

UNIVERZA V LJUBLJANI  
BIOTEHNIŠKA FAKULTETA  
ŠTUDIJ BIOTEHNOLOGIJE

Eva EH

**INTEGRACIJA mikroRNA (miRNA) V GENSKE MREŽE PRI  
KRONI NI LIMFOCITNI LEVKEMIJI**

DIPLOMSKO DELO  
Univerzitetni študij

**INTEGRATION OF microRNAs (miRNAs) INTO GENE NETWORKS  
ASSOCIATED WITH CHRONIC LYMPHOCYTIC LEUKEMIA**

GRADUATION THESIS  
University studies

Ljubljana, 2009

*Everything is connected ...  
no one thing can change by itself.*

*Vse je povezano ...  
nobena stvar se ne spremeni sama po sebi.*

Paul Hawken

Diplomsko delo je zaključek univerzitetnega študija biotehnologije. Opravljeno je bilo na Katedri za genetiko, animalno biotehnologijo in imunologijo Oddelka za zootehniko Biotehniške fakultete Univerze v Ljubljani.

Študijska komisija med oddel nega dodiplomskega študija biotehnologije je na seji dne 4.6.2009 za mentorico dela imenovala doc. dr. Tanjo Kunej.

Recenzent: prof. dr. Peter DOV

Komisija za oceno in zagovor:

Predsednica: prof. dr. Branka JAVORNIK  
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za agronomijo  
Članica: doc. dr. Tanja KUNEJ  
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko  
Član: prof. dr. Peter DOV  
Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Datum zagovora:

Naloga je rezultat lastnega raziskovalnega dela.

Podpisana se strinjam z objavo svoje naloge v polnem tekstu na spletni strani Digitalne knjižnice Biotehniške fakultete. Izjavljam, da je naloga, ki sem jo oddala v elektronski obliki, identična tiskani verziji.

Eva EH

KLJU NA DOKUMENTACIJSKA INFORMACIJA

ŠD Dn  
DK UDK 616:575(043.2)=163.6  
KG medicina/molekularna genetika/kroni na limfocitna levkemija/miRNA/genske mreže/sistemska biologija/genska regulacija/kandidatni geni/podatkovne zbirke  
KK AGRIS /  
AV EH, Eva  
SA KUNEJ, Tanja (mentorica)  
KZ SI-1000 Ljubljana, Jamnikarjeva 101  
ZA Univerza v Ljubljani, Biotehniška fakulteta, Študij biotehnologije  
LI 2009  
IN INTEGRACIJA mikroRNA (miRNA) V GENSKE MREŽE PRI KRONI NI LIMFOCITNI LEVKEMIJI  
TD Diplomsko delo (univerzitetni študij)  
OP XI, 91 str., 9 pregl., 15 sl., 3 pril., 198 vir.  
IJ sl  
JI sl/en  
AI Razumevanje genske regulacije je v biologiji raka še vedno velik izziv. eprav je v uporabi veliko razli nih metodoloških pristopov k identifikaciji sprememb na ravni genov, pa z nobenim od njih ne pridobimo zadovoljivega vpogleda v medsebojne odnose razli nih molekulskih mehanizmov. Na primeru kroni ne limfocitne levkemije (KLL) je bilo pred kratkim prvi dokazano, da igrajo pri razvoju in napredovanju raka pomembno vlogo interakcije z regulatornimi molekulami mikroRNA (miRNA). Namen diplomskega dela je bil združiti vsa obstoje a znanja o genskih lokusih in miRNA, povezanih s KLL, prikazati integrativni genomski pristop ter na tej osnovi analizirati domnevne vzroke in posledice deregulacije miRNA pri KLL. Izdelali smo podatkovno zbirko, ki vsebuje ca. 2000 genskih lokusov za KLL, na osnovi katere smo identificirali 241 mo nejših kandidatnih genov in analizirali njihove glavne biološke funkcije. Na podlagi genske lokacije v genih za KLL smo predpostavili novo vlogo pri KLL za 15 miRNA. Odkrili smo 67 novih domnevnih povezav med geni za KLL in miRNA za KLL. Pokazali smo genetsko variabilnost 24 genov za miRNA, šestih tar miRNA s skupno 20 SNP-ji in treh komponent z vlogo v utiševalnem mehanizmu, ki predstavljajo potencialne ozna evalce za KLL. Vizualizirali smo genske mreže z molekulami miRNA pri KLL, pokazali vlogo 37 novih vozliš v genskih mrežah za KLL ter potrdili že poznano vpletjenost nekaterih genov. Ugotovitve diplomskega dela bodo prispevale k boljšemu razumevanju KLL, prikazan sistemski pristop pa bo mogo e aplicirati pri analizi drugih poligenskih bolezni in kompleksnih lastnosti.

## KEY WORDS DOCUMENTATION

DN Dn  
DC UDC 616:575(043.2)=163.6  
CX medicine/molecular genetics/chronic lymphocytic leukemia/microRNA/gene networks/systems biology/gene regulation/candidate genes/databases  
CC AGRIS /  
AU EH, Eva  
AA KUNEJ, Tanja (supervisor)  
PP SI-1000 Ljubljana, Jamnikarjeva 101  
PB University of Ljubljana, Biotechnical Faculty, Academic Study Programme in Biotechnology  
PY 2009  
TI INTEGRATION OF microRNAs (miRNAs) INTO GENE NETWORKS ASSOCIATED WITH CHRONIC LYMPHOCYTIC LEUKEMIA  
DT Graduation Thesis (University studies)  
NO XI, 91 p., 9 tab., 15 fig., 3 ann., 198 ref.  
LA sl  
AL sl/en  
AB Understanding the genetic regulation remains a major challenge in cancer biology. Although many different methodological approaches for the identification of changes at the genetic level have been used, none of them offers a satisfactory insight into the mutual relations between different molecular mechanisms. Recently, interactions with regulatory molecules microRNAs (miRNAs) have been recognised to play important role in cancer development and progression, initially shown in chronic lymphocytic leukemia (CLL). The aim of the present graduate thesis was to combine all the existing knowledge of the genetic loci and microRNAs associated with CLL, demonstrate integrative genomic approach, and on this basis analyze the putative causes and consequences of miRNA deregulation in CLL. For this purpose a database containing ca. 2000 genetic loci for CLL was constructed, on the basis of which we identified 241 strong candidate genes and analyzed their main biological functions. Based on the genetic location in genes for CLL, a new role in CLL for 15 miRNAs was assumed. We inferred 67 new links between the genes and miRNAs for CLL. Genetic variability in 24 miRNA genes, six miRNA targets with a total of 20 SNPs, and in three components of silencing machinery, which represent potential markers for CLL, was demonstrated. Additionally, gene regulatory networks in CLL with integrated miRNAs were visualized. They revealed 37 new nodes, with a role in CLL and confirmed the already known involvement of certain genes. Conclusions of this thesis provide evidence that systematic approach can be used in the analysis of other polygenic disorders and complex characteristics, and will also improve our understanding of gene regulation in CLL.

KAZALO VSEBINE

str.

