



UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA
ŠTUDIJ BIOTEHNOLOGIJE

Anamarija HABIČ

**DOLOČANJE FUNKCIONALNIH
POLIMORFIZMOV V ZELO OHRANJENIH
GENOMSKIH REGIJAH PRI ČLOVEKU
Z BIOINFORMACIJSKIMI ORODJI**

DIPLOMSKO DELO

Univerzitetni študij - 1. stopnja

Ljubljana, 2017

UNIVERZA V LJUBLJANI
BIOTEHNIŠKA FAKULTETA
ŠTUDIJ BIOTEHNOLOGIJE

Anamarija HABIČ

**DOLOČANJE FUNKCIONALNIH POLIMORFIZMOV
V ZELO OHRANJENIH GENOMSKIH REGIJAH PRI ČLOVEKU
Z BIOINFORMACIJSKIMI ORODJI**

DIPLOMSKO DELO
Univerzitetni študij - 1. stopnja

**IDENTIFICATION OF FUNCTIONAL POLYMORPHISMS
IN ULTRACONSERVED REGIONS IN HUMAN
USING BIOINFORMATICS TOOLS**

B. SC. THESIS
Academic Study Programmes

Ljubljana, 2017

Diplomski seminar je zaključek univerzitetnega študija – 1. stopnja Biotehnologija.

Študijska komisija 1. in 2. stopnje Študija biotehnologije je za mentorja diplomskega seminarja imenovala izr. prof. dr. Tanjo Kunej in za somentorja izr. prof. dr. Janeza Konca.

Komisija za oceno in predstavitev:

Predsednik: prof. dr. Mojca NARAT
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Članica: izr. prof. dr. Polona JAMNIK
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za živilstvo

Članica: izr. prof. dr. Tanja KUNEJ
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Član: izr. prof. dr. Janez KONC
Kemijski inštitut, Ljubljana

Datum predavitve: 12. 9. 2017

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

ŠD	Du1
DK	UDK601.4:577.2:575.112(043.2)
KG	bioinformatika/genom/UCR/BLAT/BioMart/gen/ polimorfizem/fenotip/bolezen
AV	HABIČ, Anamarija
SA	KUNEJ, Tanja (mentor)/KONC, Janez (somentor)
KZ	SI-1000 Ljubljana, Jamnikarjeva 101
ZA	Univerza v Ljubljani, Biotehniška fakulteta, Študij biotehnologije, Univerzitetni študijski program prve stopnje Biotehnologija
LI	2017
IN	DOLOČANJE FUNKCIONALNIH POLIMORFIZMOV V ZELO OHRANJENIH GENOMSKIH REGIJAH PRI ČLOVEKU Z BIOINFORMACIJSKIMI ORODJI
TD	Diplomsko delo (Univerzitetni študij - 1. stopnja)
OP	VI, 20, [17] str., 2 pregl., 3 sl., 3 pril., 81 vir.
IJ	sl
JI	sl/en
AI	Zelo ohranjene genomske regije (angl. <i>ultraconserved regions</i>; UCRs) so definirane kot zaporedja DNA, ki so stoodstotno ohranjena v ortolognih regijah genomov različnih vrst. Novejše raziskave kažejo, da so v UCR-jih prisotni tudi polimorfizmi. Večina UCR-jev še ni bila funkcionalno anotirana, zato sta njihov vpliv na fenotip oz. morebitna vpletenost v razvoj bolezni še precej neznana. Namen diplomskega dela je zato bil: 1. z bioinformacijskimi orodji na novo določiti genomske lokacije predhodno odkritih UCR-jev v skladu z najnovejšo različico genoma človeka, 2. preveriti, s katerimi geni se prekrivajo UCR-ji, 3. identificirati polimorfizme znotraj UCR-jev in 4. preveriti, ali so kateri izmed njih povezani s fenotipom/bolezni. Z analizo z orodjem BioMart smo ugotovili, da je 25 % UCR-jev medgenskih, ostali pa se vsaj delno prekrivajo z geni. Znotraj ohranjenih regij smo določili 30139 polimorfizmov, od katerih jih ima v genomskem brskalniku Ensembl 183 anotirano povezavo s fenotipom. Za 37 polimorfizmov je bilo možno pridobiti znanstvene članke, v katerih so poročali povezavo s fenotipom. Polimorfizmi so povezani z različnimi boleznimi, med drugim z družinsko adenomatozno polipozo, adolescentno idiopatsko skoliozo, amiotrofično lateralno sklerozo, mišičnimi distrofijami in spastično paraplegijo. Rezultati so osnova za nadaljnje odkrivanje funkcije UCR-jev in identifikacijo pomembnih regij znotraj teh odsekov.

KEY WORDS DOCUMENTATION

ND Du1

DC UDC601.4:577.2:575.112(043.2)

CX bioinformatics/genome/UCR/BLAT/BioMart/gene/ polymorphism/phenotype/disease

AU HABIČ, Anamarija

AA KUNEJ, Tanja (supervisor)/ KONC, Janez (co-advisor)

PP SI-1000 Ljubljana, Jamnikarjeva 101

PB University of Ljubljana, Biotechnical Faculty, Academic Study Programme in Biotechnology

PY 2017

TI IDENTIFICATION OF FUNCTIONAL POLYMORPHISMS IN ULTRACONSERVED REGIONS IN HUMAN USING BIOINFORMATICS TOOLS

DT B. Sc. Thesis (Academic Study Programmes)

NO VII, 20, [17] p., 2 tab., 3 fig., 3 ann., 81 ref.

LA sl

AL sl/en

AB Ultraconserved regions (UCRs) are defined as DNA sequences, which are absolutely conserved between orthologous genomic regions of multiple species. However, recent studies have proved the presence of polymorphisms within UCRs. Most of the UCRs have not yet been functionally annotated, therefore their effect on phenotype and involvement in disease development remain substantially unknown. The aims of the thesis were to: 1. remap previously reported UCRs according to the latest human genome release using bioinformatics tools, 2. identify genes overlapping UCRs, 3. identify polymorphisms within UCRs and 4. check whether any associations between UCR polymorphisms and phenotype/diseases exist. Using BioMart data mining tool we showed that 25 % of UCRs are intergenic, while the rest overlap genes. Our analysis performed by BioMart identified 30,139 polymorphisms within UCRs. Among these, 183 have been annotated to be associated with phenotype according to the Ensembl genome browser. For 37 among 183 polymorphisms it was possible to obtain published literature reporting associations with phenotype. Polymorphisms are associated with various diseases, for example with familial adenomatous polyposis, adolescent idiopathic scoliosis, amyotrophic lateral sclerosis, muscle dystrophies and spastic paraplegia. Our results serve as a basis for further investigation of UCRs' functions and identification of important regions within these segments.

KAZALO VSEBINE

	Str.
KLJUČNA DOKUMENTACIJSKA INFORMACIJA	III
KEY WORDS DOCUMENTATION	IV
KAZALO VSEBINE	V
KAZALO PREGLEDNIC	VI
KAZALO SLIK	VI
KAZALO PRILOG	VI
OKRAJŠAVE IN SIMBOLI	VI
1 UVOD	1
2 PREGLED OBJAV	1
3 MATERIALI IN METODE	3
3.1 IZBIRA UCR-JEV, DOLOČITEV NJIHOVIH GENOMSKIH LOKACIJ IN IZRAČUN NEKATERIH STATISTIČNIH PARAMETROV	3
3.2 DOLOČITEV GENOV, ZNOTRAJ KATERIH LEŽIJO UCR-JI	4
3.3 DOLOČITEV POLIMORFIZMOV ZNOTRAJ UCR-JEV	4
3.3.1 Določitev s fenotipom povezanih polimorfizmov znotraj UCR-jev	4
4 REZULTATI	5
4.1 STATISTIČNI PARAMETRI IN GENOMSKA RAZPOREDITEV UCR-JEV	5
4.2 GENI, ZNOTRAJ KATERIH SE NAHAJAJO UCR-JI	6
4.3 POLIMORFIZMI ZNOTRAJ UCR-JEV	7
4.3.1 S fenotipom povezani polimorfizmi znotraj UCR-jev	8
5 RAZPRAVA	11
5.1 STATISTIČNI PARAMETRI IN GENOMSKA RAZPOREDITEV UCR-JEV	11
5.2 GENI, ZNOTRAJ KATERIH SE NAHAJAJO UCR-JI	11
5.3 POLIMORFIZMI ZNOTRAJ UCR-JEV	12
5.3.1 S fenotipom povezani polimorfizmi znotraj UCR-jev	12
5.4 OMEJITVE RAZISKAVE	13
6 SKLEPI	13
7 VIRI	14
ZAHVALA	
PRILOGE	

KAZALO PREGLEDNIC

Preglednica 1: Število genov določenega biotipa, v katerih se nahajajo UCR-ji	6
Preglednica 2: Povezave med polimorfizmi in fenotipi, za katere je bilo možno pridobiti znanstveni članek, v katerem so poročali povezavo	9

KAZALO SLIK

Slika 1: Prikaz razmerij med številom UCR-jev na kromosomu in dolžino kromosoma	5
Slika 2: Prikaz razporeditve UCR-jev glede na njihovo dolžino	6
Slika 3: Prikaz števila UCR-jev glede na število polimorfizmov, ki jih vsebujejo	7

KAZALO PRILOG

Priloga A: Shema poteka dela	
Priloga B: Geni, razporejeni glede na število UCR-jev, ki jih vsebujejo	
Priloga C: Preglednica povezav med polimorfizmi in fenotipi, ki so anotirane v genomskem brskalniku Ensembl in smo jih dobili v izpisu rezultatov analize z orodjem BioMart	

OKRAJŠAVE IN SIMBOLI

AS	protismerni prepis (angl. <i>antisense</i>)
bp	bazni par (angl. <i>base pair</i>)
FAP	družinska adenomatozna polipoza (angl. <i>familial adenomatous polyposis</i>)
GWAS	raziskava na ravni celotnega genoma (angl. <i>genome-wide association study</i>)
ITM	indeks telesne mase
lincRNA	dolga medgenska nekodirajoča RNA (angl. <i>long intergenic noncoding RNA</i>)
lncRNA	dolga nekodirajoča RNA (angl. <i>long noncoding RNA</i>)
miRNA	mikro RNA (angl. <i>microRNA</i>)
misc RNA	neuvrščena RNA (angl. <i>miscellaneous RNA</i>)
ncRNA	nekodirajoča RNA (angl. <i>non-coding RNA</i>)
rRNA	ribosomalna RNA (angl. <i>ribosomal RNA</i>)
SNP	polimorfizem posameznega nukleotida (angl. <i>single nucleotide polymorphism</i>)
TEC	predvidena protein-kodirajoča regija, ki jo je potrebno še eksperimentalno validirati (angl. <i>to be experimentally confirmed</i>)
T-UCR	prepisana zelo ohranjena genomska regija (angl. <i>transcribed-ultraconserved region</i>)
UCR	zelo ohranjena genomska regija (angl. <i>ultraconserved region</i>)
UTR	neprevedeno območje (angl. <i>untranslated region</i>)

1 UVOD

Le dober odstotek človekovega genoma kodira proteine, ki so neposredno povezani z različnimi funkcijami, vendar obstaja vedno več dokazov, da to niso edine pomembne regije v DNA. Rezultati asociacijskih študij na ravni celotnega genoma so v zadnjih letih pokazali, da se tudi v nekodirajočih regijah nahaja veliko funkcionalno pomembnih elementov, ki sodelujejo pri uravnavanju izražanja genov. Eno izmed zelo uporabnih orodij pri odkrivanju genomskih regij s pomembno vlogo je primerjalna genomika. Evolucijska ohranjenost določene regije nakazuje, da nanjo deluje negativna selekcija, za kar najverjetneje obstaja razlog – pomembna vloga te regije. S primerjalnimi analizami genomov so ugotovili, da je negativni selekciji podvrženih vsaj 5 % genomov sesalcev (Siepel in sod., 2005).

Zelo ohranjene regije v genomu človeka so s poravnavo genoma človeka, miši in podgane prvi določili Bejerano in sod. (2004). Odkrili so 481 medvrstno ohranjenih zaporedij DNA, ki so daljša od 200 bp (angl. *ultraconserved regions*, UCRs). Sklepajo, da gre za regije z več funkcijami in da je to vzrok za njihovo izjemno ohranjenost. Drug dobro znan primer ekstremno ohranjenih regij so npr. tiste, ki kodirajo rRNA. Le-ta mora oblikovati ustrezno strukturo in hkrati vezati tudi proteine.

Čeprav je od prve identifikacije zelo ohranjenih regij v genomu človeka minilo že več kot desetletje, njihova vloga še vedno ostaja precej neznana. Za nekatere ohranjene regije že vemo, da so vpletene v pojav bolezni pri človeku, večina pa jih še ni bila funkcionalno anotirana, zato sta njihov vpliv na fenotip oz. morebitna vloga pri razvoju bolezni še nepojasnjena. Namen diplomskega dela je zato bil: 1. na novo določiti genomske lokacije 2189 vsaj 200 bp dolgih predhodno odkritih UCR-jev (Stephen in sod., 2008) v skladu z najnovejšo različico genoma človeka, 2. preveriti, s katerimi geni se prekrivajo UCR-ji, 3. z bioinformacijskimi orodji identificirati polimorfizme znotraj teh regij in 4. preveriti, ali so kateri izmed teh polimorfizmov povezani z boleznimi oz. s fenotipskimi lastnostmi.

2 PREGLED OBJAV

Zelo ohranjene genomske regije (angl. *ultraconserved regions*; UCRs) so zaporedja DNA, dolga vsaj 200 bp, ki so stodstotno identična (brez insercij in delecij) v ortolognih regijah genomov človeka, miši in podgane. Bejerano in sod. (2004) so odkrili 481 takih odsekov. Velika večina se jih v skoraj identični obliki pojavlja tudi v ortolognih genomskih regijah kokoši in psa, približno dve tretjini pa je zelo ohranjenih tudi v genomu ribe fugu. Iz teh rezultatov sledi, da so bile te regije pri mnogih vrstah podvržene negativni selekciji več kot tristo ali celo štiristo milijonov let.

UCR-ji se nahajajo na vseh kromosomih, razen na kromosomih 21 in Y, in so pogosto združeni v gruče. 111 UCR-jev je vsaj delno eksonskih, 256 je neeksonskih, 114 pa potencialno eksonskih. Eksonski UCR-ji se večinoma nahajajo znotraj genov, povezanih s procesiranjem RNA, neeksonski, ki ležijo v intronih ali pa v medgenskih regijah, pa so najverjetneje povezani z uravnavanjem prepisovanja na ravni DNA. Precej jih npr. leži navzgor od genov, povezanih z zgodnjim razvojem, in bi lahko bili oddaljeni ojačevalci teh

genov (Bejerano in sod., 2004). Ohranjene regije imajo prav tako lahko vlogo transkripcijskih koaktivatorjev (Feng in sod., 2006) ali pa sodelujejo pri alternativnem izrezovanju intronov (Ni in sod., 2007) ali epigenetskih modifikacijah (Lee in sod., 2006).

Po objavah genomov novih vrst je postalo jasno, da UCR-ji, ki so jih določili Bejerano in sod. (2004), niso popolnoma ohranjeni pri drugih placentalnih sesalcih in da obstaja tudi veliko krajših ohranjenih genomskih odsekov. Woolfe in sod. (2005) so s poravnavo nekodirajočih regij človekovega genoma in genoma ribe napihivalke identificirali skoraj 1400 zelo ohranjenih zaporedij, daljših od 100 bp. Stephen in sod. (2008) so s poravnavo genomov človeka, psa, krave, miši in podgane določili 13736 UCR-jev, dolgih vsaj 100 bp, ki so identični pri vsaj treh od izbranih petih placentalnih sesalcev. Od teh zaporedij jih je 2189 dolgih vsaj 200 bp.

Ahituv in sod. (2007) so proučevali vlogo štirih nekodirajočih UCR-jev, ki delujejo kot ojačevalci in ležijo v bližini genov, katerih izražanje se izrazito odraža na fenotipu. Izvedli so poskus pri miših in ugotovili, da preživijo in ostanejo fenotipsko normalni in plodni tudi osebkii z eno manjkajočo kopijo in celo osebkii brez obeh kopij UCR. Iz tega sledi, da proučene štiri regije niso nujno potrebne za preživetje in njihova odsotnost v laboratorijskih pogojih ne povzroči očitnih fenotipskih sprememb.

Calin in sod. (2007) so poročali, da se mnogi UCR-ji nahajajo na fragilnih mestih (angl. *fragile sites*) v genomu oz. v regijah, ki so povezane z rakom. Pokazali so, da se velik delež UCR-jev prepisuje – med njimi so tudi UCR-ji, ki se nahajajo v medgenskih regijah. Prepise, ki predstavljajo novo skupino nekodirajočih RNA (angl. *non-coding RNA*; ncRNA), so poimenovali T-UCR-ji (angl. *transcribed-ultraconserved regions*). Le-ti bi lahko delovali kot protismerni inhibitorji protein-kodirajočih genov ali drugih ncRNA, npr. mikro RNA (angl. *micro RNA*; miRNA), ki imajo pomembno vlogo v procesu tumorigeneze (Calin in sod., 2007; Lujambio in sod., 2010; Scaruffi, 2011).

Calin in sod. (2007) so primerjali normalna in rakava tkiva in ugotovili, da je izražanje T-UCR-jev spremenjeno pri različnih levkemijah in karcinomih – za specifično bolezen so značilni specifični vzorci izražanja (angl. *expression profile*). Dokazano je bilo, da je izražanje UCR-jev spremenjeno pri kronični limfocitni levkemiji, kolorektalnem in hepatocelularnem raku (Calin in sod., 2007), nevroblastomu (Mestdagh in sod., 2010), raku prostate (Hudson in sod., 2013), adenokarcinomu trebušne slinavke (Jiang in sod., 2016), raku sečnega mehurja (Olivieri in sod., 2016) in raku dojke (Marini in sod., 2017). Scaruffi in sod. (2009) so ugotovili, da je na podlagi analize vzorca izražanja T-UCR-jev in miRNA z RT-qPCR (angl. *reverse transcription quantitative real-time PCR*) mogoče napovedati, ali bo bolnik z nevroblastomom preživel le kratek čas ali dolgoročno. Z dodatno validacijo na večjem številu vzorcev bi lahko razvili prognostični test, ki bi napovedal raven tveganja pri posameznem bolniku.

Lujambio in sod. (2010) so ugotovili, da so v tumorskih tkivih CpG (5'-citozin-fosfat-gvanin-3') otoki T-UCR-jev pogosto hipermetilirani. Honma in sod. (2017) so to epigenetsko spremembo dokazali na primeru raka na želodcu. Z metodo bisulfitnega sekvenciranja so odkrili hipermetilirano regijo navzgor od ohranjene regije Uc.160+. Izražanje Uc.160+ v

obolelem mukoznem tkivu je bilo močno znižano, kar bi lahko pomenilo, da Uc.160+ zavira rast tumorjev.

