją turinčiose ląstelėse ši srovė statistiškai reikšmingai sumažėjo ($6,64 \pm 0,49$ pA/pF; žr. 3.4.5.1 B paveikslą). Tai signalizuoja apie KCNQ1 baltymo formuojamo kalio kanalo funkcijos sutrikimus, dėl kurių ląstelėse, tikėtina, sutrinka veikimo potencialo trukmė.



3.4.5.1 paveikslas. Elektrofiziologiniai matavimai „patch-clamp“ metodu HEK293T ląstelėse. A – pavyzdys, kaip elektrofiziologiniams matavimams atrenkamos HEK293T ląstelės pagal fluorescencijos intensyvumą. Ląstelės transfektuotos plazmidėmis, koduojančiomis žaliai fluorescuojantį baltymą. B – ląstelės transfektuotos skirtingomis plazmidžių kombinacijomis. KCNQ1_WT – nepakitusį KCNQ1 baltymą koduojanti plazmidė, KCNQ1_MUT – tiriamajam nustatytą *KCNQ1* variantą koduojanti plazmidė, KCNE1 – KCNQ1 baltymo veiklai svarbų subvienetą koduojanti plazmidė, K – netransfekuotos ląstelės. Srovės tankis matuotas 2 s pulso pabaigoje iki 30 mV. Grafikas žymi vidurkių $\pm$ standartinis nuokrypis, statistinis reikšmingumas * $p \leq 0,05$, *** $p \leq 0,005$. Nuotraukos, elektrofiziologiniai matavimai atlikti ir grafikas sukurtas A. Shelest, prof. A. Alaburdos.

KCNQ1 genas koduoja širdies veikimo potencialo repoliarizacijos fazei svarbų, nuo įtampos priklausomą kalio kanalą (angl. *Potassium Voltage-Gated Channel Subfamily Q Member 1*) [201]. Tai membraninio, porą formuojančio baltymų kanalo subvienetas, kuris yra aktyvuojamas uždelstos išlyginamosios srovės [201]. Pastarajame kanale KCNQ1 sudaro kompleksą su KCNE1 (angl. *Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 1*), o šis dviejų subvienetų monomeras formuoja tetramerą, kuris reguliuoja širdies veikimo potencialo trukmę [201]. KCNQ1/KCNE1 kompleksas sukuria išorinę kalio jonų srovę (kadangi lemia kalio jonų judėjimą iš ląstelės), kuri yra repoliarizuojanti jonų srovė ir nutraukia širdies veikimo potencialą, o sutrikimai KCNQ1 subvienete lemia būtent veikimo potencialo trukmės sutrikimus [201]. KCNQ1 yra sudarytas iš trijų pagrindinių sričių: N gale esančios S0 spiralės, šešių transmembraninių spiralių (S1–S6) ir keturių C gale esančių viduląstelių spiralių (A–D) [202]. Shamgar ir kolegų (2006) [203] atliktas tyrimas, kuriame elektrofiziologiniais matavimais tirti kiti KCNQ1 baltymo pokyčiai (pvz. p.Ser373Pro, p.Ala371Thr) nustatyti tiriamiesiems su ilgo QT sindromu. Mokslininkams pavyko nustatyti pakitusių baltymų nulemtus srovės pokyčius CHO (ang. *Chinese Hamster Cells*) ląstelėse, dėl kurių sutrinka nuo įtampos priklausomas KCNQ1 kanalų inaktyvavimas [203]. Tai eksperimentas, kurio pagrindu atlikti matavimai mūsų tiriamojo DNR pokyčio raišką turinčiose ląstelėse. Ši tyrimo strategija ir HEK293T sukurtas raiškos modelis pagrindžia mūsų tiriamajam nustatyto varianto funkcinę ir klinikinę reikšmę, kadangi pakitusį KCNQ1 baltymą turinčiose ląstelėse nustatyta sumažėjusi išeinanti srovė, signalizuojanti apie potencialo pokyčius ląstelių membranose.

Atlikti bioinformatiniai ir molekuliniai tyrimai bei literatūros analizė pagrindžia *KCNQ1* gene nustatyto varianto NM_000218.3:c.1111G>C klinikinę ir funkcinę reikšmę. Nustatyti KCNQ1 formuojamo kalio kanalo funkcijos pokyčiai HEK293T ląstelėse atliekant elektrofiziologinius matavimus leidžia daryti prielaidą apie sutrikusį kalio jonų judėjimą ląstelės membrana ir pakitusią veikimo potencialo trukmę kaip galimą patogenezės mechanizmą.

Šiame rezultatų ir jų aptarimo poskyryje aprašyti galimai patogeniniai ir patogeniniai VNV bei nedidelė iškrita analizuoti ne tik molekulinio, bet ir ląstelių procesų aspektu. Kiekvienam iš penkių atvejų taikytos kiek skirtingos molekulinio ir/ar funkcinio eksperimentinio tyrimų strategijos, kurias daugiausiai nulemia geno koduojamo produkto funkcija. Minėtais tyrimais atskleisti molekuliniai patogenezės mechanizmai ir specifiniai, tik tam tikro DNR varianto sukelti molekulių ir ląstelių pokyčiai bei sutrikimai. Šiems penkiems individualiems atvejams sukurtų tyrimų strategijų įgyvendinimas ir gauti rezultatai iliustruoja ne tik molekulinio procesų kompleksiskumą, bet ir

kiekvieno DNR pokyčio bei apskritai tiriamojo genetinio ir epigenetinio fono įtaką paveldimo sutrikimo patogenezei ir jos mechanizmo nustatymui.

Moksliniame darbe taikytos tyrimų strategijos ir metodai yra potencialūs diagnostiniai įrankiai klinikinėje praktikoje ir personalizuotai medicinai. Remiantis skirtingų šalių diagnostinių laboratorijų patirtimi, šių tyrimu perkėlimas į klinikinius diagnostinius algoritmus yra sudėtingas procesas, reikalaujantis spręsti kompleksinius metodologinius, analitinius ir praktinius iššūkius. Viena iš kliūčių yra standartizuotų protokolų, skirtų mėginių apdorojimui, tyrimo strategijoms ir duomenų generavimui, taikymas globaliu mastu siekiant valdyti rezultatų palyginamumą ir atkuriamumą. Egzomo, genomo ir transkriptomo sekoskaita sukuria didelius duomenų masyvus, kurių apdorojimas, saugojimas ir klinikinės reikšmės interpretacija reikalauja sudėtingos bioinformatinės infrastruktūros bei aukštos kvalifikacijos specialistų. Taip pat labai svarbus ir daug laiko bei žmogiškųjų išteklių reikalaujantis etapas yra funkcinės variantų analizės klinikinis validavimas, kuriuo siekiama parodyti tyrimo strategijos analitinį tikslumą, atkuriamumą ir klinikinę naudą diagnozuojant, prognozuojant ligos eigą ar siekiant parinkti tinkamiausią gydymą. Naujų tyrimų ir diagnostinių algoritmų diegimą riboja ir būtinos pirminės finansinės investicijos į naujam analizės algoritmui reikalingą įrangą, galima aukšta funkcinių tyrimų savikaina. Neabejotinai, integruoti didelio masto ir funkcinis tyrimus riboja specialistų, gebančių interpretuoti daugiasluoksnius duomenis klinikiniame kontekste, trūkumas. Vis dėlto, atsižvelgiant į transkriptomo ar funkcinę analizę taikančias užsienio diagnostines laboratorijas bei į bendrą medicinines bei žmogaus genetikos ir genomikos kontekstą pasaulyje, Lietuvos asmens sveikatos priežiūros įstaigos yra potencialiai tinkamos taikyti naujas tyrimų strategijas įveikus pagrindinius iššūkius.

IV IŠVADOS

1. Naujos kartos sekoskaita nustatyti vieno nukleotido variantai žinomuose ligos genuose *RLIM*, *NAA15*, *WRN*, *DYNC1H1*, *DNMT3A* ir *KCNQ1* bei nedidelė iškrita *FAS* gene vertinti kaip neaiškos klinikinės reikšmės ar galimai patogeniniai DNR pokyčiai, tinkami molekuliniams ir/ar funkciniais tyrimams.
2. Atlikus genomo struktūrinių persitvarkymų charakterizavimą, nustatyti klinikinę reikšmę turintys genai kandidatai:
 - 2.1 Į analizuotų mikrolecijų sritis patenkantys baltymus koduojantys genai *KIT*, *ANKRD17*, *XYLT1*, iškritusio *MED13L* geno egzoniui bei mikroduplikacijos srityje esantis miRNR klasteris *MIR17HG*, tikėtina, lemia tiriamųjų klinikinio fenotipo vystymąsi ir yra potencialūs kandidatai tolesniems funkciniais tyrimams ląstelėse ar modeliniuose organizmuose, siekiant įrodyti genų lemiamą molekulinio patogenezės mechanizmo poveikį fenotipui;
 - 2.2 Chromotripsio lemti trūkio taškai ir susidariusių fragmentų persitvarkymai lemia 42–jų genų raiškos pokyčius, iš kurių *NTNG2* genas galimai lemia dalį tiriamajam būdingų klinikinių bruožų pasireiškimą.
3. Taikant bioinformatinę analizę ir molekulinius tyrimus įvertinta atrinktų VNV reikšmė baltymų lygmeniu:
 - 3.1 Baltymo struktūros ir sąveikos modeliavimu bei DNR srities metilinimo įvertinimu nustatyti *RLIM* geno c.1721T>A varianto lemiami koduojamo baltymo sąveikos ir, tikėtina, funkcijos pokyčiai, galimai sutrikdantys ubikvitilinimą tiriamojo ląstelėse;
 - 3.2 Baltymo struktūros analize ir imunofermentiniu baltymo kiekio įvertinimu nustatytas *NAA15* geno c.382C>T varianto lemiamas koduojamo baltymo raiškos sumažėjimas, tikėtina, sutrikdantis N galo acetilinimo procesus tiriamojo ląstelėse.
4. Taikant molekulinius ir/ar funkcinis tyrimus, įvertinta atrinktų VNV ir nedidelės iškritos įtaka ląstelėse vykstantiems procesams:
 - 4.1 Eksperimentiniais tyrimais fibroblastuose su *WRN* gene nustatyta DNR pokyčiu c.1578del, lemiančiu baltymo sutrumpėjimą, nustatytas pagreitinėjęs ląstelių replikacinis senėjimas;
 - 4.2 Genų ir baltymų raiškos, konfokaline mikroskopija bei genų raiškos slopinimu nustatyta *DYNC1H1* geno c.6405+1G>C varianto, lemiančio baltymo sutrumpėjimą, įtaka *DYNC1H1* baltymo kiekio

- sumažėjimui, KIFC1 geno raiškos padidėjimui ir pakitusiam Goldžio komplekso membranų atsistatymui;
- 4.3 Eksperimentiniais, fluorescencinės mikroskopijos ir tėkmės citometrijos tyrimais nustatyta FAS geno nedidelės iškritos c.16_30+1del, lemiančios baltymo sutrumpėjimą, įtaka sumažėjusios FAS receptoriaus indukuojamos apoptozės tendencijai;
 - 4.4 In vitro baltymo tyrimais, didelio masto genų raiškos ir lcDNR sekoskaitos tyrimais nustatyta DNMT3A geno c.2324C>A varianto, lemiančio aminorūgšties pokytį, įtaka DNMT3A metiltransferazės sumažėjusiam aktyvumui ir pakitusiai genų raiškai;
 - 4.5 Eksperimentiniais ir elektrofiziologiniais HEK293T ląstelių tyrimais nustatytas KCNQ1 geno c.1111G>C variantas, lemiantis aminorūgšties pokytį, neigiamas poveikis nuo potencialo priklausomai ištekanti srovei ir veikimo potencialui.

LITERATŪROS SĄRAŠAS

1. Lee CE, Singleton KS, Wallin M, Faundez V. Rare Genetic Diseases: Nature's Experiments on Human Development. *iScience*. 2020;23(5):101123.
2. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 2020;28(2):165-73.
3. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet*. 2018;19(5):253-68.
4. Posey JE, O'Donnell-Luria AH, Chong JX, Harel T, Jhangiani SN, Coban Akdemir ZH, et al. Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet Med*. 2019;21(4):798-812.
5. Might M, Crouse AB. Why rare disease needs precision medicine-and precision medicine needs rare disease. *Cell Rep Med*. 2022;3(2):100530.
6. 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(5):405-24.
7. Hartman P, Beckman K, Silverstein K, Yohe S, Schomaker M, Henzler C, et al. Next generation sequencing for clinical diagnostics: Five year experience of an academic laboratory. *Mol Genet Metab Rep*. 2019;19:100464.
8. Hartin SN, Means JC, Alaimo JT, Younger ST. Expediting rare disease diagnosis: a call to bridge the gap between clinical and functional genomics. *Mol Med*. 2020;26(1):117.
9. Rehm HL, Alaimo JT, Aradhya S, Bayrak-Toydemir P, Best H, Brandon R, et al. The landscape of reported VUS in multi-gene panel and genomic testing: Time for a change. *Genet Med*. 2023;25(12):100947.
10. Punnett RC. Linkage in the sweet pea (*Lathyrus odoratus*). *Journal of Genetics*. 1923;13(1):101-23.
11. Pulst SM. Genetic linkage analysis. *Arch Neurol*. 1999;56(6):667-72.
12. Bailey-Wilson JE, Wilson AF. Linkage analysis in the next-generation sequencing era. *Hum Hered*. 2011;72(4):228-36.
13. Marwaha S, Knowles JW, Ashley EA. A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med*. 2022;14(1):23.
14. Seaby EG, Pengelly RJ, Ennis S. Exome sequencing explained: a practical guide to its clinical application. *Brief Funct Genomics*. 2016;15(5):374-84.

15. Bick D, Fraser PC, Gutzeit MF, Harris JM, Hambuch TM, Helbling DC, et al. Successful Application of Whole Genome Sequencing in a Medical Genetics Clinic. *J Pediatr Genet.* 2017;6(2):61-76.
16. Gambin T, Yuan B, Bi W, Liu P, Rosenfeld JA, Coban-Akdemir Z, et al. Identification of novel candidate disease genes from de novo exonic copy number variants. *Genome Med.* 2017;9(1):83.
17. Licata L, Via A, Turina P, Babbi G, Benevenuta S, Carta C, et al. Resources and tools for rare disease variant interpretation. *Front Mol Biosci.* 2023;10:1169109.
18. Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol Hepatol Bed Bench.* 2014;7(1):17-31.
19. Liu D, Zhou D, Sun Y, Zhu J, Ghoneim D, Wu C, et al. A Transcriptome-Wide Association Study Identifies Candidate Susceptibility Genes for Pancreatic Cancer Risk. *Cancer Res.* 2020;80(20):4346-54.
20. Wu C, Bradley J, Li Y, Wu L, Deng HW. A gene-level methylome-wide association analysis identifies novel Alzheimer's disease genes. *Bioinformatics.* 2021;37(14):1933-40.
21. Seaby EG, Rehm HL, O'Donnell-Luria A. Strategies to Uplift Novel Mendelian Gene Discovery for Improved Clinical Outcomes. *Front Genet.* 2021;12:674295.
22. Kanzi AM, San JE, Chimukangara B, Wilkinson E, Fish M, Ramsuran V, et al. Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. *Front Genet.* 2020;11:544162.
23. Hieter P, Andrews B, Fowler D, Bellen H. Highlighting rare disease research with a GENETICS and G3 series on genetic models of rare diseases. *Genetics.* 2023;224(4).
24. Saudi Mendeliome G. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol.* 2015;16(1):134.
25. Bean LJH, Funke B, Carlston CM, Gannon JL, Kantarci S, Krock BL, et al. Diagnostic gene sequencing panels: from design to report-a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2020;22(3):453-61.
26. Fernandez-Marmiesse A, Gouveia S, Couce ML. NGS Technologies as a Turning Point in Rare Disease Research , Diagnosis and Treatment. *Curr Med Chem.* 2018;25(3):404-32.
27. Jiman OA, Taylor RL, Lenassi E, Smith JC, Douzgou S, Ellingford JM, et al. Diagnostic yield of panel-based genetic testing in syndromic inherited retinal disease. *Eur J Hum Genet.* 2020;28(5):576-86.

28. Antonarakis SE, Krawczak M, Cooper DN. The Nature and Mechanisms of Human Gene Mutation. In: Valle DL, Antonarakis S, Ballabio A, Beaudet AL, Mitchell GA, editors. *The Online Metabolic and Molecular Bases of Inherited Disease*. New York, NY: McGraw-Hill Education; 2019.
29. Dillon OJ, Lunke S, Stark Z, Yeung A, Thorne N, Melbourne Genomics Health A, et al. Exome sequencing has higher diagnostic yield compared to simulated disease-specific panels in children with suspected monogenic disorders. *Eur J Hum Genet*. 2018;26(5):644-51.
30. King DA, Sifrim A, Fitzgerald TW, Rahbari R, Hobson E, Homfray T, et al. Detection of structural mosaicism from targeted and whole-genome sequencing data. *Genome Res*. 2017;27(10):1704-14.
31. Demidov G, Laurie S, Torella A, Piluso G, Scala M, Morleo M, et al. Structural variant calling and clinical interpretation in 6224 unsolved rare disease exomes. *Eur J Hum Genet*. 2024;32(8):998-1004.
32. Tan TY, Dillon OJ, Stark Z, Schofield D, Alam K, Shrestha R, et al. Diagnostic Impact and Cost-effectiveness of Whole-Exome Sequencing for Ambulant Children With Suspected Monogenic Conditions. *JAMA Pediatr*. 2017;171(9):855-62.
33. Retterer K, Jussola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. 2016;18(7):696-704.
34. Slavotinek A, Rego S, Sahin-Hodoglugil N, Kvale M, Lianoglou B, Yip T, et al. Diagnostic yield of pediatric and prenatal exome sequencing in a diverse population. *NPJ Genom Med*. 2023;8(1):10.
35. Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. *Clinical Medicine*. 2017;17(6):545-51.
36. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74.
37. Bagnall RD, Ingles J, Dinger ME, Cowley MJ, Ross SB, Minoche AE, et al. Whole Genome Sequencing Improves Outcomes of Genetic Testing in Patients With Hypertrophic Cardiomyopathy. *J Am Coll Cardiol*. 2018;72(4):419-29.
38. Eichinger J, Elger BS, Kone I, Filges I, Shaw D, Zimmermann B, et al. The full spectrum of ethical issues in pediatric genome-wide sequencing: a systematic qualitative review. *BMC Pediatr*. 2021;21(1):387.
39. Weisschuh N, Mazzola P, Zuleger T, Schaeferhoff K, Kühlewein L, Kortüm F, et al. Diagnostic genome sequencing improves diagnostic yield: a prospective single-centre study in 1000 patients with inherited eye diseases. *Journal of Medical Genetics*. 2024;61(2):186.

40. Ewans LJ, Minoche AE, Schofield D, Shrestha R, Puttick C, Zhu Y, et al. Whole exome and genome sequencing in mendelian disorders: a diagnostic and health economic analysis. *Eur J Hum Genet.* 2022;30(10):1121-31.
41. Lee H, Huang AY, Wang LK, Yoon AJ, Renteria G, Eskin A, et al. Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet Med.* 2020;22(3):490-9.
42. Barros-Silva D, Marques CJ, Henrique R, Jeronimo C. Profiling DNA Methylation Based on Next-Generation Sequencing Approaches: New Insights and Clinical Applications. *Genes (Basel).* 2018;9(9).
43. Fu MP, Merrill SM, Sharma M, Gibson WT, Turvey SE, Kobor MS. Rare diseases of epigenetic origin: Challenges and opportunities. *Front Genet.* 2023;14:1113086.
44. Barbosa M, Joshi RS, Garg P, Martin-Trujillo A, Patel N, Jadhav B, et al. Identification of rare de novo epigenetic variations in congenital disorders. *Nat Commun.* 2018;9(1):2064.
45. Zhou H, Brockington M, Jungbluth H, Monk D, Stanier P, Sewry CA, et al. Epigenetic allele silencing unveils recessive RYR1 mutations in core myopathies. *Am J Hum Genet.* 2006;79(5):859-68.
46. Cropley JE, Martin DI, Suter CM. Germline epimutation in humans. *Pharmacogenomics.* 2008;9(12):1861-8.
47. Oussalah A, Siblini Y, Hergalant S, Chery C, Rouyer P, Cavicchi C, et al. Epimutations in both the TESK2 and MMACHC promoters in the Epi-cblC inherited disorder of intracellular metabolism of vitamin B(12). *Clin Epigenetics.* 2022;14(1):52.
48. Castillejo A, Hernández-Illán E, Rodríguez-Soler M, Pérez-Carbonell L, Egoavil C, Barberá VM, et al. Prevalence of MLH1 constitutional epimutations as a cause of Lynch syndrome in unselected versus selected consecutive series of patients with colorectal cancer. *Journal of Medical Genetics.* 2015;52(7):498.
49. Sadikovic B, Levy MA, Kerkhof J, Aref-Eshghi E, Schenkel L, Stuart A, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med.* 2021;23(6):1065-74.
50. Chen YC, Liu T, Yu CH, Chiang TY, Hwang CC. Effects of GC bias in next-generation-sequencing data on de novo genome assembly. *PLoS One.* 2013;8(4):e62856.
51. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* 2008;18(11):1851-8.

52. Yu S-Y, Xi Y-L, Xu F-Q, Zhang J, Liu Y-S. Application of long read sequencing in rare diseases: The longer, the better? *European Journal of Medical Genetics*. 2023;66(12):104871.
53. Niehaus A, Azzariti DR, Harrison SM, DiStefano MT, Hemphill SE, Senol-Cosar O, et al. A survey assessing adoption of the ACMG-AMP guidelines for interpreting sequence variants and identification of areas for continued improvement. *Genet Med*. 2019;21(8):1699-701.
54. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-43.
55. Gudmundsson S, Singer-Berk M, Watts NA, Phu W, Goodrich JK, Solomonson M, et al. Variant interpretation using population databases: Lessons from gnomAD. *Hum Mutat*. 2022;43(8):1012-30.
56. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG, ClinGen Sequence Variant Interpretation Working G. Updated recommendation for the benign stand-alone ACMG/AMP criterion. *Hum Mutat*. 2018;39(11):1525-30.
57. Fowler DM, Rehm HL. Will variants of uncertain significance still exist in 2030? *Am J Hum Genet*. 2024;111(1):5-10.
58. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36(10):928-30.
59. Tang H, Thomas PD. Tools for Predicting the Functional Impact of Nonsynonymous Genetic Variation. *Genetics*. 2016;203(2):635-47.
60. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535-48 e24.
61. Ha C, Kim JW, Jang JH. Performance Evaluation of SpliceAI for the Prediction of Splicing of NF1 Variants. *Genes (Basel)*. 2021;12(9).
62. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2019;35(11):1978-80.
63. Lunke S, Bouffler SE, Patel CV, Sandaradura SA, Wilson M, Pinner J, et al. Integrated multi-omics for rapid rare disease diagnosis on a national scale. *Nat Med*. 2023;29(7):1681-91.
64. Kurosaki T, Maquat LE. Nonsense-mediated mRNA decay in humans at a glance. *J Cell Sci*. 2016;129(3):461-7.
65. Hug N, Longman D, Caceres JF. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res*. 2016;44(4):1483-95.

66. Mao B, Yang J, Zhao X, Jia X, Shi X, Zhao L, et al. Identification and functional characterization of a novel heterozygous splice-site mutation in the calpain 3 gene causes rare autosomal dominant limb-girdle muscular dystrophy. *Exp Ther Med*. 2024;27(3):97.
67. Hejla D, Huynh S, Samra S, Richmond PA, Dalmann J, Del Bel KL, et al. Naturally occurring splice variants dissect the functional domains of BHC80 and emphasize the need for RNA analysis. *Am J Med Genet A*. 2024;194(6):e63548.
68. Tian Y, Xie H, Yang S, Shangguan S, Wang J, Jin C, et al. Possible Catch-Up Developmental Trajectories for Children with Mild Developmental Delay Caused by NAA15 Pathogenic Variants. *Genes (Basel)*. 2022;13(3).
69. Sanoguera-Miralles L, Valenzuela-Palomo A, Bueno-Martinez E, Esteban-Sanchez A, Lorca V, Llinares-Burguet I, et al. Systematic Minigene-Based Splicing Analysis and Tentative Clinical Classification of 52 CHEK2 Splice-Site Variants. *Clin Chem*. 2024;70(1):319-38.
70. Dionnet E, Defour A, Da Silva N, Salvi A, Lévy N, Krahn M, et al. Splicing impact of deep exonic missense variants in CAPN3 explored systematically by minigene functional assay. *Hum Mutat*. 2020;41(10):1797-810.
71. Torriano S, Erkilic N, Faugere V, Damodar K, Hamel CP, Roux AF, et al. Pathogenicity of a novel missense variant associated with choroideremia and its impact on gene replacement therapy. *Hum Mol Genet*. 2017;26(18):3573-84.
72. Chen XM, Guo K, Li H, Lu QF, Yang C, Yu Y, et al. A novel mutation KCNQ1p.Thr312del is responsible for long QT syndrome type 1. *Heart Vessels*. 2019;34(1):177-88.
73. Palmer EE, Carroll R, Shaw M, Kumar R, Minoche AE, Leffler M, et al. RLIM Is a Candidate Dosage-Sensitive Gene for Individuals with Varying Duplications of Xq13, Intellectual Disability, and Distinct Facial Features. *Am J Hum Genet*. 2020;107(6):1157-69.
74. Liu N, Xiao J, Gijavanekar C, Pappan KL, Grinton KE, Shayota BJ, et al. Comparison of Untargeted Metabolomic Profiling vs Traditional Metabolic Screening to Identify Inborn Errors of Metabolism. *JAMA Netw Open*. 2021;4(7):e2114155.
75. Alston CL, Stenton SL, Hudson G, Prokisch H, Taylor RW. The genetics of mitochondrial disease: dissecting mitochondrial pathology using multi-omic pipelines. *J Pathol*. 2021;254(4):430-42.
76. Chantada-Vazquez MDP, Bravo SB, Barbosa-Gouveia S, Alvarez JV, Couce ML. Proteomics in Inherited Metabolic Disorders. *Int J Mol Sci*. 2022;23(23).

77. Vangipuram M, Ting D, Kim S, Diaz R, Schule B. Skin punch biopsy explant culture for derivation of primary human fibroblasts. *J Vis Exp*. 2013(77):e3779.
78. Li Y, Polak U, Clark AD, Bhalla AD, Chen YY, Li J, et al. Establishment and Maintenance of Primary Fibroblast Repositories for Rare Diseases-Friedreich's Ataxia Example. *Biopreserv Biobank*. 2016;14(4):324-9.
79. Cordone V, Pecorelli A, Amicarelli F, Hayek J, Valacchi G. The complexity of Rett syndrome models: Primary fibroblasts as a disease-in-a-dish reliable approach. *Drug Discovery Today: Disease Models*. 2020;31:11-9.
80. Claeys L, Bravenboer N, Eekhoff EMW, Micha D. Human Fibroblasts as a Model for the Study of Bone Disorders. *Front Endocrinol (Lausanne)*. 2020;11:394.
81. Fiorillo C, Moro F, Yi J, Weil S, Brisca G, Astrea G, et al. Novel dynein DYNC1H1 neck and motor domain mutations link distal spinal muscular atrophy and abnormal cortical development. *Hum Mutat*. 2014;35(3):298-302.
82. Seco-Cervera M, Spis M, García-Giménez JL, Ibañez-Cabellos JS, Velázquez-Ledesma A, Esmoris I, et al. Oxidative stress and antioxidant response in fibroblasts from Werner and atypical Werner syndromes. *Aging (Albany NY)*. 2014;6(3):231-45.
83. Mirabelli P, Coppola L, Salvatore M. Cancer Cell Lines Are Useful Model Systems for Medical Research. *Cancers (Basel)*. 2019;11(8).
84. Anderson RH, Francis KR. Modeling rare diseases with induced pluripotent stem cell technology. *Mol Cell Probes*. 2018;40:52-9.
85. Liu GH, Barkho BZ, Ruiz S, Diep D, Qu J, Yang SL, et al. Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome. *Nature*. 2011;472(7342):221-5.
86. Li H, Yang Y, Hong W, Huang M, Wu M, Zhao X. Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal Transduct Target Ther*. 2020;5(1):1.
87. Dow LE. Modeling Disease In Vivo With CRISPR/Cas9. *Trends Mol Med*. 2015;21(10):609-21.
88. Muller M, Lee CM, Gasiunas G, Davis TH, Cradick TJ, Siksnys V, et al. *Streptococcus thermophilus* CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther*. 2016;24(3):636-44.
89. Li T, Yang Y, Qi H, Cui W, Zhang L, Fu X, et al. CRISPR/Cas9 therapeutics: progress and prospects. *Signal Transduct Target Ther*. 2023;8(1):36.