Ključna dokumentacijska informacija (KDI) .....	III
Key Words Documentation (KWD).....	IV
Kazalo vsebine .....	V
Kazalo preglednic.....	VIII
Kazalo slik.....	IX
Kazalo prilog .....	X
Okrajšave in simboli .....	XI

<b>1</b>	<b>UVOD .....</b>	<b>1</b>
1.1	NAMEN DELA IN HIPOTEZE	3
<b>2</b>	<b>PREGLED OBJAV .....</b>	<b>5</b>
2.1	SISTEMSKA BIOLOGIJA IN GENSKE MREŽE	5
2.2	miRNA	10
<b>2.2.1</b>	<b>miRNA in rak</b>	<b>11</b>
2.3	LEVKE MIJE	12
2.4	KRONI NA LIMFOCITNA LEVKEMIJA	14
<b>2.4.1</b>	<b>Genetski in epigenetski vzroki kroni ne limfocitne levkemije</b>	<b>18</b>
2.5	ŠTUDIJSKI PRISTOPI DOSEDANJIH GENETSKIH ANALIZ KRONI NE LIMFOCITNE LEVKEMIJE	22
<b>2.5.1</b>	<b>Živalski modeli</b>	<b>22</b>
<b>2.5.2</b>	<b>Citogenetske in molekularno-citogenetske študije</b>	<b>23</b>
<b>2.5.3</b>	<b>Študije genske povezanosti</b>	<b>24</b>
<b>2.5.4</b>	<b>Asociacijske in mutacijske študije</b>	<b>25</b>
<b>2.5.5</b>	<b>Ekspresijske študije</b>	<b>26</b>
<b>2.5.6</b>	<b>Epigenetske študije</b>	<b>30</b>
<b>2.5.7</b>	<b>Študije miRNA</b>	<b>30</b>
<b>3</b>	<b>MATERIALI IN METODE .....</b>	<b>32</b>
3.1	ZBIRANJE LITERATURE	32
3.2	IZDELAVA PODATKOVNE ZBIRKE GENSKIH LOKUSOV, POVEZANIH S KRONI NO LIMFOCITNO LEVKEMIJO	33

3.3	IZBOR MO NEJŠIH KANDIDATNIH GENOV ZA KRONI NO LIMFOCITNO LEVKEMIJO	33
3.4	FUNKCIJSKA ANALIZA GENOV	34
3.5	BIOINFORMACIJSKE ANALIZE miRNA V POVEZAVI S KRONI NO LIMFOCITNO LEVKEMIJO	35
3.5.1	<b>miRNA geni</b>	35
3.5.2	<b>miRNA tar e</b>	36
3.5.3	<b>Komponente utiševalnega mehanizma miRNA</b>	37
3.6	GENSKE MREŽE Z miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI	38
3.6.1	<b>Ingenuity Pathway Analysis (IPA 7.5 software, Ingenuity Systems®, www.ingenuity.com )</b>	38
3.6.2	<b>MetaCore™ (GeneGo, St Joseph, MI, www.genego.com )</b>	39
3.7	SHEMA INTEGRATIVNEGA GENOMSKEGA PRISTOPA	40
4	<b>REZULTATI .....</b>	<b>41</b>
4.1	IZDELAVA PODATKOVNE ZBIRKE GENSKIH LOKUSOV, POVEZANIH S KRONI NO LIMFOCITNO LEVKEMIJO	41
4.2	IZBOR MO NEJŠIH KANDIDATNIH GENOV ZA KRONI NO LIMFOCITNO LEVKEMIJO	43
4.3	FUNKCIJSKA ANALIZA GENOV	44
4.4	BIOINFORMACIJSKE ANALIZE miRNA	47
4.5	GENSKE MREŽE Z miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI	54
5	<b>RAZPRAVA IN SKLEPI .....</b>	<b>60</b>
5.1	UPORABLJENI INTEGRATIVNI PRISTOPI	60
5.1.1	<b>Izdelava podatkovne zbirke genskih lokusov za kroni no limfocitno levkemijo</b>	60
5.1.2	<b>Izbor mo nejni kandidatnih genov</b>	63
5.1.3	<b>Funkcijska analiza genov</b>	64
5.1.4	<b>Bioinformacijske analize miRNA</b>	64
5.1.5	<b>Genske mreže</b>	65
5.2	INTERPRETACIJA POVEZAV miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI	68
5.3	SKLEPI IN PRIHODNJE DELO	70
6	<b>POVZETEK.....</b>	<b>72</b>

7	<b>VIRI.....</b>	<b>73</b>
	<b>ZAHVALA</b>	
	<b>PRILOGE</b>	

## KAZALO PREGLEDNIC

	str.
Preglednica 1: Štirje glavni razredi levkemij	13
Preglednica 2: Pogostnost kromosomskih nepravilnosti pri 325 bolnikih s kroni no limfocitno levkemijo	19
Preglednica 3: Rezultati analize izbranih mo nejših kandidatnih genov za kroni no limfocitno levkemijo (KLL) glede na miRNA z orodjem miRNAPath	46
Preglednica 4: Rezultati analize miRNA pri kroni ni limfocitni levkemiji (KLL) glede na število genov, ki jih regulirajo, z orodjem miRNAPath	46
Preglednica 5: Geni, povezani s KLL, v katerih se nahajajo zapisi za miRNA ter odgovarjajo i polimorfizmi dаних miRNA in polimorfizmi njihovih tar glede na podatkovno zbirko Patrocles	48
Preglednica 6: Geni povezani s KLL, ki imajo funkcijo pri utiševalnem mehanizmu miRNA	49
Preglednica 7: Geni, ki so predhodno že bili povezani s KLL, ter predstavljajo validirane tar e miRNA, za katere smo odkrili, da so prav tako že bile povezane s KLL	49
Preglednica 8: Polimorfni tar ni geni v KLL, z odgovarjajo imi SNP-ji	53
Preglednica 9: Novih 37 kandidatnih genov glede na izrisano gensko mrežo z miRNA s programom IPA	56