Čeprav so UCR-ji po definiciji identični v referenčnih genomih različnih vrst, se znotraj teh regij pri predstavnikih določene vrste lahko pojavljajo polimorfizmi. Ker na UCR-je deluje negativna selekcija, obstaja precejšnja verjetnost, da so polimorfizmi znotraj teh regij povezani s fenotipskimi lastnostmi. To še posebej velja za tiste polimorfizme, ki se nahajajo v genomskih regijah, ki so domnevno povezane z boleznimi (npr. v bližini kandidatnih genov ali znotraj vezanih lokusov, ki so v vezavnem neravnovesju) (Chiang in sod., 2012). Bejerano in sod. (2004) so ob pregledu 106767 bp ohranjenih zaporedij odkrili le šest eksperimentalno potrjenih SNP-jev, kar je približno dvajsetkrat manj, kot bi jih pričakovali znotraj tako dolgega zaporedja. Chen in sod. (2007) so v podatkovni zbirki dbSNP znotraj UCR-jev odkrili 102 SNP-ja, od katerih jih je bilo 24 potrjenih s strani vsaj dveh raziskovalnih skupin. Genotipizirali so naključen vzorec več kot 600 fenotipsko normalnih ljudi. Rezultati so pokazali, da za vsakega od analiziranih 12 SNP-jev, ki se nahajajo znotraj UCR-jev, obstaja vsaj en osebek, ki je homozigot za redek alel – to stanje očitno ni letalno in ne povzroča nujno opaznih fenotipskih sprememb. Wojcik in sod. (2010) so sekvencirali 28 UCR-jev pri 95 Kavkazijcih (skupaj 9634 bp ohranjenih zaporedij). Odkrili so šest SNP-jev, kar pomeni, da je SNP povprečno prisoten na vsakih 1572 bp. Gostota SNP-jev v ohranjenih regijah je torej približno šestkrat manjša od povprečja v genomu, ki znaša en SNP na vsakih 279 bp. Istih 28 UCR-jev so sekvencirali tudi v vzorcih periferne krvi 35 bolnikov s kronično limfocitno levkemijo in v tumorskem tkivu 39 bolnikov z rakom debelega črevesa in danke. V povprečju so mutacijo v tumorski DNA odkrili na vsakih 90 bp. Somatski polimorfizmi znotraj analiziranih UCR-jev so torej pri bolnikih z rakom približno trikrat pogostejši kot pri zdravih ljudeh.

Yang in sod. (2008) so pokazali, da obstaja povezava med dvema polimorfizmoma (rs9572903 in rs2056116) znotraj UCR-jev in tveganjem za pojav raka dojke, Lin in sod. (2012) pa so odkrili povezavo med znotraj UCR-jev ležečima rs7849 in rs1399685 in dovzetnostjo za pojav raka debelega črevesa in danke. SNP rs606231388 v uc.325, ki se nahaja približno 150 kbp navzdol od gena *PAX6*, znotraj cis-regulatornega elementa, imenovanega SIMO, povzroča aniridijo. Ta točkovna mutacija prepreči vezavo proteina *PAX6* na avtoregulatorno vezavno mesto in tako prepreči delovanje ojačevalca, kar spremeni izražanje *PAX6* in povzroči bolezen (Bhatia in sod., 2013). Bao in sod. (2016) so predlagali, da bi polimorfizme znotraj ohranjenih regij lahko uporabili kot prognostične biooznačevalce za oceno odziva na zdravljenje raka prostate in oceno verjetnosti preživetja te bolezni.

3 MATERIALI IN METODE

3.1 IZBIRA UCR-JEV, DOLOČITEV NJIHOVIH GENOMSKIH LOKACIJ IN IZRAČUN NEKATERIH STATISTIČNIH PARAMETROV

Izhodišče za analize je bil seznam 13736 UCR-jev, dolgih vsaj 100 bp, ki so identični v genomih vsaj treh od petih placentalnih sesalcev (človek, pes, krava, miš, podgana) (Stephen in sod., 2008). S seznama smo izbrali vse UCR-je, ki so dolgi vsaj 200 bp (tj. 2189 UCR-jev), in jim z orodjem BLAT (angl. *BLAST-like Alignment Tool*), dostopnim na spletni strani

genomskega brskalnika Ensembl (<http://www.ensembl.org/Multi/Tools/Blast>), določili nove genomske lokacije glede na najnovejšo objavljeno različico človekovega genoma (GRCh38). S pomočjo programa Microsoft Excel smo določili povprečno dolžino analiziranih UCR-jev in najdaljše UCR-je ter jih glede na dolžine razporedili v razrede.

3.2 DOLOČITEV GENOV, ZNOTRAJ KATERIH LEŽIJO UCR-JI

Seznam genov, s katerimi se prekrivajo analizirani UCR-ji, smo dobili z orodjem BioMart, ki je namenjeno podatkovnemu rudarjenju in je dostopno na spletni strani brskalnika Ensembl (<http://www.ensembl.org/biomart>). Izbrali smo podatkovno zbirko *Ensembl Genes 87*, za podatkovni niz (angl. *dataset*) pa *Homo sapiens genes (GRCh38.p7)*. Orodje omogoča vnos in analizo do 500 genomskih lokacij hkrati (zavihek *Filters/REGION/Multiple Chromosomal Regions*). V zavihku *Attributes* oz. v podzavihku *GENES* smo izbrali informacije, ki smo jih želeli pridobiti (*Gene ID*, *Associated gene name*, *Description*, *Chromosome/scaffold name*, *Gene start (bp)*, *Gene end (bp)*, *Gene type*). Rezultat analize je bila Excelova preglednica genov, znotraj katerih se nahajajo UCR-ji, vendar brez pripisanih UCR-jev. Da bi ugotovili, v katerih genih se nahaja največ UCR-jev, smo UCR-je glede na njihove genomske lokacije s pomočjo programske kode, napisane v programskem jeziku Python, razporedili v preglednico z geni in prešteli, koliko UCR-jev se nahaja v posameznem genu.

3.3 DOLOČITEV POLIMORFIZMOV ZNOTRAJ UCR-JEV

Polimorfizme, ki se nahajajo znotraj vsaj 200 bp dolgih UCR-jev, smo prav tako identificirali z orodjem BioMart. Za podatkovno zbirko smo izbrali *Ensembl Variation 86*, za podatkovni niz (angl. *dataset*) pa *Homo sapiens Short Variants (SNPs and indels excluding flagged variants) (GRCh38.p7)*. V polje *Filters/REGION/Multiple Chromosomal Regions* smo vnesli predhodno določene genomske lokacije UCR-jev (do 250 lokacij hkrati). V zavihku *Attributes* oz. v podzavihku *VARIANT ASSOCIATED INFORMATION* smo izbrali informacije, ki smo jih želeli pridobiti (*Variant name*, *Variant source*, *Chromosome name*, *Chromosome position start (bp)*, *Chromosome position end (bp)*). Kot rezultat smo dobili Excelovo preglednico s polimorfizmi in pripadajočimi informacijami, pri čemer pa polimorfizmi niso bili porazdeljeni v pripadajoče UCR-je. S pomočjo programske kode v programskem jeziku Python smo polimorfizme glede na njihovo lokacijo porazdelili v UCR-je in prešteli, koliko polimorfizmov se nahaja znotraj posameznega UCR-ja.

3.3.1 Določitev s fenotipom povezanih polimorfizmov znotraj UCR-jev

Tudi s fenotipom povezane polimorfizme znotraj UCR-jev smo poiskali s pomočjo orodja BioMart. Nastavitve so bile enake kot pri identifikaciji vseh znanih polimorfizmov znotraj teh regij, le da smo v zavihku *Attributes* oz. v podzavihku *VARIANT ASSOCIATED INFORMATION* pod *Variant Information* obkljukali še *Variant alleles* in *Clinical significance*; pod *Phenotype annotation* smo izbrali *Study description*, *Source name*, *Associated gene with phenotype*, *Phenotype description*, *Associated variant risk allele*; pod *Variant citations* pa *Title*, *Authors*, *Year* in *PubMed ID*.

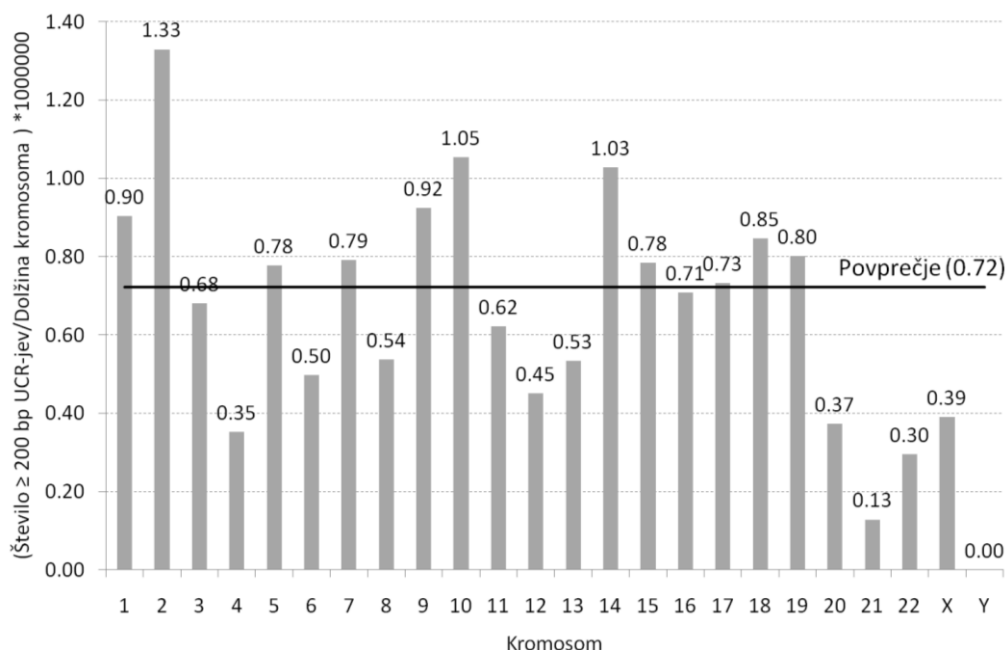
Dobljene povezave med polimorfizmi in fenotipi smo preverili v literaturi, navedeni v izpisu rezultatov orodja BioMart. Če raziskave niso bile navedene, smo jih poskušali ročno poiskati v navedenih podatkovnih zbirkah: Uniprot (zbirka proteinskih zaporedij in informacij o funkcijah proteinov), OMIM (zbirka človeških genov in genetskih bolezni, angl. *Online Mendelian Inheritance in Man*), ClinVar (arhiv poročil o povezavah med polimorfizmi in fenotipi), IIBDGC (povzetek rezultatov GWAS v povezavi s Chronovo boleznijo in ulceroznim kolitisom; angl. *International Inflammatory Bowel Disease Genetics Consortium*), GIANT (povzetek rezultatov GWAS v povezavi z antropometričnimi lastnostmi; angl. *Genome-scale Integrated Analysis of gene Networks in Tissues*).

4 REZULTATI

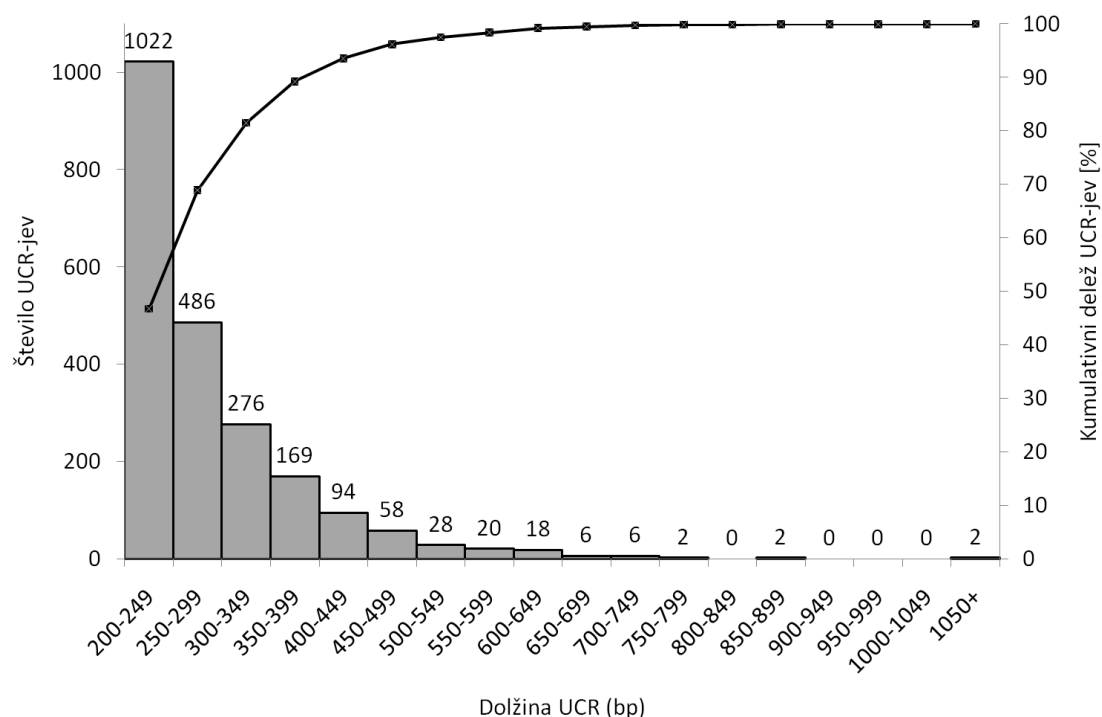
Glavni koraki in rezultati raziskave so povzeti na shemi poteka dela v prilogi A.

4.1 STATISTIČNI PARAMETRI IN GENOMSKA RAZPOREDITEV UCR-JEV

UCR-ji se nahajajo na vseh kromosomih, razen na kromosomu Y. Upoštevajoč dolžine kromosomov se v največji gostoti nahajajo na kromosomu 2, najredkejši pa so na kromosomu 21 (slika 1). Povprečna dolžina vseh 2189 analiziranih UCR-jev, dolgih vsaj 200 bp, je 287 bp. Skoraj polovica UCR-jev (46,7 % oz. 1022 UCR-jev) je dolgih 200–249 bp, 98,4 % UCR-jev je krajših od 600 bp (slika 2). Dva UCR-ja sta daljša od 1000 bp: UCR 13336 obsega 1087 bp, UCR 13347 pa 1085 bp; oba se nahajata na kromosomu X znotraj gena *POLAI* (angl. *DNA polymerase alpha 1, catalytic subunit*). UCR 2360 na kromosomu 11 je dolg 882 bp in se prekriva z dvema genoma: *ELP4* (angl. *elongator acetyltransferase complex subunit 4*) in *Z83001.1* (protismerni prepis). UCR 10729 je dolg 871 bp in se nahaja na kromosomu 5, znotraj gena *CPEB4* (angl. *cytoplasmic polyadenylation element binding protein 4*). Vsi ostali UCR-ji so krajši od 800 bp.



Slika 1: Prikaz razmerij med številom UCR-jev na kromosomu in dolžino kromosoma



Slika 2: Prikaz razporeditve UCR-jev glede na njihovo dolžino

4.2 GENI, ZNOTRAJ KATERIH SE NAHAJAJO UCR-JI

541 od 2189 UCR-jev (24,7 %) je medgenskih, preostalih 1648 (75,3 %) pa se vsaj delno nahaja znotraj eksonov in/ali intronov 874 genov različnih tipov (preglednica 1). Od 874 genov jih 537 vsebuje po le en UCR, v preostalih genih pa jih je več (priloga B). 24 genov vsebuje deset ali več UCR-jev, od tega je znotraj 11 genov ohranjenih regij več kot 15. Znotraj gena *NPAS3* (angl. *neuronal PAS domain protein 3*) jih je 19, *LRMDA* (angl. *leucine-rich melanocyte differentiation associated protein*) jih vsebuje 21, največ, tj. 28 UCR-jev, pa se nahaja v genu *FOXP2* (angl. *forkhead box P2*). *FOXP2* je transkripcijski represor, pomemben za diferenciacijo pljučnega epitela. Sodeluje pri razvoju živčevja, prebavnega in krvnožilnega sistema, pa tudi pri razvoju jezika in govora (UniProt).

Preglednica 1: Število genov določenega biotipa, v katerih se nahajajo UCR-ji

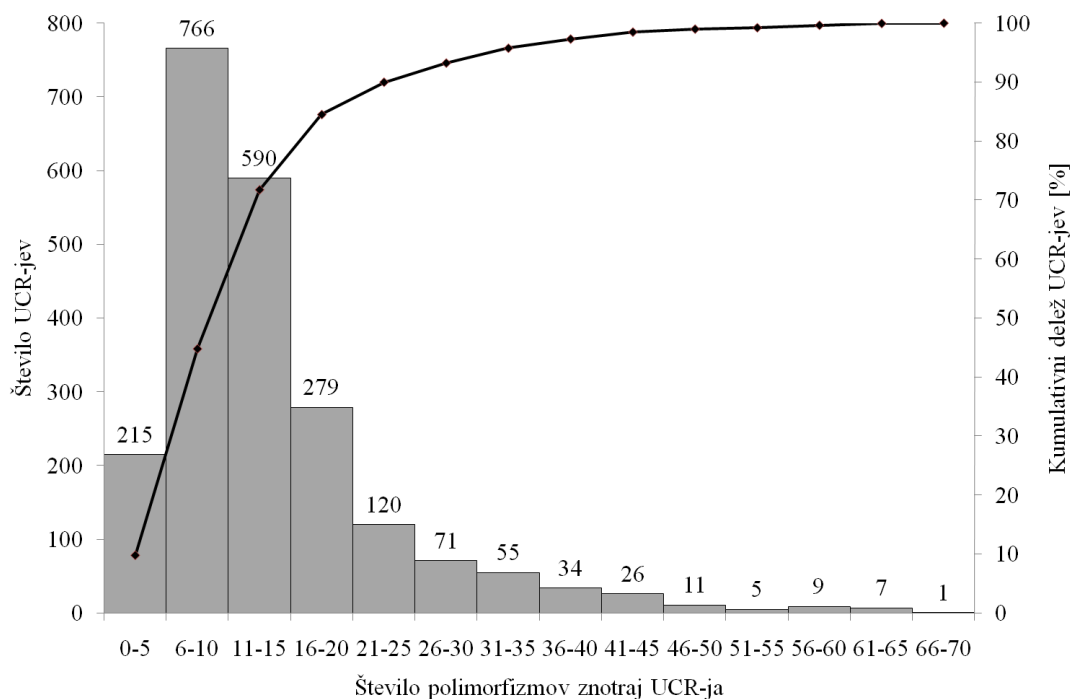
Biotip gena	Število genov tega biotipa, v katerih se nahajajo UCR-ji
Protein-kodirajoč	627
lincRNA	119
Protismerni prepis (AS)	90
Procesiran prepis	16
Ostale lincRNA: prekrivanje v smerni orientaciji (angl. <i>sense overlapping</i>); intronski, smerni (angl. <i>sense intronic</i>); 3' prekrivajoča ncRNA; nekodirajoč; dvosmerna lincRNA (angl. <i>bidirectional lincRNA</i>)	10
Pseudogeni: polimorfen psevdogen; prepisan psevdogen	5
ncRNA: misc RNA; rRNA	4
TEC (angl. <i>to be experimentally confirmed</i>)	3

lincRNA – dolga medgenska nekodirajoča RNA (angl. *long intergenic noncoding RNA*); lincRNA – dolga nekodirajoča RNA (angl. *long noncoding RNA*); misc RNA – nevrščena RNA (angl. *miscellaneous RNA*)

4.3 POLIMORFIZMI ZNOTRAJ UCR-JEV

Znotraj 2189 UCR-jev, dolgih vsaj 200 bp, smo z orodjem BioMart identificirali 30139 polimorfizmov, vendar se nekateri od njih nahajajo na istih mestih v genomu, saj so bila v izpisu rezultatov analize zanje navedena različna imena iz različnih podatkovnih zbirk (dbSNP, HGMD-PUBLIC, ESP, PhenCode in ClinVar). Nekateri insercije in delecije obsegajo daljša zaporedja DNA, znotraj katerih se nahajajo tudi drugi polimorfizmi.