90. Ward T, Tai W, Morton S, Impens F, Van Damme P, Van Haver D, et al. Mechanisms of Congenital Heart Disease Caused by NAA15 Haploinsufficiency. *Circ Res.* 2021;128(8):1156-69.
91. Bustos F, Segarra-Fas A, Chaugule VK, Brandenburg L, Branigan E, Toth R, et al. RNF12 X-Linked Intellectual Disability Mutations Disrupt E3 Ligase Activity and Neural Differentiation. *Cell Rep.* 2018;23(6):1599-611.
92. Li Y, Glass Z, Huang M, Chen ZY, Xu Q. Ex vivo cell-based CRISPR/Cas9 genome editing for therapeutic applications. *Biomaterials.* 2020;234:119711.
93. Barker. Phenol-Chloroform Isoamyl Alcohol (PCI) method for DNA Extraction. 1998.
94. 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(1):134.
95. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003;31(13):3784-8.
96. Frederick M. Ausubel RB, Robert E. Kingston, David D. Moore, J.G. Seidman, John A. Smith, Kevin Struhl. *Current Protocols in Molecular Biology*: John Wiley & Sons Inc; 2003.
97. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods.* 2012;9(7):676-82.
98. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010;26(5):589-95.
99. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-8.
100. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
101. 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 Research.* 2018;46(W1):W545-W53.
102. Andrews S. FASTQC. A quality control tool for high throughput sequence data 2010 [
103. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.

104. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
105. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923-30.
106. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014;15(12):550.
107. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*. 2012;16(5):284-7.
108. Arasimavičius J, Dilytė E, Fau - Utkus A, Utkus A, Fau - Preikšaitienė E, Preikšaitienė E. A Familial 4q12 Deletion Involving KIT Gene Causes Piebaldism. (1847-6538 (Electronic)).
109. Hoo JJ, Haslam RH, van Orman C. Tentative assignment of piebald trait gene to chromosome band 4q12. *Hum Genet*. 1986;73(3):230-1.
110. Giebel LB, Spritz RA. Mutation of the KIT (mast/stem cell growth factor receptor) protooncogene in human piebaldism. *Proc Natl Acad Sci U S A*. 1991;88(19):8696-9.
111. Funderburk SJ, Crandall BF. Dominant piebald trait in a retarded child with a reciprocal translocation and small intercalary deletion. *Am J Hum Genet*. 1974;26(6):715-22.
112. Yamamoto Y, Nishimoto H, Ikemoto S. Interstitial deletion of the proximal long arm of chromosome 4 associated with father-child incompatibility within the Gc-system: probable reduced gene dosage effect and partial piebald trait. *Am J Med Genet*. 1989;32(4):520-3.
113. Bondanza S, Bellini M, Roversi G, Raskovic D, Maurelli R, Paionni E, et al. Piebald trait: implication of kit mutation on in vitro melanocyte survival and on the clinical application of cultured epidermal autografts. *J Invest Dermatol*. 2007;127(3):676-86.
114. Roskoski R, Jr. Signaling by Kit protein-tyrosine kinase--the stem cell factor receptor. *Biochem Biophys Res Commun*. 2005;337(1):1-13.
115. Oiso N, Fukai K, Kawada A, Suzuki T. Piebaldism. *J Dermatol*. 2013;40(5):330-5.
116. Maldžienė Ž A-O, Vaitėnienė EM, Aleksūnienė B, Utkus A, Preikšaitienė E. A case report of familial 4q13.3 microdeletion in three individuals with syndromic intellectual disability. (1755-8794 (Electronic)).
117. Quintela I, Barros F, Fernandez-Prieto M, Martinez-Regueiro R, Castro-Gago M, Carracedo A, et al. Interstitial microdeletions including the chromosome band 4q13.2 and the UBA6 gene as possible causes of

- intellectual disability and behavior disorder. *Am J Med Genet A*. 2015;167a(12):3113-20.
118. Apte SS. A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: functions and mechanisms. *J Biol Chem*. 2009;284(46):31493-7.
 119. Le Goff C, Somerville RP, Kesteloot F, Powell K, Birk DE, Colige AC, et al. Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis. *Development*. 2006;133(8):1587-96.
 120. Hou SC, Chan LW, Chou YC, Su CY, Chen X, Shih YL, et al. Ankrd17, an ubiquitously expressed ankyrin factor, is essential for the vascular integrity during embryogenesis. *FEBS Lett*. 2009;583(17):2765-71.
 121. Deng M, Li F, Ballif BA, Li S, Chen X, Guo L, et al. Identification and functional analysis of a novel cyclin e/cdk2 substrate ankrd17. *J Biol Chem*. 2009;284(12):7875-88.
 122. Chopra M, McEntagart M, Clayton-Smith J, Platzer K, Shukla A, Girisha KM, et al. Heterozygous *ANKRD17* loss-of-function variants cause a syndrome with intellectual disability, speech delay, and dysmorphism. *The American Journal of Human Genetics*. 2021;108(6):1138-50.
 123. Siavriene E, Petraityte G, Mikstiene V, Maldziene Z, Sasnauskiene A, Zitkute V, et al. Molecular and Functional Characterisation of a Novel Intragenic 12q24.21 Deletion Resulting in MED13L Haploinsufficiency Syndrome. *Medicina (Kaunas)*. 2023;59(7).
 124. 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(2):129-36.
 125. Yi Z, Zhang Y, Song Z, Pan H, Yang C, Li F, et al. Report of a de novo c.2605C > T (p.Pro869Ser) change in the MED13L gene and review of the literature for MED13L-related intellectual disability. *Ital J Pediatr*. 2020;46(1):95.
 126. Soutourina J. Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology*. 2018;19(4):262-74.
 127. Utami KH, Winata CI Fau - Hillmer AM, Hillmer Am Fau - Aksoy I, Aksoy I Fau - Long HT, Long Ht Fau - Liany H, Liany H Fau - Chew EGY, et al. Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. (1098-1004 (Electronic)).

128. Hamada N, Iwamoto I, Nagata KA-O. MED13L and its disease-associated variants influence the dendritic development of cerebral cortical neurons in the mammalian brain. (1471-4159 (Electronic)).
129. Angus SP, Nevins JR. A role for Mediator complex subunit MED13L in Rb/E2F-induced growth arrest. *Oncogene*. 2012;31(44):4709-17.
130. Schiano C, Luongo L, Maione S, Napoli C. Mediator complex in neurological disease. *Life Sci*. 2023;329:121986.
131. Siavriene E, Preiksaitiene E, Maldziene Z, Mikstiene V, Rancelis T, Ambrozaityte L, et al. A de novo 13q31.3 microduplication encompassing the miR-17 ~ 92 cluster results in features mirroring those associated with Feingold syndrome 2. *Gene*. 2020;753:144816.
132. He L, Thomson Jm Fau - Hemann MT, Hemann Mt Fau - Hernando-Monge E, Hernando-Monge E Fau - Mu D, Mu D Fau - Goodson S, Goodson S Fau - Powers S, et al. A microRNA polycistron as a potential human oncogene. (1476-4687 (Electronic)).
133. 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. *Nature Genetics*. 2011;43(10):1026-30.
134. Kannu P, Campos-Xavier Ab Fau - Hull D, Hull D Fau - Martinet D, Martinet D Fau - Ballhausen D, Ballhausen D Fau - Bonafé L, Bonafé L. Post-axial polydactyly type A2, overgrowth and autistic traits associated with a chromosome 13q31.3 microduplication encompassing miR-17-92 and GPC5. (1878-0849 (Electronic)).
135. Hemmat M, Rumple 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. (1755-8166 (Print)).
136. Maldžienė Ž, Bulanovaitė E, Aleksūnienė B, Utkus A, Preiksaitiene E. 16p13.11-p12.3 Microdeletion identified in a patient with sagittal craniosynostosis and developmental delay. (1473-5717 (Electronic)).
137. Nagamani SCS, Erez A, Bader P, Lalani SR, Scott DA, Scaglia F, et al. Phenotypic manifestations of copy number variation in chromosome 16p13.11. *European Journal of Human Genetics*. 2011;19(3):280-6.
138. Taieb M, Ghannoum D, Barré L, Ouzzine M. Xylosyltransferase I mediates the synthesis of proteoglycans with long glycosaminoglycan chains and controls chondrocyte hypertrophy and collagen fibers organization of in the growth plate. *Cell Death & Disease*. 2023;14(6):355.
139. van Koningsbruggen S, Knoester H, Bakx R, Mook O, Knegt L, Cobben JM. Complete and partial XYLT1 deletion in a patient with neonatal short limb skeletal dysplasia. *Am J Med Genet A*. 2016;170A(2):510-4.

140. Bertelsen B, Nazaryan-Petersen L, Sun W, Mehrjouy MM, Xie G, Chen W, et al. A germline chromothripsis event stably segregating in 11 individuals through three generations. *Genetics in Medicine*. 2016;18(5):494-500.
141. Middelkamp S, van Heesch S, Braat AK, de Ligt J, van Iterson M, Simonis M, et al. Molecular dissection of germline chromothripsis in a developmental context using patient-derived iPS cells. (1756-994X (Electronic)).
142. Maher CA, Wilson RK. Chromothripsis and human disease: piecing together the shattering process. (1097-4172 (Electronic)).
143. Kloosterman WP, Guryev V Fau - van Roosmalen M, van Roosmalen M Fau - Duran KJ, Duran KJ Fau - de Bruijn E, de Bruijn E Fau - Bakker SCM, Bakker Sc Fau - Letteboer T, et al. Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. (1460-2083 (Electronic)).
144. Liu P, Erez A Fau - Nagamani SCS, Nagamani Sc Fau - Dhar SU, Dhar Su Fau - Kołodziejaska KE, Kołodziejaska Ke Fau - Dharmadhikari AV, Dharmadhikari Av Fau - Cooper ML, et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. (1097-4172 (Electronic)).
145. Arya P, Hodge JC, Matlock PA, Vance GH, Breman AM. Two Patients with Complex Rearrangements Suggestive of Germline Chromoanagenesis. *Cytogenetic and Genome Research*. 2021;160(11-12):671-9.
146. Dias CM, Punetha J, Zheng C, Mazaheri N, Rad A, Efthymiou S, et al. Homozygous Missense Variants in NTNG2, Encoding a Presynaptic Netrin-G2 Adhesion Protein, Lead to a Distinct Neurodevelopmental Disorder. *Am J Hum Genet*. 2019;105(5):1048-56.
147. Pan Y, Liu G, Fang M, Shen L, Wang L, Han Y, et al. Abnormal expression of netrin-G2 in temporal lobe epilepsy neurons in humans and a rat model. *Exp Neurol*. 2010;224(2):340-6.
148. Ostendorff HP, Bossenz M, Mincheva A, Copeland NG, Gilbert DJ, Jenkins NA, et al. Functional Characterization of the Gene Encoding RLIM, the Corepressor of LIM Homeodomain Factors. *Genomics*. 2000;69(1):120-30.
149. Zhang L, Huang H, Zhou F, Schimmel J, Pardo CG, Zhang T, et al. RNF12 controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting Smad7 for degradation. *Mol Cell*. 2012;46(5):650-61.
150. Gontan C, Mira-Bontenbal H, Magaraki A, Dupont C, Barakat TS, Rentmeester E, et al. REX1 is the critical target of RNF12 in imprinted X

- chromosome inactivation in mice. *Nature Communications*. 2018;9(1):4752.
151. Bertelsen B, Tümer Z, Ravn K. Three new loci for determining x chromosome inactivation patterns. *J Mol Diagn*. 2011;13(5):537-40.
 152. Middleton AJ, Zhu J, Day CL. The RING Domain of RING Finger 12 Efficiently Builds Degradative Ubiquitin Chains. *J Mol Biol*. 2020;432(13):3790-801.
 153. Liszczak G, Goldberg JM, Foy H, Petersson EJ, Arnesen T, Marmorstein R. Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nat Struct Mol Biol*. 2013;20(9):1098-105.
 154. Arnesen T, Van Damme P, Polevoda B, Helsens K, Evjenth R, Colaert N, et al. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc Natl Acad Sci U S A*. 2009;106(20):8157-62.
 155. Lyon GJ. From Molecular Understanding to Organismal Biology of N-Terminal Acetyltransferases. *Structure*. 2019;27(7):1053-5.
 156. Cheng H, Dharmadhikari AV, Varland S, Ma N, Domingo D, Kleyner R, et al. Truncating Variants in NAA15 Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies. *Am J Hum Genet*. 2018;102(5):985-94.
 157. Friedrich K, Lee L, Leistritz DF, Nürnberg G, Saha B, Hisama FM, et al. WRN mutations in Werner syndrome patients: genomic rearrangements, unusual intronic mutations and ethnic-specific alterations. *Hum Genet*. 2010;128(1):103-11.
 158. Oshima J, Sidorova JM, Monnat RJ, Jr. Werner syndrome: Clinical features, pathogenesis and potential therapeutic interventions. *Ageing Res Rev*. 2017;33:105-14.
 159. Grandori C, Wu KJ, Fernandez P, Ngouenet C, Grim J, Clurman BE, et al. Werner syndrome protein limits MYC-induced cellular senescence. *Genes Dev*. 2003;17(13):1569-74.
 160. Tam H-Y, Liu J, Yiu T-C, Leung AO-W, Li C, Gu S, et al. Amelioration of premature aging in Werner syndrome stem cells by targeting SHIP/AKT pathway. *Cell & Bioscience*. 2025;15(1):10.
 161. Gray MD, Shen JC, Kamath-Loeb AS, Blank A, Sopher BL, Martin GM, et al. The Werner syndrome protein is a DNA helicase. *Nat Genet*. 1997;17(1):100-3.
 162. Kang SM, Yoon MH, Lee SJ, Ahn J, Yi SA, Nam KH, et al. Human WRN is an intrinsic inhibitor of progerin, abnormal splicing product of lamin A. *Sci Rep*. 2021;11(1):9122.

163. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153(6):1194-217.
164. Valieva Y, Ivanova E, Fayzullin A, Kurkov A, Igrunkova A. Senescence-Associated β -Galactosidase Detection in Pathology. *Diagnostics (Basel)*. 2022;12(10).
165. Szczałuba K, Szymańska K, Rydzanicz M, Ciara E, Walczak A, Piekutowska-Abramczuk D, et al. A de novo loss-of-function DYNC1H1 mutation in a patient with parkinsonian features and a favourable response to levodopa. *Clin Genet*. 2018;93(5):1107-8.
166. Di Donato N, Timms AE, Aldinger KA, Mirzaa GM, Bennett JT, Collins S, et al. Analysis of 17 genes detects mutations in 81% of 811 patients with lissencephaly. *Genet Med*. 2018;20(11):1354-64.
167. Fang YL, Li N, Zhi XF, Zheng J, Liu Y, Pu LJ, et al. Discovery of specific mutations in spinal muscular atrophy patients by next-generation sequencing. *Neurol Sci*. 2021;42(5):1827-33.
168. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet*. 2016;17(1):19-32.
169. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet*. 2018;59(3):253-68.
170. Reck-Peterson SL, Redwine WB, Vale RD, Carter AP. The cytoplasmic dynein transport machinery and its many cargoes. *Nat Rev Mol Cell Biol*. 2018;19(6):382-98.
171. Palmer KJ, Hughes H, Stephens DJ. Specificity of cytoplasmic dynein subunits in discrete membrane-trafficking steps. *Mol Biol Cell*. 2009;20(12):2885-99.
172. Hirokawa N, Noda Y, Tanaka Y, Niwa S. Kinesin superfamily motor proteins and intracellular transport. *Nat Rev Mol Cell Biol*. 2009;10(10):682-96.
173. Thyberg J, Moskalewski S. Relationship between the Golgi complex and microtubules enriched in dephosphorylated or acetylated alpha-tubulin: studies on cells recovering from nocodazole and cells in the terminal phase of cytokinesis. *Cell Tissue Res*. 1993;273(3):457-66.
174. Pfister KK, Shah PR, Hummerich H, Russ A, Cotton J, Annuar AA, et al. Genetic analysis of the cytoplasmic dynein subunit families. *PLoS Genet*. 2006;2(1):e1.
175. Xu Y, Takeda S, Nakata T, Noda Y, Tanaka Y, Hirokawa N. Role of KIFC3 motor protein in Golgi positioning and integration. *J Cell Biol*. 2002;158(2):293-303.

176. She ZY, Pan MY, Tan FQ, Yang WX. Minus end-directed kinesin-14 KIFC1 regulates the positioning and architecture of the Golgi apparatus. *Oncotarget*. 2017;8(22):36469-83.
177. Koussounadis A, Langdon SP, Um IH, Harrison DJ, Smith VA. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci Rep*. 2015;5:10775.
178. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*. 2016;165(3):535-50.
179. Perl K, Ushakov K, Pozniak Y, Yizhar-Barnea O, Bhonker Y, Shivatzki S, et al. Reduced changes in protein compared to mRNA levels across non-proliferating tissues. *BMC Genomics*. 2017;18(1):305.
180. Price S, Shaw PA, Seitz A, Joshi G, Davis J, Niemela JE, et al. Natural history of autoimmune lymphoproliferative syndrome associated with FAS gene mutations. *Blood*. 2014;123(13):1989-99.
181. Strasser A, Jost PJ, Nagata S. The many roles of FAS receptor signaling in the immune system. *Immunity*. 2009;30(2):180-92.
182. Chang DW, Xing Z, Pan Y, Algeciras-Schimmich A, Barnhart BC, Yaish-Ohad S, et al. c-FLIP(L) is a dual function regulator for caspase-8 activation and CD95-mediated apoptosis. *Embo j*. 2002;21(14):3704-14.
183. Santiago B, Galindo M, Palao G, Pablos JL. Intracellular regulation of Fas-induced apoptosis in human fibroblasts by extracellular factors and cycloheximide. *J Immunol*. 2004;172(1):560-6.
184. Fu Q, Fu TM, Cruz AC, Sengupta P, Thomas SK, Wang S, et al. Structural Basis and Functional Role of Intramembrane Trimerization of the Fas/CD95 Death Receptor. *Mol Cell*. 2016;61(4):602-13.
185. Dodi AE, Ajayi IO, Chang C, Beard M, Ashley SL, Huang SK, et al. Regulation of fibroblast Fas expression by soluble and mechanical pro-fibrotic stimuli. *Respir Res*. 2018;19(1):91.
186. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet*. 2018;19(2):81-92.
187. Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem*. 2005;280(14):13341-8.
188. Wang B, Weidenfeld J, Lu MM, Maika S, Kuziel WA, Morrissey EE, et al. Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation. *Development*. 2004;131(18):4477-87.
189. Yagi M, Kabata M, Tanaka A, Ukai T, Ohta S, Nakabayashi K, et al. Identification of distinct loci for de novo DNA methylation by DNMT3A

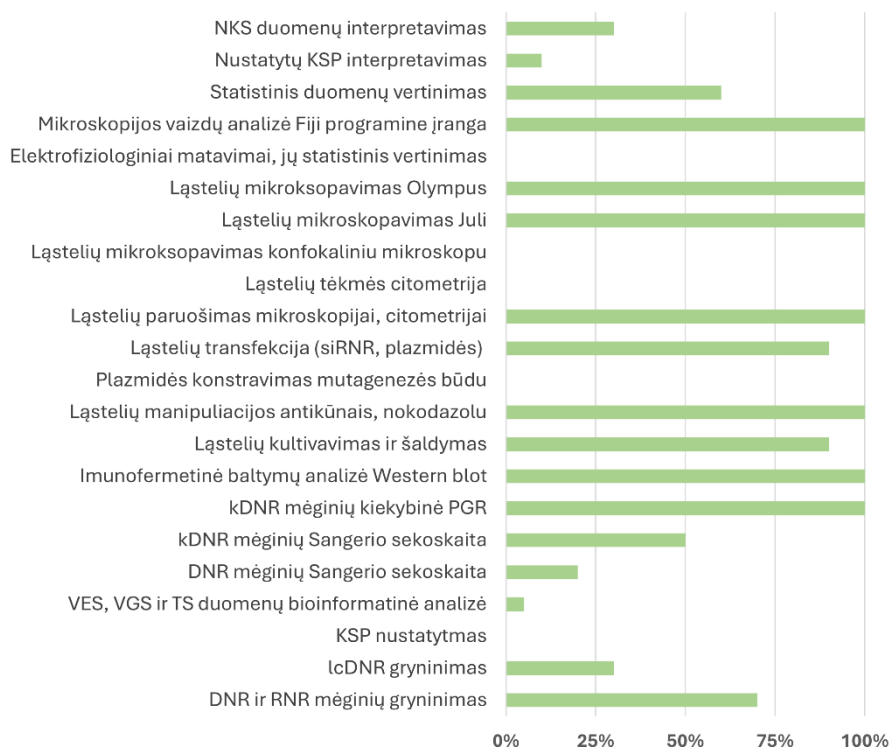
- and DNMT3B during mammalian development. *Nat Commun.* 2020;11(1):3199.
190. Benrick A, Jirholt P, Wernstedt I, Gustafsson M, Scheller J, Eriksson AL, et al. A non-conservative polymorphism in the IL-6 signal transducer (IL6ST)/gp130 is associated with myocardial infarction in a hypertensive population. *Regul Pept.* 2008;146(1-3):189-96.
 191. Boczek NJ, Ye D, Johnson EK, Wang W, Crotti L, Tester DJ, et al. Characterization of SEMA3A-encoded semaphorin as a naturally occurring Kv4.3 protein inhibitor and its contribution to Brugada syndrome. *Circ Res.* 2014;115(4):460-9.
 192. Suryawanshi H, Clancy R, Morozov P, Halushka MK, Buyon JP, Tuschl T. Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. *Cardiovascular Research.* 2019;116(8):1446-57.
 193. Behrmann A, Zhong D, Li L, Xie S, Mead M, Sabaeifard P, et al. Wnt16 Promotes Vascular Smooth Muscle Contractile Phenotype and Function via Taz (Wwtr1) Activation in Male LDLR^{-/-} Mice. *Endocrinology.* 2023;165(2).
 194. Cai Z, Luo W, Zhan H, Semenza GL. Hypoxia-inducible factor 1 is required for remote ischemic preconditioning of the heart. *Proceedings of the National Academy of Sciences.* 2013;110(43):17462-7.
 195. Smith AM, LaValle TA, Shinawi M, Ramakrishnan SM, Abel HJ, Hill CA, et al. Functional and epigenetic phenotypes of humans and mice with DNMT3A Overgrowth Syndrome. *Nature Communications.* 2021;12(1):4549.
 196. Cisneros-Villanueva M, Hidalgo-Pérez L, Rios-Romero M, Cedro-Tanda A, Ruiz-Villavicencio CA, Page K, et al. Cell-free DNA analysis in current cancer clinical trials: a review. *Br J Cancer.* 2022;126(3):391-400.
 197. Moss J, Magenheimer J, Neiman D, Zemmour H, Loyfer N, Korach A, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nature Communications.* 2018;9(1):5068.
 198. Katsman E, Orlanski S, Martignano F, Fox-Fisher I, Shemer R, Dor Y, et al. Detecting cell-of-origin and cancer-specific methylation features of cell-free DNA from Nanopore sequencing. *Genome Biol.* 2022;23(1):158.
 199. Ghosh S, Nunziato DA, Pitt GS. KCNQ1 assembly and function is blocked by long-QT syndrome mutations that disrupt interaction with calmodulin. *Circ Res.* 2006;98(8):1048-54.

200. Varghese A, TenBroek EM, Coles J, Sigg DC. Endogenous channels in HEK cells and potential roles in HCN ionic current measurements. *Progress in Biophysics and Molecular Biology*. 2006;90(1):26-37.
201. Peroz D, Rodriguez N, Choveau F, Baró I, Mérot J, Loussouarn G. Kv7.1 (KCNQ1) properties and channelopathies. *J Physiol*. 2008;586(7):1785-9.
202. Sachyani D, Dvir M, Strulovich R, Tria G, Tobelaim W, Peretz A, et al. Structural basis of a Kv7.1 potassium channel gating module: studies of the intracellular c-terminal domain in complex with calmodulin. *Structure*. 2014;22(11):1582-94.
203. Shamgar L, Ma L, Schmitt N, Haitin Y, Peretz A, Wiener R, et al. Calmodulin is essential for cardiac IKS channel gating and assembly: impaired function in long-QT mutations. *Circ Res*. 2006;98(8):1055-63.

PRIEDAI

1-as priedas.

Disertantės indėlis į šiame darbe aprašytus mokslinių tyrimų etapus



2-as priedas. PGR, kPGR ir Sangerio sekoskaitos PGR naudoti pradmenys

Genas	Tiesioginis	Atvirkštinis
<i>DYNC1H1 (DNR)</i>	5'-GACTTCGGTCTTCGGGCTTT-3'	5'-TGGTGAACAAAGGGTGCGAG-3'
<i>DYNC1H1 (kDNR)</i>	5'-ACCCCAACTACGACAAGACC-3'	5'-ATGCTCTTCCCCTTCCCGA-3'
<i>DYNC1H1 (kDNR, ilgojo fragmento)</i>	5'-AACTACGACAAGACCTCTGCC-3'	5'-TATCAGAATCTACGGGGAGAGGG-3'
<i>DNMT3A (kDNR)</i>	5'-GGCCCATTCGATCTGGTGAT-3'	5'-CTCACTTTGCTGAACCTGGCT-3'
<i>FAS (kDNR)</i>	5'-TTGGTGACCCGCTCAGTA-3'	5'-ACTTCTAAGCCATGTCCTTCATCA-3'
<i>CFLIP (kDNR)</i>	5'-CGGACTATAGAGTGCTGATGGC-3'	5'-CTTACTTGCCTCGGCCCAT-3'
<i>HPRT (kDNR)</i>	5'-TGGCGTCGTGATTAGTGATG-3'	5'-ACCCTTTCCAAATCCTCAGC-3'
<i>SLC9A6 (DNR)</i>	5'-CTTGCTATATTCCACGAGCTTCA-3'	5'-GCCTCATTCTTCCAAATGTTCC-3'
<i>SLC9A6 (kDNR)</i>	5'-CTTGCTATATTCCACGAGCTTCA-3'	5'-GCTTCAGCCAAGAGGAAGGT-3'
<i>PDE4D (DNR)</i>	5'-GAGCATGGCGGGGTATGG-3'	5'-TTATCTGGCGTGACTAGGGGA-3'
<i>PCCA (kDNR, del)</i>	5'-TCCGATGTGTATCCTGATGGC-3'	5'-TCTTGACCTTCTGCTACCGC-3'
<i>PCCA (kDNR)</i>	5'-GACAACAATGGCGGGGTCT-3'	5'-ACGACATCTTCTGCTGCCAA-3'
<i>CAPN3 (kDNR)</i>	5'-GCTTCGCCATCTACGAGGT-3'	5'-GCTTGTTTTGCCCTTTGCCCT-3'
<i>PIGN (DNR, del)</i>	5'-CCCTGTTTGGCTCAGGGATA-3'	5'-CTGGCCTGCCTACTTTGCT-3'
<i>PIGN (DNR)</i>	5'-AGGGAGCTGTTTTCTGCT-3'	5'-CTGGGCTGCCTCTTATGGAC-3'

3-ias priedas. PGR sąlygos, optimizuotos *DYNC1H1*, *SLC9A6* kDNR tyrimui

Reakcijos ciklas (iš viso 30 ciklų)	Temperatūra, laikas
Pradinė denatūracija	98 °C, 30 s
30 ciklų	
Denatūracija	98 °C, 2 s
Pradmenų prijungimas	64 °C, 10 s
Sintezė	72 °C, 15 s
Galutinė sintezė	72 °C, 60 s

4-as priedas. PGR mišinio sudėtis, optimizuota *DYNC1H1*, *SLC9A6* kDNR tyrimui

Medžiaga	Tūris (μl)
Dejonizuotas vanduo	4
<i>2x Phusion Flash PCR MasterMix</i>	10
Tiesioginis pradmuo (10 pmol/μl)	0,5
Atvirkštinis pradmuo (10 pmol/μl)	0,5
kDNR	5

5-as priedas. PGR produkto paruošimas kapiliarinei elektroforezei

PGR produktų valymas fermentais ExoI ir FastAP

PGR produktai valomi nuo reakcijoje likusių nukleotidų ir pradmenų. Šiam žingsniui naudojama Egzonukleazė I (angl. *Exonuclease I*; Thermo Fisher Scientific) ir FastAP™ *Thermosensitive Alkaline Phosphatase* (Thermo Fisher Scientific) šarminė fosfatazė pagal gamintojo rekomendacijas ir jo nurodytas reakcijos sąlygas.