## KAZALO SLIK

	str.
Slika 1: Brezlestvi na (angl. scale-free) genska mreža	6
Slika 2: Spletna stran, ki omogoča hiter vpogled v mrežo genskih povezav med obolenji in geni za bolezni, t.i. »diseasome« ( <a href="http://diseasome.eu">http://diseasome.eu</a> )	9
Slika 3: Vennov diagram presev nih množic za gene miRNA pri kroni ni limfocitni levkemiji	35
Slika 4: Vennovi diagrami: a) Presevne množice podatkovne zbirke genov za kroni no limfocitno levkemijo (KLL), validiranih tar miRNA in miRNA za KLL, b) Presevna množica podatkovne zbirke genov za KLL in polimorfnih tar miRNA, c) Presevna množica miRNA, povezanih s KLL in miRNA s polimorfnimi tarami, d) Presevna množica podatkovne zbirke genov za KLL, in tar nih genov miRNA, ki se nahajajo v genih, povezanih s KLL	37
Slika 5: Presevna množica podatkovne zbirke genov za KLL in polimorfnih komponent utiševalnega mehanizma	38
Slika 6: Shema integrativnega genomskega pristopa	40
Slika 7: Število lokusov, povezanih s kroni no limfocitno levkemijo, odkritih z različnimi študijskimi pristopi	41
Slika 8: NCBI Map Viewer je brskalno orodje, ki omogoča iskanje in grafični prikaz genomskeh informacij glede na citogenetski lokus za mnogo različnih vrst. Za kroni no limfocitno levkemijo je posebej veliko zadetkov za napake na kromosomih 6 in 13.	42
Slika 9: Rezultati analize funkcij močnejših kandidatnih genov za kroni no limfocitno levkemijo (KLL) z orodji podatkovne zbirke DAVID. a) Prve štiri anotirane skupine genov, b) Najbolj zastopane kromosomalne lokacije genov za KLL, c) Anotacija s pojmi bolezni, d) Anotacija z glavnimi biološkimi potmi iz baze KEGG, e) Glavni sekvenčni motivi genov za KLL	45
Slika 10: Deleži genov za kroni no limfocitno levkemijo (KLL), glede na vlogo, povezano z miRNA	47
Slika 11: Primerjava deleža polimorfnih miRNA glede na 706 miRNA v registru podatkovne zbirke miRBase ter 197 miRNA, za katere smo odkrili povezanost s KLL	53
Slika 12: Genska mreža pri KLL z integriranimi miRNA, izrisana s programom Ingenuity Pathway Analysis	55
Slika 13: Genska mreža pri KLL z integriranimi miRNA, izrisana z algoritmom »Shortest path« v programu MetaCore, z uporabo orodja Metalink	57
Slika 14: Genska mreža pri KLL, izrisana z združevanjem podmrež s programom MetaCore ter orodjem MetaLink	58
Slika 15: Legenda k sliki 13 in sliki 14	59

## KAZALO PRILOG

- Priloga A: Prese na množica izbora mo nejših kandidatnih genov za kroni no limfocitno levkemijo
- Priloga B: Polimorfne miRNA, povezane s kroni no limfocitno levkemijo (KLL)
- Priloga C: miRNA, ki so predhodno že bile povezane s KLL, in imajo glede na podatkovno zbirko Patrocles polimorfne tar e

## OKRAJŠAVE IN SIMBOLI

### **Okrajšava Pomen**

aCGH	CGH na mreži (angl. array CGH)
ALL	akutna limfoidna levkemija
AML	akutna mieloidna levkemija
BCR	B-celi ni receptor
BIND	angl. Biomolecular Interaction Network Database
CD	ozna evalec pripadnosti (angl. Cluster of Differentiation)
CDR	regije, ki dolo ajo komplementarnost (angl. complementarity determining regions)
CGH	komparativna genomska hibridizacija
cCGH	klasi na CGH
cDNA	komplementarna DNA
Cy	cianin
DAVID	angl. The Database for Annotation, Visualization and Integrated Discovery
DIP	angl. Database of Interacting Proteins
FDR	delež napa no pozitivnih zadetkov (angl. false discovery rate)
FISH	fluorescen na <i>in situ</i> hibridizacija
FPRP	verjetnost lažno pozitivnih zadetkov (angl. false positive report probability)
GEO	angl. Gene Expression Omnibus
GMOD	angl. Generic Model Organism Database
HCDR	regije, ki dolo ajo komplementarnost na težki verigi imunoglobulina (angl. heavy chain complementarity determining regions)
HUGO	angl. Human Genome Organisation
Ig	imunoglobulin
IgVH	variabilna regija težke verige imunoglobulina (angl. immunoglobulin heavy chain variable region)
IPA	angl. Ingenuity Pathway Analysis
KEGG	angl. Kyoto Encyclopedia of Genes and Genomes
KLL	kroni na limfocitna levkemija
KML	kroni na mieloidna levkemija
LDT	as podvojitve števila limfocitov (angl. lymphocyte doubling time)

LOD	angl. log 10 of the odds
MINT	angl. Molecular Interactions Database
MGI	angl. Mouse Genome Informatics
miRAGE	zaporedna analiza genskega izražanja (SAGE) za genomsko analizo miRNA
miRNA	mikroRNA
MM	multipli mielom
NCBI	angl. The National Center for Biotechnology Information
NHL	Ne-Hodgkinov limfom
NPL	neparametri na vezava
NZB	sev miši NZB (angl. New Zealand Black)
OS	skupni as preživetja (angl. overall survival)
PFS	as preživetja brez napredovanja bolezni (angl. progression free survival)
RAKE	angl. RNA-primed Array-based Klenow Extension
RGD	angl. Rat Genome Database
RISC	z RNA inducirani utiševalni kompleks (angl. RNA-induced silencing complex)
RLGS	angl. Restriction Landmark Genomic Scanning
ROMA	predstavitevna analiza oligonukleotidne mikromreže (angl. representational oligonucleotide microarray analysis)
SLL	limfom majhnih limfocitov (angl. small lymphocytic lymphoma)
SNP	polimorfizem posameznih nukleotidov (angl single-nucleotide polymorphism)
SOM	samoorganizacijske mreže (angl. self- organising maps)
SVM	metoda podpornih vektorjev (support vector machine)
TFS	as preživetja brez zdravljenja (angl. treatment free survival)
TS	angl. tail strength
T-UCR	prepisan UCR (angl. transcribed UCR)
UCR	ultra-ohranjena regija (angl. ultraconserved region)
UTR	neprevedljiva regija (angl. untranslated region)
VDJ	genski segmenti V (angl. variable), D (angl. diversity), J (angl. joining)
WHO	svetovna zdravstvena organizacija (angl. World Health Organisation)

## 1 UVOD

Že iz asov Isaaca Newtona je v znanosti ukoreninjen analitičen in razmišljanja, s katerim sistem razumevamo na posamezne elemente. Če bi poskusili obravnavati celotno stvarnost naenkrat, nas bi ta preplavila. Vendar pa ko v ta na in razmišljanja postavimo sebe v odnos do sveta v katerem živimo, prenehamo smatrati takšna stališča kot uporabna in spoznamo, da svet ni sestavljen samo iz predmetov in gradnikov, temveč obstajajo tudi odnosi in procesi med temi in je torej bolj ustrezan holističen oz. relativističen pogled na svet.

Do nedavnega je bil v medicini običajen pristop k odkrivanju genskih osnov bolezni preko evanje posameznih kandidatnih genov. Ta vrsta raziskav je pripomogla k razumevanju, diagnostiki in terapiji širokega spektra bolezni. Navkljub uspehom pa ima reduktionističen pristop tudi mnogo omejitev, še posebej pri preku evanju poligeničnih obolenj, kot je rak, kjer je identifikacija vzročnih genov s standardnimi testi za posamezne lokuse zaradi interakcij z genetskimi in okoljskimi faktorji otežena.

V zadnjem desetletju je poznavanje zaporedij DNA ter boljše razumevanje genoma in genskega izražanja ponudilo nove strategije za identifikacijo polimorfizmov povezanih z boleznimi. S hitrim napredkom so tehnike določanja zaporedja DNA, mikromrež in druge nove tehnologije omogočile prepoznavanje ključnih regulatornih elementov in genskega izražanja na globalni ravni genoma ter z vedno večjo dostopnostjo postale glavno gonilo moderne biologije. Že danes so ustvarile veliko količino podatkov, za katero lahko prikujujemo, da bo v prihodnosti še večja in bo tako sprožila učinek snežne kepe.