Znotraj vsakega od analiziranih 2189 UCR-jev se nahaja vsaj en polimorfizem. V 215 UCR-jih (9,8 %) je polimorfizmov do največ pet. Od tega sedem UCR-jev vsebuje po le en polimorfizem, 20 UCR-jev po dva polimorfizma, 49 UCR-jev po tri polimorfizme, 50 UCR-jev po štiri polimorfizme in 89 UCR-jev po pet polimorfizmov. Skoraj polovica UCR-jev (44,8 %) vsebuje do največ deset polimorfizmov. UCR-jev z zelo velikim številom polimorfizmov je malo – več kot 50 polimorfizmov vsebuje 22 UCR-jev. Število UCR-jev glede na število polimorfizmov, ki jih vsebujejo, je prikazano na sliki 3.



Slika 3: Prikaz števila UCR-jev glede na število polimorfizmov, ki jih vsebujejo

UCR 5114, ki je dolg 554 bp in leži znotraj gena *NFAT5* (angl. *nuclear factor of activated T-cells 5*) na 16. kromosomu, vsebuje največ, tj. 70 polimorfizmov. Največje število polimorfizmov glede na dolžino vsebuje UCR 5138 – dolg je 206 bp, znotraj ohranjenega zaporedja pa je 64 polimorfizmov. Prekriva se s protein-kodirajočim *ZFX3* (angl. *zinc finger homeobox 3*) in s protismernim prepisom *AC004943.2*.

Po le en polimorfizem vsebujejo naslednji UCR-ji: UCR 114 (326 bp, znotraj *SRSF10*, angl. *serine and arginine rich splicing factor 10*, na kromosomu 1), UCR 118 (317 bp, prav tako znotraj *SRSF10*), UCR 13480 (248 bp, znotraj *RRAGB*, angl. *Ras related GTP binding B*, na kromosomu X), UCR 13357 (226 bp, v medgenski regiji na kromosomu X), UCR 5472 (223

bp, znotraj *WNT3*, angl. *Wnt family member 3*, na kromosomu 17), UCR 13518 (203 bp, znotraj *OGT*, angl. *O-linked N-acetylglucosamine (GlcNAc) transferase*, na kromosomu X) in UCR 594 (200 bp, znotraj *FUBP1*, angl. *far upstream element binding protein 1*, na kromosomu 1).

4.3.1 S fenotipom povezani polimorfizmi znotraj UCR-jev

Orodje BioMart je v genomskem brskalniku Ensembl za 183 od 30139 (0,6 %) polimorfizmov, ki se nahajajo znotraj UCR-jev, našlo anotirane povezave s fenotipom (priloga C). Opis povezanega fenotipa je bil naveden le pri 76 od 183 polimorfizmov. Za 35 polimorfizmov je povezava s fenotipom anotirana v podatkovni zbirki ClinVar, vendar opis povezanega fenotipa v izpisu rezultatov orodja BioMart ni bil naveden. Za 94 polimorfizmov znotraj UCR-jev je bila povezava z boleznimi najdena v podatkovni zbirki HGMD (angl. *Human Gene Mutation Database*) (www.hgmd.cf.ac.uk). Zbirka HGMD je zbirka mutacij v genomu človeka, ki so povezane z boleznimi. Ker je zbirka zasebna, je bilo v izpisu rezultatov analize z orodjem BioMart navedeno le, da je povezava pri določenem polimorfizmu anotirana, ni pa bilo navedeno, s katero boleznijo je polimorfizem povezan. Javna različica podatkovne zbirke je dostopna registriranim uporabnikom akademskih/neprofitnih organizacij, ki pa brez pisnega dovoljenja upraviteljev ne smejo javno objavljati podatkov iz zbirke. Za precejšnje število polimorfizmov tako vemo, da so povezani z boleznimi, vendar ne moremo navesti, s katerimi. Takšnih je 71 polimorfizmov, za 23 polimorfizmov pa smo reference dobili v drugih podatkovnih zbirkah.

Polimorfizmi, ki imajo v genomskem brskalniku Ensembl anotirano povezavo s fenotipom, se nahajajo znotraj skupno 48 UCR-jev. V kar 16 UCR-jih se nahaja po več polimorfizmov z anotiranimi povezavami s fenotipom. V šestih UCR-jih je takih polimorfizmov po deset ali več. Največ jih vsebuje UCR 10358 v genu *APC* (angl. *APC, WNT signaling pathway regulator*) na kromosomu 5 – z analizo z orodjem BioMart smo v njem identificirali 35 polimorfizmov, ki imajo v genomskem brskalniku Ensembl anotirano povezavo s fenotipom. Povezani so z dednim sindromom predispozicije za tumorje in raka, družinskim rakom debelega črevesa in danke (ClinVar), družinsko adenomatozno polipozo (angl. *familial adenomatous polyposis*; FAP) (Fodde in sod., 1992) in blažjo obliko te bolezni (Curia in sod., 1998; Soravia in sod., 1999). FAP je dedni sindrom, za katerega je značilen pojav polipov v črevesu, pri čemer je število polipov odvisno od lokacije mutacij v genu *APC*. Polipi so sprva benigni, po določenem času pa se pri večini bolnikov razvije kolorektalni rak (Debinski in sod., 1996). Chang in sod. (2016) so na podlagi rezultatov analize HRM (angl. *High Resolution Melting*) in sekvenciranja zaključili, da mutacije v genu *APC* ne igrajo pomembne vloge pri razvoju raka dojke v tajvanski populaciji. Za nekatere polimorfizme povezani fenotipi niso navedeni, ker so povezave anotirane v zasebni podatkovni zbirki HGMD.

UCR 13372, ki se nahaja znotraj gena za distrofin (*DMD*) na kromosomu X, glede na rezultate analize z orodjem BioMart vsebuje 21 s fenotipom povezanih polimorfizmov. V zbirki ClinVar so anotirane njihove povezave z Duchennovo mišično distrofijo, Beckerjevo mišično distrofijo in dilatativno kardiomiopatijo 3B (razširjenim srcem). Tuffery-Giraud in sod. (2005) so pri bolniku z blažjo obliko Beckerjeve mišične distrofije ugotovili, da je vzrok za bolezen nesmiselna mutacija rs863224976, ki skrajša protein DMD na C-koncu. Roberts in

sod. (1992) ter Hofstra in sod. (2004) so odkrili, da lahko Duchennovo mišično distrofijo povzroči rs398123839, delecija adenina na mestu X:31169542.

V UCR 13568 znotraj *PLP1* (angl. *proteolipid protein 1*) na kromosomu X se nahaja 20 polimorfizmov z anotiranimi povezavami s fenotipom. Večina anotacij je iz podatkovne zbirke HGMD, zato opisi fenotipov niso na voljo. Gorman in sod. (2007) so pri bolniku kot vzrok za spastično paraplegijo odkrili tranzicijo citozina v timin (rs132630295) v eksonu 3B gena *PLP1*. Saugier-Weber in sod. (1994) so v istem eksonu pri bolniku z enako boleznijo identificirali še eno tranzicijo citozina v timin (rs132630287), ki povzroči spremembo histidina v tirozin. Hodes in sod. (1997) so pri družini z atipično obliko bolezni Pelizaeus-Merzbacher prav tako v tretjem eksonu *PLP1* našli nesmiselno mutacijo rs132630292.

Za 11 od 183 polimorfizmov, ki naj bi bili povezani s fenotipi, so bile reference raziskav, v katerih naj bi poročali navedene povezave, navedene že v izpisu rezultatov orodja BioMart (skupaj 38 virov). Vse vire smo ročno pregledali, da smo preverili navedene povezave. Z nadaljnjim ročnim pregledovanjem virov v različnih podatkovnih zbirkah smo s študijami podkrepili dodatnih 26 povezav med polimorfizmi in fenotipi. Skupno smo torej vire, ki potrjujejo povezave, navedene v izpisu rezultatov analize z orodjem BioMart, našli za 37 od 183 polimorfizmov (preglednica 2). Dejstvo, da polimorfizem leži v ohranjeni regiji, so raziskovalci omenili le v primeru treh polimorfizmov: rs10496382 (Chiang in sod., 2012), rs121434591 (Senderek in sod., 2009) in rs11190870 (Shen in sod., 2011). Pri ostalih primerih raziskovalci niso imeli podatka, da se polimorfizem nahaja znotraj UCR.

Preglednica 2: Povezave med polimorfizmi in fenotipi, za katere je bilo možno pridobiti znanstveni članek, v katerem so poročali povezavo

Polimorfizem	UCR ID	Gen*	Fenotip, povezan s polimorfizmom	Viri
rs17105335	371	<i>AGBL4</i>	Amiotrofična lateralna skleroza	Cronin in sod., 2008
rs2020906	6629	<i>FBXO11, MSH6</i>	Verjetno ni povezave s sindromom Lynch	Hansen in sod., 2014
rs10496382	7038	/	Ni povezave z višino	Chiang in sod., 2012
rs13382811	7246	<i>ZEB2</i>	Kratkovidnost (huda)	Khor in sod., 2013
rs104893634	7789	<i>HOXD10, HOXD9, HOXD-AS2 (AS)</i>	Vertikalni kongenitalni talus	Shrimpton in sod., 2004; Dobbs in sod., 2006
rs2307121	10019	<i>ADAMTS6</i>	Struktura roženice	Lu in sod., 2013
rs587777277	10277	<i>NR2F1, NR2F1-AS1 (AS)</i>	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs587777275	10277	<i>NR2F1, NR2F1-AS1 (AS)</i>	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs587777274	10277	<i>NR2F1, NR2F1-AS1 (AS)</i>	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs387906239	10358	<i>APC</i>	FAP, blažja; rak ščitnice	Soravia in sod., 1999
rs3797704	10358	<i>APC</i>	Ni povezave z rakom dojke	Chang in sod., 2016
rs387906232	10358	<i>APC</i>	FAP	Fodde in sod., 1992
rs387906237	10358	<i>APC</i>	FAP, blažja	Curia in sod., 1998
rs121434591	10453	<i>MATR3; ENSG00000280987 ENSG00000015479</i>	Distalna miopatija	Senderek in sod., 2009

se nadaljuje

nadaljevanje preglednice 2

Polimorfizem	UCR ID	Gen*	Fenotip, povezan s polimorfizmom	Viri
rs587777300	10453	<i>MATR3</i> : <i>ENSG00000280987</i> <i>ENSG00000015479</i>	Amiotrofična lateralna skleroza 21	Johnson in sod., 2014
rs863223403	12957	<i>HNRNPK</i>	Sindrom Au-Kline	Au in sod., 2015
rs121917900	1446	<i>ERCC6</i>	Cockayne sindrom B	Mallery in sod., 1998
rs75462234	1766	<i>PAX2</i>	Papilorenalni sindrom	Schimmenti in sod., 1999
rs77453353	1766	<i>PAX2</i>	Papilorenalni sindrom	Amiel in sod., 2000
rs76675173	1766	<i>PAX2</i>	Papilorenalni sindrom	Schimmenti in sod., 1997
rs587777708	1766	<i>PAX2</i>	Fokalna segmentna glomeruloskleroza 7	Barua in sod., 2014
rs11190870	1798	/	Adolescentna idiopatska skolioza (huda); ni povezave z rakom dojke	Takahashi in sod., 2011; Gorman in sod., 2012 Gao in sod., 2013; Jiang in sod., 2013; Miyake in sod., 2013; Londono in sod., 2014; Chettier in sod., 2015; Grauers in sod., 2015; Shen in sod., 2016 Buchert in sod., 2014
rs724159963	2195	<i>FAR1</i>	Motnja peroksisomalne maščobne acil-CoA reduktaze 1	
rs16932455	2242	<i>SOX6</i>	Občutljivost na zdravilo kapecitabin	O'Donnell in sod., 2012
rs997295	4568	<i>MAP2K5</i>	Potovalna slabost, indeks telesne mase (ITM)	Guo in sod., 2013; De in sod., 2015; Hromatka in sod., 2015 Al Turki in sod., 2014
rs587777373	4731	<i>NR2F2</i>	Kongenitalni defekti srca, več vrst	
rs398123839	13372	<i>DMD</i>	Duchennova mišična distrofija	Roberts in sod., 1992; Hofstra in sod., 2004
rs863224976	13372	<i>DMD</i>	Beckerjeva mišična distrofija	Tuffery-Giraud in sod., 2005
rs132630295	13568	<i>PLP1</i>	Spastična paraplegija 2, X-vezana	Gorman in sod., 2007
rs132630287	13568	<i>PLP1</i>	Spastična paraplegija 2, X-vezana	Saugier-Veber in sod., 1994
rs132630292	13568	<i>PLP1</i>	Bolezen Pelizaeus-Merzbacher, atipična	Hodes in sod., 1997
rs137852350	13607	<i>GRIA3</i>	Duševna zaostalost, X-vezana 94	Wu in sod., 2007
rs122459149	13666	<i>FHL1</i>	Mišična distrofija Emery-Dreifuss 6, X-vezana	Gueneau in sod., 2009; Knoblauch in sod., 2010
rs122458141	13666	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic	Windpassinger in sod., 2008; Schoser in sod., 2009 Schoser in sod., 2009
rs786200914	13666	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic	
rs267606811	13667	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic	Schoser in sod., 2009
rs62621672	13736	<i>MECP2</i>	Ni povezave z Rettovim sindromom	Zahorakova in sod., 2007

* če polimorfizem leži znotraj gena, ki ni protein-kodirajoč, je biotip gena zapisan v oklepaju

5 RAZPRAVA

5.1 STATISTIČNI PARAMETRI IN GENOMSKA RAZPOREDITEV UCR-JEV

Bejerano in sod. (2004) so s poravnavo genomov človeka, miši in podgane identificirali UCR-je na vseh kromosomih, razen na kromosomih 21 in Y. Naša analiza genomske razporeditve UCR-jev, ki so jih objavili Stephen in sod. (2008), je pokazala, da se šest UCR-jev, dolgih vsaj 200 bp, nahaja tudi na kromosomu 21. Kljub temu je razmerje med številom UCR-jev in dolžino kromosoma pri tem kromosomu znatno nižje kot pri ostalih (slika 1).

Zanimivo je, da se oba najdaljša UCR-ja (UCR 13336 in UCR 13347) nahajata znotraj istega gena – *POLAI*. Ležita znotraj 35. (predzadnjega) oz. 36. (zadnjega) introna tega gena, zato je malo verjetno, da sta neposredno povezana s funkcijo tega gena. Dolge UCR-je na 3'-koncu *POLAI* so identificirali že Bejerano in sod. (2004). Ker ležijo bližje 3'-koncu gena *ARX* (angl. *aristless related homeobox*) kot pa preostanku 303 kbp dolgega gena *POLAI*, je možno, da je njihova vloga povezana z *ARX* – morda predstavljajo gručo ojačevalcev tega gena. *ARX* je transkripcijski dejavnik, ki sodeluje pri razvoju možganov. Povezan je s številnimi boleznimi, med drugim z mentalno zaostalostjo, epilepsijo, avtizmom in cerebralnimi malformacijami (Sherr, 2003).

Tretji najdaljši UCR, UCR 2360 (882 bp), leži na kromosomu 11, in sicer znotraj protein-kodirajočega gena *ELP4* in gena *Z83001.1*, ki kodira protismerni prepis *ELP4*. Tudi v tem primeru se ohranjena regija nahaja v intronu na 3'-koncu gena. V bližini te regije je na protismerni verigi lociran gen *PAX6*, ki je eden od glavnih regulatornih razvojnih genov s pleiotrofnimi učinki - *PAX6* je transkripcijski dejavnik, ki igra pomembno vlogo v razvoju oči, nosu, centralnega živčnega sistema in trebušne slinavke (UniProt). Znano je, da na izražanje tega gena v veliki meri vplivajo cis-elementi. Kot že omenjeno v poglavju s pregledom objav, so Bhatia in sod. (2013) ugotovili, da mutacija v UCR-ju (UCR 2354, 107 bp) v devetem intronu gena *ELP4* spremeni avtoregulatorno vezavno mesto za *PAX6* in s tem poruši uravnavanje izražanja gena *PAX6*. Predpostavljamo, da bi se lahko kakšna pomembna regulatorna regija nahajala tudi znotraj UCR 2360.

5.2 GENI, ZNOTRAJ KATERIH SE NAHAJAJO UCR-JI

Analiza z orodjem BioMart je pokazala, da se UCR-ji, ki so jih objavili Stephen in sod. (2008), ne prekrivajo le s protein-kodirajočimi geni (629 genov), ampak v veliko primerih tudi z geni za lincRNA (preglednica 1). UCR-ji se prekrivajo z 238 tovrstnimi geni; v kar 119 primerih gre za lincRNA, 90 je protismernih prepisov (angl. *antisense*), 16 pa procesiranih prepisov. lincRNA so vključene v mnoge biološke procese, njihovo izražanje pa je spremenjeno pri mnogih mendelskih boleznih, raku, srčno-žilnih obolenjih in nevroloških boleznih (Maass in sod., 2014). Raziskave kažejo, da imajo SNP-ji in različice v številu kopij (angl. *copy number variation*; CNV) v lincRNA lahko vlogo pri pojavu rakavih obolenj (Kunej in sod., 2014; Gupta in Tripathi, 2017). Tudi v naši raziskavi smo v genomskem brskalniku Ensembl odkrili nekatere s fenotipom povezane polimorfizme, ki se nahajajo znotraj genov za lincRNA ali protismerne prepise (priloga C), vendar se vsi nahajajo v takšnih regijah, ki poleg lincRNA ali protismernih prepisov kodirajo tudi proteine.