Sekoskaitos PGR

Ciklinė sekoskaitos PGR atliekama naudojant *BigDye® Terminator v.3.1 Cycle Sequencing Kit* (Thermo Fisher Scientific) rinkinį vadovaujantis gamintojo protokolu.

Sekoskaitos PGR produkto valymas etanoliumi ir natrio acetatu

Sekoskaitos PGR produktams valyti pirmiausia paruošiamas reakcijos mišinys (1 lentelė).

1 lentelė. Etanolio ir natrio acetato mišinys. Tūris vienam mėginiui.

Medžiaga	Tūris (µl)
Dejonizuotas vanduo	14,5
96 % etanolis	62,5
Natrio acetatas	3
Bendras mišinio tūris	80

Mišinys išpilstomas po 80 µl į mėgintuvėlius su 7 µl sekoskaitos PGR produkto. Mėgintuvėliai inkubuojami tamsoje 8 min, o po inkubacijos centrifuguojami 7 min 13 000 aps./min. Dekantavus mėgintuvėlius, įpilama po 300 µl šalto etanolio (70 %) ir mėgintuvėliai purtomi 2 min. Mėgintuvėliai dar kartą centrifuguojami 7 min 13 000 aps./min ir pakartotinai dekantuojami. Dekantuoti mėginiai paliekami išdžiūti 20-30 min kambario temperatūroje arba 5 min +37 °C temperatūroje. Sekoskaitos produktams išdžiūvus, mėginiai užpilami 8 µl *Hi-Di™* formamido (*Hi-Di™ Formamide, ThermoFisher Scientific*, JAV) ir pipetavus pernešami į 96 šulinėlių sekoskaitos plokštelę (*Nerbe Plus, GmbH, Vokietija*). Plokštelę uždengus gumine tarpine (*Nerbe Plus, GmbH, Vokietija*), ji nucentrifuguojama 1 min 1 500 aps./min *Universal 32* (*Andreas Hettich GmbH & Co, Vokietija*) centrifugoje. Toliau atliekama kapiliarinė elektroforezė genetiniu analizatoriumi *ABI PRISM® 3130xl Genetic Analyzer* (*Hitachi, JAV*). Nuskaitytos sekos analizuojamos *Sequence Analysis*

v.5.1 (*Hitachi*, JAV) ir *Chromas* v.2.4.4 (*Technelysium Pty Ltd*) programomis. Naudojant *ExPASy LALIGN* bioinformacinį įrankį, sekos lyginamos su referentine tiriamo geno seka.

6-as priedas. Kiekybinei PGR naudoti TaqMan™ zondai

Žymens pavadinimas, taikiny	TaqMan™ žymens numeris	Gamintojas
DYNC1H1_1, <i>DYNC1H1</i> geno 15–16 egzonai	Hs01038546_m1	Thermo Fisher Scientific
DYNC1H1_2, <i>DYNC1H1</i> geno 51–52 egzonai	Hs00322286_m1	Thermo Fisher Scientific
KIFC1, <i>KIFC1</i> geno 3–4 egzonai	Hs00382558_m1	Thermo Fisher Scientific
KIFC3, <i>KIFC3</i> geno 9–10 egzonai	Hs00194304_m1	Thermo Fisher Scientific
GAPDH, <i>GAPDH</i> geno 8-as egzonas	Hs03929097_g1	Thermo Fisher Scientific
ACTB, <i>ACTB</i> geno 1-as egzonas	Hs99999903_m1	Thermo Fisher Scientific

7-as priedas. Western blot metodo protokolas

Baltymų išskyrimas

Atkeltos monosluoksniu kultivuojamos ląstelės centrifuguojamos 1000×g 7 min. 4°C. Po centrifugavimo mėginiai 1 kartą praplaunami šaltu PBS tirpalu ir dar kartą nusodinami centrifuguojant (1000×g 7 min. 4°C). Ląstelių nuosėdos lizuojamos RIPA ląstelių lizės buferiu (Sigma-Aldrich) ant ledo 30 min. Ląstelių lizatai centrifuguojami 14000×g 10 min 4°C. Baltymų koncentracija supernatante nustatoma BCA metodu.

Baltymų koncentracijos nustatymas

Baltymų koncentracijai naudotas rinkinys *Pierce™ BCA Protein Assay Kit* (*Thermo Fisher Scientific*, JAV). Matavimai atlikti pagal gamintojo pateiktą protokolą. Tirpalo sugertis matuojama spektrofotometru *Asys UVM340*, bangos ilgis 562 nm. Nežinomas baltymo kiekis nustatomas pagal kalibracinę kreivę, gautą su BCA reagentu inkubuojant 0 – 900 µg/mL JSA turintį tirpalą.

Baltymų elektroforezė denatūruojančiomis sąlygomis

Baltymų lizatas, kuriame yra 50 µg baltymo, praskiedžiamas PBS tirpalu iki 37,5 µl ir sumaišomas su 12,5 µl 4× baltymų užnešimo dažo (40% glicerolio, 8% NDS, 0,04% bromfenolio mėlio, 0,4 M DTT ir 250,4 mM Tris-HCl, pH 6,8). Mėginiai inkubuojami 5 min 95°C termomikseryje *Thermomixer comfort*.

NDS-poliakrilamidinis gelis paruošiamas ant polimerinio 10% arba 12% skiriamąjo gelio užpolimerinus 4% koncentruojamąjį gelį. Pilnai užsipolimerinęs gelis įstatomas į baltymų elektroforezės aparatą *Adjustable Height Dual Gel Unit*, pripildytą $1 \times$ NDS elektroforezės buferio. Į kraštinius gelyje suformuotus šulinėlius įnešamas baltymų dydžio standartas *Spectra Multicolor Broad Range Protein Ladder* (Thermo Fisher Scientific), kiti užpildomi paruoštais mėginiais. Elektroforezė vykdoma 1 val., esant 80 V įtampai, ir 3 val. (arba 5 val.), esant 120 V įtampai.

Baltymų pernešimas ant nitroceliuliozinės membranos

Išfrakcionuoti baltymai iš NDS-poliakrilamidinio gelio pernešami ant nitroceliuliozinės membranos (Thermo Fisher Scientific) naudojant pusiau sauso pernešimo aparatą V20-SDB. Ant pernešimo buferiu sudrėkintos anodinės aparato dalies dedami 3 pernešimo buferyje sudrėkinti *Whatman*TM 3 mm lapeliai. Ant jų dedama tokio paties dydžio pernešimo buferyje sudrėkinta nitroceliuliozinė membrana, o ant jos – gelis su išfrakcionuotais baltymais. Sluoksniavimas baigiamas uždėjus dar 3 pernešimo buferyje sudrėkintus *Whatman*TM 3mm lapelius. Išstūmus tarp sluoksnių esančius oro burbulus, uždedamas pernešimo buferiu sudrėkintas katodinis pernešimo aparato dangtis. Pernešimas vykdomas 1 val. esant 1 mA/cm^2 srovei.

Baltymų vaizdinimas

Nitroceliuliozinė membrana su perneštais baltymais blokuojama TBS/T buferyje turinčiame 5% JSA purtant 1 val. Po to membrana sukarpoma juostelėmis taip, kad jose būtų po vieną detektuojamą baltymą. Membranos juostelės inkubuojamos su atitinkamais pirminiais antikūnais, praskiestais TBS/T buferyje su 5% JSA purtant per naktį 4°C.

Po inkubacijos su pirminių antikūnų tirpalu membranos plaunamos TBS/T buferiu purtant 1 kartą 15 min ir 2 kartus po 5 min. Po plovimo membranos 2 val. inkubuojamos tamsoje purtant 4°C su antriniais antikūnais, praskiestais TBS/T buferyje su 5% JSA santykiu 1:5000. Baltymai nustatomi naudojant su krienų peroksidaze konjuguotą antrinį antikūną prieš triušio IgG (#31460, Thermo Fisher Scientific) arba su krienų peroksidaze konjuguotą antrinį antikūną prieš pelės IgG (#31430, Thermo Fisher Scientific). Po inkubacijos membranos plaunamos TBS/T buferiu purtant 1 kartą 15 min ir 2 kartus po 5 min. Prieš baltymų vaizdinimą membranos nuplaunamos dejonizuotu vandeniu ir išdžiovinamos.

Baltymai išryškinami membranas drėkinant *Pierce*TM *ECL Western Blotting Substrate* (Thermo Fisher Scientific) chemiliuminescenciniu mišiniu apie 1 min

ir jas eksponuojant ant fotojuostos *Medical X-ray Film General Purpose Blue* (Kodak). Baltymo juostelės intensyvumas įvertinamas ImageJ 1.52a programa.

8-as priedas. Moksliniame darbe naudoti pirminiai antikūnai

Antikūno pavadinimas	Gamintojas, numeris	Šeimininkas	Epitopas	Pagal gamintoją aptinkamas ties
anti-MED13L	<i>Bethyl, A302-420A-3</i>	Triušis	550–600 a.r.	~ 268 kDa
anti-NAA15	<i>ThermoFisher, PA5-84025</i>	Triušis	730–750 a.r.	95–130 kDa
anti-WRN	<i>SigmaAldrich, W0393-200UL</i>	Pelė	1074–1432 a.r.	160–250 kDa
anti-DYNC1H1	<i>EMD Milipore, ABT266</i>	Triušis	2557–2739 a.r.	~ 500 kDa
anti-GM130	<i>Invitrogen, 703794</i>	Triušis	69–157 a.r.	~ 140 kDa
anti-KIFC1	<i>Invitrogen, MA5-42583</i>	Triušis	1–100 a.r.	~ 75 kDa

9-as priedas. Naudotos duomenų bazės ir bioinformatiniai įrankiai

Pavadinimas	Nuoroda
<i>NCBI</i>	https://www.ncbi.nlm.nih.gov/
<i>ENSEMBL</i>	https://www.ensembl.org/index.html
<i>UCSC Genome Browser</i>	https://genome.ucsc.edu/
<i>GeneCards®</i>	https://www.genecards.org/
<i>GTExPortal</i>	https://www.gtexportal.org/home/
<i>OMIM®</i>	https://omim.org/
<i>UniProt</i>	https://www.uniprot.org/
<i>Orphanet</i>	https://www.orpha.net/
<i>St. Jude Cloud ProteinPaint</i>	https://viz.stjude.cloud/tools/proteinpaint
<i>Varsome</i>	https://varsome.com/
<i>Franklin</i>	https://franklin.genoox.com/clinical-db/home
<i>DECIPHER</i>	https://www.deciphergenomics.org/
<i>NCBI Primer-BLAST</i>	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
<i>gnomAD</i> (v.2.1.1 ir v.4.1.0)	https://gnomad.broadinstitute.org/
<i>1000 Genomes Project</i>	https://www.internationalgenome.org/
<i>MutationTaster</i>	https://www.mutationtaster.org/
<i>spliceAI</i>	https://spliceailookup.broadinstitute.org/
<i>Human Splicing Finder</i> v.3.1	http://www.umd.be/HSF3/

9-as priedas (Tęsinys). Naudotos duomenų bazės ir bioinformatiniai įrankiai

Pavadinimas	Nuoroda
<i>EMBOSS Transeq</i>	https://www.ebi.ac.uk/jdispatcher/st/emboss_transeq
<i>InterPro</i>	https://www.ebi.ac.uk/interpro/entry/pfam/#table
<i>VectorBuilder Sequence Alignment</i>	https://en.vectorbuilder.com/tool/sequence-alignment.html
<i>HGMD</i>	https://www.hgmd.cf.ac.uk/ac/index.php
<i>Clustal Omega</i>	https://www.ebi.ac.uk/jdispatcher/msa/clustalo

10-as priedas. Tiriamieji, kuriems atliekama NKS duomenų analizė.

Tiriamąojo (probando) numeris	Tiriami nariai	šėimos	Atlikti molekuliniai genetiniai tyrimai	Pagrindiniai tiriamąojo klinikiniai požymiai / įtariamas paveldimas sutrikimas
Tiriamasis 37 Tiriamasis 38	Mama, tėvas		Egzomo sekoskaita	Lėtinis pankreatitas, broliui autizmas, intelektinė negalia
Tiriamasis 39			Egzomo sekoskaita	Trombocitopenija, leukemija
Tiriamasis 40	Mama, tėvas		Egzomo sekoskaita	Psichomotorinės raidos atsilikimas, corpus callosum ir smegenėlių hipoplazija, prieširdžių pertvaros defektas, abipusis neurosensorinis klausos sutrikimas, kalbos sutrikimas, dismorfinės savybės
Tiriamasis 41 Tiriamasis 42	Mama, tėvas		Egzomo sekoskaita	Miomos, sesei miomos
Tiriamasis 43 Tiriamasis 44	Mama, tėvas		Egzomo sekoskaita	Atsilikusi raida, iškrypęs stuburas, asimetrinė krūtinės ląsta, sąnarių hipermobilumas, alkūnių hiperekstenzija, plokščiapėdystė, abipusis kriptorchizmas, kirkšninių ir babinė išvaržos, hidrocelė, dėmesio trūkumas, dismorfinės savybės
Tiriamasis 45	Mama, tėvas		Egzomo sekoskaita	Aklumas, necukrinis diabetas, nutukimas, aukštas ūgis, mokymosi sunkumai, elgesio problemos, stereotipiniai judesiai

10-as priedas (tęsinys). Tiriamieji, kuriems atliekama NKS duomenų analizė.

Tiriamąjo (probando) numeris, lytis	Tiriami šeimos nariai	Atlikti molekuliniai genetiniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai / įtariamas paveldimas sutrikimas
Tiriamasis 46	Mama, tėvas	Egzomo sekoskaita	Akių hipertelorizmas, įgimta ptozė, mikrotija, invertuoti speneliai, kriptorchizmas, rankų ir kojų kampomelija, suplokštėję distaliniai pirštų galai, trumpas platus, distalinis nykščio pirštakaulis, trumpi pirmieji kojų pirštai, kojų polidaktilija, odos sindaktilija, klubų kontraktūros, mišrus klausos sutrikimas, oligodontija, osteodisplazija
Tiriamasis 47	Mama, tėvas	Egzomo sekoskaita	Klausos sutrikimas
Tiriamasis 48		Egzomo sekoskaitos duomenų peranalizavimas	Ehlers-Danlos sindromas
Tiriamasis 49		Egzomo sekoskaitos duomenų peranalizavimas	Progerija
Tiriamasis 50	Mama, tėvas	Genomo sekoskaita	Balintos kavos dėmės
Tiriamasis 51	Mama tėvas, du broliai	Sangerio sekoskaita (segregacijai)	Rothmund-Thomson sindromas, inkstų hipoplazija
Tiriamasis 52	Mama, tėvas	Egzomo sekoskaita	Ehlers-Danlos sindromas, cistos žastikauliuose
Tiriamasis 53	Dukra	Egzomo sekoskaita	–

10-as priedas (tęsinys). Tiriamieji, kuriems atliekama NKS duomenų analizė.

Tiriamąojo (probando) numeris	Tiriami šeimos nariai	Atlikti molekuliniai genetiniai tyrimai	Pagrindiniai tiriamojo klinikiniai požymiai / įtariamas paveldimas sutrikimas
Tiriamasis 54	Mama, tėvas	Egzomo sekoskaita	–
Tiriamasis 55	Mama, tėvas, brolis, sesuo	Egzomo sekoskaita	Sunkus raidos atsilikimas
Tiriamasis 56	Mama, tėvas	Egzomo sekoskaita	Mikrocefalija, raidos atsilikimas
Tiriamasis 57	Mama, tėvas	Egzomo sekoskaita	Autizmas, raidos atsilikimas
Tiriamasis 58	Mama, tėvas	Genomo sekoskaita	Mielodisplazijos sindromas
Tiriamasis 59	Mama, tėvas	Sangerio sekoskaita (segregacijai)	–
Tiriamasis 60	Mama, tėvas	Egzomo sekoskaita	Vaisiaus inkstų agenezė
Tiriamasis 61	Mama, tėvas	Egzomo sekoskaita	Sunkus raidos atsilikimas
Tiriamasis 62	Mama, tėvas, sesuo	Genomo sekoskaita	Skeleto displazija, būdingi veido požymiai, VAN DEN ENDE-GUPTA sindromas
Tiriamasis 63	Mama, tėvas	Genomo sekoskaita	Vidutinė intelektinė negalia
Tiriamasis 64			
Tiriamasis 65	Mama, tėvas	Genomo sekoskaita	Raidos atsilikimas, cholestazė, epilepsija
Tiriamasis 66	Mama, tėvas	Egzomo sekoskaita	Raidos atsilikimas, mikrocefalija
Tiriamasis 67	Dukra	Egzomo sekoskaitos duomenų peranalizavimas	Proteinurija
Tiriamasis 68		Genomo sekoskaita	Porfirija
Tiriamasis 69		Sangerio sekoskaita	BTK geno išsamesnis tyrimas
Tiriamasis 70	Dukra	Egzomo sekoskaita	Inkstų ir akių anomalijos

SUMMARY OF DOCTORAL DISSERTATION

LIST OF ABBREVIATIONS

aCGH – array comparative genome hybridization
ACMG – American College of Medical Genetics and Genomics
AMP – Association of Molecular Pathology
cDNA – copy DNA
cfDNA – cell-free DNA
CGH – comparative genomic hybridization
ClinVar – NCBI database that collects data on the pathogenicity of genome variants
CNV – copy number variant
ENCODE – Encyclopedia of DNA Elements
gnomAD – Genome Aggregation database
HGMD – The Human Gene Mutation Database
iPSC – induced pluripotent stem cells
IRDiRC – The International Rare Diseases Research Consortium
lncRNA – long non-coding RNA
mtDNA – mitochondrial DNA
NCBI – National Center for Biotechnology Information
NGS – next-generation sequencing
NMD – nonsense-mediated mRNA decay
PCR – polymerase chain reaction
PI – propidium iodide
qPCR – quantitative PCR
SNV – single-nucleotide variant
SV – structural variant
VUS – variant of uncertain significance
WES – whole exome sequencing
WGS – whole genome sequencing
X-gal – (5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside) β galactosidase substrate

INTRODUCTION

Research study problem and relevance

Rare diseases are a broad group of health disorders characterized by heterogeneity in several aspects: geographical distribution, disease aetiology, and the variety and severity of phenotypic features. These diseases typically occur at an early age, often cause chronic health conditions, and can affect the nervous system, varying according to the individual's biological characteristics [1, 2]. The unifying feature for all rare diseases is their low prevalence. In Europe, a rare disease is defined as a disease that affects fewer than one in 2,000 people of the general population (Orphanet database). The low prevalence of these diseases is one of the primary reasons that has hampered the process of understanding these health conditions, as it has led to fewer and disparate efforts worldwide to investigate the causes and development of rare disorders, to develop new therapies and health care strategies, and to facilitate the “diagnostic odyssey” (a term for describing the long and challenging journey that many rare disease patients and their families undertake to receive an accurate diagnosis), which can often take about 5 years of a person's life [2]. However, it is estimated that around 30 million people in 48 European countries may currently have a rare disorder [2]. Thus, from the perspective of both the individual patient and society, rare diseases are a growing problem and a significant diagnostic challenge for both the global and Lithuanian healthcare systems, as they require a considerable amount of financial, time, and human resources. It is also important to emphasize that the long road to the confirmed diagnosis creates a significant emotional burden for the patient and their relatives.

Over the past two decades, the development and refinement of molecular genetics (such as next-generation sequencing (NGS)) and cytogenetics (such as comparative genomic hybridization (CGH)) technologies have fundamentally changed the sciences of biology, genetics, and medicine. This has undoubtedly led to a significant breakthrough in the field of rare diseases: it has stimulated research into their genetic characterization and attracted more interest not only from scientists but also from the general public. According to the literature analyzed during the writing of this dissertation, approximately 72–80% of rare disorders are caused by genetic (hereditary) factors [2, 3]. The majority of rare hereditary diseases are monogenic, i.e., they are determined by single-nucleotide variants (SNVs) of a specific gene, with minor deletions/insertions or structural variants (SVs) and copy number variations (CNVs) being less common [4, 5]. According to data from the International Rare Diseases Research Consortium (IRDiRC), approximately 4,500 genes have already been linked to rare inherited disorders; however, the impact of most variants identified in these genes on the

function of the encoded protein and related biological processes has not been studied. Therefore, a common problem is that when a unique and rare DNA sequence change is identified, its clinical significance or pathogenicity is unproven. This process is dynamic and is carried out in accordance with the guidelines described by the American College of Medical Genetics and Genomics (ACMG), which are based on bioinformatic analysis, accumulated knowledge about the genetic diversity of populations and diseases, and various scientific studies of biological processes [6]. Although NGS, when targeting the exome or genome, achieves a diagnostic efficiency of 24-50%, approximately 22% of identified variants are classified as variants of uncertain clinical significance [7-9]. This is most often due to a lack of knowledge about the functionality of the genetic region, as well as the molecular pathogenesis mechanism caused by the variant identified in the gene. It is worth emphasizing that a fundamentally sound molecular diagnosis can only be obtained by identifying DNA changes that belong to the class of pathogenic variants. Therefore, even in cases where likely pathogenic or likely benign variants are identified in known disease genes, additional functional studies aimed at thoroughly testing the effect of the DNA variant on molecular and biological processes are relevant and valuable.

Thus, the most advanced molecular genetics and cytogenetics technologies have enabled large-scale research into the genetic characteristics of rare hereditary diseases, as well as the refinement of the variants' classification. However, sequencing (exome, genome) and bioinformatic analysis of the results alone do not guarantee a definitive diagnosis. Combining genetic, bioinformatic, and functional analysis more often results in a confirmed molecular diagnosis, which is essential for effective genetic counseling. Understanding the pathogenesis mechanism of a disease not only allows us to confirm or rule out the diagnosis of a rare congenital disorder, but it is also important from an epistemological perspective. Research into rare genetic diseases can provide knowledge and insights into complex molecular processes that are relevant to more common or multifactorial diseases. A broader and deeper understanding of the processes underlying the development of rare inherited disorders is crucial for the development of new biological therapies for treating these pathologies.

Scientific novelty and significance

In this study, compared to similar scientific research in Lithuania, not only CGH, gene group, and exome sequencing or exome data reanalysis were used to determine the origin of suspected rare hereditary diseases, but also genome sequencing was performed for some of the affected individuals. The sequencing strategies employed enabled the identification of previously undescribed CNVs,

SVs, and SNVs in known disease genes. For CNV, SV, and SNV that had not been described in scientific literature and were potentially pathogenic or of uncertain clinical significance, the quantitative PCR method was employed, along with transcriptome sequencing, which is not widely used in the field of rare genetic diseases in Lithuania. These studies enabled a broader analysis of molecular processes and an investigation of pathogenesis mechanisms that are rarely described in the literature, which were determined by DNA changes in specific genes identified in this work. Different functional research strategies were used in the analysis of DNA sequence changes, employing methods that allow the study of various molecular processes in cells (confocal microscopy, flow cytometry, immunoblotting, gene expression reduction using siRNA) – to our knowledge, this has not been previously applied in the field of rare genetic diseases in Lithuania.

The genotypic, molecular, and/or functional characterization of selected gene sequence changes in patients with rare hereditary disorders, performed in this study, contributes to the understanding of the genetic diversity and molecular biology complexity of these disorders. The scientific research described demonstrates the potential of functional studies of rare hereditary diseases in terms of scientific output, which is not yet fully exploited in Lithuania. Fundamental research into the pathogenesis mechanism underlying a rare hereditary disease is an ongoing process that can form the basis for new scientific discoveries, the development of innovative treatment strategies, and targeted biological therapies. Such a research model can also be adapted for research into other rare diseases.

The study goal of the dissertation work was to evaluate the pathogenicity of DNA sequence variants identified in genes associated with rare diseases in affected individuals by performing bioinformatic, molecular, and/or functional studies at the levels of RNA, proteins, and cellular processes.

Towards this goal, the following specific tasks have been formulated:

1. To assess the pathogenicity of DNA sequence variants identified by next-generation sequencing using databases, literature, and in silico analysis, and to select candidate genes for functional studies.
2. Analyze genomic structural variants and their significance for identifying candidate genes.
3. Determine the significance of selected DNA sequence variants at the protein level using bioinformatic and molecular methods.
4. Evaluate the impact of selected DNA sequence variants at the level of cellular processes by performing molecular and/or functional studies.

Statements to be defended:

1. The process of selecting candidate genes for functional studies is heavily influenced by the assessment of the pathogenicity of variants found in the analyzed disease genes, the novelty of the DNA change, and fundamental knowledge, along with results from existing functional studies on the gene-encoded protein and its role in cells.
2. Genomic structural rearrangements and the clinical phenotype of the subject significantly narrow the search for candidate genes and may indicate the influence of regulatory genomic elements on the development of hereditary diseases.
3. It is recommended to use bioinformatic methods, such as protein structure modeling tools and conservation analysis algorithms, in addition to molecular methods for measuring protein quantity and function, since their combination helps uncover the biological significance of DNA variants and potential mechanisms of pathogenesis.
4. Molecular and functional studies of DNA variants, including expression analysis at the RNA and protein levels and in vitro studies of molecular processes involving altered proteins, offer the opportunity to assess unique aspects of the cellular pathogenesis mechanisms.

1. METHODS

1.1 Basis of the research study and the study's participants

The primary research for the dissertation was conducted as part of the scientific project “Integrative Research into the Genetic Causes of Rare Hereditary Diseases“ funded by the Research Council of Lithuania (acronym ATGC; S-MIP-21-15; project budget 149,994 EUR; project leader prof. Dr. Eglė Preikšaitienė). Part of this research was conducted using data from the scientific project “Identification of the Genetic Architecture of Intellectual Disabilities and Congenital Anomalies” (acronym INGENES; S-MIP-17-19/LSS-150000-1179; project budget 99,918 EUR; project leader prof. Dr. Eglė Preikšaitienė). Some of the INGENES project participants were invited to participate in the ATGC scientific project for additional or further molecular and/or functional studies. In addition, a small part of the dissertation research was funded by the Research Promotion Fund of Vilnius University after winning a competition for funding ideas from young scientists (project “Genome and transcriptome analysis in the pathogenesis of rare hereditary diseases”; MSF-JM-2/2020; project budget 9970.24 EUR; project leader PhD student Gunda Petraitytė). The research was conducted at the Department of Human and Medical Genetics, Faculty of

Medicine, Vilnius University, and the Department of Biochemistry and Molecular Biology, Faculty of Medicine, Vilnius University. Permission to conduct the research was obtained from the Vilnius Regional Biomedical Research Ethics Committee (INGENES No. 158200-17-962-469 and ATGC No. 2021/9-1373-849). The individuals included in the study have signed informed consent forms to participate in INGENES and ATGC biomedical research.