Največ truda in finančnih vlaganj je bilo namenjenih ravno biologiji raka, vendar kljub veliki dostopnosti informacij, te kompleksne bolezni še vedno ne razumemo. Ozko grlo raziskav dandanes namreč ne predstavlja več pomanjkanje analitskih podatkov, temveč pomanjkanje učinkovitih metod za identifikacijo in validacijo pomembnejših vzročnih dejavnikov, njihovih funkcij ter interakcij.

Za svoje enje s tem izzivom je potrebna integracija vseh znanih podatkov različnih študijskih in multidisciplinarnih pristopov v smiselnega celota. V ta namen se v bioznanostih uporablja risanje mrež, ki prikazujejo kompleksnost bioloških odnosov in omogočajo informativen sistemski vpogled v molekularne mehanizme, s tem pa tudi postavljanje novih vprašanj, hipotez in zasnovno laboratorijskih eksperimentov.

Ker je rak v osnovi genska bolezen, pri kateri pride do spremenjenega izražanja tako kodirajočih kot nekodirajočih regij, smo se odločili raziskati gensko mrežo z enim največjih razredov genskih regulatorjev: mikroRNA (miRNA), dolgo neopažena skupina nekodirajočih RNA, ki popolnoma spremeni razumevanje genske regulacije in privredila do spoznanja, da molekule RNA tvorijo kompleksne genske interakcije ter so zavestno in prostorsko specifiko ključno vplivajo na ekspresijo tistih genov.

Kot modelno bolezen smo uporabili kronično limfocitno levkemijo (KLL), ki smo jo izbrali zaradi njene poligenske narave in dobro raziskane vpleteneosti miRNA v razvoju bolezni. Povezava med miRNA in raki je bila namreč prvič dokazana ravno pri KLL (Calin in sod., 2002). Novejša odkritja pa kažejo, da so molekule miRNA vpletene v patogenezo vseh do sedaj analiziranih tipov raka (Aqeilan in sod., 2009).

Vzroki tistih genov s pomembnejšim vplivom na nagnjenost k KLL še niso odkrili, eprav obstaja veliko dokazov za gensko osnovo KLL. Zelo verjeten razlog za to so spremembe uravnavanja genske ekspresije več lokusov z manjšimi prispevki. Zato nas je v diplomskem delu zanimalo, kakšna je pri tem vloga miRNA.

Pri KLL obstajajo torej jasne indikacije o deregulaciji genov miRNA, hkrati pa smo ob pregledu obstoječe literature zaznali pomanjkanje analiz te bolezni z metodami sistemskih biologij. Zato smo se odločili preveriti, ali lahko z integracijo vseh do sedaj znanih genskih vzrokov v regulatorne genske mreže identificiramo funkcije in interakcije ključnih genskih faktorjev ali popolnoma novih genov kandidatov za KLL ter novih regulatornih poti miRNA pri tej bolezni. Rezultati raziskave nakazujejo tudi pot k razvoju novih diagnostik, prognostičnih in terapevtskih markerjev za KLL.

## 1.1 NAMEN DELA IN HIPOTEZE

Namen diplomskega dela je bila izdelava podatkovne zbirke vseh do sedaj poznanih genetskih vzrokov za KLL, ugotovljenih z vsemi do sedaj uporabljenimi študijskimi pristopi. Na osnovi teh smo narisali genske mreže, ki vklju ujejo miRNA. S prekrivanjem lokusov, povezanih s KLL v ve študijah z razli nimi študijskimi pristopi ali povezanih z ve razli nimi klini nimi stanji in analizo regulatornih mrež smo dolo ili pomembnejše kandidatne lokuse. Z uporabo bioinformacijskih metod smo opredelili glavne funkcije genov za KLL, poiskali nove kandidatne lokuse za KLL ter ugotovili genetsko variabilnost na vseh ravneh z miRNA-posredovane regulacije: na ravni 1) genov za miRNA, 2) njihovih tar ter 3) komponent utiševalnega kompleksa. Na modelu KLL smo razvili integrativni genomski pristop, ki vklju uje informacije, pridobljene z razli nimi študijskimi pristopi, in ga je mogo e aplicirati pri analizi drugih poligenskih bolezni in fenotipov.

V diplomskem delu smo zastavili slede o hipotezo:

Z integrativnimi genomskimi pristopi lahko :

- identificiramo mo nejše že odkrite in popolnoma nove kandidatne gene za KLL, ki predstavljajo potencialne diagnosti ne, prognosti ne in terapevtske markerje
- odkrijemo nove povezave med geni in miRNA pri KLL
- odkrijemo genetsko variabilnost miRNA uravnavanja
- pridobimo boljši vpogled v razumevanje genske regulacije pri KLL

V prvem poglavju diplomskega dela so podane osnove za razumevanje širšega konteksta podro ja sistemski biologije, genskih mrež, molekul miRNA, KLL in študijskih pristopov k preu evanju genetike te bolezni. V drugem poglavju so opisane metode in postopki, ki smo jih uporabili za konstrukcijo podatkovne zbirke, genskih mrež ter bioinformacijske analize. V poglavju z rezultati se nahajajo opis izdelane podatkovne zbirke, ugotovitve do katerih smo prišli z analizo genskih mrež in bioinformacijskih analiz, nove domnevne povezave miRNA z geni pri KLL ter odkriti polimorfizmi razli nih ravni miRNA

mehanizmov pri KLL. V etrtem poglavju pa so predstavljene naše interpretacije uporabljenih študijskih pristopov, odkritih funkcij, polimorfizmov in povezav med miRNA ter geni, vpletenimi v razvoj KLL.

## 2 PREGLED OBJAV

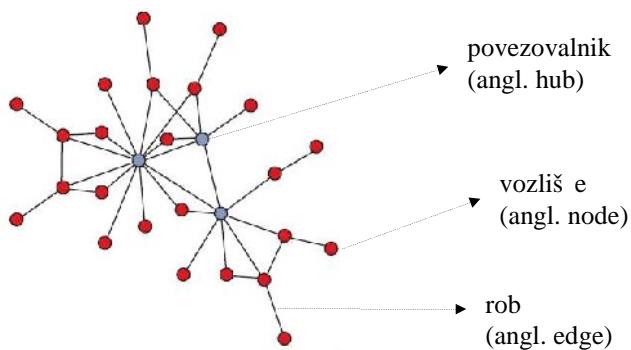
### 2.1 SISTEMSKA BIOLOGIJA IN GENSKE MREŽE

Glavni cilj sistemске biologije je, tako kot namen drugih bioloških ved, razumevanje bioloških procesov. Vendar pa sistemski biologiji za namene procesiranja in integracije podatkov, analize, simulacij in predikcij temelji na celostnem pristopu. Trenutno obstajata dve prevladajoči definiciji sistemski biologije. Prva pravi, da je sistemski biologija integracija različnih podatkov tehnologij »-omik« in predstavlja vidik bioinformatike za razvoj orodij združevanja in integracije podatkov. Druga interpretacija pa poudarja, da gre pri sistemski biologiji za dinamične interakcije genskih produktov, proteinov in celic, ki privedejo do struktur in funkcij celic ter višjih ravni organizacije, kot so tkiva, organi itd. in k tem pristopa z matematičnim modeliranjem in simulacijami (Han, 2008).