Potrebno se je zavedati, da UCR-ji, locirani v določenem genu, ne vplivajo nujno na prepisovanje, prevajanje in/ali funkcijo produkta tega gena. Možno je, da gre za ojačevalce drugih genov, ki se nahajajo v bližini – primer so v prejšnjem poglavju omenjene ohranjene regije v genih *POLAI* (Bejerano in sod., 2004) in *ELP4* (Bhatia in sod., 2013). Zelo zanimivo je, da se znotraj nekaterih genov, npr. v protein-kodirajočih *FOXP2*, *LRMDA* (*C10orf11*), *NPAS3*, *ZEB2* in *AUTS2* ter v genu za lincRNA *LINC01122*, nahaja zelo veliko UCR-jev (priloga B). Potrebno pa je upoštevati, da je njihovo število verjetno odvisno tudi od dolžine gena.

5.3 POLIMORFIZMI ZNOTRAJ UCR-JEV

Z analizo *in silico* smo znotraj 2189 UCR-jev, ki so dolgi vsaj 200 bp, odkrili kar 30139 polimorfizmov (SNP-jev, insercij in delecij). Skupno smo pregledali 628364 bp ohranjenih regij v genomu človeka, kar pomeni, da je v povprečju polimorfizem prisoten na vsakih 20,8 bp zaporedja. Potrebno je upoštevati, da je to le zelo približna ocena dejanskega povprečja, saj so nekateri polimorfizmi zaradi različnih imen v različnih podatkovnih zbirkah na našem seznamu podvojeni, nekatere insercije in delecije pa obsegajo tudi daljša nukleotidna zaporedja. Poleg tega niso vsi polimorfizmi eksperimentalno potrjeni.

UCR-jev z zelo velikim številom polimorfizmov je malo. Kljub temu so te regije precej zanimive, saj je težko pojasniti njihovo medvrstno ohranjenost. Zlasti zanimivi pa so se nam zdeli tudi UCR-ji, ki vsebujejo zelo malo polimorfizmov. Zelo nizka variabilnost teh regij tudi na ravni vrste najverjetneje ni naključna, ampak povezana z njihovo funkcijo. Primerne bi bile kot kandidati za eksperimentalno preverjanje funkcije ohranjenih regij.

5.3.1 S fenotipom povezani polimorfizmi znotraj UCR-jev

Z uporabo orodja BioMart smo v 2189 analiziranih UCR-jih identificirali 183 s fenotipom povezanih polimorfizmov (priloga C), kar predstavlja približno 0,6 % od vseh polimorfizmov, ki se nahajajo znotraj teh regij. Skoraj vsi polimorfizmi z anotirano povezavo s fenotipom (179 od 183 polimorfizmov; 97,8 %), ki smo jih odkrili z našo analizo, se nahajajo znotraj genskih UCR-jev – v medgenskih regijah smo identificirali le štiri. rs11190870 v UCR 1798 je povezan z adolescentno idiopatsko skoliozo (Takahashi in sod., 2011), rs10496382 v UCR 7038 pa ni povezan z višino (Chiang in sod., 2012). Za CR087569 v UCR 9337 je povezava s fenotipom anotirana v zasebni zbirki HGMD, za rs5927163 v UCR 13386 pa v podatkovni zbirki ClinVar, vendar opis povezanega fenotipa ni naveden.

V kar 16 UCR-jih se nahaja po več kot en polimorfizem, ki ima v genomskem brskalniku Ensembl anotirano povezavo s fenotipom. V šestih UCR-jih je takih polimorfizmov deset ali več. Neenakomerna razporeditev s fenotipom povezanih polimorfizmov po UCR-jih nakazuje, da imajo vsaj nekatere od ohranjenih regij velik vpliv na fenotip in da so lahko polimorfizmi v teh regijah v nekaterih primerih vzrok za genetsko pogojene bolezni. Največ, tj. 35 polimorfizmov, ki imajo v genomskem brskalniku Ensembl anotirano povezavo s fenotipom, se nahaja v UCR 10358, ki je edini UCR v genu *APC* na kromosomu 5, daljši od 200 bp. Polimorfizmi v njem so večini povezani s FAP oz. podobnimi bolezenskimi stanji.

APC je tumor supresor. Raziskave kažejo, da se velika večina (> 98 %) mutacij v genu *APC* nahaja v prvi polovici kodirajočega zaporedja in povzroči skrajšano obliko proteina APC (Laurent-Puig in sod., 1998).

Na podlagi rezultatov, ki smo jih pridobili v diplomski nalogi, bo mogoče nadalje identificirati pomembne regulatorne regije v UCR-jih oz. vloge teh zaporedij. Za dejansko karakterizacijo bo v prihodnosti potrebno vloge teh regij preveriti tudi eksperimentalno. Naši rezultati lahko služijo kot pomoč pri izbiri regij za tovrstne raziskave. Dobro razumevanje vloge UCR-jev ni pomembno zgolj z vidika boljšega razumevanja genske regulacije, ampak tudi zato, ker raziskave kažejo, da so vsaj nekatere od teh regij vključene v razvoj bolezni in zato predstavljajo potencialne tarče za razvoj novih metod zdravljenja.

5.4 OMEJITVE RAZISKAVE

Na koncu je potrebno omeniti še nekatere pomanjkljivosti naše raziskave. Analizirali smo le UCR-je, ki so dolgi vsaj 200 bp, pomembne regije, ki vplivajo na fenotip, pa se zagotovo nahajajo tudi v krajših UCR-jih. Z analizo z orodjem BioMart nismo identificiranih vseh znanih s fenotipom povezanih polimorfizmov znotraj UCR-jev, ker v genomskem brskalniku Ensembl vse znane povezave še niso anotirane. Poleg tega tudi nismo našli referenc, ki bi potrjevale vse povezave, izpisane v rezultatih analize. Pri nekaterih polimorfizmih, za katere smo našli in pregledali originalne raziskave, se je izkazalo, da je bil opis povezanega fenotipa v izpisu rezultatov orodja BioMart napačen. Možno je, da je napačen pri še kakšnem polimorfizmu, za katerega povezave nismo mogli neposredno preveriti v literaturi. Ročno iskanje referenc za specifičen polimorfizem po različnih podatkovnih zbirkah je zaradi zelo različnih načinov poimenovanj in zaradi različnih različic genomov precej težavno.

6 SKLEPI

V diplomskem delu smo na novo določili genomske lokacije 2189 znanih UCR-jev pri človeku, ki so jih predhodno identificirali Stephen in sod. (2007). Statistično smo analizirali njihove dolžine in genomsko razpršenost. Ugotovili smo, da se nahajajo na vseh kromosomih, razen na kromosomu Y. Povprečna dolžina analiziranih UCR-jev je 287 bp; le štirje UCR-ji so daljši od 800 bp.

Z analizo z orodjem BioMart smo ugotovili, da 25 % UCR-jev leži v medgenskih regijah, 75 % pa se jih prekriva z geni: največ je protein kodirajočih genov (627), sledijo lincRNA (119 genov) in protismerni prepisi (90 genov). Znotraj nekaterih genov se nahaja več UCR-jev: 11 genov vsebuje po več kot 15 ohranjenih regij. Na podlagi primerov iz literature smo pokazali, da ohranjena regija znotraj določenega gena ni nujno povezana s tem genom, ampak lahko npr. deluje kot ojačevalec drugega gena, ki se nahaja v bližini.

Z orodjem BioMart smo znotraj 2189 UCR-jev, katerih skupna dolžina je 628364 bp, odkrili 30139 polimorfizmov (SNP-jev, insercij, delecij, INDEL-ov). Ugotovili smo, da vsak izmed analiziranih UCR-jev vsebuje vsaj en polimorfizem. V 9,8 % UCR-jev je polimorfizmov do največ pet. Od 30139 polimorfizmov znotraj UCR-jev smo z orodjem BioMart identificirali

183 polimorfizmov, za katere je v genomskem brsklaniku Ensembl anotirana povezava s fenotipom. Štirje se nahajajo v medgenskih regijah, preostali pa znotraj genov. Vire, v katerih smo dobljene povezave lahko preverili, smo našli za 37 polimorfizmov. Potrdili smo, da so povezani s precej različnimi fenotipskimi lastnostmi in boleznimi, med drugim z amiotrofično lateralno sklerozo, kratkovidnostjo, strukturo roženice, sindromom optične atrofije Bosch-boonstra-schaaf, FAP, distalno miopatijo, papilorenalnim sindromom, adolescentno idiopatsko skoliozo, občutljivostjo na kapecitabin, ITM-jem, kongenitalnimi defekti srca, Duchenovo in Beckerjevo mišično distrofijo, spastično paraplegijo, duševno zaostalostjo in miopatijo z atrofijo posturalnih mišic. S tem smo uspešno anotirali funkcijo 23 UCR-jev.

7 VIRI

- Ahituv N., Zhu Y., Visel A., Holt A., Afzal V., Pennacchio L. A., Rubin E. M. 2007. Deletion of ultraconserved elements yields viable mice. *Plos Biology*, 5, 9: 1906-1911
- Al Turki S., Manickaraj A. K., Mercer C. L., Gerety S. S., Hitz M. P., Lindsay S., D'Alessandro L. C. A., Swaminathan G. J., Bentham J., Arndt A. K. in sod. 2014. Rare variants in NR2F2 cause congenital heart defects in humans. *American Journal of Human Genetics*, 94, 4: 574-585
- Amiel J., Audollent S., Joly D., Dureau P., Salomon R., Tellier A. L., Auge J., Bouissou F., Antignac C., Gubler M. C. in sod. 2000. PAX2 mutations in renal-coloboma syndrome: mutational hotspot and germline mosaicism. *European Journal of Human Genetics*, 8, 11: 820-826
- Au P. Y. B., You J., Caluseriu O., Schwartzentruber J., Majewski J., Bernier F. P., Ferguson M., Valle D., Parboosingh J. S., Sobreira N. in sod. 2015. GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability, unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. *Human Mutation*, 36, 10: 1009-1014
- Bao B. Y., Lin V. C., Yu C. C., Yin H. L., Chang T. Y., Lu T. L., Lee H. Z., Pao J. B., Huang C. Y., Huang S. P. 2016. Genetic variants in ultraconserved regions associate with prostate cancer recurrence and survival. *Scientific Reports*, 6: 22124, doi: 10.1038/srep22124: 8 str.
- Barua M., Stellacci E., Stella L., Weins A., Genovese G., Muto V., Caputo V., Toka H. R., Charoonratana V. T., Tartaglia M. in sod. 2014. Mutations in PAX2 associate with adult-onset FSGS. *Journal of the American Society of Nephrology*, 25, 9: 1942-1953
- Bateman A., Martin M. J., O'Donovan C., Magrane M., Alpi E., Antunes R., Bely B., Bingley M., Bonilla C., Britto R. in sod. 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, 45, D1: D158-D169
- Bejerano G., Pheasant M., Makunin I., Stephen S., Kent W. J., Mattick J. S., Haussler D. 2004. Ultraconserved elements in the human genome. *Science* 304, 5675: 1321-1325
- Bhatia S., Bengani H., Fish M., Brown A., Divizia M. T., de Marco R., Damante G., Grainger R., van Heyningen V., Kleinjan D. A. 2013. Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *American Journal of Human Genetics*, 93, 6: 1126-1134

- Bosch D. G. M., Boonstra F. N., Gonzaga-Jauregui C., Xu M. F., de Ligt J., Jhangiani S., Wiszniewski W., Muzny D. M., Yntema H. G., Pfundt R. in sod. 2014. NR2F1 mutations cause optic atrophy with intellectual disability. *American Journal of Human Genetics*, 94, 2: 303-309
- Braconi C., Valeri N., Kogure T., Gasparini P., Huang N. Y., Nuovo G. J., Terracciano L., Croce C. M., Patel T. 2011. Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 2: 786-791
- Buchert R., Tawamie H., Smith C., Uebe S., Innes A. M., Al Hallak B., Ekici A. B., Sticht H., Schwarze B., Lamont R. E. in sod. 2014. A peroxisomal disorder of severe intellectual disability, epilepsy, and cataracts due to fatty acyl-CoA reductase 1 deficiency. *American Journal of Human Genetics*, 95, 5: 602-610
- Calin G. A., Liu C. G., Ferracin M., Hyslop T., Spizzo R., Sevignani C., Fabbri M., Cimmino A., Lee E. J., Wojcik S. E. in sod. 2007. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell*, 12, 3: 215-229
- Chang Y. S., Lin C. Y., Yang S. F., Ho C. M., Chang J. G. 2016. Analysing the mutational status of adenomatous polyposis coli (APC) gene in breast cancer. *Cancer Cell International*, 16: 23, doi: 10.1186/s12935-016-0297-6 str.
- Chen C. T., Wang J. C., Cohen B. A. 2007. The strength of selection on ultraconserved elements in the human genome. *American Journal of Human Genetics*, 80, 4: 692-704
- Chettier R., Nelson L., Ogilvie J. W., Albertsen H. M., Ward K. 2015. Haplotypes at LBX1 have distinct inheritance patterns with opposite effects in adolescent idiopathic scoliosis. *PLoS One*, 10, 2: e0117708, doi: 10.1371/journal.pone.0117708: 11 str.
- Chiang C. W., Liu C. T., Lettre G., Lange L. A., Jorgensen N. W., Keating B. J., Vedantam S., Nock N. L., Franceschini N., Reiner A. P. in sod. 2012. Ultraconserved elements in the human genome: association and transmission analyses of highly constrained single-nucleotide polymorphisms. *Genetics*, 192, 1: 253-66
- Cronin S., Berger S., Ding J. H., Schymick J. C., Washecka N., Hernandez D. G., Greenway M. J., Bradley D. G., Traynor B. J., Hardiman O. 2008. A genome-wide association study of sporadic ALS in a homogenous Irish population. *Human Molecular Genetics*, 17, 5: 768-774
- Curia M. C., Esposito D. L., Aceto G., Palmirotta R., Crognale S., Valanzano R., Ficari F., Tonelli F., Battista P., Mariani-Costantini R. in sod. 1998. Transcript dosage effect in familial adenomatous polyposis: Model offered by two kindreds with exon 9 APC gene mutations. *Human Mutation*, 11, 3: 197-201
- De R., Verma S. S., Drenos F., Holzinger E. R., Holmes M. V., Hall M. A., Crosslin D. R., Carrell D. S., Hakonarson H., Jarvik G. in sod. 2015. Identifying gene-gene interactions that are highly associated with Body Mass Index using Quantitative Multifactor Dimensionality Reduction (QMDR). *BioData Mining*, 8: 41, doi: 10.1186/s13040-015-0074-0: 14 str.

- Debinski H. S., Love S., Spigelman A. D., Phillips R. K. S. 1996. Colorectal polyp counts and cancer risk in familial adenomatous polyposis. *Gastroenterology*, 110, 4: 1028-1030
- Dobbs M. B., Gurnett C. A., Pierce B., Exner G. U., Robarge J., Morcuende J. A., Cole W. G., Templeton P. A., Foster B., Bowcock A. M. 2006. HOXD10 M319K mutation in a family with isolated congenital vertical Talus. *Journal of Orthopaedic Research*, 24, 3: 448-453
- Feng J. C., Bi C. M., Clark B. S., Mady R., Shah P., Kohtz J. D. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes & Development*, 20, 11: 1470-1484
- Fodde R., Vanderluijt R., Wijnen J., Tops C., Vanderklift H., Vanleeuwen Cornelisse I., Griffioen G., Vasen H., Khan P. M. 1992. 8 novel inactivating germ line mutations at the APC gene identified by denaturing gradient gel-electrophoresis. *Genomics*, 13, 4: 1162-1168
- Gao W., Peng Y., Liang G., Liang A., Ye W., Zhang L., Sharma S., Su P., Huang D. 2013. Association between common variants near LBX1 and adolescent idiopathic scoliosis replicated in the Chinese Han population. *PLoS One* 8, 1: e53234, doi: 10.1371/journal.pone.0053234: 7 str.
- Gorman K. F., Julien C., Moreau A. 2012. The genetic epidemiology of idiopathic scoliosis. *European Spine Journal*, 21, 10: 1905-1919
- Gorman M. P., Golomb M. R., Walsh L. E., Hobson G. M., Garbern J. Y., Kinkel R. P., Darras B. T., Urion D. K., Eksioğlu Y. Z. 2007. Steroid-responsive neurologic relapses in a child with a proteolipid protein-1 mutation. *Neurology*, 68, 16: 1305-1307
- Grauers A., Wang J., Einarsson E., Simony A., Danielsson A., Åkesson K., Ohlin A., Halldin K., Grabowski P., Tenne M. in sod. 2015. Candidate gene analysis and exome sequencing confirm LBX1 as a susceptibility gene for idiopathic scoliosis. *Spine Journal*, 15, 10: 2239-2246
- Gueneau L., Bertrand A. T., Jais J. P., Salih M. A., Stojkovic T., Wehnert M., Hoeltzenbein M., Spuler S., Saitoh S., Verschueren A. in sod. 2009. Mutations of the FHL1 gene cause Emery-Dreifuss muscular dystrophy. *American Journal of Human Genetics*, 85, 3: 338-353
- Guo Y., Lanktree M. B., Taylor K. C., Hakonarson H., Lange L. A., Keating B. J. 2013. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. *Human Molecular Genetics*, 22, 1: 184-201
- Gupta S. C., Tripathi Y. N. 2017. Potential of long non-coding RNAs in cancer patients: From biomarkers to therapeutic targets. *International Journal of Cancer*, 140, 9: 1955-1967
- Hansen M. F., Neckmann U., Lavik L. A., Vold T., Gilde B., Toft R. K., Sjørsen W. 2014. A massive parallel sequencing workflow for diagnostic genetic testing of mismatch repair genes. *Molecular Genetics & Genomic Medicine*, 2, 2: 186-200
- Hodes M. E., Blank C. A., Pratt V. M., Morales J., Napier J., Dlouhy S. R. 1997. Nonsense mutation in exon 3 of the proteolipid protein gene (PLP) in a family with an unusual form of Pelizaeus-Merzbacher disease. *American Journal of Medical Genetics*, 69, 2: 121-125
- Hofstra R. M. W., Mulder I. M., Vossen R., de Koning-Gans P. A. M., Kraak M., Ginjaar I. B., van der Hout A. H., Bakker E., Buys C., van Ommen G. J. B. in sod. 2004. DGGE