The participants included in the scientific study meet one main criterion: the primary participant, the proband, is suspected of having a rare genetic disorder. Based on this criterion, 70 probands and 125 family members with or without signs of the rare disease, i.e., if possible, both parents and siblings, were included in the study, totaling 195 subjects. Unrelated healthy individuals without signs of rare genetic pathologies (n=22) were included in the study for molecular and functional testing.

To meet the objectives of the dissertation, various molecular and/or functional experimental studies were conducted using different methods, which are described below.

1.2. DNA, cell-free DNA, and RNA purification

Genomic DNA from peripheral blood leukocytes was purified using the standard phenol-chloroform-isoamyl alcohol method [93]. A NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) was used to determine DNA concentration and purity.

Cell-free DNA (cfDNA) was purified from the blood plasma of individuals with the *DNMT3A* gene variant. The MagMAX™ Cell-Free DNA Isolation Kit (Applied Biosystems, USA) was used for purification, following the manufacturer's protocol. A NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) was used to determine DNA concentration and purity.

Most of the total RNA from the participants was purified from peripheral blood samples collected in Tempus™ Blood RNA Tubes (ThermoFisher Scientific, USA). Purification was performed using the Tempus Spin RNA Isolation Kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. RNA concentration and purity were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA).

The total RNA from some participants was purified from primary fibroblast cultures. For RNA purification, $1-3 \times 10^6$ cells were collected. The cells were collected by washing their monolayer with a 0.02% EDTA solution. Then, the cells were transferred to a 0.05% trypsin solution, and the collected cells were centrifuged (7 min, $1000 \times g$, 4°C). The settled cells were washed once with PBS solution and centrifuged again (7 min, $1000 \times g$, 4°C). Next, RNA purification

was performed using the RNeasy® Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. RNA concentration and purity were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA).

1.3 Synthesis of cDNA

Copy DNA (cDNA) was synthesized from total RNA samples using the High-Capacity RNA-to-cDNA Kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol, with the reverse transcription PCR (RT-PCR) conditions specified therein applied. 500 ng of RNA was used for cDNA synthesis. The ProFlex PCR System (Thermo Fisher Scientific, USA) or Mastercycler (Eppendorf, Germany) thermocycler heating blocks were used to perform RT-PCR.

1.4 Fibroblast culture cultivation and cryopreservation

The primary fibroblasts of the participants were obtained from skin biopsies. These cells were cultivated to form a monolayer in DMEM medium with 10% FBS (Fetal Bovine Serum; Gibco, UK), with the addition of 0.4% amphotericin B (Gibco, UK), or in AmnioMAX C-100 medium with 14.2% AmnioMAX C-100 medium supplement (Gibco, UK), with the addition of 0.4% amphotericin B. Cultured at 37°C, 5% CO₂, and 95% humidity. Cells were passaged two or three times per week when the culture reached 90-100% confluence. For transfer, the cells were first washed with a 0.02% EDTA solution, then transferred to a 0.05% trypsin solution, and diluted in a 1:2 ratio.

For long-term storage of fibroblasts in liquid nitrogen, a freezing medium consisting of 90% FBS and 10% DMSO (Dimethyl Sulfoxide; OriGen Biomedical, USA). Cells are frozen in 1 mL of freezing medium and stored in cryotubes (Thermo Fisher Scientific, USA).

1.5 PCR

PCR primers were designed using the publicly available Primer-BLAST tool [94]. The pairs of primers used to amplify DNA and cDNA fragments are listed in Appendix 2.

DNA and cDNA fragments were amplified using the PCR method with the 2× Phusion Flash High-Fidelity PCR MasterMix (Thermo Fisher Scientific, USA). The reaction mixture also included deionized water, primer pair, and DNA/cDNA sample. The PCR mixture was, in most cases, prepared according to the manufacturer's protocol; in some cases, the composition of the mix was optimized (changes are listed in Appendix 3). PCR conditions were set according

to the manufacturer's recommendations for the polymerase, with some optimizations made in certain cases (changes are indicated in Appendix 4). A ProFlex PCR System thermal cycler heating block (Thermo Fisher Scientific, USA) was used to perform the PCR.

Electrophoresis was used to evaluate the length, quantity, and specificity of the PCR products. A 1.5% agarose Top Vision Agarose (Thermo Fisher Scientific, USA) gel with 10 mg/ml ethidium bromide (EtBr) (Roth, Germany; 2–3 µl EtBr per 100 ml agarose gel), 3x loading dye DNA LoadingDye (Thermo Fisher Scientific, USA), and DNA molecular weight and size standard 100 bp GeneRuler (Thermo Fisher Scientific, USA). Electrophoresis was performed in an EC 105 horizontal electrophoresis system (E-C Apparatus Corporation, USA). The gel was visualized in a UV transilluminator (E.A.S.Y. 442K, Herolab, Germany) and photographed using the EasyWin 32 gel analysis software (Herolab, Germany).

1.6 Purification of PCR products from agarose gel

PCR products were purified from agarose gel in the case of the *DYNCH1* gene study using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol and in accordance with the specified safety recommendations.

1.7 Sanger sequencing

The preparation of PCR products for sequencing is described in detail in Appendix 5. Sanger sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. Capillary electrophoresis was performed using the ABI PRISM® 3130xl Genetic Analyzer (Hitachi, USA). The electropherograms were analyzed using the Sequence Analysis v.5.1 program (Thermo Fisher Scientific, USA), and the resulting electropherograms were visualized and analyzed using the Chromas v.2.4.4 program (Technelysium Pty Ltd). The ExpASy LALIGN [95] bioinformatic tool was used to compare the obtained sequences with reference gene sequences.

1.8 Quantitative PCR

Commercial oligonucleotide hydrolysis probes, TaqMan™ (Thermo Fisher Scientific, USA), were used to perform the quantitative PCR (qPCR) method. A list of probes used for specific and housekeeping genes is provided in Appendix 6. A TaqMan™ Universal PCR Master Mix (ThermoFisher Scientific, USA) was

used for the qPCR mixture. The reaction mixture was prepared, and the reaction conditions were set according to the manufacturer's protocol. The StepOne™ Real-Time PCR System (Applied Biosystems™, USA) was used to perform qPCR. The same instrument's software was used for qPCR data analysis. Differences in gene expression were determined using the standard $\Delta\Delta C_t$ calculation method. qPCR results were presented as fold changes in gene expression, normalized to the housekeeping gene *GAPDH* (or *ACTB*) and compared to the expression of the gene of interest in cells from healthy control individuals.

The efficiency of *CFLIP/CFLAR* gene expression reduction was verified using the qPCR method with Maxima SYBR Green/ROX qPCR MasterMix (2X) (Thermo Fisher Scientific, JAV) reagent according to the manufacturer's recommendations and protocol. The forward and reverse primers used in this reaction are presented in Appendix 2. qPCR results were normalized according to the housekeeping gene *HPRT*.

1.9 Next-generation sequencing

The genomic DNA samples of the participants were sent to the sequencing service provider *Novogene* for human exome and/or genome NGS. The Illumina NovaSeq 6000 platform was used for exome sequencing, performing paired-end sequencing of 150 bp fragments with 50× sequencing coverage. Genome sequencing was also performed using the Illumina NovaSeq 6000, which performed paired-end sequencing of 150 bp fragments with 30× sequencing coverage. The RNA samples from the participants were sent to the same sequencing service provider for transcriptome sequencing (target mRNA). The Illumina NovaSeq 6000 was also used for transcriptome sequencing, performing paired-end sequencing of 150 bp fragments, which generated ≥ 20 million paired reads per RNA sample.

1.10 Western blot analysis

Protein quantification was performed by Western blot according to Ausubel et al. (2003) [96]. The exact protocol used in the study is provided in Appendix 7. A list of primary antibodies used in the analysis is provided in Appendix 8. Proteins on the transferred membrane were detected using a secondary antibody conjugated with horseradish peroxidase against rabbit IgG (Thermo Fisher Scientific, USA) or against mouse IgG (Thermo Fisher Scientific, USA).

1.11 Cell transfection with small interfering RNAs

Small interfering RNAs (siRNAs) were used to suppress gene expression temporarily. To evaluate the transfection efficiency of these molecules using the qPCR method, cells were seeded in a 6-well plate at a density of 1×10^6 cells/ml, and for microscopic observation, they were seeded into a 24-well plate at a density of 4×10^4 cells/ml, corresponding to the volume of one well of an 8-well microscopy plate. Cells for the experiment were cultured in a medium without amphotericin B.

Lipofectamine™ RNAiMAX Transfection Reagent (Thermo Fisher Scientific, USA) was used for transfection. Transfection was performed according to the manufacturer's protocol. To temporarily suppress *KIFC1* protein-encoding gene expression in the cells, a siRNA specific for the *KIFC1* transcript (s7907; Thermo Fisher Scientific, USA) was used. To temporarily suppress the *KIFC3* protein-encoding gene expression in the cells, a siRNA specific for the *KIFC3* transcript (s7843; Thermo Fisher Scientific, USA) was used. In both cases, siRNA (#4404021; Thermo Fisher Scientific), which is not specific to any human transcript, was used for comparison.

1.12 Cell transfection with plasmids

For cell transfection with plasmids, HEK293T/17 (ATCC) cells were seeded into 24-well plates (Biolite 24-well Multidish, Thermo Scientific) at a density of 80,000 cells/ml the day before transfection. The reagent Lipofectamine® 3000 (Thermo Fisher Scientific) was used for transfection according to the manufacturer's recommendations and protocol. The plasmid pIRES_EGFP_GA_KCNE1 (Addgene, # 173160) encoding the KCNE1 protein and the plasmid pIRES(A7Mut)_CyOFP1_KCNQ1 (Addgene, #173162) encoding the KCNQ1 protein, donated by Al George, were transfected in different combinations (KCNQ1_WT; KCNQ1_MUT; KCNQ1_WT+KCNE1; KCNQ1_MUT+KCNE1) while maintaining the plasmid DNA concentration recommended by the transfection reagent manufacturer. After transfection, the cells were incubated (37°C, 5% CO₂) for 24–48 hours. Electrophysiological measurements were then performed.

1.13 Flow cytometry

For cytometry, cells were cultured in 12-well plates at a density of 40,000 cells/mL. For analysis by cytometer, the medium, along with the cells, was collected in a test tube. The cells were washed once with PBS buffer solution.

The cell suspension was centrifuged at 1000 rpm for 7 min at 4 °C. The supernatant was removed, and the remaining cells were stained with the Tali™ Apoptosis Kit - Annexin V Alexa Fluor™ 488 & Propidium Iodide (Thermo Fisher Scientific, USA) reagent according to the manufacturer's protocol and recommendations. Fluorescence was measured in at least 900 cells using a BDA FACSymphony™ A1 Cell Analyzer (BD Biosciences) flow cytometer and BD FACSDiva™ software. Data were analyzed using Flowing Software 2.5.1. Cells were classified into the following groups: annexin V and propidium iodide (PI) positive, annexin V positive and PI negative, annexin V negative and PI positive, and annexin V and PI negative. Among the samples, dying cells that fall into the annexin V positive and PI-negative cell group were evaluated.

1.14 Senescence of fibroblasts analysis

To determine fibroblasts' senescence, cells were seeded in 60 mm diameter dishes at a density of 5×10^5 cells/ml. The Senescence Cells Histochemical Staining Kit (Sigma-Aldrich, USA) was used to determine senescence. The study was conducted in accordance with the manufacturer's protocol and recommendations. After the specified incubation period, the cells were examined under an inverted microscope (Olympus CKX41, using a 10× objective lens, NA 0.25), photographing at least 10 random locations in a single culture dish.

1.15 Immunofluorescence microscopy

The Golgi complex was visualized using confocal microscopy after exposure of fibroblasts to nocodazole. Cells were seeded into 8-well plates (C8-1-N, Cellvis, Germany) at a density of 3×10^5 cells/ml. Twenty-four hours after seeding, the cells were treated with a 5 µg/ml concentration of nocodazole solution (Fisher Scientific, Germany). Cell cultures at different time points of nocodazole exposure were fixed with a 4% formaldehyde solution for 20 min, then washed three times for 5 min with PBS solution. The cell membranes were permeabilized by incubating them for 15 min with a 0.1% Triton X-100 solution, followed by three 5-minute washes with PBS solution. A blocking solution (2% BSA solution in PBS) was applied to the fixed and permeabilized cells and incubated for 60 min. The cells were then incubated overnight at 4°C with a blocking solution diluted (1:200) with primary antibody against DYNC1H1 or GM130 protein. After removing the primary antibody solution, the cells were washed three times for 5 min with PBS solution. After washing, the cells were incubated for one hour in the dark in blocking solution containing a 1:200 dilution of Alexa Fluor® 647-conjugated secondary antibody against rabbit IgG (A21244; Invitrogen, USA). After removing the secondary antibody solution, the cells were

washed three times for 5 min with PBS solution and then incubated for 15 min with DAPI dye diluted 1:1000 in PBS solution. After incubation, the cells were washed three times for 5 min with PBS solution. DYNC1H1 and GM130 immunofluorescence were recorded, and GM130 localization was evaluated using a Leica SP8 confocal microscope with a 63× objective lens and a numerical aperture (NA) of 1.4, at an excitation wavelength (λ_{ex}) of 604–644 nm and an emission wavelength (λ_{em}) of 672–712 nm. In the case of FAS receptor activation, apoptosis was evaluated using an inverted fluorescence microscope Olympus IX-81 with a 40× objective, NA 0.9, λ_{ex} = 480 nm, λ_{em} = 530 nm, and equipped with a Hamamatsu ORCA-FLASH 4.0LT camera. Microscopy images were analyzed and evaluated using the Fiji 2.15.0 software [97].

1.16 Electrophysiological measurements of cells using the patch-clamp method

Electrophysiological measurements using the patch-clamp method were performed 24–48 hours after transfection. Cells cultured on glass slides were transferred to the Nikon FN-S2N microscope, equipped with a recording camera (using a 40× objective, NA 0.80), which was used to select fluorescent cells.

Electrical currents were recorded at room temperature using a Multiclamp 700B (Molecular Devices) amplifier controlled by pCLAMP 10.5 (Molecular Devices) software. Recording microelectrodes (“patch”) were pulled from borosilicate glass capillaries using a Flaming/Brown micropipette (P-1000; Sutter Instrument Co.) with a resistance of 2–4 M Ω and filled with a special cell solution (mM): 140 KCl, 1 Na₂-ATP, 2 EGTA, 10 HEPES, 0.1 CaCl₂, 1 MgCl₂, and 10 glucose, pH 7.4. The extracellular solution consisted of the following (mM): 140 NaCl, 4 KCl, 2 CaCl₂, 1 MgCl₂, 10 HEPES, and 10 glucose, at pH 7.4.

The cell membrane current was recorded from a membrane potential of –80 mV to a depolarization potential of +30 mV, then in 2-second increments of 20 mV (up to +60 mV). The values obtained were calculated by normalizing the recorded currents according to cell capacitance (pA/pF).

1.17 Bioinformatic analysis

Bioinformatic analysis involves preparing primary exome, genome, and transcriptome data for interpretation and evaluation. This process also includes an initial assessment of the pathogenicity and impact of variants at both the RNA and protein levels, using in silico analysis. The databases and bioinformatics tools employed in the dissertation are listed in Appendix 9.

The primary exome and genome data were aligned to the human genome reference sequence provided in the UCSC (hg38) database using the BWA (Burrows-Wheeler Aligner) program [98]. The GATK [99] tool was used for the

NGS data analysis algorithm. The ANNOVAR v.2018Apr16 [100] and VarAFT v.2.16 [101] tools were used to annotate the identified variants.

The raw transcriptomic data were evaluated using FastQC v0.11.9 [102], and low-quality bases and adapter sequences were trimmed using Trimmomatic v0.39 [103]. The cleaned sequences were aligned to the human genome (GRCh38, Ensembl version 111) using STAR v2.7.11a [104]. Gene-level count calculations were performed using featureCounts v2.0.6 [105]. Differential gene expression analysis was performed using DESeq2 v1.42.0 [106]. In this analysis, proband 1 (P1) and proband 2 (P2) were compared with age-matched male controls (K), and the mother (M) was compared with age-matched female controls (K). Log2 expression change correction was performed using the lfcShrink function with the apeglm option in the DESeq2 tool. Altered gene expression was determined using the following threshold: FDR-corrected p-value < 0.1 and expression change > 1.5. Gene ontology (GO) term enrichment analysis was performed using clusterProfiler v4.10.1 [107]. All differentially expressed genes were included in the GO analysis. GO term enrichment analysis was performed using the enrichGO function in the clusterProfiler package. The parameters were configured to include all three main GO ontologies: biological process, cell component, and molecular function, specifying "ont = 'ALL'". The "pAdjustMethod" method was set to the Benjamini-Hochberg method. The "orgDb" parameter was set to "org.Hs.eg.db" to match the human gene ID ontology vocabulary. Other parameters were set to default values. GO terms were considered significant with a "pvalueCutoff" of less than 0.05. All statistical analyses were performed using R version 4.3.2.

1.18 Statistical data analysis

Experimental data were processed, and some graphs were plotted using SigmaPlot 14.0 (Systat Software), while others were plotted using MS Office Excel. The data presented in the graphs were obtained from 2–3 independent biological experiments. At least two technical replicates were used to evaluate relative gene expression in each experiment. The normal distribution of the data was checked using the Shapiro-Wilk normality test. If the data were normally distributed, parametric statistics were used – specifically, the t-test or the Kruskal-Wallis test – to compare differences. If the data were not normally distributed, non-parametric statistics were used – the Mann-Whitney U test. Gene expression, microscopy images, flow cytometry, and electrophysiological measurement data were considered statistically significant when the p-value ≤ 0.05.

2. RESULTS AND THEIR DISCUSSION

To identify genomic changes and, when possible, evaluate how these changes manifest at the molecular and/or functional level of cellular processes, it is crucial to plan the research strategy carefully, choose the most effective methods given the conditions, and consider potential research risks. The primary tasks of the experimental study presented in this work involve assessing whether the functional research objectives can be achieved in the fibroblast culture under investigation, as not all genes (or proteins) are expressed in these cells. It is also important to remember that the cellular process being studied and its regulation may vary between cell types (e.g., fibroblasts and nerve cells). Additionally, the individual genomic, epigenomic, and metabolomic characteristics of subjects influence the biological variation observed in quantitative (or qualitative) studies of patient and control cells. As a result, control subjects are selected based on sex and age. These specific conditions for functional studies influence both the choice of methods and the complexity and sophistication of their execution.

The aim of this study is to evaluate the pathogenicity of DNA sequence variants identified in genes associated with rare diseases in research participants, as well as their significance at the RNA, protein, and cellular process levels, through bioinformatic, molecular, and/or functional studies. First, likely pathogenic or pathogenic DNA sequence variants identified by NGS and Sanger sequencing, selected through detailed database, literature, and *in silico* analysis, were selected for further molecular and/or functional studies. Next, we describe CNV in the genome and candidate genes that may be related to the development of clinical symptoms. Finally, molecular and/or functional studies were conducted to assess the significance of the selected SNVs and small deletions in cellular processes and to investigate the potential mechanisms of pathogenesis.

2.1 Description of the research study participants and the selected candidate genes for experimental studies

During the research study, 70 participants (probands) were included. Of these, 36 (and, in some cases, their family members) underwent molecular, cytogenetic, and/or NGS testing to determine the genetic cause of the suspected rare hereditary condition. The project on which this research and data analysis are based is ongoing. Therefore, during and after the writing of the dissertation, NGS data analysis will be performed for the other 34 participants included in the study (their data are presented in Appendix 10).

Of the 36 individuals mentioned above, 64% (23/36) are male and 36% (13/36) are female (see Figure 2.1). Twenty (55% of subjects) were found to have

variants of uncertain significance, likely pathogenic, pathogenic SNVs, or small deletions/insertions in known disease genes (see Table 2.1 and Table 2.3). DNA changes were detected by Sanger sequencing in 2 participants (2/20), by gene panels in 2 participants (2/20), and by exome sequencing in the remaining individuals (16/20). DNA changes were detected in heterozygous, homozygous, and compound heterozygous states. SNVs account for 43,5% of all detected variants (10/23), splice site variants – 26% (6/23), and small deletions/insertions – 30,5% (7/23) (see Figure 2.1).

Six participants (17%) were found to have CNVs (see Tables 2.2 and 2.3), which determined their clinical phenotype. CNVs were detected in four individuals (4/6) using the SNP-CGH method, in one proband (1/6) using the aCGH method, and in 1 participant (1/6) using genome sequencing. In total, seven different CNVs were identified: two cases of duplication (2/7), four cases of deletion (4/7), and one case of chromothripsis (1/7) (see Figure 2.1). The latter is described in more detail in section 3.2.

After analyzing the exome sequencing data, no clinically significant DNA changes that could be the cause of the rare genetic phenotype were identified in 10 subjects (see Table 2.4 and Figure 2.1).

Of the 36 participants, additional molecular and/or functional studies were conducted during the dissertation on seven individuals with DNA changes in known disease genes that had not been previously described in the literature (see Table 3.3). The DNA variants in the *DYNC1H1*, *FAS*, *DNMT3A*, and *KCNQ1* genes that have been investigated through molecular and/or functional studies are described in more detail in section 3.3.

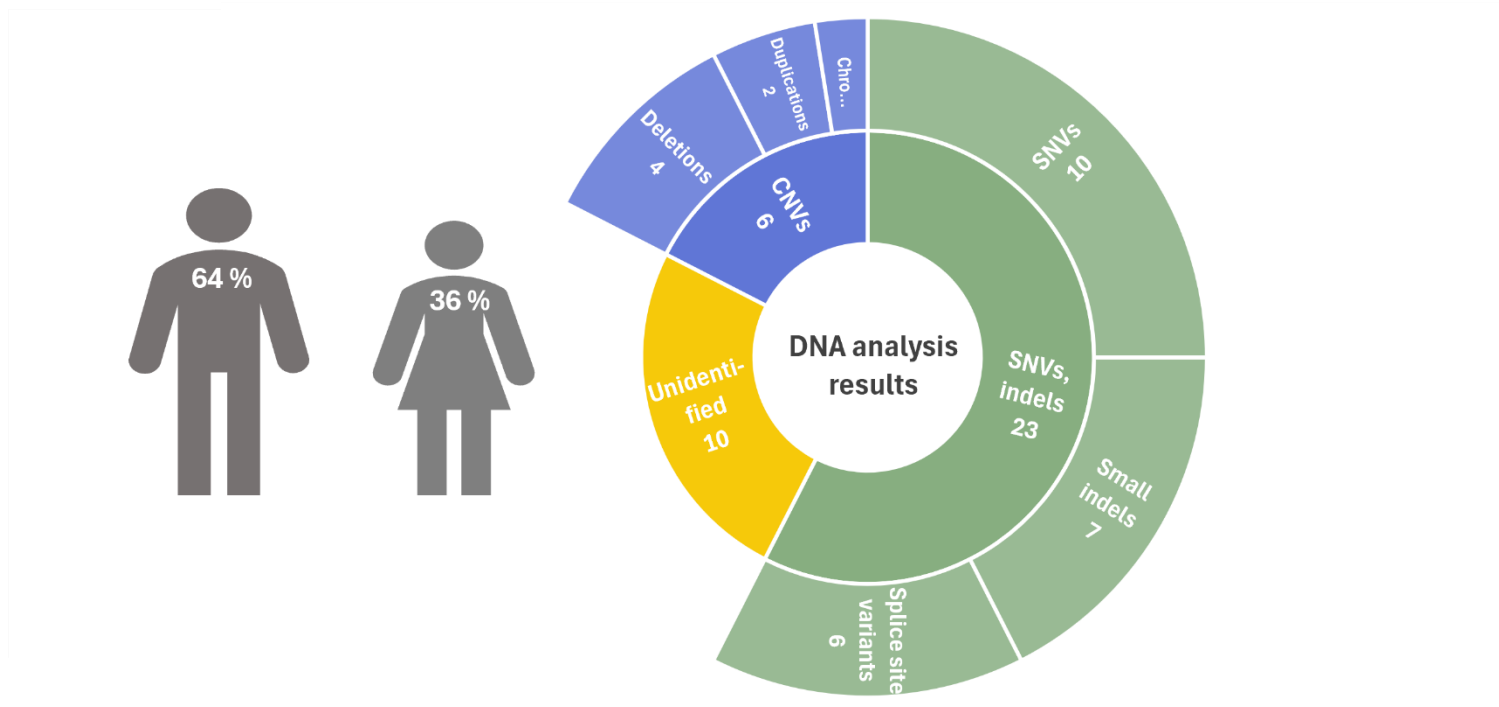


Figure 2.1. DNA variants identified and analyzed in the participants of the dissertation research. A total of 36 DNA samples from participants were studied during the dissertation research. Twenty-three SNVs or small indels and six CNVs were identified. No DNA changes were identified in the remaining individuals using molecular genetics analysis methods.

Table 2.1. Study individuals with DNA SNVs that cause a rare inherited disease.

Subject's (proband's) code, sex	Analyzed family members	Method of variant identification	Gene	Genotype, identified variant	Additional molecular tests performed	Main clinical features of the subject/rare inherited disorder
IN17-39-P, male	Mother, father	Sanger sequencing	<i>PDE4D</i> MIM#600129	Heterozygous NM_001104631.2:c.581G > C	-	Acrodysostosis (MIM#614613)
IN17-03-P1, female	Mother, father	Exome sequencing	<i>SLC27A4</i> MIM#604194	Hemizygous NM_005094.3:c.1528C>T	Sanger sequencing	Ichtiosis (MIM#608649), hearing impairment, and visual impairment
IN17-09-P1, male	Mother, father	Gene panel NGS, exome sequencing	<i>MAST1</i> MIM#612256	Heterozygous NM_014975:c.1549G>A	Sanger sequencing	Psychomotor developmental delay, ventriculomegaly, Dandy-Walker anomaly
IN17-19-P1, male	Brother	Sanger sequencing	<i>SLC6A8</i> MIM#300036	Hemizygous NM_:c.1248C>A	-	Intellectual disability
IN17-15-P1, female	Mother, father	Exome sequencing	<i>CHD7</i> MIM#608892	Heterozygous NC_000008.11(NM_017780.4):c.5535-1G>A	Sanger Sequencing (DNA, cDNA)	<i>CHARGE</i> syndrome (MIM#214800)

Table 2.1 (continuation). Study individuals with DNA SNVs that cause a rare inherited disease.

Subject's (proband's) code, sex	Analyzed family members	Method of variant identification	Gene	Genotype, identified variant	Additional molecular tests performed	Main clinical features of the subject/rare inherited disorder
IN17-22-P1, female	Mother, father, grandmother	Exome sequencing	<i>GLI3</i> MIM#165240	Heterozygous NC_000007.14(NM_000168.6):c.47 3+3A>T	Sanger sequencing (DNA, cDNA)	Greig cephalopolysyn- dactyly syndrome (MIM#175700)
IN17-37-P, female	Mother, father	Exome sequencing	<i>ARID1B</i> MIM#614556	Heterozygous NC_000006.12(NM_020732.3):c.50 25+2T>C	Sanger sequencing (DNA, cDNA)	<i>Coffin-Siris</i> syndrome (type 1; MIM#135900)
IN17-47-P, male	Mother, father	Gene panel NGS	<i>CAPN3</i> MIM#114240	Compound heterozygous NM_000070.2:c.598_612delTTCT GGAGTGCTCTG NC_000015.10(NM_000070.2):c.17 46-20C>G	Sanger sequencing sekoskaita	Muscular dystrophy of the limbs and lower back (type 1; MIM#253600)
IN17-48-P, male	Mother, father	Gene panel NGS	<i>TGFBR2</i> MIM#190182	Heterozygous NC_000003.12(NM_001024847.2): c.1600-2A>G	Sanger sequencing, qPCR	<i>Loeys-Dietz</i> syndrome (type 2; MIM#610168)
IN17-54-P, female	Mother, father	Exome sequencing	<i>PCCA</i> MIM#232000	Compound heterozygous NG_008768.1(NM_000282.4):c.30 0+2dup NG_008768.1(NM_000282.4):c.(17 46+1_1747-1)_(1845+1_1846-1)del	Sanger sequencing (DNA, cDNA)	Propionic acidemia (MIM# 606054)

Table 2.1 (continuation). Study individuals with DNA SNVs that cause a rare inherited disease.