Za im bolj jasen vpogled v strukturo in delovanje bioloških sistemov oz. bioloških mrež so potrebni podatki, ki so točni in tehnično nepristranski, zato je integracija informacij nujna. Ta obsega različne pristope, od preprostih analiz s presekom različnih podatkov (von Mering in sod., 2002; Han in sod., 2004; Said in sod., 2004) do bolj sofisticiranih sistemov takovanja, ki temeljijo na verjetnosti (Jansen in sod., 2003; Lee in sod., 2004; Rhodes in sod., 2005). Na primer, Gunsalus in sod. (2005) so z integracijo podatkov transkriptoma, interaktoma in fenoma identificirali molekularne mehanizme, ki delujejo v zgodnji embriogenezi nematode *Caenorhabditis elegans* ter nato še eksperimentalno potrdili biološke funkcije desetih do tedaj še neopisanih proteinov. Mani in sod. (2008) so z uporabo informacij o interakcijah protein-protein, protein-DNA in posttranslacijskih interakcijah z visoko zanesljivostjo identificirali tarce tumorigeneze B-celičnih limfomov.

Ker je dani fenotip rezultat interakcij na različnih ravneh, je pomembno predstaviti sistem v obliki mreže. Tako poznamo več vrst bioloških mrež: proteinske interakcijske mreže, metabolne mreže, signalne mreže, genske interakcijske in regulatorne mreže... Biološke mreže imajo »emergentne« lastnosti, ki niso linearne, temveč so produkt kombinatornih in inkov interakcij med komponentami mreže. Na primer, lastnosti kot so topologija, stabilna stanja mreže ali tok informacij ne moremo zaznati na ravni posameznih genov.

Genske mreže so grafi ni modeli, ki predstavljajo odnose med molekularnimi zna ilnostmi genov in jih sestavljajo vozliš a (angl. nodes), ki predstavljajo gene, in robovi (angl. edges) med vozliš i, ki predstavljajo odnose med geni. Robovi so lahko enosmerni, kar predstavlja recipro en odnos med dvema genoma, ali usmerjeni, kar kaže na vzro ni odnos med dvema genoma. Topologija oz. arhitektura ve ine bioloških mrež je brezlestvi na (angl. scale-free; slika 1), kar pomeni, da mreža vsebuje majhen delež vozliš , ki imajo veliko povezav in predstavljajo povezovalnike (angl. hubs). Ve ina ostalih vozliš ima samo nekaj povezav.



Slika 1: Brezlestvi na (angl. scale-free) genska mreža (Barabási in Oltvai, 2004: 105).

Povezovalniki so obi ajno definirani za ta namen (*ad hoc*) kot vozliš a, ki imajo posebne topološke in funkcionalne lastnosti v organizaciji mreže. Vendar pa ravno zaradi brezlestvi ne narave mrež v literaturi ne obstaja dogovor o natan ni definiciji povezovalnika. Na primer, Han in sod. (2004) so kot povezovalnike identificirali vozliš a z ve kot petimi povezavami, Ekman in sod. (2006) kot vozliš a z ve kot osmimi povezavami, Aragues in sod. (2007) pa kot vozliš a z vsaj dvajsetimi povezavami. Možno je, da so posebne lastnosti povezovalnikov posledica ravno same definicije (Vallabhajosyula in sod., 2009). Posledi no je analiza grafov deloma subjektivna. Zaradi o itne potrebe po tem, da se postavijo natan nejše definicije povezovalnikov, ki jih bo možno aplicirati na razli ne mreže, so Vallabhajosyula in sod. (2009) predlagali tri objektivne kriterije za identifikacijo povezovalnikov. Prvi kriterij temelji na definiciji povezljivosti, saj naj se vozliš a z ve jim številom povezav ne bi v ve ji meri povezovala med seboj (Maslov in sod., 2002; Vázquez in sod., 2003). Drug kriterij temelji na opažanju, da imajo vozliš a z veliko povezavami pogosto esencialno biološko funkcijo. Tretji kriterij za identifikacijo povezovalnikov pa je osnovan na možnosti grupiranja glede

na biološke lastnosti v t.i. »date« ali »party« povezovalnike, ki imajo različne funkcije. Prvi imajo visoko, drugi pa nizko ko-ekspresijo z bližnjimi sosedji v mreži.

Okvare vozlišč z manj povezavami imajo za delovanje celotne mreže manj drastične posledice kot pa okvare večjih povezovalnikov, ki lahko popolnoma onemogojo delovanje celotne mreže (Wang, 2008a). V splošnem so povezovalniki v genskih regulatornih mrežah transkripcijski faktorji, ki nadzirajo funkcije večjega števila genov ter njihove odzive na notranje in zunanje signale. Wang in Purisima (2005) ter Batada in sod. (2006) so odkrili, da imajo transkripcijski faktorji v genskih regulatornih mrežah kvasovki in bakterije *Escherichia coli* hitrejšo razgradnjo mRNA kot pa vozlišča, ki ne predstavljajo povezovalnikov. Iz tega naj bi sledilo, da transkripcijski faktorji, ki predstavljajo povezovalnike, omogočajo hitrejši odziv mreže na zunanje dražljaje. Ravasi in sod. (2009) so v okviru globalnega projekta genomske mreže (angl. The Genome Network Project, [http://genomenetwork.nig.ac.jp/index\\_e.html](http://genomenetwork.nig.ac.jp/index_e.html)) kot prvi pobudniki konstruirali celotno mrežo transkripcijskih faktorjev v celici na primeru diferenciacije celic ne-linijne mieloidne levkemije. Odkrili so, da je mreža transkripcijskih faktorjev redundantna in prilagodljiva – noben transkripcijski faktor ni nujen niti zadosten za diferenciacijski proces.

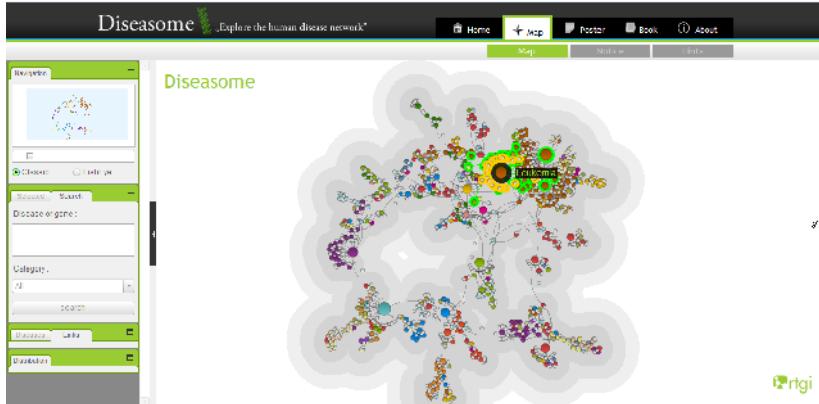
Druge lastnosti brezleskega mrež so še porazdelitev stopnje (število povezav na eno vozlišč), stopnja centralnosti (angl. degree centrality), ki je povezana z esencialnostjo in ohranjenostjo genov (Jeong in sod., 2001; Fraser in sod., 2002) in vmesnost (angl. betweenness), ki predstavlja število najkrajših poti med katerimkoli parom vozlišč, ki potekajo skozi dano vozlišč (Yu in sod., 2007). Vmesnost ima boljšo napovedno sposobnost esencialnosti gena kot stopnja centralnosti.