- based whole-gene mutation scanning of the dystrophin gene in Duchenne and Becker muscular dystrophy patients. *Human Mutation*, 23, 1: 57-66
- Honma R., Goto K., Sakamoto N., Sekino Y., Sentani K., Oue N., Yasui W. 2017. Expression and function of Uc.160+, a transcribed ultraconserved region, in gastric cancer. *Gastric Cancer* (v tisku)
- Hromatka B. S., Tung J. Y., Kiefer A. K., Do C. B., Hinds D. A., Eriksson N. 2015. Genetic variants associated with motion sickness point to roles for inner ear development, neurological processes and glucose homeostasis. *Human Molecular Genetics*, 24, 9: 2700-2708
- Hudson R. S., Yi M., Volfovsky N., Prueitt R. L., Esposito D., Volinia S., Liu C. G., Schetter A. J., Van Roosbroeck K., Stephens R. M. in sod. 2013. Transcription signatures encoded by ultraconserved genomic regions in human prostate cancer. *Molecular Cancer*, 12: 13, doi: 10.1186/1476-4598-12-13: 13 str.
- Jiang H., Qiu X., Dai J., Yan H., Zhu Z., Qian B., Qiu Y. 2013. Association of rs11190870 near LBX1 with adolescent idiopathic scoliosis susceptibility in a Han Chinese population. *European Spine Journal*, 22, 2: 282-286
- Jiang J. M., Azevedo-Pouly A. C. P., Redis R. S., Lee E. J., Gusev Y., Allard D., Sutaria D. S., Badawi M., Elgamal O. A., Lerner M. R. in sod. 2016. Globally increased ultraconserved noncoding RNA expression in pancreatic adenocarcinoma. *Oncotarget*, 7, 33: 53165-53177
- Johnson J. O., Pioro E. P., Boehringer A., Chia R., Feit H., Renton A. E., Pliner H. A., Abramzon Y., Marangi G., Winborn B. J. in sod. 2014. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature Neuroscience*, 17, 5: 664-666
- Khor C. C., Miyake M., Chen L. J., Shi Y., Barathi V. A., Qiao F., Nakata I., Yamashiro K., Zhou X., Tam P. O. in sod. 2013. Genome-wide association study identifies ZFX1B as a susceptibility locus for severe myopia. *Human Molecular Genetics*, 22, 25: 5288-5294
- Knoblauch H., Geier C., Adams S., Budde B., Rudolph A., Zacharias U., Schulz-Menger J., Spuler A., Ben Yaou R., Nurnberg P. in sod. 2010. Contractures and hypertrophic cardiomyopathy in a novel FHL1 mutation. *Annals of Neurology*, 67, 1: 136-140
- Kunej T., Obsteter J., Pogacar Z., Horvat S., Calin G. A. 2014. The decalog of long non-coding RNA involvement in cancer diagnosis and monitoring. *Critical Reviews in Clinical Laboratory Sciences*, 51, 6: 344-357
- Laurent-Puig P., Beroud C., Soussi T. 1998. APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Research*, 26, 1: 269-270
- Lee T. I., Jenner R. G., Boyer L. A., Guenther M. G., Levine S. S., Kumar R. M., Chevalier B., Johnstone S. E., Cole M. F., Isono K. in sod. 2006. Control of developmental regulator's by polycomb in human embryonic stem cells. *Cell*, 125, 2: 301-313
- Lin M., Eng C., Hawk E. T., Huang M., Greisinger A. J., Gu J., Ellis L. M., Wu X., Lin J. 2012. Genetic variants within ultraconserved elements and susceptibility to right- and left-sided colorectal adenocarcinoma. *Carcinogenesis*, 33, 4: 841-847
- Londono D., Kou I., Johnson T. A., Sharma S., Ogura Y., Tsunoda T., Takahashi A., Matsumoto M., Herring J. A., Lam T. P. in sod. 2014. A meta-analysis identifies

- adolescent idiopathic scoliosis association with LBX1 locus in multiple ethnic groups. *Journal of Medical Genetics*, 51, 6: 401-406
- Lu Y., Vitart V., Burdon K. P., Khor C. C., Bykhovskaya Y., Mirshahi A., Hewitt A. W., Koehn D., Hysi P. G., Ramdas W. D. in sod. 2013. Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nature Genetics*, 45, 2: 155-163
- Lujambio A., Portela A., Liz J., Melo S. A., Rossi S., Spizzo R., Croce C. M., Calin G. A., Esteller M. 2010. CpG island hypermethylation-associated silencing of non-coding RNAs transcribed from ultraconserved regions in human cancer. *Oncogene*, 29, 48: 6390-6401
- Maass P. G., Luft F. C., Bähring S. 2014. Long non-coding RNA in health and disease. *Journal of Molecular Medicine*, 92, 4: 337-346
- Mallery D. L., Tanganelli B., Colella S., Steingrimsdottir H., van Gool A. J., Troelstra C., Stefanini M., Lehmann A. R. 1998. Molecular analysis of mutations in the CSB (ERCC6) gene in patients with Cockayne syndrome. *American Journal of Human Genetics*, 62, 1: 77-85
- Marini A., Lena A. M., Panatta E., Ivan C., Han L., Liang H., Annicchiarico-Petruzzelli M., Di Daniele N., Calin G. A., Candi E. in sod. 2017. Ultraconserved long non-coding RNA uc.63 in breast cancer. *Oncotarget*, 8, 22: 35669-35680
- Mestdagh P., Fredlund E., Pattyn F., Rihani A., Van Maerken T., Vermeulen J., Kumps C., Menten B., De Preter K., Schramm A. in sod. 2010. An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours. *Oncogene*, 29, 24: 3583-3592
- Miyake A., Kou I., Takahashi Y., Johnson T. A., Ogura Y., Dai J., Qiu X., Takahashi A., Jiang H., Yan H. in sod. 2013. Identification of a susceptibility locus for severe adolescent idiopathic scoliosis on chromosome 17q24.3. *PLoS One*, 8, 9: e72802, doi: 10.1371/journal.pone.0072802: 5 str.
- Ni J. Z., Grate L., Donohue J. P., Preston C., Nobida N., O'Brien G., Shiue L., Clark T. A., Blume J. E., Ares M. 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & Development*, 21, 6: 708-718
- O'Donnell P. H., Stark A. L., Gamazon E. R., Wheeler H. E., McIlwee B. E., Gorsic L., Im H. K., Huang R. S., Cox N. J., Dolan M. E. 2012. Identification of novel germline polymorphisms governing capecitabine sensitivity. *Cancer*, 118, 16: 4063-4073
- Olivieri M., Ferro M., Terreri S., Durso M., Romanelli A., Avitabile C., De Cobelli O., Messere A., Bruzzese D., Vannini I. in sod. 2016. Long non-coding RNA containing ultraconserved genomic region 8 promotes bladder cancer tumorigenesis. *Oncotarget*, 7, 15: 20636-20654
- Roberts R. G., Bobrow M., Bentley D. R. 1992. Point mutations in the dystrophin gene. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 6: 2331-2335

- Saugier-Veber P., Munnich A., Bonneau D., Rozet J. M., Lemerrer M., Gil R., Boespflugtanguy O. 1994. X-linked spastic paraplegia and Pelizaeus-Merzbacher disease are allelic disorders at the proteolipid protein locus. *Nature Genetics*, 6, 3: 257-262
- Scaruffi P. 2011. The Transcribed-Ultraconserved Regions: A Novel Class of Long Noncoding RNAs Involved in Cancer Susceptibility. *The scientific world journal*, 11: 340-352
- Scaruffi P., Stigliani S., Moretti S., Coco S., De Vecchi C., Valdora F., Garaventa A., Bonassi S., Tonini G. P. 2009. Transcribed-ultra conserved region expression is associated with outcome in high-risk neuroblastoma. *BMC Cancer* 9: 441, doi: 10.1186/1471-2407-9-441: 9 str.
- Schimmenti L. A., Cunliffe H. E., McNoe L. A., Ward T. A., French M. C., Shim H. H., Zhang Y. H., Proesmans W., Leys A., Byerly K. A. in sod. 1997. Further delineation of renal-coloboma syndrome in patients with extreme variability of phenotype and identical PAX2 mutations. *American Journal of Human Genetics*, 60, 4: 869-878
- Schimmenti L. A., Shim H. H., Wirtschafter J. D., Panzarino V. A., Kashtan C. E., Kirkpatrick S. J., Wargowski D. S., France T. D., Michel E., Dobyns W.B. 1999. Homonucleotide expansion and contraction mutations of PAX2 and inclusion of Chiari 1 malformation as part of Renal-Coloboma syndrome. *Human Mutation*, 14, 5: 369-376
- Schoser B., Goebel H. H., Janisch I., Quasthoff S., Rother J., Bergmann M., Muller-Felber W., Windpassinger C. 2009. Consequences of mutations within the C terminus of the FHL1 gene. *Neurology* 73, 7: 543-551
- Senderek J., Garvey S. M., Krieger M., Guergueltcheva V., Urtizberea A., Roos A., Elbracht M., Stendel C., Tournev I., Mihailova V. in sod. 2009. Autosomal-dominant distal myopathy associated with a recurrent missense mutation in the gene encoding the nuclear matrix protein, matrin 3. *American Journal of Human Genetics*, 84, 4: 511-518
- Shen H., Lu C., Jiang Y., Tang J., Chen W., Zhang H., Zhang Q., Wang J., Liang J., Hu Z. 2011. Genetic variants in ultraconserved elements and risk of breast cancer in Chinese population. *Breast Cancer Research and Treatment*, 128, 3: 855-861
- Sherr E. H. 2003. The ARX story (epilepsy, mental retardation, autism, and cerebral malformations): one gene leads to many phenotypes. *Current Opinion in Pediatrics*, 15, 6: 567-571
- Shrimpton A. E., Levinsohn E. M., Yozawitz J. M., Packard D. S., Cady R. B., Middleton F. A., Persico A. M., Hootnick D. R. 2004. A HOX gene mutation in a family with isolated congenital vertical talus and Charcot-Marie-Tooth disease. *American Journal of Human Genetics*, 75, 1: 92-96
- Siepel A., Bejerano G., Pedersen J. S., Hinrichs A. S., Hou M. M., Rosenbloom K., Clawson H., Spieth J., Hillier L. W., Richards S. in sod. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research*, 15, 8: 1034-1050
- Smedley D., Haider S., Durinck S., Pandini L., Provero P., Allen J., Arnaiz O., Awedh M. H., Baldock R., Barbiera G. in sod. 2015. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research* 43, W1: W589-W598

- Soravia C., Sugg S. L., Berk T., Mitri A., Cheng H., Gallinger S., Cohen Z., Asa S. L., Bapat B. V. 1999. Familial adenomatous polyposis-associated thyroid cancer - A clinical, pathological, and molecular genetics study. *American Journal of Pathology*, 154, 1: 127-135
- Stephen S., Pheasant M., Makunin I. V., Mattick J. S. 2008. Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock. *Molecular Biology and Evolution*, 25, 2: 402-408
- Takahashi Y., Kou I., Takahashi A., Johnson T. A., Kono K., Kawakami N., Uno K., Ito M., Minami S., Yanagida H. in sod. 2011. A genome-wide association study identifies common variants near LBX1 associated with adolescent idiopathic scoliosis. *Nature Genetics*, 43, 12: 1237-1240
- Tuffery-Giraud S., Saquet C., Thorel D., Disset A., Rivier F., Malcolm S., Claustres M. 2005. Mutation spectrum leading to an attenuated phenotype in dystrophinopathies. *European Journal of Human Genetics*, 13, 12: 1254-1260
- Windpassinger C., Schoser B., Straub V., Hochmeister S., Noor A., Lohberger B., Farra N., Petek E., Schwarzbraun T., Ofner L. in sod. 2008. An X-linked myopathy with postural muscle atrophy and generalized hypertrophy, termed XMPMA, is caused by mutations in FHL1. *American Journal of Human Genetics*, 82, 1: 88-99
- Wojcik S. E., Rossi S., Shimizu M., Nicoloso M. S., Cimmino A., Alder H., Herlea V., Rassenti L. Z., Rai K. R., Kipps T. J. in sod. 2010. Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer. *Carcinogenesis*, 31, 2: 208-215
- Woolfe A., Goodson M., Goode D. K., Snell P., McEwen G. K., Vavouri T., Smith S. F., North P., Callaway H., Kelly K. in sod. 2005. Highly conserved non-coding sequences are associated with vertebrate development. *Plos Biology*, 3, 1: 116-130
- Wu Y., Arai A. C., Rumbaugh G., Srivastava A. K., Turner G., Hayashi T., Suzuki E., Jiang Y. W., Zhang L. L., Rodriguez J. in sod. 2007. Mutations in ionotropic AMPA receptor 3 alter channel properties and are associated with moderate cognitive impairment in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 46: 18163-18168
- Yang R., Frank B., Hemminki K., Bartram C. R., Wappenschmidt B., Sutter C., Kiechle M., Bugert P., Schmutzler R. K., Arnold N. in sod. 2008. SNPs in ultraconserved elements and familial breast cancer risk. *Carcinogenesis*, 29, 2: 351-355
- Yates A., Akanni W., Amode M. R., Barrell D., Billis K., Carvalho-Silva D., Cummins C., Clapham P., Fitzgerald S., Gil L. in sod. 2016. Ensembl 2016. *Nucleic Acids Research*, 44, D1: D710-D716
- Zahorakova D., Rosipal R., Hadac J., Zumrova A., Bzduch V., Misovicova N., Baxova A., Zeman J., Martasek P. 2007. Mutation analysis of the MECP2 gene in patients of Slavic origin with Rett syndrome: novel mutations and polymorphisms. *Journal of Human Genetics*, 52, 4: 342-348

ZAHVALA

Najlepše se zahvaljujem mentorici izr. prof. dr. Tanji Kunej za vse njene nasvete, komentarje, predloge in vedno hitro odgovorjena e-poštna sporočila. Brez njene vsestranske pomoči in vzpodbud diplomsko delo ne bi bilo takšno, kot je. Hvala tudi somentorju izr. prof. dr. Janezu Koncu.

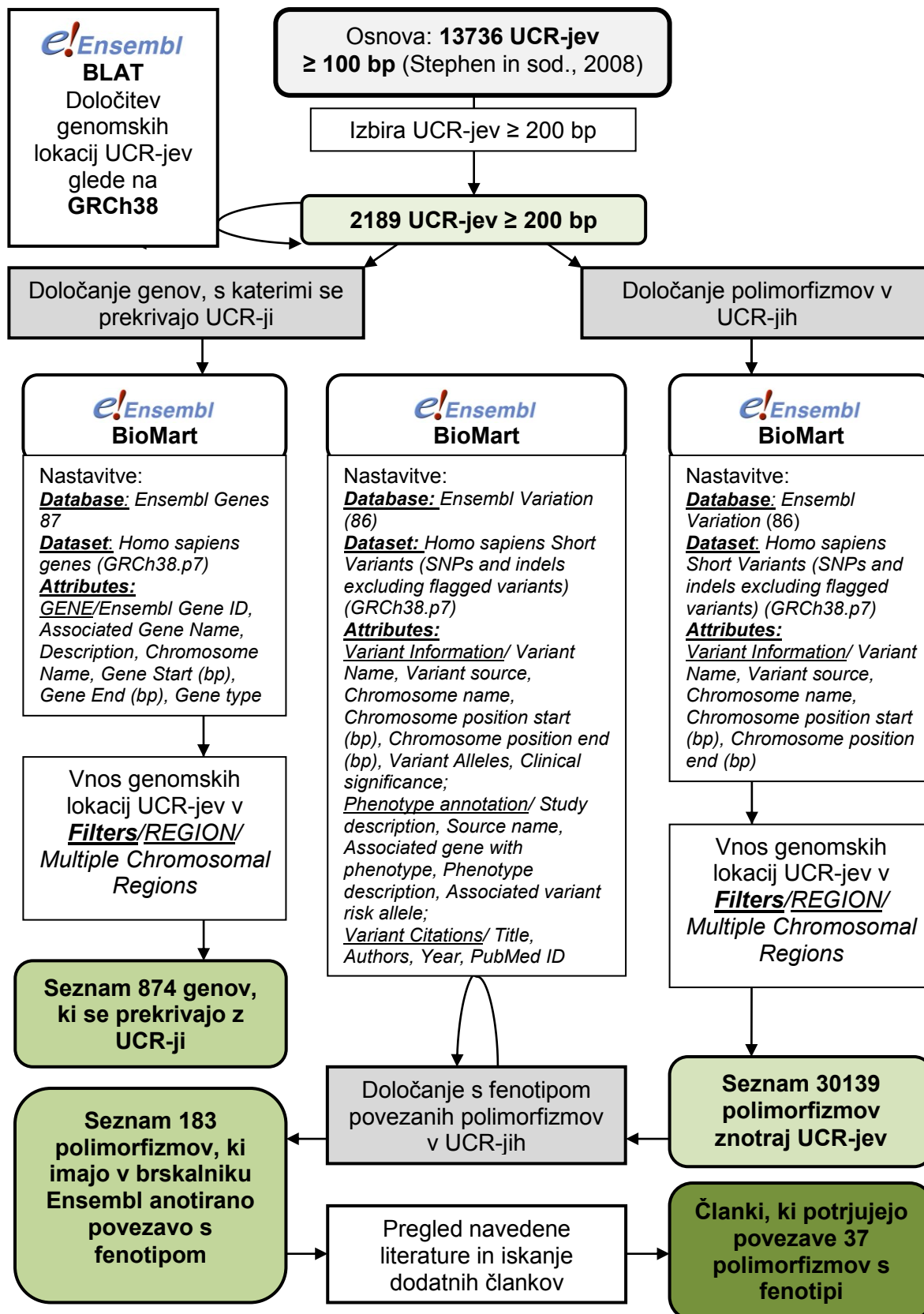
Velika zahvala gre tudi sestri Andreji za podporo v času izdelave naloge, predvsem pa za njeno pomoč pri vseh korakih, ki vključujejo programiranje. Njeno znanje mi je zelo olajšalo delo, prihranilo veliko časa in veliko prispevalo h kakovosti in preglednosti končnih rezultatov. Hvala tudi ostalim družinskim članom, ki so me ves čas študija podpirali in verjeli vame.

Hvala fantu za vso pomoč in vzpodbudne besede v trenutkih, ko sem jih najbolj potrebovala.

Zahvaljujem se tudi prijateljem, s katerimi smo skupaj preživljali čas v klopeh, laboratorijih in še kje, včasih skupaj obupovali, predvsem pa si vedno pomagali in si stali ob strani. Študentska leta nam bodo vsem ostala v lepem spominu.