Subject's (proband's) code, sex	Analyzed family members	Method of variant identification	Gene	Genotype, identified variant	Additional molecular tests performed	Main clinical features of the subject/rare inherited disorder
IN17-44-P, male	Mother, father	Exome sequencing	<i>POMK</i> MIM#615247	Homozygous NM_032237.5:c.136C>T	Sanger sequencing	Distroglycanopathy with brain abnormalities (MIM# 615249)
IN17-55-P, male	-	Exome sequencing	<i>PHF21A</i> MIM#608325	Heterozygous NM_016621.3:c.959_967delin sTTACA	Sanger sequencing	Intellectual disability, behavioral disorders (MIM# 618725)
ATGC21-1-P, male	Mother, father	Exome sequencing	<i>NAA15</i> MIM#608000	Heterozygous NM_057175.5:c.382C>T	Sanger sequencing, analysis by Western blot	Intellectual disability, skeletal abnormalities (MIM#617787)
ATGC21-35-P, male	Mother, father, sister	Exome sequencing	<i>PIGN</i> MIM#606097	Compound heterozygous NM_176787.5:c.1247_1251del NM_176787.5:c.1942G>T	Sanger sequencing	Congenital anomalies, hypotonia (MIM# 614080)

Table 2.2. Study individuals with identified DNA variants – CNVs, causing a rare hereditary disorder.

Subject's (proband's) code, sex	Analyzed family members	CNV identification method	CNV	Additional molecular and/or functional tests performed	Main clinical features of the subject/rare inherited disorder
IN17-05-P1, female	-	Array CGH	arr[hg19] 13q31.3(91,166,748-92,010,901)×3	Exome sequencing, qPCR	Intellectual disability, tall stature, skeletal abnormalities, congenital heart defect
IN17-20-P1, female	Mother, father, two sisters	SNP-CGH	arr[hg19] 4q13.3(72,647,749–74,208,199)x1	FISH	Intellectual disability, skeletal abnormalities, congenital heart defect
IN17-21-P1, male	Mother, father, brother	SNP-CGH	arr[hg19] 1q21.1q21.2(146,476,526-147,820,342)x3,4q12(53,531,672-56,238,263)x1	-	Psychomotor developmental delay, cryptorchidism, piebaldism
IN17-26-P1, female	-	SNP-CGH	arr[hg19] 16p13.1p12.3 (15,493,046-18,069,465)×1	-	Psychomotor developmental delay, autism, craniosynostosis, cerebellar hypoplasia

Table 2.2 (continuation). Study individuals with identified DNA variants – CNVs, causing a rare hereditary disorder.

Subject's (proband's) code, sex	Analyzed family members	CNV identification method	CNV	Additional molecular and/or functional tests performed	Main clinical features of the subject/rare inherited disorder
IN17-28-P1, male	-	VNP-LGH	arr[hg19] 12q24.21(116,523,305- 116,621,185)x1	Sanger sequencing (cDNA), qPCR, CRISPR-Cas9 editing, analysis by Western blot, transcriptome sequencing	MED13L haploinsufficiency syndrome (MIM#616789)

Table 2.3. Study individuals with DNA variants selected for additional molecular and/or functional studies.

Subject's (proband's) code, sex	Analyzed family members	Method of variant identification	Gene(s)	Genotype, identified variant/CNV	Additional molecular and/or functional tests performed	Main clinical features of the subject/rare inherited disorder
ATGC21–34–P, male	-	Genome sequencing	-	Chromothripsis	Conventional karyotyping, FISH, transcriptome sequencing	Intellectual disability, epilepsy, visual impairment, tall stature, skeletal abnormalities
ATGC21–3–P, male	Mother, father, brother	Exome sequencing	<i>DYNCH1</i> MIM#600112	Heterozygous NC_000014.9(NM_001376.5):c.6405+1G>C	Sanger sequencing (DNA, cDNA), qPCR, suppression of expression using siRNA, fluorescence microscopy, analysis by Western blot	Intellectual disability and changes in the cerebral cortex
ATGC21–4–P, male	Mother, father	Exome sequencing	<i>RLIM</i> MIM#300379	Hemizygous NM_016120.4:c.1721T>A	Sanger sequencing, protein structure modeling, X chromosome inactivation study	Tonne-Kalscheuer syndrome (MIM#300978)

Table 2.3 (continuation). Study individuals with DNA variants selected for additional molecular and/or functional studies.

Subject's (proband's) code, sex	Analyzed family members	Method of variant identification	Gene(s)	Genotype, identified variant/CNV	Additional molecular and/or functional tests performed	Main clinical features of the subject/rare inherited disorder
ATGC21-12-P1, male	Mother, father, brother	Exome sequencing	<i>DNMT3A</i> MIM#602769	Heterozygous NM_022552.4:c.2324C>A	Sanger sequencing (DNA, cDNA), protein structure modeling, transcriptome sequencing	Tatton-Brown-Rahmane syndrome (MIM#615879)
ATGC21-15-P, male	-	Exome sequencing	<i>WRN</i> MIM#604611	Homozygous NM_000553.6:c.1578del	Cell senescence study, analysis by Western blot	Werner syndrome (MIM#277700)
ATGC21-18-P, female	Mother, father	Exome sequencing	<i>FAS</i> MIM#134637	Heterozygous NC_000010.11(NM_000043.6):c.16_30+1del	Sanger sequencing (DNA, cDNA)	Autoimmune lymphoproliferative syndrome (type 1A; MIM#601859)
ATGC21-19-P, male	Mother, father	Exome sequencing	<i>KCNQ1</i> MIM#607542	Heterozygous NM_000218.3:c.1111G>C	Cell transfection, electrophysiological measurements	Long QT syndrome (MIM#192500)

Table 2.4. Study individuals with no DNA variants identified that could potentially cause a suspected rare inherited disorder, following cytogenetic, NGS, and/or other molecular testing.

Subject's (proband's) code, sex	Analyzed family members	Cytogenetic and molecular studies performed	Gene candidates	Molecular diagnosis	Main clinical features of the subject/rare inherited disorder
IN17-01-P1, female	-	SNP-CGH, Sanger sequencing	<i>HYLS1</i> <i>POMK</i>	No clinically significant DNA variants identified	Intrauterine growth restriction, congenital anomalies of the central structures of the brain
IN17-04-P1, male	-	SNP-CGH, Sanger sequencing	-	No clinically significant DNA variants identified	Hypotonia, delayed motor development, ventricular septal defect, patent ductus arteriosus
IN17-06-P1, male	Mother, father	SNP-CGH, gene panel NGS, exome sequencing, Sanger sequencing	<i>NRCAM</i>	No clinically significant DNA variants identified	Autism
IN17-07-P1, male	Mother, father	Array CGH, FISH, exome sequencing, Sanger sequencing	<i>RAC1</i>	No clinically significant DNA variants identified	Psychomotor developmental delay, astigmatism
IN17-08-P1, male	Mother, father	SNP-CGH, exome sequencing	<i>HERC2</i>	No clinically significant DNA variants identified	Microphthalmia of the left eye, coloboma, and syndactyly of 3-4 fingers of the hand
IN17-11-P, female	Mother, father	SNP-CGH, exome sequencing	-	No clinically significant DNA variants identified	Specific mixed developmental disorder, congenital brain cysts, hyperopia

Table 2.4 (continuation). Study individuals with no DNA variants identified that could potentially cause a suspected rare inherited disorder, following cytogenetic, NGS, and/or other molecular testing.

Subject's (proband's) code, sex	Analyzed family members	Cytogenetic and molecular studies performed	Gene candidates	Molecular diagnosis	Main clinical features of the subject/rare inherited disorder
IN17-14-P1, male	Mother, father	Exome sequencing, Sanger sequencing	<i>ANKRD11</i>	No clinically significant DNA variants identified	Delayed psychomotor development, vascular cysts, moderate ventriculomegaly, ptosis
IN17-30-P, male	Mother, father	SNP-CGH, gene panel NGS, exome sequencing, Sanger sequencing, qPGR	<i>RIPK4</i>	No clinically significant DNA variants identified	Congenital heart defect, anorectal atresia, horseshoe kidney, visual impairment, microcephaly, thumb aplasia
IN17-40-P, male	Mother, father, sister	SNP-CGH, exome sequencing, Sanger sequencing	<i>TENM1</i>	No clinically significant DNA variants identified	Severe intellectual disability
IN17-42-P, female	Mother, father	SNP-CGH, exome sequencing, Sanger sequencing	<i>BLM</i> , <i>FANCD2</i>	No clinically significant DNA variants identified	Aplastic anemia

3.2. Characterization of the chromothripsis and analysis of its effect on gene expression

This research study analyzed the DNA of a unique individual (ATGC21–34–P, see Table 2.3) diagnosed with intellectual disability, epilepsy, ataxia, tremor, scoliosis, arachnodactyly, strabismus, hyperopia, and tall stature. Due to suspicion of a rare hereditary disorder, a conventional karyotype analysis was performed, which revealed a *de novo* balanced translocation between the short arms of chromosomes 2 and 12 (t(2;12)(p16.2;p12.2)). A aCGH test confirmed that the translocation was balanced. To determine whether any genes were interrupted in the translocation region, a FISH test was performed. The *NRXN1* (MIM#600565) gene, located on the second chromosome near the breakpoints identified by cytogenetic testing, remained unchanged. It was decided to refine the breakpoints of the chromosomal change by performing whole genome sequencing of the proband's DNA sample, purified from a peripheral blood sample. This method revealed a complex and rare chromosomal rearrangement called chromothripsis. Two breakpoints were identified on the short arm of chromosome 2, separated by ~276 kb. Six breakpoints were identified on a ~4.11 Mb fragment of the short arm of chromosome 12. The scheme of rearrangement of fragments formed during chromothripsis, which includes translocations, inversions, and insertions, is shown in Figure 3.2.1. No genetic information was lost during the rearrangements – the alteration is balanced. The fragment of the short arm of chromosome 2 and the nearest sequences around the breakpoints do not contain any genes. However, the following gene sequences are located in the rearrangement region of the short arm of chromosome 12: *LMO3*, *SKP1P2*, *LINC02378*, *RERGL*, *PIK3C2G*, *PLCZ1*, *PLEKHA5*, *AEBP2*, *LINC02398*. The breakpoints in this chromosome interrupt only two genes: the first breakpoint is located in the *MGST1* gene, and the last one is in the *PDE3A* (MIM#123805) gene. Only one gene (*PDE3A*) in the OMIM database is known to be a disease-causing gene. The other seven encode proteins but are not associated with known disorders, and two (*LINC02378*, *LINC02398*) are lncRNA genes whose impact on pathologies is unknown.

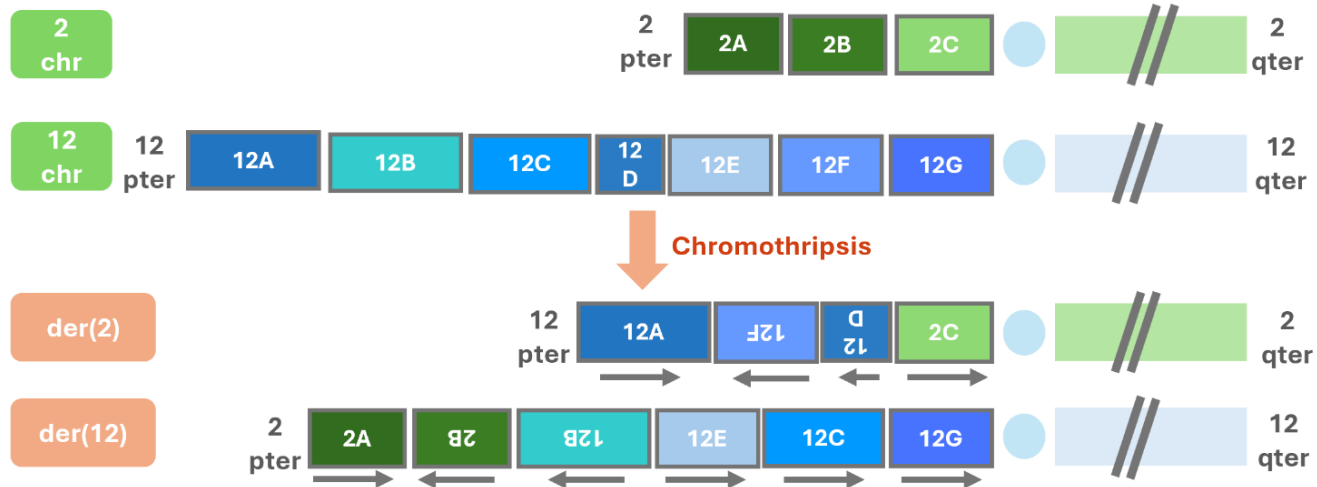


Figure 3.2.1 Chromothripsis scheme. Chromosomal rearrangements (translocations, insertions, inversions) that occurred between chromosomes 2 (2 chr) and 12 (12 chr). Three fragments of different lengths (2A–2C) formed in the short arm of chromosome 2 (2 pter), and seven fragments of different lengths (12A–12G) formed in the short arm of chromosome 12 (12 pter). No rearrangements were found in the long arms of the chromosomes (2 qter, 12 qter). The gray arrows below the fragments indicate the 5'→3' direction. The rows der(2) and der(12) show the derivative (altered) chromosomes 2 and 12.

To evaluate how breakpoints and fragments formed during rearrangements that caused chromothripsis influence gene expression within and outside the region, since regulatory genomic elements are also likely located in the region, the transcriptome sequencing of the proband's RNA sample (purified from a peripheral blood sample) was performed. It was determined that the expression levels of 42 genes in the proband's sample exhibited statistically significant differences when compared to three samples from healthy control subjects of comparable age (see Figure 3.2.2). The downregulated group includes 22 genes, while the upregulated group includes 20 genes. Of the 42 genes, 30 encode proteins, 11 are non-disease-related lncRNA genes, and one transcript is designated as a new protein gene (see Figure 3.2.2, ENSG00000285816). Of the protein-coding genes, nine are associated with diseases in the OMIM database: *NUP210L*, *PRDM5*, *NTNG2*, *RNF213*, *CFD*, *DYSF*, *CSF2RB*, *NADK2*, *SERPINA1*. The expression of the *NUP210L* gene is increased, while the expression of all other listed genes is decreased in the proband's RNA sample (see Figure 3.2.2). The phenotypes determined by the latter genes are autosomal recessive. Only three genes (*DNAH6*, *SLC15A5*, *DYSF*) are located on the short arms of chromosomes 2 and 12, which are involved in the chromothripsis process – only *DYSF* (MIM#603009) is a known disease gene. The phenotypic features characteristic of neurodevelopmental disorder (MIM#618718) associated with the *NTNG2* (MIM#618689) gene only partially overlap with those of our proband and may potentially contribute to intellectual disability, tremor, and ataxia. However, homozygous variants in the *NTNG2* gene cause general growth retardation and short stature, whereas our proband is tall. The phenotypes determined by other genes do not overlap with those of our individual.

After performing a GO (Gene Ontology) term enrichment analysis, the genes belonging to the upregulated group are primarily associated with cell and microtubule movement, immune system-related activities, and several aspects of glutathione metabolism (see Figure 3.2.3 B), while genes with lower expression compared to healthy controls enriched immune system-related GO terms (see Figure 3.2.3 C). Given the clinical features of the subject, immune system dysfunction is not directly related to the patient's phenotype. Still, microtubule-mediated movement within the cell, which is essential for the transport of various cellular cargoes and, therefore, affects multiple cellular processes, is a potential molecular link to the clinical phenotype characteristics of the subject.

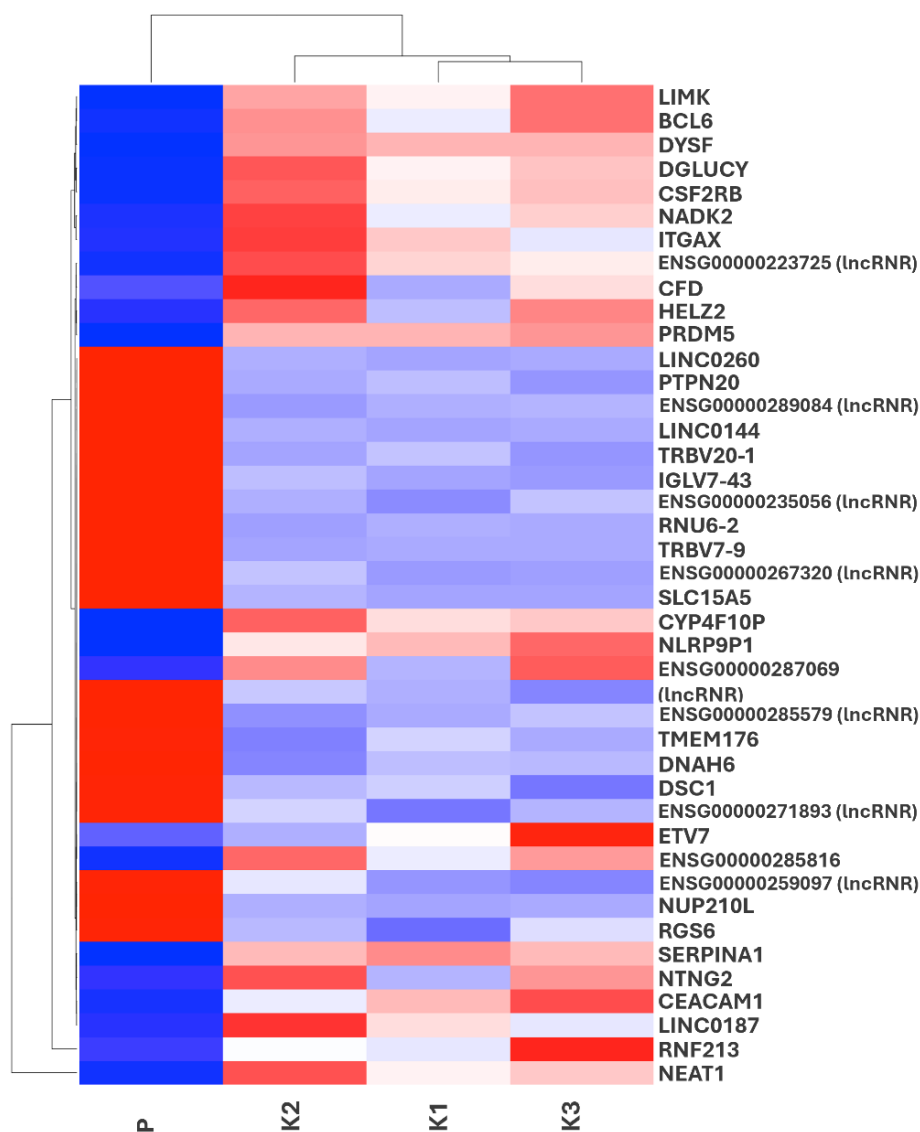


Figure 3.2.2. Changes in gene expression in a blood RNA sample from the individual with chromothripsis. Heatmap represents genes with statistically significant changes in expression. Each row corresponds to a gene, and each column represents a participant. P – proband, K1–K3 – healthy controls. Red – upregulation, blue – downregulation, white – no change. Graph created by Dr. A. Konovalovas.

The chromothripsis identified in this study is a rare chromosomal rearrangement, as it most often occurs in cancer patients as a somatic change in

cells, and cases of constitutional chromothripsis are rarely identified [140, 141]. Chromothripsis describes a process in which multiple breakpoints form fragments of varying lengths, the rearrangement of which involves different chromosomal changes, including inversions, insertions, translocations, and deletions [142]. The literature describes congenital cases of this molecular phenomenon involving different chromosomes in individuals with various congenital anomalies and/or intellectual disabilities, neurological disorders, and recurrent miscarriages [140, 141, 143-145].

The clinical significance of chromothripsis usually depends on and is assessed based on the genes located in the rearranged regions. In our proband's case, one of the deletion points was identified in the *PDE3A* gene region, which is part of the translocation between chromosome 12 and chromosome 2. In the OMIM database, this gene is associated with autosomal dominant inheritance of hypertension and brachydactyly syndrome (MIM#112410). The main phenotypic features characteristic of this syndrome are short stature, early-onset hypertension, and short and thick bones of the hands and feet (type E brachydactyly). These clinical features do not correspond to the phenotype of our proband, and some of the features mentioned (short stature, short toes and fingers) are opposite. The remaining genes in the chromothripsis regions are not associated with specific disorders in the OMIM database. The group of DNA sequences of unclear clinical significance also includes two lncRNAs that have not yet been functionally studied. Therefore, it can be assumed that chromosomal rearrangements occur in regulatory genomic sequences (e.g., gene promoters, repressors, enhancers, or distant regulatory elements) that may influence the development of specific phenotypic traits.

Middelkamp and colleagues (2017) [141] conducted transcriptome sequencing of nerve cells derived from an individual with various congenital anomalies (craniosynostosis, facial dysmorphisms, growth retardation) and intellectual disability, in whom germline chromothripsis was identified, to assess the functional effects of the chromosomal rearrangements on gene expression. The researchers successfully identified altered expression of genes located in the chromothripsis region and in nearby sequences (a total of 67 protein-coding genes) [141]. Sequencing of the RNA transcriptome purified from the blood of our proband also revealed altered expression of 42 genes, 30 of which are protein-coding genes, but only nine of which are known disease genes in the OMIM database. Interestingly, only three of the OMIM genes fall within the chromothripsis region of our proband. In terms of phenotype, the clinical features of the individual partially overlap with several characteristics typical of neurodevelopmental disorders associated with the *NTNG2* gene: intellectual

disability, tremor, and ataxia. However, the latter disorder is characterized by general growth retardation and short stature, while our patient is tall. It is also important to note that the expression of this gene in the proband's RNA sample was found to be lower than in healthy individuals, and the disorder caused by *NTNG2* is of autosomal recessive inheritance. Therefore, it is likely that the function of the protein encoded by both alleles of the gene would have to be impaired for most of the clinical features associated with *NTNG2* to manifest, particularly severe intellectual disability and language development disorder. The gene encodes G2, which provides guidance for axons and is crucial for the formation of synapses and their plasticity during the development of the central nervous system. A scientific study by Dias et al. (2019), which analyzed *NTNG2* homozygous variants, emphasizes that sufficient gene expression is necessary for neurotypical development [146]. Some of the patients described by researchers, in addition to the symptoms mentioned above, also have epilepsy and strabismus – our individual also has these symptoms. To elucidate the mechanism of pathogenesis, Dias and colleagues used siRNA molecules to reduce *NTNG2* expression (achieving twofold lower expression compared to unaffected cells) in mouse neuroblastoma N2A cells. They found that in cultures with lower gene expression, the number, length, and dendrite development area of neurites (dendrites and/or axon precursors) were reduced [146]. Interestingly, a study by Pan and colleagues (2010) found increased *NTNG2* expression in temporal neocortex biopsies from individuals with epilepsy, as well as in cells of the rat hippocampus and adjacent cerebral cortex [147]. In contrast, our proband has epilepsy and exhibits lower expression of this gene. The studies described suggest that the neurological symptoms of the proband could be caused by a dysfunction of the protein encoded by *NTNG2*. However, based on these assumptions, the alteration in gene expression should be validated through an alternative method (qPCR), and the potential impact of other genes with modified expression – those not linked to diseases or encoding lncRNAs – on the

development of the patient’s clinical symptoms should not be disregarded.

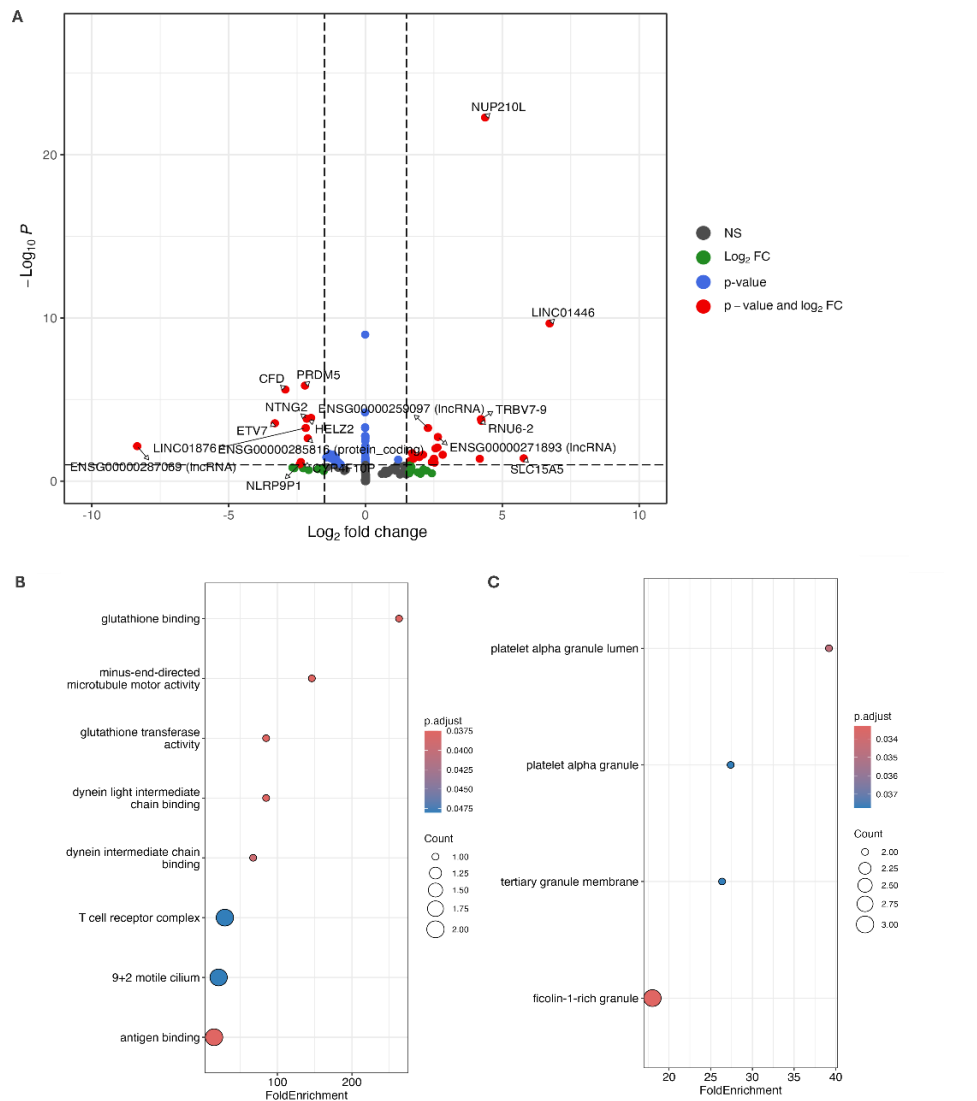


Figure 3.2.3. Analysis of the effect of chromothripsis on gene expression. A – graph showing statistically significant changes in gene expression between the proband with chromothripsis and healthy control individuals. The X-axis represents the change in gene expression on a log₂ scale, while the Y-axis shows the negative decimal logarithm of FDR. Significantly altered genes (FDR < 0.1) and genes with more than a 1.5-fold change in expression are shown above the horizontal dashed line and beyond the vertical dashes. B – dot plot of GO term enrichment for upregulated genes. C – dot plot of GO term enrichment for downregulated genes. Plots created by Dr. A. Konovalovas.