Vsi in bioloških mrež predstavljajo t.i. mreže tipa »small-world« (Han, 2008). To pomeni, da tvorijo lokalne soseske, kjer je dolžina poti med vozliščema kratka, vozlišča pa imajo podobne funkcije (Watts in sod., 1998). Nadaljnja topološka lastnost mrež je vzorec lokalnih povezav oz. motiv mreže. To so npr. povratne in naprejšnje (angl. feedforward) zanke (Milo in sod., 2002; Wuchty in sod., 2003). V povratnih zankah naj bi se nahajali transkripcijski faktorji s hitro razgradljivo mRNA, ki jim omogočajo hitro odzivanje na

notranje in zunanje stimuluse (Wang, 2008a). Naprejšnje (angl. feedforward) zanke pa se odzivajo samo na stanovitne endogene signale.

Trenutno je zelo malo empiri nega znanja o evoluciji genskih mrež (Shiu and Borevitz, 2006). Motivi mrež med evolucijsko sorodnimi organizmi niso ohranjeni kot cele enote, kot tudi niso motivi paralognih genov znotraj enega organizma (Chen in Wu, 2007; Wang, 2008a). Vendar pa imajo nekateri daljno sorodni organizmi ohranjene ortologne motive mrež. Barrett in Palsson (2006) menita, da je to posledica podobnih življenjskih slogov organizmov. V razli nih okoljih naj bi se s prilagajanjem regulatornih interakcij oblikovali razli ni vzorci genskega izražanja ortolognih genov sorodnih organizmov (Bjornstad in Harvill, 2005; Dekel in Alon, 2005; Fong in sod., 2005; Babu in Aravind, 2006). Tekom evolucije so transkripcijske mreže izredno plasti ne, saj lahko vklju ujejo nove gene in nove regulacije. Evolucija na robovih (angl. edges) mreže poteka veliko hitreje kot pa na kodirajo ih regijah genov (Alon, 2007). Zaradi pomembne funkcije povezovalnikov, bi sklepali, da so povezovalniki evolucijsko ohranjeni, kar meni tudi veliko avtorjev (Jordan in sod., 2003; Saeed in Deane, 2006). Vendar so mnoge analize pokazale, da ni razlike v ohranjenosti povezovalnikov in drugih transkripcijskih faktorjev v genomu. Chen in Wu (2007) razlagata, da so se mehanizmi regulacije ortolognih genov z regulatornimi povezovalniki razvili neodvisno pri razli nih organizmih. Razlog naj bi bil v tem, da na afiniteto vezave in specifi nost transkripcijskega faktorja ter njegovega tar nega mesta lahko vpliva že relativno majhna sprememba v vezavnem mestu. Tako naj bi DNA-vezavna domena hitro razvila tar na mesta za nove transkripcijske interakcije (Babu in sod., 2006). Ker je prispevek ortolognega transkripcijskega faktorja k fitnesu razli nih organizmov v razli nih okoljih raznolik, so se tekom evolucije v organizmih razvili razli ni regulatorni povezovalniki.

Odnosi med genskimi mutacijami in boleznimi so zelo zapleteni. Dana bolezen je lahko povezana z mutacijami v dolo enih genih. Mutacije v drugih genih pa so lahko povezane z ve boleznimi. Prvi korak k razumevanju kompleksnih poligenskih obolenj na osnovi mrež predstavlja t.i. »diseasome«. Goh in sod. (2007) so rekonstruirali mrežo bolezni loveka in mrežo genov bolezni na osnovi funkcijске anotacije, ekspresije in podatkov o interakcijah proteinov (slika 2).



Slika 2: Spletna stran, ki omogoča hiter vpogled v mrežo genskih povezav med obolenji in geni za bolezni, t.i. »diseasome« (<http://diseasome.eu/>).

Poleg tega so avtorji prišli do zanimivega zaključka, ko so preučili ali geni za bolezni predstavljajo povezovalnike. Gene za bolezni so glede na letalen oz. ne-letalen učinek iznajmljiva (angl. knock-out) ortologa na fenotip miške razdelili v razreda esencialnih in ne-esencialnih genov. Esencialni geni so bili v mreži povezani s povezovalniki. Geni za bolezni, ki so ne-esencialni (ki predstavljajo 78% vseh genov za bolezni), niso imeli več povezav kot geni, ki niso povezani z boleznijo. Do podobnih sklepov so prišli Lu in sod. (2008), ki so preučili spremembe v genskem izražanju v mišjem modelu astme. Geni, katerih ekspresija je bila v primeru bolezni najbolj spremenjena, so imeli malo povezav, medtem ko so geni, ki kodirajo povezovalnike, ohranili stabilno raven izražanja. Ta odkritja so v nasprotju z odkritji drugih avtorjev. Cui in sod. (2007) so poročali, da so mutacije, ki prispevajo k razvoju raka bolj verjetne v vozliščih, ki predstavljajo signalne proteine, kot pa v vozliščih s pasivnimi interakcijami. Vozlišča, ki imajo več povezav z mutiranimi geni, naj bi bila bolj verjetno povezana z rakom ali naj bi predstavljala biomarkerje tumorigeneze. Awan in sod. (2007) so odkrili, da v raku mutirani geni v večji meri predstavljajo jedrne proteine, Jonsson in Bates (2006) pa da imajo z rakom povezana vozlišča več povezav in sodelujejo v večjem številu bioloških procesov. Eprav že obstajajo ideje, da bi ekspresijo določenih motivov mrež uporabili kot markerje za bolezni (Chuang in sod., 2007), pa je očitno, da bo pred tem potrebno razrešiti dileme glede funkcionalnega in topološkega položaja genov, ki vplivajo na razvoj bolezni, v genskih mrežah.

Pred kratkim so bile kot glavni posttranskripcijski regulatorji odkrite molekule miRNA, vendar pa njihovo delovanje v povezavah med geni še vedno ni dobro poznano. Cui in sod. (2006) so analizirali porazdelitev tar miRNA v celi ni signalni mreži na razli nih ravneh - od lokalnih motivov mrež do mrež na globalni ravni in odkrili, da miRNA preferen no uravnava pozitivne regulatorne motive, gosto povezana vozliš a in transkripcijske faktorje. Na ta na in lahko miRNA prekinejo obstoje e signale in omogo ijo hiter odziv na nove signale. Po drugi strani, miRNA manj pogosto ciljajo negativne regulatorne motive, proteine za osnovne celi ne mehanizme ali ligande.

## 2.2 miRNA

miRNA so kratke, 19-25 nukleotidov dolge molekule RNA, za katere je bilo nedavno dokazano, da igrajo z inhibicijo translacije ali aktivacijo razgradnje tar nih mRNA klju no vlogo v regulaciji genskega izražanja (Bartel in sod., 2004). Nastanejo s procesiranjem daljših primarnih transkriptov (dolgih od nekaj sto do nekaj tiso nukleotidov) do struktur lasni nih zank (dolgih 60-110 nukleotidov) in nadaljnji koraki zorenja pod vplivom encima Drosha v jedru ter encimskega kompleksa Dicer v citoplazmi. miRNA regulirajo gensko izražanje na sekven no-specifi en na in. Po vklju itvi v ribonukleoproteinski kompleks RISC (angl. RNA-induced silencing complex), ki ga sestavljajo proteini, kot so Dicer in lani družine Argonavtov, se lahko miRNA z delno komplementarnostjo preko prvih 2-8 nukleotidov zaporedja (angl. seed sequence) veže na 3'- neprevedene regije (UTR, angl. untranslated regions) mRNA. Posledica vezave je prizadeta translacija in stabilnost mRNA molekul (Valencia-Sanchez in sod., 2006; Filipowicz in sod., 2008), kar vodi v znižano izražanje na ravni proteinov.