PRILOGA A

Shema poteka dela



PRILOGA B

Geni, razporejeni glede na število UCR-jev, ki jih vsebujejo

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
1 (537)	PRDM16 (5), HNRNPR (105), RP4-799D16.1 (124), ARID1A (134), AHDC1 (150), ZMYM4 (193), ZMYM4-AS1 (193), AGO1 (199), THRAP3 (201), GRIK3 (206), POU3F1 (222), RP5-884C9.2 (224), HIVEP3 (282), YBX1 (289), IPO13 (293), TMEM53 (347), MKNK1 (359), SSBP3 (426), RP11-776H12.1 (454), RP11-436K8.1 (457), CACHD1 (557), ZRANB2-AS2 (580), NEGR1 (580), RP11-93N20.1 (588), ADGRL2 (608), LINC01364 (667), ZNF326 (727), RP4-665J23.1 (765), RP4-586O15.1 (786), RP11-147C23.1 (799), MIR137HG (834), TRIM33 (894), CSDE1 (898), BCL9 (912), CHTOP (927), ASH1L (939), DNM3 (1018), DNM3OS (1018), RP1-35C21.2 (1027), CEP350 (1029), LHX4 (1030), EDEM3 (1038), SERTAD4 (1097), RCOR3 (1103), RP11-323I1.1 (1126), PROX1-AS1 (1129), RP11-53A1.3 (1158), RP11-152L7.2 (1251), ARID4B (1289), RP11-271F18.1 (1355), CELF2 (1365), SKIDA1 (1389), C10orf67 (1400), DRGX (1442), ERCC6 (1446), SIRT1 (1501), HNRNPH3 (1505), TET1 (1509), RP11-399K21.11 (1557), ZNF503-AS2 (1558), KCNMA1 (1663), KCNMA1-AS1 (1663), RP11-90J7.3 (1679), LINC00856 (1679), IDE (1713), HHEX (1715), HIF1AN (1739), RP11-107I14.5 (1808), SHTN1 (1932), EMX2OS (1939), EMX2 (1939), RAB11FIP2 (1958), RP11-354M20.3 (1958), LINC01163 (2021), RP11-234G16.5 (2082), GLRX3 (2129), FAM160A2 (2138), STK33 (2166), RP11-685M7.5 (2172), EIF4G2 (2172), TEAD1 (2187), FAR1 (2195), RP11-531H8.1 (2213), RP11-531H8.2 (2216), KCNC1 (2287), LGR4 (2302), MPPED2 (2317), IMMP1L (2341), PAX6 (2364), RCN1 (2374), THEM7P (2379), CCDC73 (2387), PHF21A (2405), CELF1 (2427), NRXN2 (2447), RBM14 (2463), RBM14-RBM4 (2463), PCF11 (2511), DLG2 (2513), GRIA4 (2539), MSANTD2 (2671), RP11-687M24.8 (2682), KIRREL3 (2698), ZBTB44 (2710), DERA (2772), PLEKHA5 (2824), RP11-114G22.1 (2840), RP11-778H2.1 (2890), KIF21A (2914), RP11-413B19.2 (2917), RP11-793H13.8 (2964), PCBP2-OT1 (2964), uc_338 (2964), ATP5G2 (2973), HOTAIR (2982), HOXC11 (2982), HOXC10 (2985), HOXC8 (2990), RP11-834C11.7 (3012), RP11-834C11.6 (3014), LINC01598 (3014), RP11-968A15.8 (3019), HNRNPA1 (3019), RP11-123K3.4 (3037), R3HDM2 (3037), CNOT2 (3075), BTG1 (3100), ANKS1B (3121), RP11-144F15.1 (3145), RFX4 (3145), RIC8B (3148), ANAPC7 (3160), PPP1CC (3162), CUX2 (3166), ATXN2 (3181), MSI1 (3219), XPO4 (3237), HMGB1 (3257), KPNA3 (3295), PCDH17 (3316), SPRY2 (3525), FGF14 (3590), RP11-384J4.2 (3664), RP11-562L8.1 (3710), STRN3 (3754), LINC00609 (3887), SLC25A21 (3908), SAMD4A (3975), TMEM260 (3990), RP11-1085N6.5 (4005), RP11-108M12.3 (4025), AP5M1 (4031), RP11-517O13.3 (4037), PPP2R5E (4054), SRSF5 (4067), RBM25 (4074), IFT43 (4090), RP11-187O7.3 (4103), NRXN3 (4117), PAPOLA (4153), BCL11B (4201), MGA (4388), CTDSPL2 (4405), MYEF2 (4412), GABPB1 (4417), GABPB1-AS1 (4417), ONECUT1 (4427), RORA (4507), ZNF609 (4521), RP11-798K3.4 (4536), SKOR1 (4585), TLE3 (4602), NEO1 (4628), NPTN (4632), CLK3 (4634), MIR9-3HG (4657), LINC01197 (4689), RP11-398J10.2 (4696), NR2F2 (4731), RBFOX1 (4812), C16orf72 (4832), RBBP6 (4856), ATXN2L (4872), RP11-24N18.1 (4872), FBXL19 (4888), C16orf97 (4973), RP11-142G1.3 (4976), RP11-297L17.4 (4996), CTD-2302M15.1 (5050), RP11-26L20.3 (5074), ZNF821 (5121), CTD-2009A10.1 (5166), AC009120.4 (5184), RP11-525K10.3 (5225), SMG6 (5261), PAFAH1B1 (5267), NCOR1 (5301), RAI1 (5303), SREBF1 (5303), ZNF207 (5349), RP11-445F12.1 (5377), LHX1 (5378), CDK12 (5418), NEUROD2 (5423), THRA (5432), NR1D1 (5432), MSL1 (5436), SMARCE1 (5445), RP5-1028K7.3 (5445), ATXN7L3 (5458), WNT3 (5472), se nadaljuje

nadaljevanje priloge B

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
	THRA1/BTR (5501), HOXB2 (5505), HOXB-AS1 (5505), HOXB5 (5515), HOXB6 (5519), HOXB9 (5527), UBE2Z (5530), LUC7L3 (5535), ANKFN1 (5551), NOG (5556), RP11-332H18.5 (5602), MED13 (5612), DDX5 (5630), RP11-318A15.7 (5674), SRSF2 (5674), BAHCC1 (5686), TAF4B (5754), KCTD1 (5757), CHST9 (5759), KLHL14 (5776), RP11-680N20.1 (5787), RP11-95O2.5 (5813), SETBP1 (5909), SLC14A2 (5917), SKOR2 (5937), RP11-157P23.2 (5950), SMAD4 (5976), TCF4-AS1 (6009), LINC01416 (6017), CTD-2008L17.1 (6017), TSHZ1 (6083), ATP9B (6134), AKAP8L (6175), AC006128.2 (6175), URI1 (6189), AC005597.1 (6209), AC008992.1 (6317), LSM14A (6379), DYRK1B (6388), POU2F2 (6400), SNRNP70 (6413), LINC00298 (6448), LINC00299 (6448), MYCN (6453), PUM2 (6467), LTBP1 (6533), FAM98A (6535), SRSF7 (6550), LRPPRC (6565), SIX3 (6618), LINC01121 (6621), RP11-358H12.1 (6640), ACTG1P22 (6652), AHSA2 (6844), WDPCP (6879), AFTPH (6884), AC008074.1 (6884), MEIS1 (6902), AC007392.4 (6928), AC007403.2 (6938), AC010987.5 (6955), EXOC6B (6974), CTNNA2 (6991), LRRTM1 (6991), KCMF1 (6996), AC013727.1 (7031), LINC01102 (7054), RP11-13J10.1 (7066), POU3F3 (7066), AC009303.1 (7088), AC093901.1 (7091), RP11-107E5.3 (7253), ZEB2-AS1 (7266), ZEB2_AS1_4 (7266), AC064865.1 (7311), EPC2 (7389), RP11-366G5.1 (7465), TANK (7509), AC009299.3 (7515), RNA5SP108 (7528), FAP (7558), AC007277.3 (7652), TLK1 (7656), SP3 (7704), LNPK (7777), EVX2 (7782), HOXD-AS2 (7792), HOXD8 (7792), LINC01117 (7800), AC092162.1 (7803), AC079305.11 (7816), HNRNPA3 (7818), UBE2E3 (7827), AC104820.2 (7829), AC098617.1 (7850), TMEFF2 (7850), SF3B1 (7854), PLCL1 (7863), AC005235.1 (7863), AC018717.1 (7876), CLK1 (7906), CDK15 (7908), PARD3B (7917), KLF7 (7928), ERBB4 (7951), AC079610.1 (7954), PAX3 (7990), ASB18 (8037), EBF4 (8050), SLC23A2 (8052), SNAP25-AS1 (8064), SNAP25 (8064), RALGAPA2 (8083), NCOA6 (8123), CPNE1 (8127), RP1-309K20.6 (8127), RBM12 (8127), RALGAPB (8145), SRSF6 (8186), RP11-15M15.1 (8230), NRIP1 (8258), AF127577.13 (8258), PAXBP1 (8281), DONSON (8294), HIRA (8321), C22orf39 (8321), MIATNB (8327), PITPNB (8354), PIK3IP1-AS1 (8374), PATZ1 (8374), MIR6820 (8389), POLR2F (8389), SOX10 (8389), DDX17 (8394), MPPED1 (8424), BRPF1 (8457), RP11-320G10.1 (8489), CTNNB1 (8561), ZDHHC3 (8564), RBM6 (8572), RBM5-AS1 (8574), SMIM4 (8593), PBRM1 (8593), IL17RD (8614), FHIT (8624), PTPRG-AS1 (8639), MAGI1 (8674), RYBP (8779), PCNP (8826), MIR568 (8844), U2SURP (8994), ZIC1 (9010), RP11-639B1.1 (9032), RP11-501O2.4 (9035), ANKUB1 (9042), PFN2 (9042), RP11-362A9.3 (9066), CCNL1 (9073), KPNA4 (9120), TNIK (9179), EIF4G1 (9275), TRA2B (9284), ETV5 (9286), TPRG1 (9300), STX18 (9332), GBA3 (9375), PCDH7 (9419), LINC00682 (9447), BEND4 (9449), CHIC2 (9465), EXOC1 (9476), ADGRL3 (9485), G3BP2 (9506), NAA11 (9529), RP11-452C8.1 (9539), RP11-8L2.1 (9573), RP11-42A4.1 (9580), GRID2 (9594), UBE2D3 (9609), CXXC4 (9618), RP11-269F21.3 (9670), SCOC (9718), HHIP-AS1 (9726), HHIP (9726), TTC29 (9743), ZFR (9898), FST (9968), GPBP1 (9984), ZSWIM6 (9996), C5orf64 (10006), CTD-2249K22.1 (10106), CTC-467M3.1 (10147), APC (10358), UBE2B (10427), PCBD2 (10432), C5orf66 (10435), FAM13B (10444), BRD8 (10446), UBE2D2 (10459), CXXC5 (10465), PURA (10478), IGIP (10486), CTC-359M8.1 (10506), POU4F3 (10506), TCERG1 (10512), FBXO38 (10519), CSNK1A1 (10524), CLINT1 (10539), RP11-32D16.1 (10552), CTC-340A15.2 (10612), FBXW11 (10719), CLK4 (10744), JARID2 (10816), ATXN1 (10821), NUP153 (10828), RP1-242N11.1 (10830), RP4-625H18.2 (10837), SOX4 (10855), C6orf62 (10862), SRSF3 (10883), FOXP4 (10903), RP11-397G17.1 (10938), TFAP2B (10957), PHF3 (11001), EPHA7 (11064), POU3F2 (11147), PNISR (11167), ASCC3 (11181), GRIK2 (11184), OSTM1 (11204), NR2E1 (11212), LACE1 (11218), MED23 (11243), MYB (11253), CITED2 (11275), RP1-67K17.3 (11284), HIVEP2 (11284), RP11-307P5.1 (11292), RP11-230C9.3 (11308), AC073094.4 (11350), NXPH1 (11361), ETV1 (11380), SP8 (11420), DNAH11 (11448), TRA2A

se nadaljuje

nadaljevanje priloge B

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
2 (166)	(11456), HNRNPA2B1 (11469), AC004540.5 (11478), KIAA0087 (11484), HOXA3 (11521), RP1-170019.22 (11521), HOXA-AS3 (11521), HOXA5 (11521), HOXA7 (11526), HOXA11 (11533), HOXA11-AS (11533), EVX1 (11537), GLI3 (11623), VSTM2A (11637), VSTM2A-OT1 (11637), CDK6 (11721), CUX1 (11744), PPP1R3A (11776), LINC01392 (11847), LINC01393 (11847), RP11-458K10.2 (11869), ST7 (11892), ST7-OT3_2 (11892), CTTNBP2 (11895), FEZF1 (11898), FEZF1-AS1 (11898), NRF1 (11919), UBE2H (11920), EZH2 (11957), SMARCD3 (11968), PEBP4 (12029), STC1 (12038), UNC5D (12072), YTHDF3 (12187), RP11-21C4.1 (12212), CYP7B1 (12223), EYA1 (12249), RP11-587H10.2 (12407), CTD-3239E11.2 (12414), TRIQK (12414), LINC00535 (12428), OSR2 (12444), VPS13B (12450), RP11-574O7.1 (12481), EIF3H (12550), EXT1 (12554), PHF20L1 (12566), ZC3H3 (12573), UHRF2 (12607), LINC00583 (12624), CCDC171 (12668), C9orf92 (12678), RNF SRRM1 (120, 121), KHDRBS1 (176, 180), SFPQ (191, 192), AGLB4 (367, 371), SRSF11 (574, 575), FUBP1 (594, 596), RP5-1052I5.2 (636, 639), LINC01140 (636, 639), DPYD-AS1 (819, 820), PRPF38B (866, 867), RSNB1 (884, 885), RPS6KC1 (1110, 1111), GPATCH2 (1235, 1237), WDR26 (1275, 1276), ZBTB18 (1317, 1321), HNRNPU (1332, 1334), ARID5B (1471, 1482), CCAR1 (1513, 1515), ADK (1536, 1537), KAT6B (1544, 1546), CPEB3 (1709, 1711), LCOR (1728, 1729), TCF7L2 (1913, 1915), LINC00867 (1972, 1974), TIAL1 (1982, 1983), RP11-282I1.1 (2007, 2011), LMO1 (2148, 2151), EMSY (2492, 2493), RP11-64D24.2 (2559, 2562), NTM (2724, 2725), MGST1 (2783, 2793), LMO3 (2783, 2793), RP11-444D3.1 (2890, 2894), PDZRN4 (2916, 2917), ARID2 (2931, 2932), PCBP2 (2964, 2968), HOXC5 (2998, 2999), FLJ12825 (3012, 3014), RP11-314P15.2 (3644, 3646), FOXG1-AS1 (3675, 3677), FOXG1 (3677, 3680), RP11-260G13.1 (3715, 3719), PRKD1 (3733, 3740), HECTD1 (3757, 3758), RP11-964E11.3 (3895, 3896), PRPF39 (3939, 3940), ARID4A (4036, 4037), TCF12 (4457, 4458), IQCH-AS1 (4547, 4551), LINC00924 (4696, 4699), SALL1 (4939, 4941), CASC16 (4993, 4996), AC007906.1 (5006, 5007), FTO (5022, 5032), ATXN1L (5120, 5121), IST1 (5120, 5121), LINC01572 (5132, 5133), RP5-991G20.1 (5137, 5138), MED1 (5411, 5412), HOXB7 (5526, 5527), MSI2 (5562, 5569), VEZF1 (5572, 5573), RP11-159D12.5 (5574, 5575), AQP4-AS1 (5757, 5759), AC012123.1 (5776, 5777), RP11-746B8.1 (5782, 5785), KIAA1328 (5809, 5811), RNF165 (5928, 5929), MEX3C (5977, 5980), HNRNPM (6148, 6150), MSH6 (6629, 6630), FBXO11 (6629, 6630), NRXN1 (6636, 6640), VRK2 (6662, 6664), USP34 (6844, 6848), XPO1 (6852, 6854), SFXN5 (6977, 6978), RP11-76I14.1 (7031, 7035), CCNT2 (7115, 7117), RP11-570L15.1 (7153, 7161), MBD5 (7376, 7380), AC093375.1 (7445, 7447), NR4A2 (7453, 7454), AC008277.1 (7487, 7490), AC009299.2 (7511, 7515), PSMD14 (7524, 7528), TBR1 (7531, 7533), MYO3B (7651, 7652), METAP1D (7670, 7673), AC013461.1 (7685, 7687), MLK7-AS1 (7685, 7687), HOXD10 (7787, 7789), IKZF2 (7954, 7959), KIZ (8096, 8097), RBM39 (8131, 8132), RP5-823G15.5 (8241, 8242), DYRK1A (8311, 8312), RBFOX2 (8381, 8382), TNRC6B (8401, 8407), RBM5 (8574, 8575), ZBTB20-AS1 (8848, 8849), RP11-197K3.1 (8858, 8859), RP11-649A16.1 (9004, 9005), TSC22D2 (9049, 9051), RP11-641D5.2 (9159, 9161), CPEB2 (9352, 9353), RP11-665G4.1 (9352, 9353), LCORL (9363, 9367), DHX15 (9406, 9407), PHOX2B (9440, 9441), RP11-227F19.1 (9440, 9441), RP11-231C18.3 (9465, 9466), LINC01088 (9525, 9529), HNRNPD (9548, 9550), HNRNPDL (9553, 9554), GRIA2 (9804, 9805), LINC01019 (9873, 9874), NIPBL (9908, 9919), SREK1 (10023, 10024), TNPO1 (10033, 10039), TMEM161B-AS1 (10126, 10128), MEF2C (10147, 10154), RP11-133F8.2 (10227, 10230), NR2F1-AS1 (10273, 10277), NR2F1 (10277, 10280), MCTP1 (10328, 10334), MATR3 (10453, 10454), MATR3 (10453, 10454), CTB-178M22.1 (10668, 10669), CPEB4 (10727, 10729), HNRNPH1 (10746, 10748), OFCC1 (10785, 10789), TFAP2A (10800, 10806), TFAP2D (10952, 10954), RP11-321N4.5 (11030, 11031), se nadaljuje