Interestingly, the analysis of GO term enrichment revealed that immune system activities were statistically significantly attributed to downregulated genes. In contrast, increased expression of the genes was statistically significantly associated with several processes related to cell external movement and cargo transport in plasma, as well as the activity of glutathione transferase, which is essential in cell division, apoptosis, and necrosis processes. However, given the proband's phenotype and the previously described candidate gene, it is challenging to establish a direct link between the altered gene expression profile and the individual's clinical symptoms. Therefore, additional bioinformatic studies are needed to analyze smaller groups of genes with increased or decreased expression to identify specific genes that potentially determine the pathogenesis mechanism specifically.

Bioinformatic and molecular studies, as well as literature analysis, support the clinical and functional significance of chromotripsy identified in the proband. The identified changes in gene expression suggest the influence of 42 protein- and lncRNA-encoding genes on the development of the subject's phenotype, emphasizing the reduced expression of the *NTNG2* gene and the reduction of the neutrino G2 it encodes as a possible candidate gene and its role in pathogenesis.

3.3 Assessment of the impact of SNVs and small deletion on cellular processes

Alterations in gene sequence and specific regulatory regions, identified using NGS, enable the prediction of potential qualitative and quantitative modifications in mRNA and/or proteins. Based on the functions of the encoded proteins, preliminary insights into changes within intracellular processes can also be inferred. To elucidate these modifications, this subsection discusses four research participants for whom the clinical and functional significance of detected DNA variants was characterized through the application of molecular and/or functional research strategies tailored to each individual case.

3.3.1 Influence of the splice site variant identified in the *DYNC1H1* gene on the expression of KIFC1 and KIFC3 proteins and the localization of the Golgi complex

In this research study, a heterozygous novel splice site variant NC_000014.9(NM_001376.5):c.6405+1G>C identified in the *DYNC1H1* gene (MIM#600112) was analyzed. The DNA change was identified in the individual (ATGC21–3–P, see Table 2.3) who presented with psychomotor developmental

delay, particularly language delay. The origin of the *de novo* variant was confirmed by testing samples from the healthy parents of the proband.

In the OMIM database, the *DYNC1H1* gene is associated with several neuromuscular disorders (MIM#614228, MIM#158600) and a neurodevelopmental disorder (MIM#614563). The latter includes phenotypic features (such as psychomotor developmental delay) that are also characteristic of the proband. It is also emphasized that the variants causing this disorder are most often of *de novo* origin. It is important to note that at the time of writing the dissertation (December 2024, only three protein-truncating variants (p.Arg728Ter; c.11941+2T>A; c.7473+105C>T) in the *DYNC1H1* gene have been described in the scientific literature [165-167]. All other DNA variants that cause neurodevelopmental or neuromuscular disorders result in a change in one amino acid at the DYNC1H1 protein level.

The identified splice site variant is not found in the *gnomAD* database. The bioinformatics tools MutationTaster, HumanSplicingFinder, and SpliceAI classify this variant as potentially disrupting the messenger RNA (mRNA) splicing process. This effect was assessed in the proband's cDNA sample by amplifying specific fragments using the PCR method and then sequencing them using the Sanger method. Using a direct primer that hybridizes at the junction of exons 29-30 and a reverse primer that binds to the sequence of exon 33, a fragment was obtained whose sequence chromatogram showed additional peaks indicating the presence of part of intron 31. To confirm the sequence of the additional fragment, the PCR products were fractionated in an agarose gel for an extended period. The additional fragment separated in the gel was excised from the agarose gel, and using the same direct primer, Sanger sequencing was performed to identify the mRNA fragment containing part of the intron 31 sequence. Additional primers were designed to specifically sequence only the altered fragment, confirming that the splice site variant in the *DYNC1H1* gene results in an intron 31 retention (see Figure 3.3.1.1). DNA variants found at canonical splice sites more frequently cause exon skipping or alternative splicing, while skipping multiple exons or retaining entire introns are less common [168, 169].

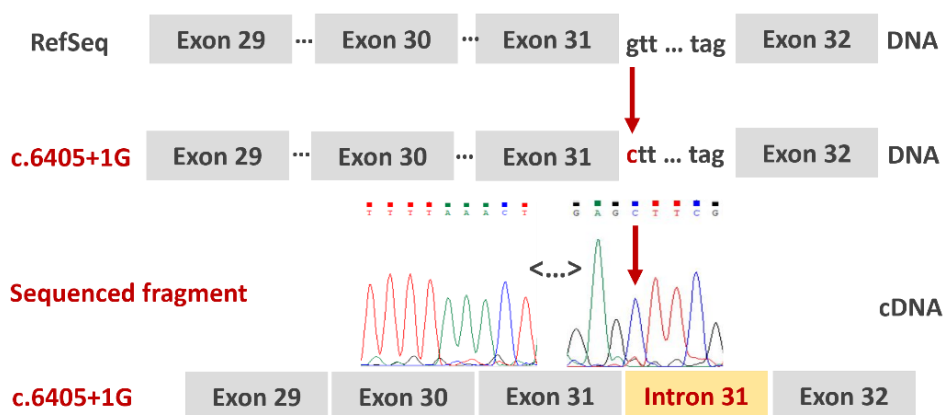
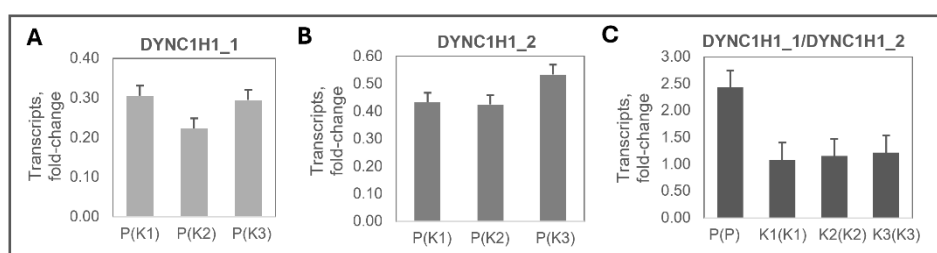


Figure 3.3.1.1. Schematic representation of the effect of the *DYNC1H1* gene splice site variant on mRNA. The identified variant results in intron retention.

The *DYNC1H1* gene encodes a subunit of a large (~1.5 MDa) protein complex called cytoplasmic dynein 1 (or dynein), heavy chain 1 (cytoplasmic dynein 1 heavy chain 1, DYNC1H1). This protein complex is a motor protein belonging to the dynein superfamily. In silico, the insertion of intron 31 causes a shift in the reading frame and the formation of a premature stop codon (NP_001367.2:p.(Ile2136LeufsTer20)). Since the position of the new stop codon in the mRNA sequence is closer to the middle of the transcript, it is likely that nonsense-mediated mRNA decay (NMD) initiated by the premature stop codon is activated. Significantly, this process degrades altered transcripts, whose translation products may be dysfunctional proteins that could have an adverse effect (dominant–negative effect) on the organism’s phenotype [65].



3.3.1.2 Changes in the expression of *DYNC1H1* gene transcripts. A – relative amount of unaltered and altered *DYNC1H1* transcripts. B – relative amount of unaltered *DYNC1H1* transcripts. C – relative amount of *DYNC1H1* gene’s start and end transcripts. P(K1) represents the comparison of the amount of transcripts in the proband with the control individual K1, P(K2) and P(K3) represent the same, respectively. DYNC1H1_1 – the first TaqMan™ probe, which targets the 15–16 exon junction. DYNC1H1_2 – the second TaqMan™ probe, which targets the 51–52 exon junction.

To assess the potential decay of the altered mRNA transcript, the quantity of *DYNC1H1* transcripts was evaluated using the qPCR method by comparing the gene expression determined in the proband's fibroblasts (P) with that of three healthy male controls (K1, K2, K3; see Figure 3.3.1.2). Since the *DYNC1H1* gene consists of 78 exons (transcript NM_001376.5) and the detected change is at the end of the exon 31, two regions were selected for analysis: the first probe (DYNC1H1_1), which targets the beginning of the transcript (exon junction 15–16), and the second probe (DYNC1H1_2), which targets the end of the transcript (exon junction 51–52), thus, the first probe can potentially identify both altered and unaltered *DYNC1H1* gene transcripts, while the second probe can only identify unaltered transcripts. It was found that the amount of *DYNC1H1* transcripts with the first target (total amount of *DYNC1H1* transcripts) in the proband's cells decreased approximately 3-fold compared to the control cells (see Figure 3.3.1.2 A). The amount of *DYNC1H1* transcripts with the second target (unchanged *DYNC1H1* transcripts) in the proband's fibroblasts decreased insignificantly, by 1.3 times, compared to the control individuals' cells (see Figure 3.3.1.2 B). These gene expression results suggest that the formation of a premature termination codon at the identified splice site variant triggered the NMD mechanism, leading to the degradation of the altered *DYNC1H1* transcript in the cells. It is worth noting that the total amount of transcripts in the proband's cells was 2.4 times higher than the amount of unaltered transcripts (see Figure 3.3.1.2 C). Thus, the proband's fibroblasts contain more *DYNC1H1* transcripts with the first exon sequence than with the final exon sequence. In the control samples, the amount of transcripts with the first and second targets did not differ.

Bioinformatic tools predict that the premature termination codon position (p.(Ile2136LeufsTer20)) is located in the motor domain of the motor protein, which is in the first part of the protein (comprising 4646 amino acids), suggesting that DYNC1H1 may lose its function. Additionally, the detected change in the expression of the *DYNC1H1* gene in the proband's fibroblasts may impact protein expression. Immunofluorescence microscopy was used to assess the amount of DYNC1H1 protein. the level of DYNC1H1 in the proband's fibroblasts is decreased by 1.7-fold ($p=0.014$) compared to K1 cells, 1.4-fold ($p=0.001$) relative to K2 cells, and is halved (2-fold reduction, $p<0.001$) compared to K3 cells (see Figure 3.3.1.3). This study suggests that translation from the altered

DYNC1H1 transcript is unlikely to occur due to the activated NMD mechanism; therefore, the shortened DYNC1H1 protein is not expected to function in cells.

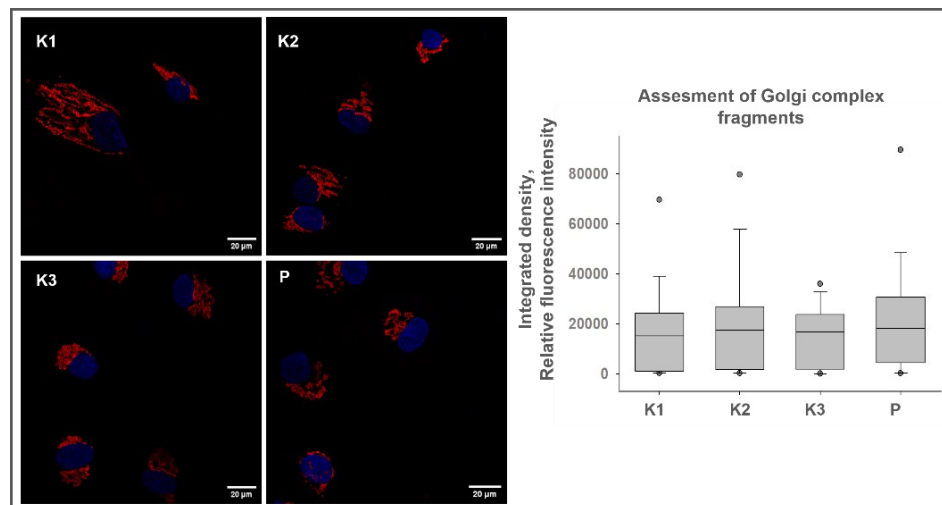


Figure 3.3.1.3. Assessment of DYNC1H1 protein levels in fibroblasts. In the microscopy images, DYNC1H1 is marked in green, and cell nuclei are marked in blue. P – fibroblasts of the proband; K1, K2, K3 – fibroblasts of control individuals. The graph shows the relative protein quantity.

Protein and lipid modification, as well as sorting, by the Golgi complex are crucial for various cellular processes. This is a dynamic organelle composed of cisterns and vesicles, whose movement, change, and localization in the cell are supported by cytoplasmic dynein 1 [170, 171]. Taking into account this function of dynein, the localization of the Golgi complex and the amount of membranes in fibroblasts were evaluated using immunofluorescence microscopy and the Golgi complex marker GM130. Interestingly, the fluorescence intensity of the GM130 marker did not differ significantly between the proband's cells and the control individuals' fibroblasts (see Figure 3.3.1.4). No differences were found in the distribution of Golgi complex membranes in fibroblasts.

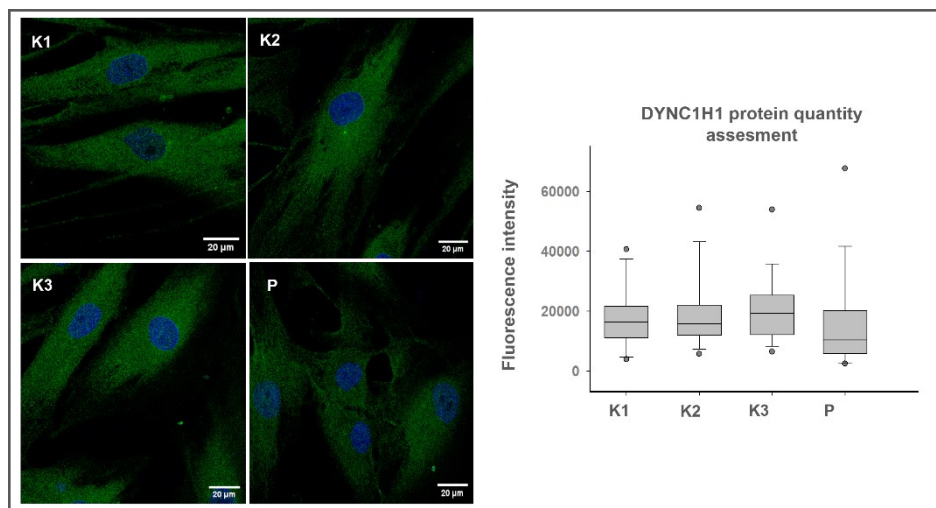
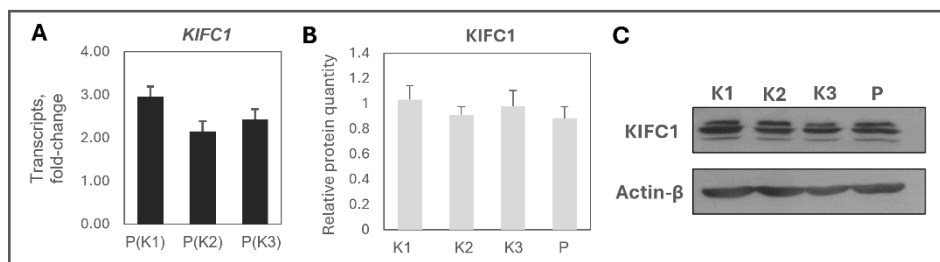


Figure 3.3.1.4. Assessment of Golgi complex fragments in fibroblasts. In the microscopy images, the protein of the Golgi complex GM130 is marked in red, and the cell nuclei are marked in blue. P – fibroblast culture of the proband; K1, K2, K3 – fibroblasts of control individuals. The integrated density shown in the graph is calculated by multiplying the area of the object by the average pixel fluorescence.

Since lower levels of DYNC1H1 protein were found in the fibroblasts of the study individual, and it is known that dynein plays a crucial role in maintaining the dynamics of the Golgi complex, it was expected that changes in the structure, location, and/or volume of the Golgi complex would be observed. It was hypothesized that a compensatory mechanism initiated by other motor proteins may exist in cells, resulting in no significant changes to organelles despite their increased activity. The literature notes that motor proteins belonging to the kinesin superfamily (KIF), KIFC1 and KIFC3, also transport cell cargo towards the minus ends of microtubules [172]. qPCR was used to evaluate the expression of genes encoding KIFC1 and KIFC3 proteins. It was found that the amount of *KIFC1* transcripts in the proband's cells increased 3-fold compared to the K1 control, 2-fold compared to the K2 control, and 2.5-fold compared to the K3 control (see Figure 3.3.1.5 A). When comparing the samples of control subjects with each other, no difference was found in the amount of *KIFC1* transcripts. It was decided to evaluate *KIFC1* expression at the protein level; however, immunoenzymatic analysis using the Western blot method revealed that the amount of KIFC1 protein in the proband's cells did not differ from that in the control cells (see Figure 3.3.1.5 B and C).



3.3.1.5 Changes in the expression of the *KIFC1* gene and its encoded motor protein in the fibroblasts of the proband. A – relative amount of *KIFC1* transcripts. B – relative *KIFC1* protein levels. P(K1) represents the comparison of the proband’s transcript levels with those of control individual K1, P(K2) and P(K3) represent the same for the other controls. C – Western blot analysis of *KIFC1* protein levels.

Based on the analyzed literature, nocodazole can be used to evaluate the importance of dynein for maintaining the localization of the Golgi complex [81, 173]. It is a compound that depolymerizes microtubules by binding to the structural subunit β -tubulin, causing the Golgi complex to disintegrate. Therefore, after exposure, the restoration of the Golgi complex depends on the functional properties of dynein – its ability to restore the structure and localization of the Golgi complex near the cell nucleus [81]. In our study, when comparing the fibroblasts of the proband with those of the healthy controls, changes in Golgi restoration were observed (assessed by the integrated density of the Golgi marker GM130, which is calculated by multiplying the area of the object by the average pixel fluorescence intensity) 3 hours and 45 minutes after nocodazole was washed off. When comparing the cells of the proband and the control individuals, the proband’s cells showed a higher integrated intensity of GM130 fluorescence, 1.5 times higher than in K1 ($p=0.016$) times higher than in K2 ($p<0.001$), and 1.3 times higher than in K3 ($p=0.047$) control individuals (see Figure 3.3.1.6). However, when comparing K2 with other control samples, a statistically significant difference was found; therefore, K1 and K3 samples are used to support the conclusion. It is possible that the structure of the Golgi complex in the fibroblasts of the proband is less compact at the specified time point and is more widely distributed near the cell nucleus.

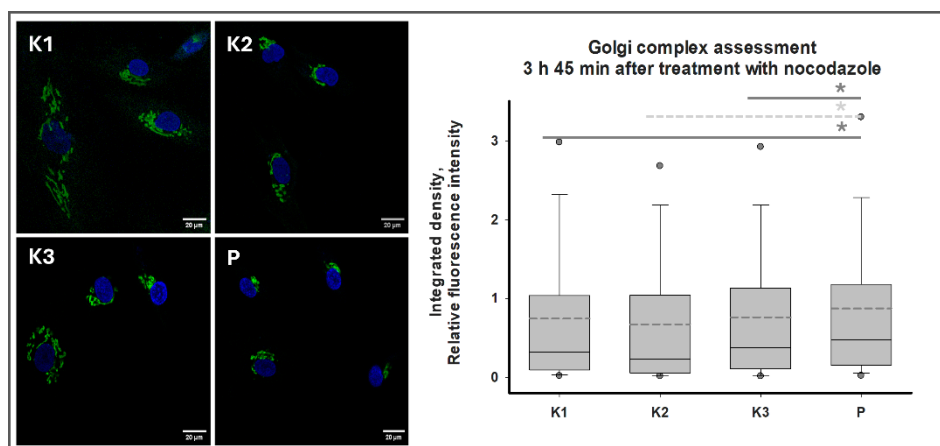


Figure 3.3.1.6. Restoration of the Golgi complex after treatment with nocodazole. The graph shows the fluorescence intensity of GM130 in the tested samples 3 hours and 45 minutes after exposure to nocodazole. The integrated density was evaluated by multiplying the area of the object by the average fluorescence intensity of the pixel. The dotted gray line in the boxes indicates the mean of the data, while the solid line indicates the median. The dotted line above the data indicates that the data for control subject K2 cannot be used for conclusions, as a statistically significant difference was found when comparing the data for K2 with other control samples. In the microscopy images, the Golgi complex protein GM130 is marked in green, and the cell nuclei are marked in blue. P – fibroblast culture of the proband; K1, K2, K3 – fibroblasts of control individuals. Statistical significance * $p \leq 0.05$.

However, the activity of KIFC1 and KIFC3 motor proteins in the study cells may partially mask the effects of DYNC1H1 dysfunction (or reduced expression) on the recovery process of the Golgi complex. Therefore, it was decided to evaluate the ability of DYNC1H1 to restore the Golgi complex after treatment with nocodazole and silencing the expression of genes encoding KIFC1 and KIFC3 proteins, which carry cargo in the direction of the microtubule ends, using siRNA. The efficiency of silencing was verified using the qPCR method; 24 hours after cell transfection, the expression of siRNA-specific genes decreased by a factor of 3. After silencing *KIFC1* and *KIFC3* in control cells, the fluorescence of Golgi complex fragments 3 hours and 45 minutes after nocodazole washout did not differ from the negative control (non-specific siRNA). However, silencing the expression of these genes in the proband's cells reduced GM130 marker fluorescence 13-fold compared to the negative control ($p < 0.001$) (see Figure 3.3.1.7). This study shows that the function of the cytoplasmic dynein 1, whose core and motor parts are composed of DYNC1H1, is impaired in the fibroblasts of the proband. Therefore, in the absence of KIFC1

and KIFC3 proteins, the Golgi complex acquires a more compact structure and localizes closer to the cell nucleus.

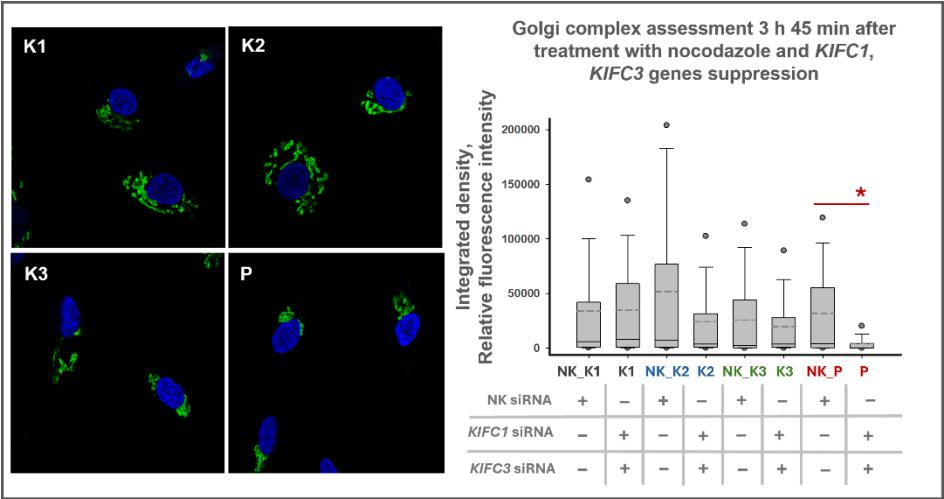


Figure 3.3.1.7. Assessment of Golgi complex recovery after gene silencing and treatment with nocodazole in the studied fibroblasts. In the microscopy images, the Golgi complex protein GM130 is labeled in green, and the cell nuclei are labeled in blue. P – fibroblasts of the proband (red font in the graph); K1 (dark gray font in the graph), K2 (blue font in the graph), K3 (green font in the graph) – fibroblasts of control individuals. The integrated density was evaluated by multiplying the area of the object by the average fluorescence intensity of the pixel. The dotted gray line in the boxes indicates the mean of the data, while the solid line indicates the median. The graph shows the integrated density in the tested samples after silencing the expression of the *KIFC1* and *KIFC3* genes, and 3 hours and 45 minutes after washing off nocodazole. Statistical significance * $p\leq0.05$.

The splice site variant was identified in the DYNCH1 protein, which transports various cargoes within cells, such as cell organelles, spindle components, viruses, nucleic acids, or other proteins through microtubules that make up the cytoskeleton toward their minus end [170]. Adenosine triphosphate is used to power this movement, as this molecular motor has ATPase activity [170]. The ATPase-active domain (motor domain), located in the globular head at the C-terminus of the protein, is formed by two identical heavy chains, which are a common structural component of dyneins [174]. The heavy chains form two more functionally distinct domains: a cargo-binding tail (at the N-terminus of the protein) and a microtubule-binding stem (at the C-terminus of the protein) [174]. The variable part of dyneins can consist of several light chains and several intermediate chains [170, 174]. The identified splice variant, which causes a

decrease in the expression of the *DYNC1H1* gene and the loss of the motor domain of the protein, also affected the efficiency of DYNC1H1 protein expression in our test cell, as indicated by confocal microscopy images.

Dynein is the primary motor protein responsible for maintaining the dynamics of the Golgi complex within the cell. Therefore, when examining the Golgi complex under a microscope, we expected to observe visual changes in this organelle resulting from altered DYNC1H1 expression. However, experimental studies described in the literature have found that KIFC1 and KIFC3, which belong to the kinesin superfamily and also move to the minus ends of microtubules, participate in maintaining the position of the Golgi complex [175, 176]. Nevertheless, studies of the RNA and protein expression of these proteins in our fibroblast cells revealed contradictory results: the amount of *KIFC1* gene transcripts increased. In contrast, the quantity of KIFC1 protein in the cells remained unchanged. It is interesting to note that an increase in the amount of transcripts in cells does not always correlate with changes in protein expression, as in the case of KIFC1 expression, and this may be influenced by various processes occurring in cells: translation efficiency, when transcripts compete with each other for a place in the translation mechanism; protein stability, due to which differences in quantity are not detected because the protein's lifespan in the cell is shorter; impaired protein localization and transport due to an increased amount of specific protein, which can accumulate in cells in places not characteristic of it and therefore be directed for degradation [177-179]. Thus, it can be assumed that the KIFC1 motor protein partially compensates for the impaired activity of the DYNC1H1 protein in the cell. Meanwhile, the amount of *KIFC3* transcripts in the proband's RNA sample did not differ from that of the control individuals. In addition, to assess changes in the volume of the Golgi complex, the expression of the Golgi complex protein GOLGB1 was determined at the transcript level; the expression did not change when comparing the proband's sample with those of control subjects.

Since the Golgi complex is a dynamic and unstable structure, studies are often performed to monitor its position and dynamics. Such experiments with nocodazole, which depolymerizes the microtubules that make up the cytoskeleton, have also been used to evaluate the effect of variants identified in the *DYNC1H1* gene that cause amino acid changes (p.Gln1194Arg and p.Glu3048Lys) on the dynamics of the Golgi complex. In this case, it was found that the Golgi structure is restored more slowly in the cells under investigation [81]. This may be due to impaired DYNC1H1 protein function, which causes the vesicles of the disrupted Golgi complex to move more slowly towards the nucleus (minus ends of microtubules). Changes in the movement and position of Golgi complex vesicles relative to the nucleus were also observed in our fibroblast

cells. This was due to both changes in the expression of the *DYNC1H1* protein and alterations in the expression of *KIFC1*. In the cells studied in this work, the Golgi complex vesicles were more compactly arranged near the nucleus, and no differences in movement speed were observed.

Thus, changes in the expression of this gene at the RNA and protein levels were detected in the fibroblasts of the proband with the *DYNC1H1* splice site variant, the detected increase in the amount of *KIFC1* transcripts, and the detected changes in Golgi complex regeneration confirm the clinical and functional significance of the variant detected in the *DYNC1H1* gene. However, it would be useful to conduct additional studies to investigate the expression of the *KIFC1* gene in more samples of control individuals, to extend immunofluorescence studies to determine KIFC1 protein levels and to monitor the dynamics of cytoplasmic dynein and microtubules after exposure to nocodazole at more frequent time points, and thus determine the exact starting point of Golgi complex recovery, where the fibroblasts of the proband and healthy individuals would differ the most. Given the neurodevelopmental phenotype of the affected individual, it would be advisable to perform these experiments in nerve cells to confirm the dysfunction of this protein and its effect on nerve cell activity.