Sedaj je znano, da so u inki miRNA na gensko ekspresijo ve plastni. Ker je možna inhibicija translacije tar z nepopolnim ujemanjem zaporedij, znanstveniki predvidevajo, da vsaka miRNA regulira ve je število genov ter je posamezna tar a regulirana z ve miRNA molekulami (Krek in sod., 2005; Lewis in sod., 2005; Rajewsky in sod., 2006). miRNA lahko pod dolo enimi celi nimi pogoji izražanje mRNA tudi aktivirajo. Primer sta miR-369-3 in let-7, ki na izhodu iz celi nega cikla s spremembou mikro-nukleoproteinov aktivirata translacijo, ki jih sicer med proliferacijo celic utišata (Vasudevan in sod., 2007).

Zanimivo novo spoznanje je tudi, da se miRNA ne vežejo samo na 3'-UTR, temve tudi na 5'-UTR, kot je bilo dokazano za miR-10a (Ørom in sod., 2008). Nadalje presene ajo rezultati raziskav na embrionalnih mati nih celicah, da miRNA vpliva na osnovne diferenciacijske procese tudi preko vezave na mesta v mRNA kodirajo ih zaporedjih (Tay in sod., 2008). Poleg tega pa so funkcije miRNA različne v različnih celičnih razdelkih. Hwang in sod. (2007) so dokazali vejo prisotnost miR-29b, ki nosi heksanukleotidne motive, v jedru, Place in sod. (2008) pa so na primeru miR-373 pokazali, da ta stimulira transkripcijo z vezavo na promotor.

Seveda ostaja mnogo vprašanj še neodgovorjenih. Kako je regulirano izražanje miRNA? Zakaj pride do deregulacije miRNA? Kje v genomu se nahajajo še neodkrite miRNA? Kateri geni so tarče za miRNA? Ali miRNA deluje kot binarno stikalo ekspresije ali služi kot pufer za bolj fino in bolj skladno izražanje genov? Odgovore na nekatera od teh vprašanj odstirajo raziskave raka.

### **2.2.1 miRNA in rak**

Prva povezava med miRNA in rakom je bila odkrita leta 2002, ko so Calin in sodelavci opisali izgubo miR-15 in miR-16 ob deleciji 13q14 pri kroni ni limfocitni levkemiji (Calin in sod., 2002, 2004a). To odkritje je sprožilo nov zagon raziskav na področju miRNA in levkemij.

Nenavadno pri tem je, da je pomen miRNA v normalni hematopoezi postal cenjen šele dve leti kasneje (Chen in sod., 2004). Nadaljnje raziskave so odkrile, da so geni miRNA pogosto locirani na fragilnih mestih genoma in sodelujejo pri razvoju raka (Calin in sod., 2004b). Nepravilna ravneni miRNA ima pri raku pomembno posledico: miRNA, ki so prekomerno izražene, lahko prispevajo k raku z zniževanjem ravni izražanja tumorskih supresorjev, kot je npr. *BCL-2* (Cimmino in sod., 2005); tiste miRNA, ki se v postopku maligne transformacije izgubijo, pa omogočajo prekomerno ekspresijo onkogenov. Vloga miRNA pri onkogenezi je tkivno-specifična in je odvisna tudi od vrste raka. miR-155 na primer deluje v levkemijah, limfomih in solidnih tumorjih epitelnega izvora kot onkogen,

pri rakih endokrinega sistema pa ima močno znižano izražanje in možno supresivno funkcijo (Calin in sod., 2006).

Deregulacijo miRNA lahko povzročijo delecije ali amplifikacije genomskega regija, kjer so locirane z rakiom povezane miRNA. Ker ležijo navzdol od glavnih onkogenov in tumorskih supresorjev, lahko le-te vplivajo na njihovo izražanje kot transkripcijski faktorji. Poleg tega so Lujambio in sod. (2008) identificirali za humane metastaze značilen hipermetilacijski profil miRNA. Prav tako so možne vzroke napake v procesiranju prekurzorskih miRNA v zrelo obliko. To so le nekateri od prvih predlaganih mehanizmov deregulacije izražanja miRNA.

Vzroki in posledice neuravnovešenega izražanja miRNA v onkogenezi še vedno niso dobro raziskani. V prihodnjih letih lahko pri akujemo identifikacijo novih interakcij miRNA tako s tamnimi geni, epigenetskimi faktorji, kot tudi drugimi nekodirajočimi RNA, kot so npr. UCR (angl. ultraconserved regions, transkripti s popolno genomsko ohranjenostjo med lovekom, podgano in mišjo), za katere so bile interakcije z miRNA in vpliv na maligni fenotip že dokazane (Calin in sod., 2007), ali pa s popolnoma novimi, še nepoznanimi nekodirajočimi transkripti.

### 2.3 LEVKEMIJE

Tumorje na splošno razvrščamo glede na tkivo iz katerega izvirajo. Večina tipov raka pri loveku izvira iz epitelnih tkiv, ki predstavljajo več kot 80% z rakiom-povezanih smrti v Zahodnem svetu. Poleg teh karcinomov poznamo še tri tipe tumorjev, ki nimajo epitelnega izvora. Prvo skupino predstavljajo tumorji različnih vezivnih tkiv mezodermalnega izvora, ki jih imenujemo sarkomi. Ti predstavljajo le 1% vseh tipov raka pri loveku. Druga skupina sestoji iz nevroektodermalnih tumorjev, ki izvirajo iz različnih celic centralnega in perifernega živnega sistema ter vključujejo gliome, glioblastome, nevroblastome, schwannome in medulloblastome. V tretji skupini pa so tumorji, ki izvirajo iz hematopoetskega tkiva, vključno s celicami imunskega sistema (Weinberg, 2007).

Levkemije in limfomi sta glavna podtipa hematoloških neoplazij. Dobesedni prevod levkemije pomeni »bela kri«, ki se nanaša na nepigmentirane rakaste celice potomke hematopoetskih celi nih linij. Levkemije tvorijo razpršene populacije tumorskih celic v krvi, kjer te lahko prosto cirkulirajo. Limfomi so tumorji, ki izvirajo iz limfoidne celi ne linije, npr. iz limfocitov B ali limfocitov T, ki se sprimejo v solidne mase in jih najbolj obi ajno najdemo v bezgavkah (Weinberg, 2007). Najbolj pogosti raki hematopoetskih celic so akutna limfoblastna levkemija (ALL), akutna mieloidna levkemija (AML), kroni na mieloidna levkemija (KML), kroni na limfocitna levkemija (KLL), multipli mielom (MM), Ne-Hodgkinov limfom (NHL) in Hodgkinov limfom (Weinberg, 2007).