nadaljevanje priloge B

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
3 (59)	11032), KLHL32 (11075, 11077), MMS22L (11080, 11081), MARCKS (11228, 11229), ARID1B (11308, 11312), QKI (11325, 11326), DGKB (11386, 11388), LINC01162 (11426, 11427), SKAP2 (11497, 11500), HOXA2 (11509, 11510), HIBADH (11544, 11546), POU6F2 (11590, 11594), IKZF1 (11635, 11636), ATXN7L1 (11763, 11766), TFEC (11884, 11890), CADPS2 (11899, 11900), RP5-1101C3.1 (11899, 11900), AGAP3 (11961, 11962), RP11-459E5.1 (12026, 12029), CHD7 (12178, 12180), RP11-32K4.1 (12200, 12201), TRPS1 (12532, 12534), PAX5 (12797, 12802), RP11-15B24.5 (12938, 12948), HNRNPK (12951, 12957), SLC44A1 (12989, 12991), PSMB7 (13125, 13127), DMD (13372, 13386), BCOR (13414, 13416), OGT (13518, 13521), STAG2 (13618, 13620), FHL1 (13666, 13667) SRSF10 (114, 118, 119), ELAVL4 (374, 386, 391), LINC00466 (538, 539, 540), PTBP2 (813, 814, 815), ESRRG (1203, 1207, 1220), RP11-370K11.1 (1312, 1314, 1315), FAM204A (1967, 1968, 1970), MGMT (2063, 2065, 2068), ZBTB16 (2552, 2555, 2559), PKNOX2 (2681, 2682, 2692), RP11-239A17.1 (2801, 2804, 2805), TAOK3 (3207, 3210, 3212), NBEA (3264, 3268, 3271), MBNL2 (3554, 3556, 3559), RALGAP1 (3875, 3878, 3879), UNC79 (4140, 4141, 4142), DPH6 (4240, 4241, 4243), DPH6-AS1 (4248, 4256, 4258), RP11-16L14.2 (4278, 4280, 4281), TOX3 (4983, 4984, 4987), CHD9 (5006, 5007, 5009), RPGRIP1L (5015, 5016, 5017), AATF (5379, 5381, 5382), SRSF1 (5574, 5575, 5576), BCAS3 (5592, 5593, 5602), CTD-2130O13.1 (5942, 5950, 5956), LINC01533 (6353, 6355, 6357), AC007131.2 (6757, 6761, 6765), AC074391.1 (6894, 6895, 6898), AC007392.3 (6924, 6927, 6928), AC016910.1 (7197, 7200, 7202), GPD2 (7456, 7460, 7461), DPP4 (7546, 7547, 7548), SLC25A12 (7663, 7665, 7666), HAT1 (7663, 7665, 7666), AGAP1 (8017, 8022, 8026), AC079135.1 (8033, 8034, 8037), RP4-678D15.1 (8233, 8235, 8240), EWSR1 (8367, 8368, 8369), SETD5 (8446, 8447, 8450), SATB1 (8500, 8509, 8512), MBNL1 (9054, 9057, 9066), RP11-290K4.2 (9082, 9083, 9084), CCDC39 (9198, 9200, 9202), OTP (10054, 10055, 10056), RP11-348J24.2 (10219, 10220, 10227), ADGRB3 (11003, 11005, 11010), SYNCRIP (11030, 11032, 11034), DLX6-AS1 (11731, 11734, 11735), AC073626.2 (11787, 11789, 11790), RP11-458K10.1 (11866, 11867, 11868), EBF2 (12044, 12049, 12052), TOX (12155, 12158, 12164), RFX3 (12592, 12593, 12596), TLE4 (12893, 12894, 12897), TLE1 (12916, 12924, 12925), RC3H2 (13072, 13073, 13076), DDX3X (13437, 13439, 13440), GRIA3 (13605, 13606, 13607)
4 (27)	PEX14 (32, 35, 38, 40), DPYD (819, 820, 822, 831), PROX1 (1140, 1142, 1144, 1147), USH2A (1160, 1165, 1177, 1184), AKT3 (1306, 1312, 1314, 1315), VTIIA (1880, 1886, 1887, 1891), DCDC1 (2331, 2335, 2337, 2338), HOXC6 (2990, 2992, 2995, 2996), RP11-834C11.14 (2995, 2996, 2998, 2999), CTD-2384A14.1 (3686, 3687, 3688, 3689), CTD-2251F13.1 (3740, 3744, 3745, 3747), NFAT5 (5112, 5113, 5114, 5115), ZFHX3 (5137, 5138, 5141, 5152), SKAP1 (5486, 5489, 5493, 5501), CCDC178 (5785, 5787, 5793, 5799), CELF4 (5812, 5813, 5820, 5822), CTC-360P9.4 (6349, 6353, 6355, 6357), BAZ2B (7486, 7487, 7490, 7496), SON (8290, 8294, 8290, 8294), RP11-4B14.3 (9227, 9230, 9239, 9240), ADAMTS6 (10014, 10015, 10019, 10020), TBCA (10063, 10068, 10070, 10073), LINC00461 (10131, 10135, 10139, 10140), ST18 (12134, 12136, 12139, 12143), ZFHX4-AS1 (12287, 12290, 12291, 12292), DENND1A (13106, 13107, 13112, 13115), RP11-343J18.2 (13207, 13208, 13211, 13212)
5 (19)	LMO4 (652, 655, 659, 660, 664), SOX5 (2857, 2864, 2865, 2869, 2872), RNF219-AS1 (3481, 3485, 3491, 3493, 3497), IQCH (4540, 4541, 4543, 4547, 4551), ZNF423 (4898, 4904, 4907, 4912, 4919), ACACA (5387, 5389, 5392, 5397, 5398), HOXB3 (5506, 5507, 5513, 5514, 5515), TCF4 (6000, 6001, 6005, 6009, 6013), TSHZ3 (6297, 6303, 6304, 6306, 6309), RP11-434H14.1 (7169, 7170, 7173, 7176, 7178), TEX41 (7288, 7302, 7304, 7309, 7310), AC092684.1 (7613, 7615, 7617, 7618, 7627), MECOM (9135, 9140, 9146, 9159, 9161), ANKRD17 (9499, 9500, 9501, 9503, 9504, 9505, 9506, 9507, 9508, 9509, 9510, 9511, 9512, 9513, 9514, 9515, 9516, 9517, 9518, 9519, 9520, 9521, 9522, 9523, 9524, 9525, 9526, 9527, 9528, 9529, 9530, 9531, 9532, 9533, 9534, 9535, 9536, 9537, 9538, 9539, 9540, 9541, 9542, 9543, 9544, 9545, 9546, 9547, 9548, 9549, 9550, 9551, 9552, 9553, 9554, 9555, 9556, 9557, 9558, 9559, 9560, 9561, 9562, 9563, 9564, 9565, 9566, 9567, 9568, 9569, 9570, 9571, 9572, 9573, 9574, 9575, 9576, 9577, 9578, 9579, 9580, 9581, 9582, 9583, 9584, 9585, 9586, 9587, 9588, 9589, 9590, 9591, 9592, 9593, 9594, 9595, 9596, 9597, 9598, 9599, 9600, 9601, 9602, 9603, 9604, 9605, 9606, 9607, 9608, 9609, 9610, 9611, 9612, 9613, 9614, 9615, 9616, 9617, 9618, 9619, 9620, 9621, 9622, 9623, 9624, 9625, 9626, 9627, 9628, 9629, 9630, 9631, 9632, 9633, 9634, 9635, 9636, 9637, 9638, 9639, 9640, 9641, 9642, 9643, 9644, 9645, 9646, 9647, 9648, 9649, 9650, 9651, 9652, 9653, 9654, 9655, 9656, 9657, 9658, 9659, 9660, 9661, 9662, 9663, 9664, 9665, 9666, 9667, 9668, 9669, 9670, 9671, 9672, 9673, 9674, 9675, 9676, 9677, 9678, 9679, 9680, 9681, 9682, 9683, 9684, 9685, 9686, 9687, 9688, 9689, 9690, 9691, 9692, 9693, 9694, 9695, 9696, 9697, 9698, 9699, 9700, 9701, 9702, 9703, 9704, 9705, 9706, 9707, 9708, 9709, 9710, 9711, 9712, 9713, 9714, 9715, 9716, 9717, 9718, 9719, 9720, 9721, 9722, 9723, 9724, 9725, 9726, 9727, 9728, 9729, 9730, 9731, 9732, 9733, 9734, 9735, 9736, 9737, 9738, 9739, 9740, 9741, 9742, 9743, 9744, 9745, 9746, 9747, 9748, 9749, 9750, 9751, 9752, 9753, 9754, 9755, 9756, 9757, 9758, 9759, 9760, 9761, 9762, 9763, 9764, 9765, 9766, 9767, 9768, 9769, 9770, 9771, 9772, 9773, 9774, 9775, 9776, 9777, 9778, 9779, 9780, 9781, 9782, 9783, 9784, 9785, 9786, 9787, 9788, 9789, 9790, 9791, 9792, 9793, 9794, 9795, 9796, 9797, 9798, 9799, 9800, 9801, 9802, 9803, 9804, 9805, 9806, 9807, 9808, 9809, 9810, 9811, 9812, 9813, 9814, 9815, 9816, 9817, 9818, 9819, 9820, 9821, 9822, 9823, 9824, 9825, 9826, 9827, 9828, 9829, 9830, 9831, 9832, 9833, 9834, 9835, 9836, 9837, 9838, 9839, 9840, 9841, 9842, 9843, 9844, 9845, 9846, 9847, 9848, 9849, 9850, 9851, 9852, 9853, 9854, 9855, 9856, 9857, 9858, 9859, 9860, 9861, 9862, 9863, 9864, 9865, 9866, 9867, 9868, 9869, 9870, 9871, 9872, 9873, 9874, 9875, 9876, 9877, 9878, 9879, 9880, 9881, 9882, 9883, 9884, 9885, 9886, 9887, 9888, 9889, 9890, 9891, 9892, 9893, 9894, 9895, 9896, 9897, 9898, 9899, 9900, 9901, 9902, 9903, 9904, 9905, 9906, 9907, 9908, 9909, 9910, 9911, 9912, 9913, 9914, 9915, 9916, 9917, 9918, 9919, 9920, 9921, 9922, 9923, 9924, 9925, 9926, 9927, 9928, 9929, 9930, 9931, 9932, 9933, 9934, 9935, 9936, 9937, 9938, 9939, 9940, 9941, 9942, 9943, 9944, 9945, 9946, 9947, 9948, 9949, 9950, 9951, 9952, 9953, 9954, 9955, 9956, 9957, 9958, 9959, 9960, 9961, 9962, 9963, 9964, 9965, 9966, 9967, 9968, 9969, 9970, 9971, 9972, 9973, 9974, 9975, 9976, 9977, 9978, 9979, 9980, 9981, 9982, 9983, 9984, 9985, 9986, 9987, 9988, 9989, 9990, 9991, 9992, 9993, 9994, 9995, 9996, 9997, 9998, 9999, 10000)

se nadaljuje

nadaljevanje priloge B

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
6 (17)	9504), MEF2C-AS1 (10154, 10157, 10164, 10171, 10172), KIAA0825 (10307, 10309, 10312, 10317, 10322), TENM2 (10658, 10660, 10662, 10668, 10669), CNTLN (12718, 12719, 12720, 12721, 12727), MVB12B (13224, 13228, 13230, 13231, 13233) PUM1 (165, 166, 167, 168, 170, 171), ERI3 (295, 301, 303, 305, 313, 316), NFIA (457, 460, 468, 482, 484, 493), PBX1 (962, 968, 969, 972, 973, 978), SDCCAG8 (1298, 1300, 1302, 1303, 1305, 1306), PAX2 (1766, 1771, 1772, 1774, 1777, 1780), HOXB-AS3 (5506, 5507, 5513, 5514, 5515, 5519), CAMKMT (6578, 6580, 6585, 6586, 6590, 6592), BCL11A (6809, 6814, 6817, 6823, 6827, 6832), AC096558.1 (7137, 7140, 7141, 7143, 7144, 7145), GTDC1 (7197, 7200, 7202, 7206, 7207, 7210), FIGN (7585, 7586, 7594, 7604, 7606, 7607), TSHZ2 (8230, 8233, 8235, 8240, 8241, 8242), SATB1-AS1 (8515, 8517, 8521, 8524, 8525, 8526), AP3B1 (10085, 10086, 10087, 10089, 10090, 10093), NFIB (12626, 12631, 12638, 12643, 12648, 12651), ELAVL2 (12745, 12746, 12748, 12749, 12754, 12756)
7 (14)	RNF220 (325, 326, 330, 333, 337, 343, 347), FAF1 (399, 403, 406, 407, 409, 411, 412), BTRC (1812, 1820, 1821, 1827, 1828, 1829, 1831), ELP4 (2347, 2351, 2353, 2356, 2357, 2359, 2360), Z83001.1 (2347, 2351, 2353, 2356, 2357, 2359, 2360), RP11-834C11.12 (2985, 2990, 2992, 2995, 2996, 2998, 2999), NOVA1 (3650, 3651, 3652, 3653, 3655, 3657, 3660), AKAP6 (3761, 3763, 3765, 3766, 3771, 3780, 3783), MIPOL1 (3919, 3920, 3921, 3923, 3926, 3927, 3928), OTX2-AS1 (4005, 4011, 4012, 4013, 4017, 4021, 4025), MEIS2 (4302, 4303, 4311, 4315, 4325, 4329, 4337), RP11-570L15.2 (7137, 7140, 7141, 7143, 7144, 7145, 7153), FAM172A (10288, 10290, 10291, 10292, 10295, 10296, 10298), ZCCHC7 (12804, 12807, 12809, 12812, 12813, 12814, 12816)
8 (5)	HOXC4 (2992, 2995, 2996, 2998, 2999, 3001, 3006, 3007), C15orf41 (4278, 4280, 4281, 4284, 4285, 4287, 4289, 4293), RANBP17 (10681, 10682, 10684, 10695, 10696, 10697, 10698, 10703), ZFH4 (12296, 12300, 12303, 12314, 12315, 12320, 12322, 12331), MAPKAP1 (13146, 13147, 13148, 13149, 13152, 13153, 13155, 13158)
9 (6)	MAP2K5 (4556, 4558, 4561, 4568, 4571, 4573, 4574, 4577, 4582), ZNF407 (6046, 6050, 6051, 6052, 6053, 6060, 6061, 6063, 6067), ZNF536 (6212, 6214, 6217, 6223, 6231, 6236, 6240, 6241, 6247), RP11-444A22.1 (6757, 6761, 6765, 6768, 6782, 6783, 6785, 6795, 6796), LRBA (9756, 9757, 9758, 9759, 9762, 9765, 9768, 9779, 9781), EBF1 (10558, 10567, 10573, 10575, 10581, 10584, 10585, 10587, 10589)
10 (7)	CASZ1 (44, 47, 48, 50, 55, 56, 57, 58, 60, 61), EBF3 (2079, 2082, 2085, 2086, 2087, 2092, 2098, 2099, 2107, 2118), ZNF521 (5710, 5712, 5717, 5718, 5719, 5721, 5727, 5730, 5731, 5734), TBC1D5 (8470, 8472, 8474, 8479, 8482, 8489, 8493, 8500, 8509, 8512), RSRC1 (9089, 9095, 9096, 9098, 9099, 9103, 9106, 9109, 9114, 9115), BNC2 (12686, 12687, 12689, 12692, 12700, 12704, 12709, 12712, 12714, 12715), POLA1 (13326, 13328, 13331, 13335, 13336, 13337, 13339, 13344, 13346, 13347)
11 (2)	OLA1 (7712, 7713, 7714, 7715, 7716, 7721, 7723, 7725, 7727, 7728, 7729), ZBTB20 (8844, 8848, 8849, 8858, 8859, 8865, 8878, 8880, 8888, 8894, 8897)
12 (1)	EHP1 (6855, 6856, 6857, 6860, 6861, 6862, 6863, 6864, 6865, 6866, 6867, 6869)
13 (1)	FOXP1 (8723, 8724, 8730, 8734, 8736, 8739, 8745, 8749, 8752, 8756, 8757, 8758, 8775)
15 (2)	RP11-436D23.1 (11082, 11083, 11084, 11087, 11096, 11098, 11104, 11107, 11110, 11111, 11112, 11119, 11123, 11128, 11130), ZFPM2 (12472, 12473, 12474, 12475, 12478, 12479, 12481, 12484, 12485, 12488, 12495, 12498, 12499, 12505, 12506)
16 (1)	PBX3 (13165, 13167, 13171, 13172, 13175, 13176, 13178, 13181, 13182, 13184, 13186, 13188, 13190, 13192, 13193, 13195)

se nadaljuje

nadaljevanje priloge B

Št. UCR-jev znotraj gena (št. genov s tem št. UCR-jev)	Ime gena (UCR ID)
17 (4)	SOX6 (2242, 2247, 2248, 2250, 2251, 2252, 2257, 2259, 2261, 2262, 2264, 2265, 2268, 2273, 2275, 2277, 2281), DACH1 (3373, 3378, 3383, 3385, 3387, 3390, 3392, 3396, 3398, 3400, 3403, 3407, 3413, 3417, 3419, 3420, 3422), ARHGAP15 (7128, 7131, 7137, 7140, 7141, 7143, 7144, 7145, 7153, 7161, 7164, 7167, 7169, 7170, 7173, 7176, 7178), SOX2-OT (9208, 9210, 9213, 9214, 9216, 9217, 9218, 9221, 9223, 9226, 9227, 9230, 9239, 9240, 9243, 9244, 9250)
18 (3)	LINC01122 (6670, 6675, 6677, 6683, 6684, 6687, 6688, 6693, 6694, 6695, 6696, 6702, 6708, 6712, 6719, 6726, 6728, 6730), ZEB2 (7222, 7226, 7227, 7228, 7229, 7231, 7237, 7238, 7239, 7241, 7246, 7248, 7253, 7254, 7258, 7262, 7263, 7266), AUTS2 (11660, 11669, 11672, 11673, 11676, 11677, 11681, 11688, 11691, 11660, 11669, 11672, 11673, 11676, 11677, 11681, 11688, 11691)
19 (1)	NPAS3 (3795, 3797, 3799, 3800, 3802, 3803, 3808, 3812, 3814, 3819, 3820, 3826, 3835, 3839, 3846, 3848, 3850, 3860, 3863)
21 (1)	C10orf11 (1565, 1571, 1588, 1589, 1591, 1592, 1597, 1600, 1609, 1615, 1616, 1619, 1620, 1623, 1625, 1632, 1633, 1635, 1638, 1639, 1645)
28 (1)	FOXP2 (11777, 11778, 11779, 11780, 11787, 11789, 11790, 11792, 11793, 11795, 11797, 11799, 11804, 11806, 11811, 11812, 11815, 11816, 11818, 11820, 11823, 11825, 11826, 11829, 11831, 11832, 11834, 11835)

PRILOGA C

Preglednica povezav med polimorfizmi in fenotipi, ki so anotirane v genomskem brskalniku Ensembl in smo jih dobili v izpisu rezultatov analize z orodjem BioMart

Polimorfizmi, ki se nahajajo na istem mestu v genomu, so združeni v istem polju preglednice. Če polimorfizem leži znotraj gena, ki ni protein kodirajoč, je biotip gena zapisan v oklepaju. Pod viri so navedeni tudi dodatni viri, ki smo jih naknadno poiskali ročno. Vpisani so dodatni opisi povezanih fenotipov, ki smo jih odkrili pri pregledu literature, popravljeni so napačni opisi fenotipov.