Bioinformatic, molecular, and functional studies, as well as a literature review, support the pathogenicity of the splice variant NC_000014.9(NM_001376.5):c.6405+1G>C identified in the *DYNC1H1* gene. These studies have identified changes in the levels of *DYNC1H1* and *KIFC1*, as well as impaired Golgi complex dynamics in the cell, suggesting that the impaired movement of other cellular cargo may be a possible mechanism of pathogenesis.

3.3.2 Effect of the small deletion in the *FAS* gene on apoptosis induction

The exome sequencing analysis identified a heterozygous variant NC_000010.11(NM_000043.6): c.16_30+1del in the *FAS* gene (MIM#134637) in the individual (ATGC21–18–P, see Table 3.3), who was suspected of having autoimmune lymphoproliferative syndrome (ALPS; MIM#601859). In the OMIM database, this gene is identified as a disease gene, where variants in DNA lead to the aforementioned immune system disorder. Interestingly, the same small deletion was found in the proband's mother, who has no signs of the disease. Based on the available literature, the clinical penetrance of ALPS is estimated to be ~60%, indicating that some individuals with changes in the *FAS* gene sequence are asymptomatic carriers of the disease [180].

Since the identified small deletion covers the canonical splice site, cDNA was synthesized using RNA samples from the proband and her mother, purified from blood samples, and sequenced using the Sanger method. A direct primer that hybridizes to the 5' untranslated region of the gene and a reverse primer that binds to the junction of exons 3 and 4 were used to amplify the fragments. It was found that the small deletion results in an alternative splice site (see Figure 3.3.2.1). *In silico*, this results in a reading frame shift and the formation of a premature termination codon (NP_000034.1:p. (Thr6Phefs17Ter)). This codon is located near the start of the *FAS* transcript (5' end), so the NMD mechanism is likely initiated, and the altered transcript is degraded in cells.

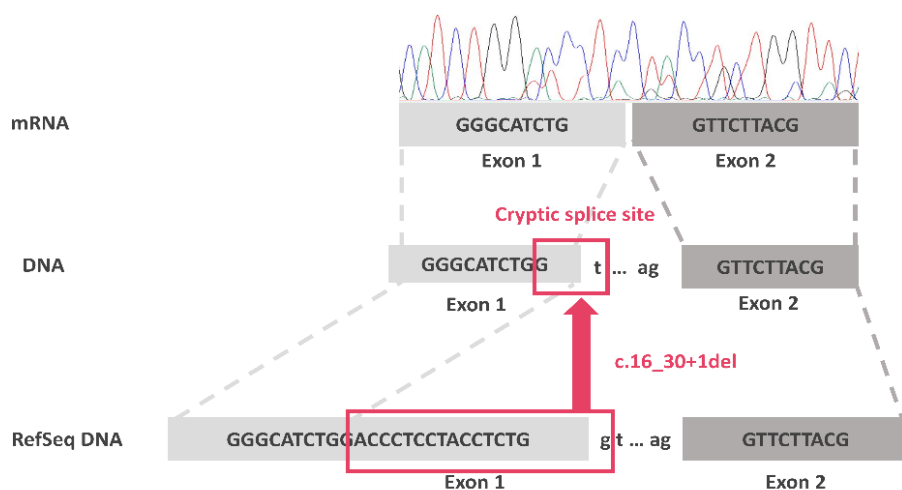


Figure 3.3.2.1. Schematic representation of the effect of a small deletion in the *FAS* gene covering the donor splice site on mRNA. The identified variant determines the formation of an alternative splice site and a shift in the reading frame.

The *FAS* gene encodes a glycoprotein of the TNF (Tumor Necrosis Factor) receptor superfamily, called the Fas cell surface death receptor. The primary function of this receptor in most cells is the regulation of programmed cell death, also known as apoptosis, which is triggered when the *FAS* receptor binds to its ligand, FasL [181]. According to the GTEx domain database, the *FAS* gene is expressed in various tissues and cells, with high expression also occurring in fibroblasts. Given the possible decrease in *FAS* protein levels and the corresponding effect on programmed cell death, it was decided to evaluate apoptosis in the fibroblasts of the proband, her mother, and healthy control individuals. Since the proband was diagnosed with ALPS, which includes symptoms such as impaired lymphocyte apoptosis and an increased number of T

cells with specific markers, it was hypothesized that apoptosis in fibroblasts is also impaired and likely less effective in the patient's cells. For the experiment, fibroblasts were incubated for 28 hours with the Fas receptor-activating antibody anti-Fas (Sigma-Aldrich), and apoptosis induction was assessed using the commercial dye CellEvent™ Caspase-3/7 Green ReadyProbes™ Reagent (Invitrogen). After exposure to anti-Fas antibody to activate apoptosis, fibroblast cultures were observed and evaluated using fluorescence microscopy. No differences were found when comparing the fibroblasts of the proband with those of her mother. The efficiency of apoptosis was also analyzed using flow cytometry with a commercial kit, the Dead Cell Apoptosis Kit with Annexin V for Flow Cytometry (Invitrogen), but no significant differences were observed.

According to the literature, studies analyzing FAS receptor-dependent apoptosis in various cell types have revealed that this programmed cell death can be inhibited by a specific protein called FLIP (Cellular FLICE-like inhibitory protein; acronyms FLIP, cFLIP, CFLAR) [182]. This protein inhibits the proteolytic activity of caspase-8 in the FAS apoptosis signaling pathway, which prevents apoptosis from being activated [182]. Interestingly, FLIP can act both as an inhibitor and a promoter of FAS-dependent apoptosis, depending on the intensity of *FLIP* expression. In cells where this protein is constantly expressed, the FAS apoptotic pathway is inhibited [182]. Experimental studies conducted on human skin fibroblasts have shown that by inhibiting *FLIP* expression with specific antisense oligonucleotides or using cycloheximide, which disrupts protein synthesis, FAS receptor-induced apoptosis is activated – fibroblasts become sensitive to FAS ligand and eventually die [183].

Based on the study described above, *CFLIP* gene expression was reduced in our proband's and her mother's fibroblasts using specific siRNA (inhibition efficiency verified by qPCR, with expression reduced 4- to 16-fold) to assess whether splicing changes at the mRNA level alter the efficiency of FAS-dependent apoptosis induction. The Tali Apoptosis Kit - Annexin V Alexa Fluor 488 & Propidium Iodide was used to identify dying cells in fibroblast cultures. In samples from the proband and his mother that were not affected by the antibody and siRNA or were only affected by the antibody, approximately 10% of dying cells, identified as Annexin V-bound cells, were detected. After exposure to FAS-activating antibody and *CFLIP* expression reduction, the percentage of apoptotic cells in the proband's samples increased approximately 3-fold compared to the proband's cells treated with target-free siRNA and anti-FAS antibody ($p=0.045$; see Figure 3.3.2.2). A similar effect was observed in the corresponding samples of mother cells – the number of apoptotic cells increased by more than 20% after downregulation of *CFLIP* expression and induction of the FAS receptor apoptotic pathway (see Figure 3.3.2.2). It is also interesting

that, after exposure to FAS antibody and inhibition of *CFLIP* expression, approximately 20% more dying maternal fibroblasts were found compared to the corresponding proband's sample. However, statistical significance was not achieved ($p = 0.086$). However, in biological terms, the results suggest a possible tendency for apoptosis to be less effective in the proband's fibroblasts. However, due to technical limitations, it was not possible to collect sample data in all experiments. Therefore, to achieve statistical significance and confirm the trend in apoptosis efficiency, it is advisable to repeat the experiments and conduct similar tests on cells from healthy control individuals. It would also be useful to examine the expression of proteins involved in the FAS signaling pathway and related to this pathway (the genes encoding them) in the cells of the proband and her mother, using genome and transcriptome sequencing methods to identify possible causes of the (non)manifestation of the disease at the signaling pathway level.

			Percentage of apoptotic cells		
anti-FAS	NK siRNA	<i>CFLIP</i> siRNA	I	II	III
-	✓	-	5.0	8.4	7.8
-	-	✓	4.3	15.7	11.0
✓	✓	-	5.2	13.1	10.2
✓	-	✓	16.7	38.5	34.7
P					
			NA	30.0	0.0
			NA	20.6	19.1
			NA	33.6	13.8
			NA	52.2	47.9
M					
			NA	30.0	0.0
			NA	20.6	19.1
			NA	33.6	13.8
			NA	52.2	47.9

Figure 3.3.2.2. Proportion of apoptotic cells after exposure to FAS-activating antibody and/or reduction of *CFLIP* expression in fibroblasts from the patient and her mother. P – fibroblast samples from the patient, M – fibroblast samples from the patient's mother. Roman numerals (I, II, III) indicate three separate experiments. NA (Not Available) indicates data that could not be collected for technical reasons. Each row of the graph represents a separate cell sample. In each sample, flow cytometry data were collected from at least 800 cells. The Tali™ Apoptosis Kit - Annexin V Alexa Fluor™ 488 & Propidium Iodide (Thermo Fisher Scientific) was used to determine regulated cell death.

Since the FAS receptor is located on the cell surface, its structure is divided into two parts: extracellular and intracellular. Thus, the N-terminus is the ligand-

binding extracellular domain (Prie-ligand Assembly Domain, PLAD), followed by the transmembrane domain and the cytoplasmic death domain (Death Domain, DD) at the C-terminus of the receptor [184]. Typically, the FAS receptor exists in a monomeric state within cells. Still, after FasL binding, it forms a homotrimer, which is necessary to initiate further intracellular signaling processes (DISC complex formation, caspase cascade) that determine apoptosis [185]. The deletion identified in our proband's DNA and her asymptomatic mother, which determined the alternative splicing, predicted the position of the premature termination codon (p.(Thr6Phefs17Ter)), suggesting that the protein residue of several amino acids in the cell should not function and should not be translated at all. Variants that cause changes in the DD domain are more frequently found in the FAS gene, resulting in the formation of an altered protein in cells that can interact with unchanged FAS proteins through the extracellular domain and cause "dominant interference", where the mutated FAS protein competes with unchanged FAS monomers and forms inactive heterotrimers [180]. Less frequently, changes in FAS are observed that decrease the amount of the receptor encoded by this gene in the cell, leading to haploinsufficiency of this protein [180]. The experiments in this study were performed not in immune system cells, but in fibroblasts from the patient and her mother. However, the differences in apoptosis efficiency observed between these two subjects suggest that the identified protein-shortening variant has a lesser effect on cell death in the proband's cells compared to those of the mother. This suggests a possible compensatory mechanism that determines the incomplete penetrance of the *FAS* gene-determined disease.

Bioinformatic, molecular, and functional studies, as well as a literature review, indicate the possible pathogenicity of the small deletion NC_000010.11(NM_000043.6):c.16_30+1del identified in the *FAS* gene. The changes in the FAS mRNA splicing process identified in these studies suggest that the decreased efficiency of FAS-induced apoptosis in cells may be a possible mechanism of disease pathogenesis. However, additional genome and transcriptome studies are needed to determine the differences in signaling pathways between the study cells showing disease symptoms and the healthy mother cells carrying the *FAS* gene deletion.

3.3.3 Analysis of the impact of the identified variant in the *DNMT3A* gene on DNA methylation in cells

In this study, the heterozygous variant NM_022552.4:c.2324C>A, which has not been described in the literature, was identified by whole exome sequencing in the *DNMT3A* gene (MIM#602769) in the individual (ATGC21-12-P1, see

Table 3.3) who exhibited cardiac pathology (aortic root dilatation, left ventricular non-compaction cardiomyopathy, arrhythmia), intellectual disability, tall stature, skeletal abnormalities, and obesity. Segregation analysis using the Sanger sequencing method revealed that the proband inherited this variant from his mother, who has cardiac pathology and related diseases. The variant was also confirmed in his brother, who has skeletal anomalies, tall stature, and cardiac disorders. In the OMIM database, *DNMT3A* is a known disease gene, where somatic DNA variants are associated with cancerous disorders (MIM#601626). In contrast, germline variants are linked to autosomal dominant microcephaly and short stature syndrome (MIM#618724) or tall stature Tatton-Brown-Rahman syndrome (MIM#615879). The latter includes phenotypic features characteristic of the individuals.

The identified variant was not found in the gnomAD database. Several bioinformatic tools (SIFT, PolyPhen) predicted the variant to be likely pathogenic. Although the SpliceAI tool predicted that the variant is unlikely to affect splicing, the MutationTaster tool suggested that this variant may impact the mRNA splicing process. This is most likely due to its position at the 5' end of exon 20, specifically at position 2, near the canonical acceptor splice site. It was decided to conduct Sanger sequencing on the test cDNA, which was synthesized from the proband's RNA sample purified from a peripheral blood sample. A direct primer that hybridizes to exon 18 and a reverse primer that hybridizes at the junction of exons 21 and 22 were employed to amplify the relevant fragments. However, no fragments of the altered cDNA sequence were detected.

The *DNMT3A* gene encodes the protein DNA methyltransferase 3 α (DNA Methyltransferase 3 Alpha), which is responsible for *de novo* DNA methylation. Its expression is found in almost all tissues of the body and is most intense in undifferentiated cells, as it creates methylation marks in unmethylated DNA positions [186]. The epigenetic DNA modification performed by DNMT3A is one of the ways in which cells regulate gene expression, as gene promoter sequences are often methylated and the expression of specific genes is suppressed due to the altered chromatin state. Therefore, this methyltransferase is particularly important during the embryonic stage of development, as well as for X chromosome inactivation and the formation of imprinting marks [186].

The DNMT3A protein consists of two main domains: the N-terminal regulatory domain (comprising the PWWP and ADD domains), which is necessary for interaction with chromatin, and the C-terminal catalytic domain, which performs the methyltransferase function [186]. The c.2324C>A variant identified at the protein level causes an amino acid change NP_072046.2:p.(Ser775Tyr) in the methyltransferase domain, which is

evolutionarily conserved. To evaluate the *in silico* impact of the identified amino acid change on protein structure and function, computer modeling studies were conducted. It was found that the changed amino acid is located in a region important for interaction with the DNMT3L protein. The latter is a protein without catalytic activity that acts as a regulatory factor for DNMT3A, promoting its catalytic activity [187]. Correct interaction between DNMT3A and DNMT3L causes conformational changes in DNA methyltransferase, which opens the catalytic part of DNMT3A and promotes DNA binding [187]. Therefore, it is likely that changes in interaction may lead to impaired DNA methylation efficiency. It is also important that Ser775 is close to the catalytic domain and the region important for ADD domain interaction (switch region). *In silico*, the change of serine to the larger amino acid tyrosine may alter the thermodynamic properties of the protein and disrupt the interaction between these domains, which is likely to have a negative impact on the function of DNMT3A (see Figure 3.3.3.1).

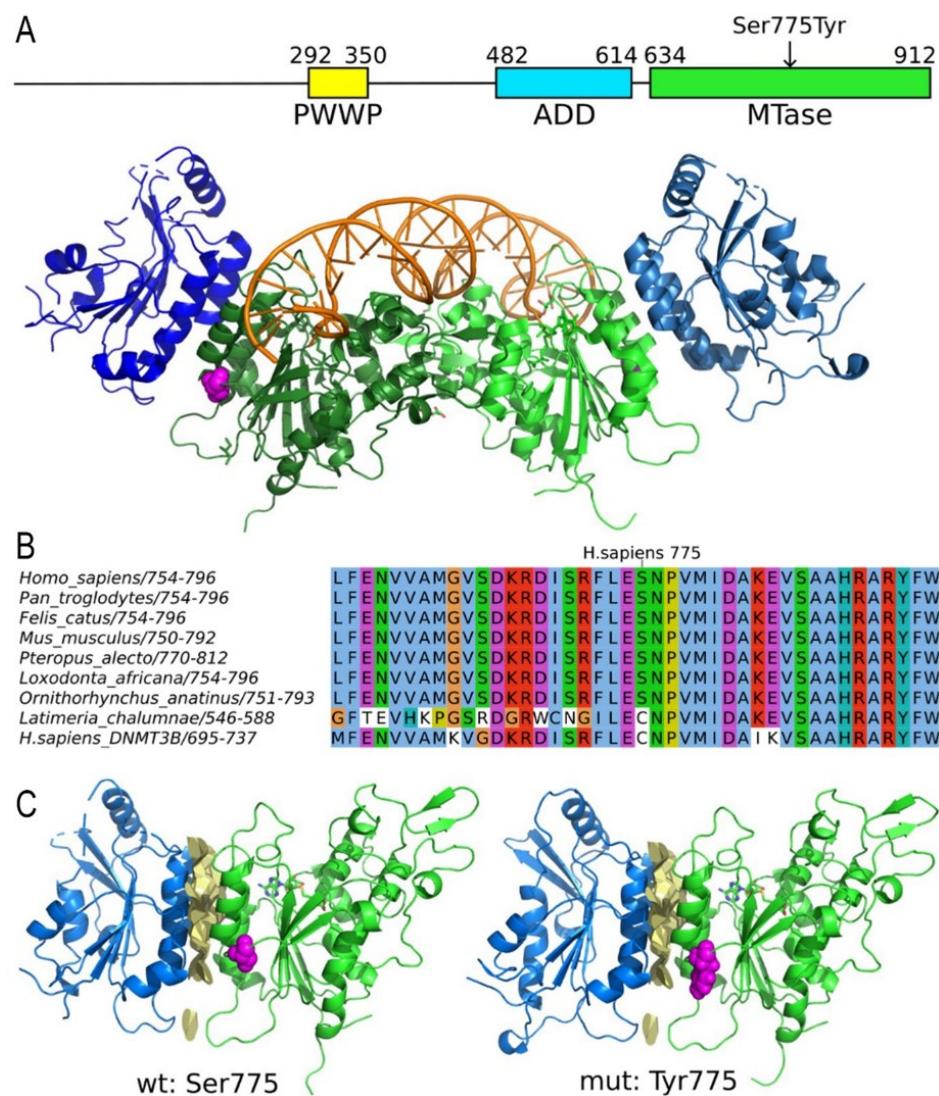


Figure 3.3.1.1. Structural changes in the DNMT3A protein due to the amino acid substitution Ser775Tyr. A – Organization and structure of DNMT3A domains in complex with DNMT3L and DNA (PDB: 5YX2). DNMT3A is green, DNMT3L is blue, DNA is orange, and the Ser775 amino acid is represented by magenta spheres. B – Multiple sequence alignment of DNMT3A proteins found in mammals shows a conserved region containing the Ser775 amino acid. C – The altered amino acid at position 775 is close to the interaction between DNMT3A and DNMT3L proteins (PDB: 5YX2). The interaction is shown in yellow. Image created by Dr. J. Dapkūnas.

Based on the results of protein computer modeling and the function of the DNMT3A, transcriptome sequencing was performed on RNA samples from the subject, his brother, and a group of healthy males of similar age ($n = 3$). It was found that when comparing the samples of the proband and his brother with healthy control individuals, the expression of 139 genes changed: the expression of 69 genes increased, while the expression of 70 genes decreased (see Figures 3.3.3.2 and 3.3.3.3). More than three-quarters of the genes with altered expression encode proteins, and there are also long non-coding RNAs (lncRNAs) and transcripts not associated with a specific function. It is worth noting that the expression of the *DNMT3A* gene in the samples studied has not changed significantly. Some of the identified genes are associated with the biology and physiology of the cardiovascular system, as reported in the analyzed literature. One of the genes with increased expression is *FOXP1* (Forkhead Box P1). This gene encodes a transcription factor crucial for the development of various tissues, including cardiac tissue [188]. Interestingly, *Foxp1* is associated with cardiac abnormalities in mouse embryonic cell studies and is described in the literature as a methylated gene of the DNMT3A protein [188, 189]. Other genes related to the cardiovascular system, whose expression is significantly increased in the proband and his brother's sample, are involved in protein synthesis (e.g., *RPL29*), proteins important for the immune system (e.g., *IL6ST*, *MS4A1*, *CD1C*), and factors involved in signaling pathways (e.g., *SEMA3A*, *WNT16*) [190-193]. Significantly reduced expression of genes encoding proteins associated with transcription regulation (e.g., *ZNF257*) and signaling pathways (e.g., *SPATS2L*, *S100P*), whose importance to the cardiovascular system has been analyzed in various studies [194]. In addition, some of the genes with altered expression (e.g., *MYADM*, *MMRNI*, *IRF7*, *LINC02446*, *CD9*, *VNN1*, *SH2D2A*) identified in our index cases were also identified by Smith and

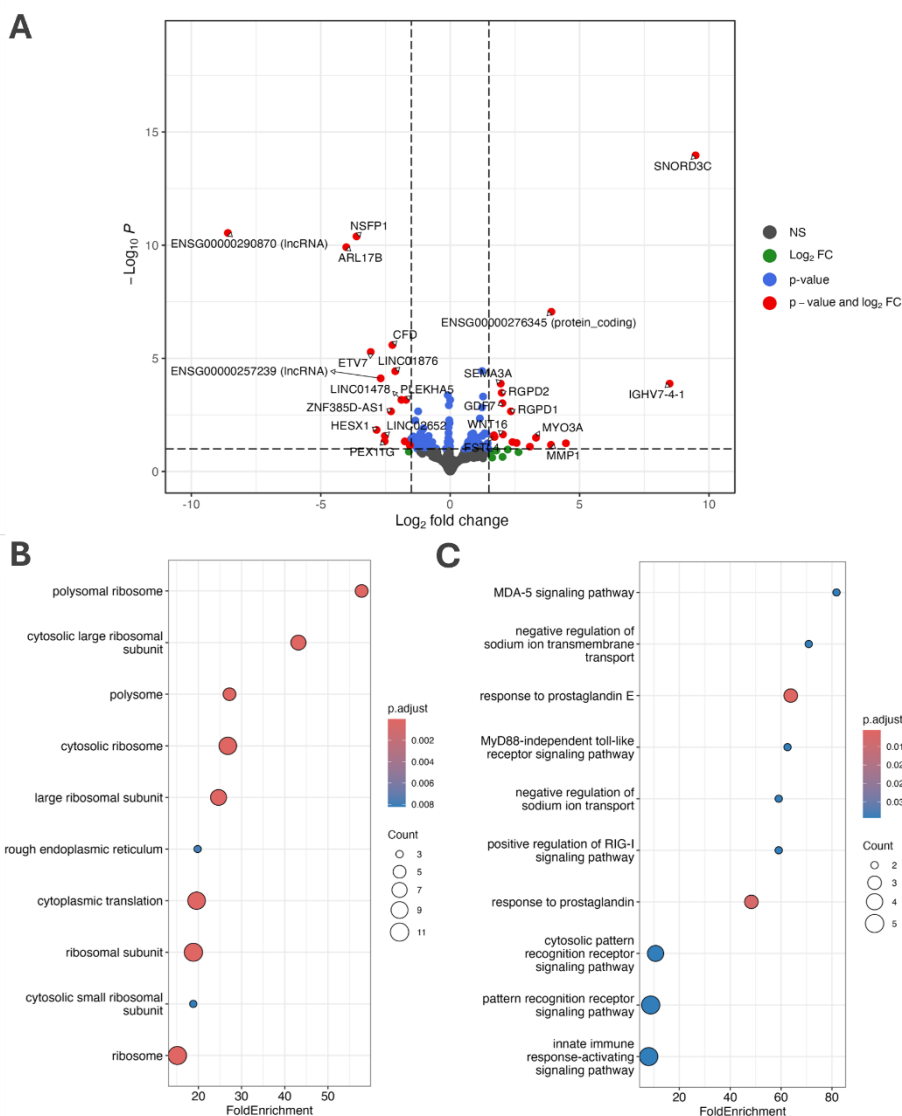


Figure 3.3.3.3. Downregulated and upregulated genes in peripheral blood samples from individuals with the *DNMT3A* gene variant. A – Volcano plot of statistically significant changes in gene expression between probands with the *DNMT3A* variant and healthy controls. The X-axis represents the change in gene expression in log2 scale, and the Y-axis shows the negative decimal logarithm of FDR. Genes with significantly altered expression (FDR < 0.1) and more than 1.5-fold change in expression are shown above the horizontal dashed line and beyond the vertical dashes. B – Dot plot of GO term enrichment for genes with statistically significant increased expression. C – Dot plot of GO term enrichment for genes with statistically significant decreased expression. Plots created by Dr. A. Konovalovas.

Colleagues (2021) [195] analyzed another DNMT3A protein change (Arg882His) in a subject diagnosed with Tatton-Brown-Rahman syndrome. Although the direct link between the identified genes and the phenotype of our subjects has not yet been experimentally investigated, there are indications of their significance for the biological and physiological processes of the cardiovascular system; therefore, more detailed functional studies are undoubtedly important to confirm the link and reveal the mechanisms of pathogenesis.

To assess the function of the potentially altered DNMT3A protein, cell-free DNA (cfDNA) analysis was performed using third-generation sequencing (Oxford Nanopore technology) on the proband's brother (only this person's sample was available for the study). This study analyzes DNA fragments wrapped around nucleosomes, the length of which may vary depending on DNA methylation [196]. A shorter mononucleosome fragment length (156 bp) was found in the subject compared to the data of control subjects stored in freely available databases (168 bp; see Figure 3.3.3.4). According to the literature, shorter fragments are characteristic of hypomethylation and higher nuclease activity in samples from cancer patients [197, 198]. In addition, third-generation sequencing enables the determination of methylation (5mC) and hydroxymethylation (hm5C) characteristics of cfDNA. In this case, the total CpG island modification determined in the cfDNA sample of the brother is determined by DNMT3A: 75.8% 5mC and 2.4% 5hmC, for a total of 78.2% (see Figure 3.3.3.4). The detected amount is similar to the methylation levels observed in peripheral blood samples from patients with the same syndrome and another amino acid-altering variant in the *DNMT3A* gene, as reported in a study by Smith et al. (2021) [195]. Thus, the *DNMT3A* gene variant identified in our individuals causes dysfunction of the protein it encodes, which affects not only methylation and hydroxymethylation processes but also the expression of other genes. These studies confirm the pathogenicity of the *DNMT3A* gene variant and potentially reveal the complex biological processes that determine the phenotype of the Tatton-Brown-Rahman syndrome under investigation.

Bioinformatic and molecular studies, as well as literature analysis, support the clinical and functional significance of the variant NM_022552.4:c.2324C>A identified in the *DNMT3A* gene. The changes in DNMT3A protein function identified in the study samples by RNA and cfDNA sequencing suggest a less efficient methylation process and impaired expression of specific genes as a possible mechanism of pathogenesis.

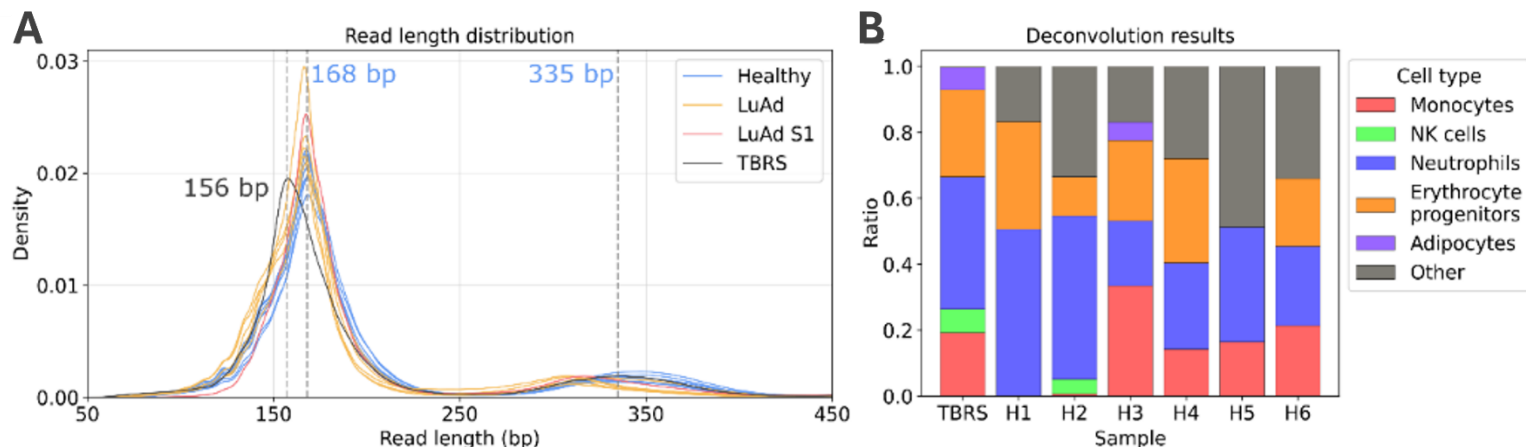


Figure 3.3.3.4. Analysis of the sequence of cfDNA nanopores in the blood of the proband's brother. A – Graph of cfDNA fragment density. The distribution of fragment sizes was compared among patients with the DNMT3A gene variant (TBRs), seven healthy individuals, and six patients with lung adenocarcinoma (LuAd). Sample S1 is a sample from a lung adenocarcinoma patient whose cfDNA library was prepared in a manner similar to that of our study subject. B – Determination of the cellular origin of lcDNA fragments using the method of Moss et al. [197]. The results of our proband's brother (P2) were compared with those of six healthy individuals (H1–H6) from the study by Katsman et al. (2022) [198]. The figure was created and the studies were conducted by Dr. M. Tomkuvienė and J. Balčiūnas.