Razlaga okrajšav: polim. – polimorfizem; delec. – delecija; inserc. – insercija; sprem. zapor. – spremenjeno zaporedje; v. – variant

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
CM129343	SNP	Coding sequence v.	199	<i>AGO1</i>	Ni javno dostopen *	HGMD-PUBLIC
rs17105335	SNP	Intron v.	371	<i>AGBL4</i>	Amiotrofična lateralna skleroza	Cronin in sod., 2008
rs1857227	SNP	Intron v.	1160	<i>USH2A</i>	Ulcerozni kolitis	IIBDGC
rs587779201	Inserc.	3 prime UTR v.	6629	<i>FBXO11, MSH6</i>	Sindrom Lynch	ClinVar
rs2020906	SNP	3 prime UTR v.	6629	<i>FBXO11, MSH6</i>	Verjetno ni povezave s sindromom Lynch	Hansen in sod., 2014
rs201802152	SNP	Splice region v.	6640	<i>NRXN1, AC009234.1</i> (AS)	Ni naveden	ClinVar
rs10496382	SNP	Downstream gene v.	7038	/	Ni povezave z višino	Chiang in sod., 2012
rs13382811	SNP	Intron v.	7246	<i>ZEB2</i>	Kratkovidnost (huda)	Khor in sod., 2013
rs368339420	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs796052719	Delec.	Frameshift v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs483352748	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs797045689	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs146031838	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs369869865	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs776228346	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs145808884	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs794727974	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs536900412	SNP	Missense v.	7376	<i>MBD5</i>	Ni naveden	ClinVar
rs778516851	SNP	Missense v.	7380	<i>MBD5</i>	Ni naveden	ClinVar
rs398124344	SNP	Missense v.	7380	<i>MBD5</i>	Ni naveden	ClinVar
rs564759063	SNP	Missense v.	7380	<i>MBD5</i>	Ni naveden	ClinVar

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
rs377062993	SNP	Missense v.	7380	<i>MBD5</i>	Ni naveden	ClinVar
rs114359726	SNP	Missense v.	7380	<i>MBD5</i>	Ni naveden	ClinVar
CM003266	SNP	Coding sequence v.	7453	<i>NR4A2</i>	Ni javno dostopen *	HGMD-PUBLIC
CM003265	SNP	Coding sequence v.	7453	<i>NR4A2</i>	Ni javno dostopen *	HGMD-PUBLIC
rs193920787	SNP	Missense v.	7715	<i>OLA1</i>	Rak prostate	ClinVar
rs104893634; CM041370	SNP; SNP	Missense v.; Coding sequence v.	7789	<i>HOXD10, HOXD9, HOXD-AS2 (AS)</i>	Vertikalni kongenitalni talus; Ni javno dostopen *	Shrimpton in sod., 2004; Dobbs in sod., 2006; HGMD-PUBLIC
CR087569	Sprem. zapor. **	Intergenic v.	9337	/	Ni javno dostopen *	HGMD-PUBLIC
CM090675	SNP	Coding sequence v.	9441	<i>PHOX2B, AC105389.3 (AS), AC105389.1 (lincRNA)</i>	Ni javno dostopen *	HGMD-PUBLIC
CM080495	SNP	Splice region v.	9441	<i>PHOX2B, AC105389.3 (AS), AC105389.1 (lincRNA)</i>	Ni javno dostopen *	HGMD-PUBLIC
rs587784054	SNP	Missense v.	9919	<i>NIPBL</i>	Ni naveden	ClinVar
rs587784055	Delec.	Frameshift v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
rs797045787	Inserc.	Frameshift v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
rs797045788	INDEL	Protein altering v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
rs587784056	Delec.	Frameshift v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
CD107703	Delec.	Coding sequence v.	9919	<i>NIPBL</i>	Ni javno dostopen *	HGMD-PUBLIC
CD107704	Delec.	Coding sequence v.	9919	<i>NIPBL</i>	Ni javno dostopen *	HGMD-PUBLIC
CH107687	Inserc.	Coding sequence v.	9919	<i>NIPBL</i>	Ni javno dostopen *	HGMD-PUBLIC
rs587784057	SNP	Missense v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
rs797045789	INDEL	Frameshift v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1	ClinVar
CD107705	Delec.	Coding sequence v.	9919	<i>NIPBL</i>	Ni javno dostopen *	HGMD-PUBLIC
rs398124474; CM107662	SNP; SNP	Stop gained; Coding sequence v.	9919	<i>NIPBL</i>	Cornelia de Lange sindrom 1; Ni javno dostopen *	ClinVar; HGMD-PUBLIC
rs587784058	SNP	Missense v.	9919	<i>NIPBL</i>	Ni naveden	ClinVar
rs2307121	SNP	Intron v.	10019	<i>ADAMTS6</i>	Struktura roženice	Lu in sod., 2013

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
rs587777277	SNP	Missense v.	10277	<i>NR2F1, NR2F1-AS1</i> (AS)	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs587777275	SNP	Missense v.	10277	<i>NR2F1, NR2F1-AS1</i> (AS)	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs587777274	SNP	Missense v.	10277	<i>NR2F1, NR2F1-AS1</i> (AS)	Bosch-boonstra-schaaf sindrom optične atrofije	Bosch in sod., 2014
rs863224903	SNP	Missense v.	10277	<i>NR2F1, NR2F1-AS1</i> (AS)	Bosch-boonstra-schaaf sindrom optične atrofije	ClinVar
rs587781451	Inserc.	Frameshift v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka	ClinVar
rs387906239	Delec.	Frameshift v.	10358	<i>APC</i>	FAP, blažja oblika	Soravia in sod., 1999
CD014918	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
CI058166	Inserc.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs863224554	SNP	Missense v.	10358	<i>APC</i>	FAP	ClinVar
rs876659673	SNP	Missense v.	10358	<i>APC</i>	Sindrom predispozicije tumorjem	ClinVar
rs748010172	SNP	Synonymous v.	10358	<i>APC</i>	Sindrom predispozicije tumorjem	ClinVar
CM042290	SNP	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs730881221	SNP	Missense v.	10358	<i>APC</i>	Ni naveden	ClinVar
rs786203194	SNP	Missense v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka	ClinVar
CD106396; CD982434	Delec.; Delec.	Coding sequence v.; Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *; Ni javno dostopen *	HGMD-PUBLIC; HGMD-PUBLIC
CD106397	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
CI994627	Inserc.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs148343173	SNP	Synonymous v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka	ClinVar
rs775126020; CM980087	SNP; SNP	Stop gained; Coding sequence v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka; Ni javno dostopen *	ClinVar; HGMD-PUBLIC
rs377665107	SNP	Missense v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka; FAP	ClinVar

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
CM045406	SNP	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs3797704	SNP	Synonymous v.	10358	<i>APC</i>	Ni ovezave z rakom dojke	Chang in sod., 2016
CI058167	Inserc.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs77058194; CM021063	SNP; SNP	Missense v.; Coding sequence v.	10358	<i>APC</i>	Družinski rak debelega črevesa in danke; Ni javno dostopen *	ClinVar; HGMD-PUBLIC
rs786201840	SNP	Missense v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka	ClinVar
CM013690	SNP	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs863225307	SNP	Stop gained	10358	<i>APC</i>	Ni naveden	ClinVar
CD106390	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs760345157	SNP	Synonymous v.	10358	<i>APC</i>	FAP	ClinVar
CD042579	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
CI920920	Inserc.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs387906232	Inserc.	Frameshift v.	10358	<i>APC</i>	FAP	Fodde in sod., 1992
rs863224455 CM990149	SNP; SNP	Stop gained; Coding sequence v.	10358	<i>APC</i>	FAP Ni javno dostopen *	ClinVar; HGMD-PUBLIC
CD106391	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs387906237	Delec.	Frameshift v.	10358	<i>APC</i>	FAP, blažja oblika	Curia in sod., 1998
CD951573	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs587782744	Delec.	Frameshift v.	10358	<i>APC</i>	Dedni sindrom predispozicije za raka	ClinVar
CD072335	Delec.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
CI058168	Inserc.	Coding sequence v.	10358	<i>APC</i>	Ni javno dostopen *	HGMD-PUBLIC
rs121434591; CM092014	SNP; SNP	Missense mutation; Coding sequence v.	10453	<i>MATR3</i> : <i>ENSG00000280987</i> , <i>ENSG00000015479</i>	Distalna miopatija 2; Ni javno dostopen *	Senderek in sod., 2009; HGMD-PUBLIC
rs587777300	SNP	Missense v.	10453	<i>MATR3</i> : <i>ENSG00000280987</i> , <i>ENSG00000015479</i>	Amiotrofična lateralna skleroza 21	Johnson in sod., 2014
rs727505245	SNP	Synonymous v.	10506	<i>POU4F3</i> , <i>AC011396.2</i> (AS)	Ni naveden	ClinVar

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
CM128407	SNP	Coding sequence v.	10957	<i>TFAP2B</i>	Ni javno dostopen *	HGMD-PUBLIC
CR105349	Sprem. zapor. **	Non coding transcript exon v.	11735	<i>DLX6, DLX6-AS1</i> (AS)	Ni javno dostopen *	HGMD-PUBLIC
rs727503937	SNP	Missense v.	11820	<i>FOXP2</i>	Ni naveden	ClinVar
CM095374	SNP	Coding sequence v.	12532	<i>TRPS1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD099281	Delec.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CI054456	Inserc.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD092515	Delec.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD013151	Delec.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD000253	Delec.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CI099366	Inserc.	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM053232	SNP	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM071714	SNP	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM099147	SNP	Coding sequence v.	12554	<i>EXT1</i>	Ni javno dostopen *	HGMD-PUBLIC
rs863223403	SNP	Missense v.	12957	<i>HNRNPK</i>	Sindrom Au-Kline	Au in sod., 2015
rs16928288	SNP	Intron v.	13149	<i>MAPKAP1</i>	Ulcerozni kolitis	IIBDGC
CM1210262	SNP	Coding sequence v.	13195	<i>PBX3</i>	Ni javno dostopen *	HGMD-PUBLIC
rs116373975	SNP	Missense v.	1446	<i>ERCC6</i>	Ni naveden	ClinVar
CM980624	SNP	Coding sequence v.	1446	<i>ERCC6</i>	Ni javno dostopen *	HGMD-PUBLIC
rs121917900	SNP	Stop gained	1446	<i>ERCC6</i>	Cockayne sindrom B	Mallery in sod., 1998
rs786205458	SNP	Missense v.	1537	<i>ADK</i>	Ni naveden	ClinVar
rs74550754	Delec.	Splice acceptor v.	1766	<i>PAX2</i>	Ni naveden	ClinVar
rs76803183	Delec.	Frameshift v.	1766	<i>PAX2</i>	Ni naveden	ClinVar
rs75999420	Delec.	Frameshift v.	1766	<i>PAX2</i>	Ni naveden	ClinVar
rs75462234	Delec.	Frameshift v.	1766	<i>PAX2</i>	Papilorenalni sindrom	Schimmenti in sod., 1999
rs77453353	Inserc.	Frameshift v.	1766	<i>PAX2</i>	Papilorenalni sindrom	Amiel in sod., 2000
rs74676646	Delec.	Inframe deletion	1766	<i>PAX2</i>	Ni naveden	ClinVar
rs76675173	Delec.	Frameshift v.	1766	<i>PAX2</i>	Papilorenalni sindrom	Schimmenti in sod., 1997
rs587777708	SNP	Missense v.	1766	<i>PAX2</i>	Fokalna segmentna glomeruloskleroza 7	Barua in sod., 2014

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
rs11190870	SNP	Intergenic v.	1798	/	Adolescentna idiopatska skolioza (huda); ni povezave z rakom dojke	Takahashi in sod., 2011; Gorman in sod., 2012 Gao in sod., 2013; Jiang in sod., 2013; Miyake in sod., 2013; Londono in sod., 2014; Chettier in sod., 2015; Grauers in sod., 2015; Shen in sod., 2016
rs724159963	SNP	Missense v.	2195	<i>FAR1</i>	Motnja peroksisomalne maščobne acil-CoA reduktaze 1	Buchert in sod., 2014
rs16932455	SNP	Intron v.	2242	<i>SOX6</i>	Občutljivost na kapecitabin	O'Donnell in sod., 2012
rs362420	SNP	3 prime UTR v.	4074	<i>RBM25</i>	Ulcerozni kolitis	IIBDGC
rs997295	SNP	Intron v.	4568	<i>MAP2K5</i>	Potovalna slabost, ITM	Guo in sod., 2013; De in sod., 2015; Hromatka in sod., 2015
rs4548812	SNP	Intron v.	4577	<i>MAP2K5</i>	ITM	GIANT
rs587777373	Inserc.	Inframe Inserc.	4731	<i>NR2F2</i>	Kongenitalni defekti srca, več vrst 4	Al Turki in sod., 2014
rs8053982	SNP	Intron v.	5113	<i>NFAT5</i>	Ulcerozni kolitis	IIBDGC
rs12602067	SNP	Intron v.	5418	<i>CDK12</i>	Ulcerozni kolitis	IIBDGC
rs148373393	SNP	Missense v.	5445	<i>SMARCE1, AC073508.2</i>	Družinski meningiom	ClinVar
CM084901	SNP	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CM040028	SNP	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CD992503	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
rs773880226	SNP	Missense v.	13372	<i>DMD</i>	Duchennova mišična distrofija	ClinVar
rs863224977	SNP	Stop gained	13372	<i>DMD</i>	Duchennova mišična distrofija	ClinVar
CD071339	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CI084915	Inserc.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CX931209	INDEL	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CD0910047	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CD941669	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CM960494	SNP	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
rs398123838	Delec.	Frameshift v.	13372	<i>DMD</i>	Dilatativna kardiomiopatija 3B	ClinVar

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
rs398123839; CD920855; CD972170; CX096771	Delec.; Delec.; Delec.; INDEL	Frameshift v.; Coding sequence v.; Coding sequence v.; Coding sequence v.	13372	<i>DMD</i>	Duchennova mišična distrofija Ni javno dostopen *	Roberts in sod., 1992; Hofstra in sod., 2004; HGMD-PUBLIC
CI071429	Inserc.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CD073610	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
rs398123837	Delec.	Frameshift v.	13372	<i>DMD</i>	Dilatativna kardiomiopatija 3B, Beckerjeva mišična distrofija, Duchennova mišična distrofija	ClinVar
CD115814	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
rs863224976; CM054661	SNP; SNP	Stop gained; Coding sequence v.	13372	<i>DMD</i>	Beckerjeva mišična distrofija Ni javno dostopen *	Tuffery-Giraud in sod., 2005; HGMD-PUBLIC
CD098838	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CI098865	Inserc.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
CD118636	Delec.	Coding sequence v.	13372	<i>DMD</i>	Ni javno dostopen *	HGMD-PUBLIC
rs5927163	SNP	Upstream gene v.	13386	/	Ni naveden	ClinVar
CD084184	Delec.	Coding sequence v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD084183	Delec.	Coding sequence v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM962400	SNP	Splice region v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD099688	Delec.	Splice region v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM051603	SNP	Coding sequence v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CM002826; rs878853076	SNP; SNP	Coding sequence v.; Missense v.	13568	<i>PLP1</i>	Ni javno dostopen *; Ni naveden	HGMD-PUBLIC; ClinVar
rs132630295; CM078425	SNP; SNP	Missense v.; Incomplete terminal codon v.	13568	<i>PLP1</i>	Spastična paraplegija 2, X-vezana; Ni javno dostopen *	Gorman in sod., 2007; HGMD-PUBLIC
CM036008	SNP	Incomplete terminal codon v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
CD004759	Delec.	Coding sequence v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC
rs132630287; CM941173	SNP; SNP	Missense v.; Coding sequence v.	13568	<i>PLP1</i>	Spastična paraplegija 2, X-vezana Ni javno dostopen *	Saugier-Weber in sod., 1994; HGMD-PUBLIC
CM983943	SNP	Coding sequence v.	13568	<i>PLP1</i>	Ni javno dostopen *	HGMD-PUBLIC

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
rs132630292; CM950965	SNP; SNP	Stop gained; Coding sequence v.	13568	<i>PLPI</i>	Bolezen Pelizaeus-Merzbacher, atipična; Ni javno dostopen *	Hodes in sod., 1997; HGMD-PUBLIC
CS114571	Sprem. zapor. **	Coding sequence v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CS910714	Sprem. zapor. **	Coding sequence v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CM991054	SNP	Coding sequence v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CM057971	SNP	Splice region v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CM952021; CS035581; CS035582	Sprem. zapor. **; Sprem. zapor. **; Sprem. zapor. **	Splice acceptor v.; Splice acceptor v.; Splice acceptor v.	13568	<i>PLPI</i>	Ni javno dostopen *; Ni javno dostopen *; Ni javno dostopen *	HGMD-PUBLIC; HGMD-PUBLIC; HGMD-PUBLIC
CS004735	Sprem. zapor. **	Splice donor v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CS004736	Sprem. zapor. **	Splice region v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
CS051696	Sprem. zapor. **	Splice region v.	13568	<i>PLPI</i>	Ni javno dostopen *	HGMD-PUBLIC
rs137852350; CM074896	SNP; SNP	Missense v.; Coding sequence v.	13607	<i>GRIA3</i>	Duševna zaostalost, X-vezana 94 Ni javno dostopen *	Wu in sod., 2007; HGMD-PUBLIC
rs122459149; CM095546	SNP; SNP	Missense v.; Coding sequence v.	13666	<i>FHL1</i>	Mišična distrofija Emery-Dreifuss 6, X-vezana; Ni javno dostopen *	Gueneau in sod., 2009; Knoblauch in sod., 2010; HGMD-PUBLIC
rs122458141; CM080366	SNP; SNP	Missense v.; Coding sequence v.	13666	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic; Ni javno dostopen *	Windpassinger in sod., 2008; Schoser in sod., 2009; HGMD-PUBLIC
rs786200914; CS094840	SNP; Sprem.	Splice donor v.; Splice donor v.	13666	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic;	Schoser in sod., 2009; HGMD-PUBLIC

se nadaljuje

nadaljevanje priloge C

Polimorfizem	Tip polim.	Najhujša posledica polim.	UCR ID	Gen, v katerem se nahaja polim.	Opis povezanega fenotipa	Vir
	zapor. **				Ni javno dostopen*	
rs267606811; CM094843	SNP; SNP	Missense v.; Coding sequence v.	13667	<i>FHL1</i>	X-vezana miopatija z atrofijo posturalnih mišic; Ni javno dostopen *	Schoser in sod., 2009; HGMD-PUBLIC
rs267608352	SNP	3 prime UTR v.	13736	<i>MECP2</i>	X-vezana dovzetnost za avtizem 3	ClinVar
rs267608345	SNP	3 prime UTR v.	13736	<i>MECP2</i>	X-vezana dovzetnost za avtizem 3	ClinVar
rs267608342	Delec.	3 prime UTR v.	13736	<i>MECP2</i>	Rettov sindrom	ClinVar
rs267608341	Inserc.	3 prime UTR v.	13736	<i>MECP2</i>	Ni naveden	ClinVar
rs267608326	SNP	3 prime UTR v.	13736	<i>MECP2</i>	Ni naveden	ClinVar
rs62621672; CR105039	SNP; Sprem. zapor. **	3 prime UTR v.; 3 prime UTR v.	13736	<i>MECP2</i>	Ni povezave z Rettovim sindromom; Ni javno dostopen *	Zahorakova in sod., 2007; HGMD-PUBLIC
rs267608344	SNP	3 prime UTR v.	13736	<i>MECP2</i>	Ni naveden	ClinVar
rs267608347	SNP	3 prime UTR v.	13736	<i>MECP2</i>	Ni naveden	ClinVar
rs267608393	Delec.	Coding sequence v.	13736	<i>MECP2</i>	Ni naveden	ClinVar
rs199963992	SNP	3 prime UTR v.	13736	<i>MECP2</i>	Rettov sindrom Ni naveden	ClinVar
rs267608637	Delec.	Coding sequence v.	13736	<i>MECP2</i>	Rettov sindrom	ClinVar
rs144008995	SNP	3 prime UTR v.	13736	<i>MECP2</i>	Ni naveden	ClinVar

* Povezava je anotirana v HGMD

** Tip polimorfizma ni naveden, podrobnejši podatki so v zasebni podatkovni zbirki HGMD