3.3.4 Analysis of the disruption of potassium ion movement in the cell membrane caused by a variant identified in the *KCNQ1* gene

This scientific study analyzed the heterozygous variant NM_000218.3:c.1111G>C, identified by exome sequencing in the *KCNQ1* gene (MIM#607542) and not described in the literature in the proband (ATGC21–19–P, see Table 3.2), who was diagnosed with long QT syndrome. In the OMIM database, this gene is a known disease gene whose DNA variants are associated with autosomal recessive inheritance Jervell and Lange-Nielsen syndrome (MIM#220400) and autosomal dominant inheritance familial atrial fibrillation (type 3; MIM#607554), short QT syndrome (type 2; MIM#609621), and long QT syndrome (type 1; MIM#192500). The phenotypic features characteristic of the latter disorder correspond to the cardiopathology of our subject.

The variant identified at the protein level (NP_000209.2:p.(Ala371Pro)) is located in the A helix at the C-terminus of KCNQ1, which is important for the interaction of the potassium channel with calmodulin, which regulates both the opening of the channel and the formation of the tetrameric structure [199]. Therefore, it is likely that the identified variant may have an adverse effect on potassium channel function or even on the entire tetrameric structure. To assess the potential dysfunction of the altered KCNQ1, a system was created to compare altered and unaltered protein expression in a stable cell line, and subsequent electrophysiological measurements were performed. The HEK293T cell line, as described in the literature, was selected for these studies. This cell line contains the *KCNQ1* gene in its genome, but the expression of this gene and the processes of the encoded protein do not occur in these cells [200]. Chemical transfection was used to introduce plasmids encoding the unmodified (KCNQ1_WT) and modified (KCNQ1_MUT) proteins, as well as a plasmid encoding the KCNE1 protein, into the cells. Their expression was verified using a fluorescence microscope. Twenty-four to forty-eight hours after transfection, electrophysiological measurements were performed on cells growing on glass slides in specialized plates using the patch-clamp method (see Figure 3.3.4.1A). Several response examples were successfully recorded. A minimal outward current was recorded in non-transfected HEK293T cells (3.78 ± 0.88 pA/pF; see Figure 3.3.4.1B), a finding also reported in other studies [200]. No statistically significant differences were found when comparing untransfected cells with KCNQ1_WT (4.31 ± 1.39 pA/pF) or KCNQ1_MUT (4.56 ± 1.62 pA/pF) plasmid-transfected cells (see Figure 3.3.4.1 B). However, after transfection of cells with a combination of KCNQ1_WT and KCNE1 plasmids, the output current increased significantly (22.25 ± 1.91 pA/pF), while in cells containing a combination of KCNQ1_MUT and KCNE1 plasmids, this current decreased

significantly (6.64 ± 0.49 pA/pF; see Figure 3.3.4.1 B). This signals disturbances in the function of the potassium channel formed by the KCNQ1 protein, which are likely to disrupt the duration of the action potential in cells.

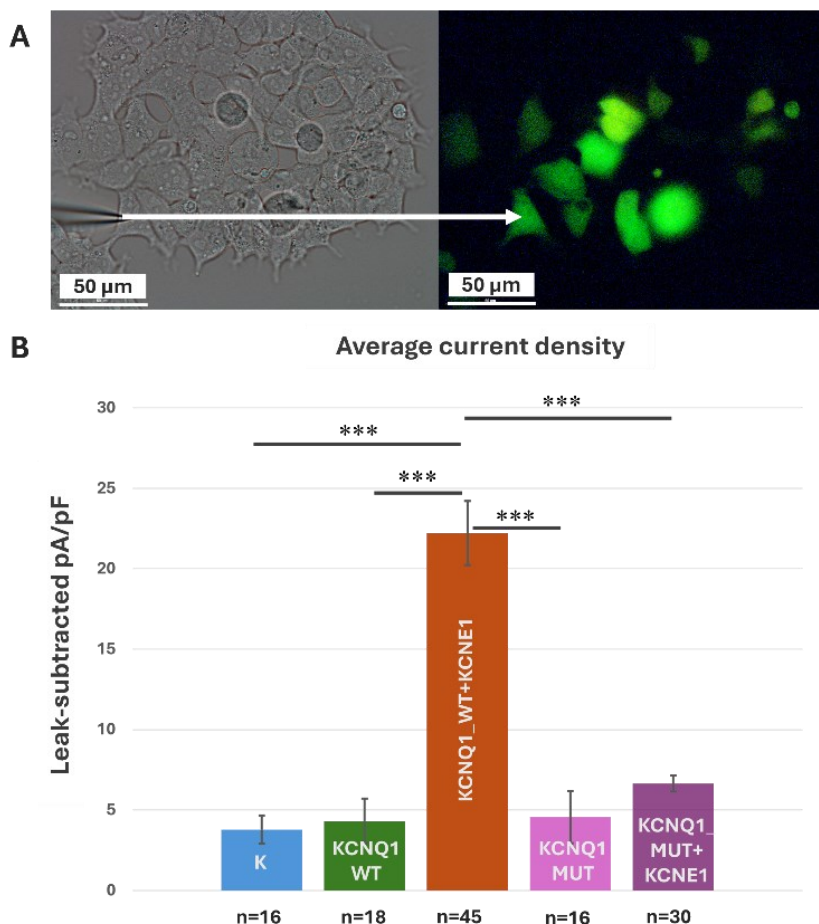


Figure 3.3.4.1. Electrophysiological measurements using the patch-clamp method in HEK293T cells. A – an example of how HEK293T cells are selected for electrophysiological measurements based on fluorescence intensity. Cells were transfected with plasmids encoding a green fluorescent protein. B – cells transfected with different combinations of plasmids. KCNQ1_WT – plasmid encoding the unmodified KCNQ1 protein, KCNQ1_MUT – plasmid encoding the *KCNQ1* variant identified in the proband, KCNE1 – plasmid encoding a subunit important for KCNQ1 protein activity, K – untransfected cells. Current density was measured at the end of a 2 s pulse up to 30 mV. The graph shows averages $\pm$ standard deviations, with statistical significance indicated by * $p \leq 0.05$ and *** $p \leq 0.005$. Photographs, electrophysiological measurements performed, and a graph created by A. Shelest, Prof. A. Alaburda.

The *KCNQ1* gene encodes a voltage-dependent potassium channel (Potassium Voltage-Gated Channel Subfamily Q Member 1) that is important for the repolarization phase of cardiac action potentials [201]. It is a subunit of a membrane-bound, pore-forming protein channel that is activated by delayed rectifier currents [201]. In the latter channel, KCNQ1 forms a complex with KCNE1 (Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 1), and this monomer of two subunits forms a tetramer that regulates the duration of the cardiac action potential. The KCNQ1/KCNE1 complex creates an outward potassium ion current (as it determines the movement of potassium ions out of the cell), which is a repolarizing ion current and terminates the cardiac action potential, while disturbances in the KCNQ1 subunit cause disturbances in the duration of the action potential [201]. KCNQ1 consists of three main regions: the S0 helix at the N-terminus, six transmembrane helices (S1–S6), and four intracellular helices (A–D) at the C-terminus [202]. A study by Shamgar and colleagues (2006) [203] used electrophysiological measurements to investigate other KCNQ1 protein changes (e.g., p.Ser373Pro, p.Ala371Thr) in subjects with long QT syndrome. The scientists were able to identify changes in the current in CHO (Chinese Hamster Cells) cells caused by the altered proteins, which disrupt the voltage-dependent inactivation of KCNQ1 channels [203]. This research strategy, along with the expression model created in HEK293T cells, substantiates the functional and clinical significance of the variant identified in our study, as a decrease in outgoing current was observed in cells with the altered KCNQ1 protein, indicating changes in the cell membrane potential.

Bioinformatic and molecular studies, as well as literature analysis, support the clinical and functional significance of the variant NM_000218.3:c.1111G>C in the *KCNQ1* gene. Changes in the function of the KCNQ1 potassium channel in HEK293T cells, as determined by electrophysiological measurements, suggest that impaired potassium ion movement across the cell membrane and altered action potential duration may be possible mechanisms of pathogenesis.

This section presents the results and their discussion, describing likely pathogenic and pathogenic SNVs, as well as a small number of cases analyzed from both molecular and cellular perspectives. Each of the four cases was approached using slightly different molecular and/or functional experimental strategies, which were primarily determined by the function of the gene-encoded product. The studies revealed potential molecular mechanisms of pathogenesis and specific molecular and cellular changes caused by a particular DNA variant. The implementation of the research strategies developed for these five individual cases and the results obtained illustrate not only the complexity of molecular processes, but also the influence of each DNA change and, in general, the genetic

and epigenetic background of the subject on the pathogenesis of the inherited disorder and the determination of its mechanism.

The research strategies and methods applied in this scientific work are potential diagnostic tools in clinical practice and personalized medicine. Based on the experience of diagnostic laboratories in various countries, transferring these studies to clinical diagnostic algorithms is a complex process that requires addressing methodological, analytical, and practical challenges. One of the obstacles is the global application of standardized protocols for sample processing, research strategies, and data generation to manage the comparability and reproducibility of results. Exome, genome, and transcriptome sequencing generate large datasets, the processing, storage, and interpretation of which require complex bioinformatic infrastructure and highly qualified specialists. Another crucial and time-consuming stage that requires significant human resources is the clinical validation of functional variant analysis, which aims to demonstrate the analytical accuracy, reproducibility, and clinical utility of the research strategy in diagnosing, predicting disease progression, or selecting the most appropriate treatment. The implementation of new tests and diagnostic algorithms is also limited by the necessary initial financial investment in the equipment required for the latest analysis algorithm, as well as the potentially high cost of functional tests. Undoubtedly, the integration of large-scale and functional tests is limited by the shortage of specialists capable of interpreting multi-layered data in a clinical context. However, considering foreign diagnostic laboratories that apply transcriptome or functional analysis, as well as the broader context of medical and human genetics and genomics worldwide, Lithuanian healthcare institutions are well-suited to implement new testing strategies once the main challenges have been overcome.

IV CONCLUSIONS

1. Next-generation sequencing identified single-nucleotide variants in known disease genes *RLIM*, *NAA15*, *WRN*, *DYNC1H1*, *DNMT3A*, and *KCNQ1*, as well as a small deletion in the *FAS* gene were classified as DNA changes of uncertain clinical significance or likely pathogenic variants and selected for molecular and/or functional studies.
2. After characterizing the structural rearrangements in the genome, candidate genes of clinical significance were identified:
 - 2.1 Protein-coding genes *KIT*, *ANKRD17*, *XYLT1*, exons of the deleted *MED13L* gene, and the miRNA cluster *MIR17HG* located within the microduplication region are likely to influence the development of the clinical phenotype of the probands and are potential candidates for subsequent functional investigations in cellular or model organism systems to elucidate their definitive roles in the molecular pathogenesis of the condition phenotype;
 - 2.2 Chromothripsis determined rearrangements and breakpoints of the resulting fragments influence the expression of 42 genes, with the *NTNG2* gene potentially contributing to the manifestation of specific clinical features observed in the proband.
3. The significance of the selected SNVs at the protein level was assessed using bioinformatic analysis and molecular studies:
 - 3.1 Protein structure and interaction modeling, along with DNA methylation analysis, determined that the c.1721T>A variant of the *RLIM* gene causes changes in the interaction and likely function of the encoded protein, which may disrupt ubiquitination in the proband's cells;
 - 3.2 Analysis of protein structure and immunoenzymatic assessment of protein levels revealed that the c.382C>T variant in the *NAA15* gene reduces the expression of the encoded protein, which likely disrupts N-terminal acetylation processes in the proband cells.
4. Molecular and/or functional studies were used to evaluate the impact of selected SNVs and small deletions on cellular processes:
 - 4.1 Experimental studies in fibroblasts with the c.1578del DNA change in the *WRN* gene, which causes protein shortening, revealed accelerated cellular replicative senescence;
 - 4.2 Gene and protein expression, confocal microscopy, and gene expression inhibition studies have established the influence of the *DYNC1H1* gene c.6405+1G>C variant, which causes a premature

- stop codon, on the decrease in DYNC1H1 protein levels, *KIFCI* gene expression, and altered Golgi complex recovery;
- 4.3 Experimental studies using fluorescence microscopy and flow cytometry showed that a small deletion in the FAS gene (c.16_30+1del), which shortens the protein, decreases the tendency of FAS receptor-induced apoptosis;
 - 4.4 *In vitro* protein studies, large-scale gene expression, and cfDNA sequencing studies have established the influence of the c.2324C>A variant of the *DNMT3A* gene, which causes an amino acid change, on the reduced activity of DNMT3A methyltransferase and altered gene expression;
 - 4.5 Experimental and electrophysiological studies in HEK293T cells of the *KCNQ1* gene c.1111G>C variant, which causes an amino acid change, have demonstrated a negative impact on the potential-dependent outward current and action potential.

PUBLIKACIJŲ SĄRAŠAS

Mokslinio darbo tema išspausdintos publikacijos:

1. **Petraitytė, Gunda**; Šiaurytė, Kamilė; Mikštienė, Violeta; Cimbališienė, Loreta; Kriaučiūnienė, Dovilė; Matulevičienė, Aušra; Utkus, Algirdas; Preikšaitienė, Eglė. A novel variant in the PDE4D gene is the cause of Acrodysostosis type 2 in a Lithuanian patient: a case report // BMC endocrine disorders. London : BMC. eISSN 1472-6823. 2021, vol. 21, art. no. 71, p. [1-6]. DOI: 10.1186/s12902-021-00741-6. IF: 3,263
2. Siavrienė, Evelina; **Petraitytė, Gunda**; Burnytė, Birutė; Morkūnienė, Aušra; Mikštienė, Violeta; Rančelis, Tautvydas; Utkus, Algirdas; Kučinskas, Vaidutis; Preikšaitienė, Eglė. Compound heterozygous c.598_612del and c.1746-20C > G CAPN3 genotype cause autosomal recessive limb-girdle muscular dystrophy-1: a case report // BMC musculoskeletal disorders. London : BMC. eISSN 1471-2474. 2021, vol. 22, iss. 1, art. no. 1020, p. [1-8]. DOI: 10.1186/s12891-021-04920-3. IF: 2,562
3. **Petraitytė, Gunda**; Mikštienė, Violeta; Siavrienė, Evelina; Cimbališienė, Loreta; Maldžienė, Živilė; Rančelis, Tautvydas; Vaitėnienė, Evelina Marija; Ambrozaitytė, Laima; Dapkūnas, Justas; Dzindzalieta, Ramūnas; Prancevičienė, Erinija; Kučinskas, Vaidutis; Utkus, Algirdas; Preikšaitienė, Eglė. Donor splice site variant in SLC9A6 causes Christianson syndrome in a Lithuanian family: a case report // Medicina. Kaunas; Basel : LSMU ; MDPI AG. ISSN 1010-660X. eISSN 1648-9144. 2022, vol. 58, no. 3, art. no. 351, p. [1-11]. DOI: 10.3390/medicina58030351. IF: 2,60
4. Siavrienė, Evelina; Maldžienė, Živilė; Mikštienė, Violeta; **Petraitytė, Gunda**; Rančelis, Tautvydas; Dapkūnas, Justas; Burnytė, Birutė; Benušienė, Eglė; Sasnauskienė, Aušra; Grikinienė, Jurgita; Griškevičiūtė, Eglė; Utkus, Algirdas; Preikšaitienė, Eglė. PIGN-related disease in two Lithuanian families: a report of two novel pathogenic variants, molecular and clinical characterisation // Medicina. Basel : LSMU ; MDPI. ISSN 1010-660X. eISSN 1648-9144. 2022, vol. 58, no. 11, art. no. 1526, p. [1-13]. DOI: 10.3390/medicina58111526. IF: 2,60
5. Siavrienė, Evelina; **Petraitytė, Gunda**; Mikštienė, Violeta; Maldžienė, Živilė; Sasnauskienė, Aušra; Žitkutė, Vilmantė; Ambrozaitytė, Laima; Rančelis, Tautvydas; Utkus, Algirdas; Kučinskas, Vaidutis; Preikšaitienė, Eglė. Molecular and functional characterisation of a novel intragenic 12q24.21 deletion resulting in MED13L haploinsufficiency syndrome // Medicina. Basel : MDPI. ISSN 1010-660X. eISSN 1648-

9144. 2023, vol. 59, iss. 7, art. no. 1225, p. [1-14]. DOI: 10.3390/medicina59071225. IF:2,40
6. Žebrauskienė, Dovilė; Sadauskienė, Eglė; Dapkūnas, Justas; Kairys, Visvaldas; Balčiūnas, Joris; Konovalovas, Aleksandras; Masiulienė, Rūta; **Petraitytė, Gunda**; Valevičienė, Nomeda Rima; Matačiūnas, Mindaugas; Barysienė, Jūratė; Mikštienė, Violeta; Tomkuvienė, Miglė; Preikšaitienė, Eglė. Aortic disease and cardiomyopathy in patients with a novel DNMT3A gene variant causing Tatton-Brown–Rahman syndrome // Clinical epigenetics. London : BioMed Central Ltd. ISSN 1868-7075. eISSN 1868-7083. 2024, vol. 16, iss. 1, art. no. 76, p. [1-15]. DOI: 10.1186/s13148-024-01686-y. IF: 4,80

Kitos publikacijos:

1. Druta, Jovita Patricija; **Petraitytė, Gunda**; Sasnauskienė, Aušra; Preikšaitienė, Eglė. Werner syndrome caused by homozygous frameshift variant c.1578del in WRN = Wernerio sindromas, kurį lemia rėmelio poslinkį sukeliantis homozigotinis WRN geno c.1578del variantas // Acta medica Lituanica. Vilnius : Vilniaus universiteto leidykla. ISSN 1392-0138. eISSN 2029-4174. 2024, vol. 31, no. 2, p. 279-285. DOI: 10.15388/Amed.2024.31.2.9.
2. **Petraitytė, Gunda**; Preikšaitienė, Eglė; Mikštienė, Violeta. Genome editing in medicine: tools and challenges = Genomo redagavimas medicinoje: įrankiai ir iššūkiai // Acta medica Lituanica. Vilnius : Vilniaus universiteto leidykla. ISSN 1392-0138. eISSN 2029-4174. 2021, vol. 28, no. 2, p. 266-280. DOI: 10.15388/Amed.2021.28.2.8.
3. Siavrienė, Evelina; **Petraitytė, Gunda**; Mikštienė, Violeta; Rančelis, Tautvydas; Maldžienė, Živilė; Morkūnienė, Aušra; Byčkova, Jekaterina; Utkus, Algirdas; Kučinskas, Vaidutis; Preikšaitienė, Eglė. A novel CHD7 variant disrupting acceptor splice site in a patient with mild features of CHARGE syndrome: a case report // BMC Medical Genetics. London : BioMed Central Ltd. ISSN 1471-2350. 2019, vol. 20, art. no. 127, p. [1-7]. DOI: 10.1186/s12881-019-0859-y.
4. Siavrienė, Evelina; Mikštienė, Violeta; Radzevičius, Darius; Maldžienė, Živilė; Rančelis, Tautvydas; **Petraitytė, Gunda**; Tamulytė, Giedrė; Kavaliauskienė, Ingrida; Šarkinas, Laurynas; Utkus, Algirdas; Kučinskas, Vaidutis; Preikšaitienė, Eglė. Novel GLI3 variant causes Greig cephalopolysyndactyly syndrome in three generations of a Lithuanian family // Molecular genetics & genomic medicine. Hoboken : Wiley. ISSN 2324-9269. eISSN 2324-9269. 2019, vol. 7, iss. 9, art. no. e878, p. [1-7]. DOI: 10.1002/mgg3.878.

PRANEŠIMŲ SĄRAŠAS

Žodiniai pranešimai tarptautiniuose moksliniuose renginiuose:

1. **Petraitytė, Gunda** (pranešėja). *Delineation of the genetic causes of rare diseases using NGS and functional studies. Baltic Congress in Laboratory Medicine*. 2022 m., Estija.

Stendiniai pranešimai, tezės, paskelbtos tarptautiniuose moksliniuose renginiuose:

1. **G. Petraitytė**, Ž. Maldžienė, V. Mikštienė, E. Siavrienė, T. Rančelis, V. Kučinskas, E. Preikšaitienė. *Molecular analysis of a novel donor splice site variant in DYNC1H1. European Human Genetics Virtual Conference 2021*. Virtuali konferencija, 2021 m.
2. **G. Petraitytė**, E. Siavrienė, Ž. Maldžienė, V. Mikštienė, I. Tavorienė, L. Griškevičius, E. Preikšaitienė. *A small deletion in FAS causes splicing disturbances in a patient with suspected ALPS. European Human Genetics Conference*. Austrija, 2022 m.
3. **G. Petraitytė**, E. Preikšaitienė. *Molecular investigation of a DNA variant's pathogenicity: a splice site variant in DYNC1H1*. Jungtinis tarptautinis renginys „Joint International Meeting: 22nd EAA Congress, 15th ISGA Congress, 5th International Conference on Evolutionary Medicine“. Lietuva, 2022 m.
4. J. Songailaitė, **G. Petraitytė**, E. Preikšaitienė. *Novel truncating variant in DYNC1H1 leading to intellectual disability and autism. The COINS*. Lietuva, 2022 m.

CURRICULUM VITAE

Vardas Pavardė **Gunda Petraitytė**
El. pašto adresas gunda.petraityte@mf.vu.lt
Adresas Santariškių g. 2, LT-08661 Vilnius



DARBINĖ VEIKLA

2025 m. spalį **Tyrėja**
2025 m. gruodį VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius

2025 m. spalį **Lektorė**
VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius

2025 m.
Dabartinis **Medicinos genetikė**
VUL Santaros klinikos Medicininės genetikos centras, Vilnius

2025 m. kovas **Jaunesnioji mokslo darbuotoja**
2025 m. rugsėjį VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius

2019 m.
2024 m. **Jaunesnioji mokslo darbuotoja**
VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius

2024 m.
2025 m. **Projekto vykdytoja**
VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius
VU Mokslo skatinimo fondo projektas „Platforma įvairių genomo redagavimo įrankių tikslumo įvertinimui“(MSF-JM-05/2024). Projekto vadovas dr. Paulius Toliušis.

2021 m.
2024 m. **Projekto vykdytoja**
VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius

Lietuvos mokslo tarybos finansuojamas projektas „Retų paveldimų ligų genetinių priežasčių integraciniai tyrimai“, akronimas ATGC (S-MIP-21-15). Projekto vadovė prof. dr. Eglė Preikšaitienė.

2021 m. **Projekto vykdytoja**
2022 m. VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius
VU Mokslo skatinimo fondo projektas „CRISPR-Cas13 technologijos taikymas tiriant atsparumą chemoterapijoms vaistams“ (MSF-JM-2/2021). Projekto vadovė Vilmantė Žitkutė.

2020 m. **Projekto vadovė**
2021 m. VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius
VU Mokslo skatinimo fondo projektas „Genomo ir transkriptomo analizė retų paveldimų ligų patogenezės tyrimuose“ (MSF-JM-2/2020).

2017 m. **Projekto specialistė**
2020 m. VU MF BMI Žmogaus ir medicininės genetikos katedra, Vilnius
Lietuvos mokslo tarybos finansuojamas projektas „Intelektinės negalios ir įgimtų anomalijų genetinės architektūros identifikavimas“, akronimas INGENES (S-MIP-17-19). Projekto vadovė doc. dr. Eglė Preikšaitienė.

IŠSILAVINIMAS

2019 m. **Medicinos mokslo krypties doktorantūros studijos**
2024 m. Vilniaus universitetas, Vilnius

2017 m. **Medicinos genetikos magistrantūros studijos**
2019 m. Vilniaus universitetas, Vilnius

2013 m. **Genetikos bakalauro studijos**
2017 m. Vilniaus universitetas, Vilnius

PADĖKA

Nuoširdžiausius padėkos žodžius tariu:

mokslinei vadovei prof. dr. Eglei Preikšaitienei už nuolatinį paskatinimą, kantrybę ir suteiktą galimybę gilintis į retų paveldimų ligų genetikos, genomikos ir funkcinį tyrimų subtilybes;

mokslinei konsultantei doc. dr. Aušrai Sasnauskienei už mokslinius ir gyvenimiškus patarimus vykdant funkcinį tyrimą ir rašant disertaciją;

visiems tiriamiesiems, jų šeimos nariams, sutikusiems dalyvauti šiame tyrime ir leidusiems analizuoti jų klinikinius ir genetinius duomenis;

disertacijos Gynimo tarybos pirmininkui prof. dr. Juozui Rimantui Lazutkai bei narėms prof. dr. Audronei Jakaitienei, dr. Daliai Seaborne, doc. dr. Sonatai Šaulytei–Trakymienei ir prof. dr. Rasai Ugenskienei už įdomius klausimus diskusijai;

recenzentėms doc. dr. Ingridai Domarkienei, dr. Violetai Jonušienei, doc. dr. Sonatai Šaulytei–Trakymienei už pastabas, padėjusias patobulinti šį darbą;

VU Medicinos fakulteto Doktorantūros mokyklos vadovei prof. dr. Janinai Tutkuvienei ir doktorantūros specialistei Rasai Janulionienei už konsultacijas ir pagalbą visais doktorantūros studijų metu iškilusiais klausimais;

VU MF BMI ŽMGK kolegėms dr. Violetai Mikštienai ir dr. Evelinai Siavrienei už nepailstamą palaikymą ir pagalbą visuose disertacinio darbo vykdymo etapuose;

VU MF BMI ŽMGK ir VU GMC mokslininkų bei specialistų komandai, vykdžiusiai INGENES ir ATGC projektus, kurios bendradarbiavimas moksliniuose tyrimuose yra neatsiejama šio darbo dalis;

visam VU MF BMI Žmogaus ir medicininės genetikos katedros kolektyvui už bendrą darbą ir patarimus;

VU GMC kolegei dr. Vilmantei Žitkutei už palaikymą, pagalbą ir bendrus interesus, padėjusius įveikti doktorantūros studijų etapą;

draugėms Gintarei, Auksei, Evelinai ir Violetai už unikalias įžvalgas, palaikymą ir turiningą laiką (net ir virtualų) kartu;

už supratimą, neišsemiamą kantrybę, įkvėpimą, humoro jausmą ir visokeriopą paramą noriu ypač padėkoti savo artimiausiems žmonėms – vyrui Martynui, broliams Aurimui ir Venantui, mamai ir tėčiui, močiutei ir tetai Rasai, taip pat augintiniui Botui, kuris yra neatskiriama šios šeimos dalis.

UŽRAŠAMS

UŽRAŠAMS

Vilniaus universiteto leidykla
Saulėtekio al. 9, III rūmai, LT-10222 Vilnius
El. p. info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Tiražas 20 egz.