Levkemije lahko glede na potek bolezni v grobem razdelimo na akutne in kroni ne (preglednica 1). Za akutne levkemije je zna ilno hitro povisanje števila nezrelih krvnih celic, ki se nakopi ijo v kostnem mozgu, v katerem posledi no ni možna produkcija zdravih krvnih celic. Akutne levkemije prizadenejo ve inoma otroke in mlajše odrasle. So najbolj pogost vzrok z rakom povezanih smrti pri otrocih v ZDA (Jemal in sod., 2005). Akutna oblika levkemije zahteva takojšnje zdravljenje, saj bolezen hitro napreduje, celice se akumulirajo v krvi in širijo v druge organe. Kroni ne levkemije napredujejo po asi, abnormalne celice pa lahko kljub njihovi nefunkcionalnosti dozorijo. Še ve, celice kroni nih levkemij so dolgožive in se po asi, a vztrajno akumulirajo v krvi in v asih v sekundarnih limfoidnih organih. Vendar pa lahko traja mesece ali leta, da bolezen napreduje. Kroni ne levkemije prizadenejo ve inoma starejše osebe in jih pogosto diagnosticiramo šele z rutinskim krvnim testom, saj se simptomi ne pokažejo, dokler ni bolezen doseglja že napredovane stopnje.

Preglednica 1: Štirje glavni razredi levkemij

	Akutna levkemija	Kroni na levkemija
Limfocitna levkemija	Akutna limfoblastna levkemija	Kroni na limfocitna levkemija
Mieloidna levkemija	Akutna mielogena levkemija	Kroni na mielogena levkemija

Levkemije lahko nadalje glede na tip prizadetih celic razdelimo na limfoblastne (limfocitne) in mielogene (mieloidne). Pri limfocitnih levkemijah pride do transformacije limfocitov ali limfocitnih predniških (angl. progenitor) celic. V mielogenih levkemijah pa

se transformirajo predniške celice, ki bi se normalno diferencirale v eritrocite, levkocite ali trombocite.

Tako razvršamo akutne levkemije glede na tip prizadetih krvnih celic na AML in ALL. Akutna mielogena levkemija (AML) prizadene mieloidno celi no linijo, ki vklju uje granulocite, monocite in trombocite, medtem ko so v primeru akutne limfoblastne levkemije (ALL) prizadete celice limfoidne linije, kar vklju uje limfocite T in limfocite B (Catovsky, 1982). Analogno, sta tipa kroni nih levkemij kroni na limfocitna levkemija (KLL) in kroni na mielogena levkemija (KML).

#### 2.4 KRONI NA LIMFOCITNA LEVKEMIJA

Kroni na limfocitna levkemija (KLL) je najbolj pogosta oblika levkemije v zahodnem svetu. Baze SEER (angl. Analysis of the Surveillance, Epidemiology, and End Results) (National Cancer Institute, SEER cancer statistics review 1975–2006, 2008) beležijo, da je incidenca v ZDA 4,43 na 100 000 na leto (moški 6,19; ženske 3,15). Ve inoma prizadene starejše osebe. Manifestira se kot nenadzorovana akumulacija zrelih, a funkcionalno nekompetentnih celic B v kostnem mozgu in krvi. Kljub histološki homogenosti je klini ni izid zelo raznolik. Preživetje bolnikov je v širokem razponu od nekaj mesecev do ve desetletij, bolezen pa je še vedno neozdravljiva. Zato predstavlja dolo itev molekularnih in celi nih markerjev, ki lahko napovedujejo napredovanje bolezni pri bolnikih s KLL, velik napredok.

Klini na diagnoza KLL je definirana s pove anim številom vsaj  $5 \times 10^9/L$  zrelih limfocitov primerenega fenotipa (Binet in sod., 2006). Te zna ilnosti lo ijo kroni no limfocitno levkemijo od plaš noceli nega limfoma in spleni nega limfoma marginalne cone, ki najbolje posnemata KLL (Matutes in sod., 1994). Pri nekaterih osebah je tumor omejen na bezgavke ali druga tkiva, ki ne vklju ujejo krvi ali kostnega mozga. V teh primerih se obolenje imenuje limfom SLL (angl. small lymphocytic lymphoma), histološko in imunofenotipsko pa je identi nen KLL (Muller-Hermelink in sod., 2001).

Za celice KLL je zna ilno, da izražajo veino membranskih antigenov, prisotnih na zrelih celicah B, vključno s CD19, CD20, CD21 in CD23. Poleg teh izražajo tudi CD5, ki je sicer bolj zna ilen za celice T. Za celice KLL so zna ilne še nekatere posebnosti. Izražajo zelo nizko raven površinskih imunoglobulinov (Ternynck in sod., 1974), ki pogosto reagirajo na lastne antigene (Caligaris-Cappio, 1996). Dalje, imajo nenavadno kinetiko nizke odzivnosti (Andreeff in sod., 1980; Nilsson, 1992). Mitogeni signali, ki obiškajo inducirajo proliferacijo normalnih celic B, so zelo šibek stimulus za celice KLL. Prav tako pomemben mehanizem akumulacije celic KLL je okvarjen mehanizem apoptoze, ki vodi v podaljšano preživetje celic (Caligaris-Cappio, 1996; Nilsson, 1992). Celice KLL niso sposobne predstavljati topnih in alo-antigenov (Ranheim in Kipps, 1993; Dazzi in sod., 1995), medtem ko normalne celice B po aktivaciji delujejo kot uinkovite antigen-predstavljivne celice (Lanzavecchia, 1985).

Velik pomen pri KLL ima B-celi ni receptor (BCR). BCR je multimerni kompleks površinskega imunoglobulinskega homodimera in nekovalentno vezanega heterodimera Ig<sub>α</sub>/Ig<sub>β</sub> (CD79A/CD79B). Znižana raven izražanja B-celi nih receptorjev na membrani limfocitov je znamenje KLL (Vuillier in sod., 2005). Mehanizem za nizko ekspresijo BCR pri KLL ni poznani. Z izjemo ene dokazane mutacije v *CD79b* (Thompson in sod., 1997), raziskave niso pokazale nobenih genetskih okvar v *BCR* (Payelle-Brogard in sod., 1999; Alfarano in sod., 1999). V nasprotju s slabo izraženostjo na membrani, pa sta transkripcija in intracelularna sinteza komponent BCR normalni (Alfarano in sod., 1999; Payelle-Brogard in sod., 2002). Združevanje komponent in transport preko endoplazmatskega retikuluma ni možen zaradi okvar procesov zvijanja in glikozilacije verig  $\mu$  ter CD79A. Zvijanje naj bi bilo pri KLL problematično tudi pri molekuli CD22 (Payelle-Brogard in sod., 2006).

Normalno je raznolikost imunoglobulinov in s tem antigenska specifičnost vezave B-celi nega receptorja posledica združevanja V (angl. Variable), D (angl. Diversity) in J (angl. Joining) genskih segmentov težke verige imunoglobulinov ter V in J segmentov lahkikh verig imunoglobulinov. Nadaljnjo diverzitetom omogoči rekombinacija ob preklopu verig in delecije ter insercije nukleotidov, kot tudi uporaba alternativnih italnih okvirjev. Za specifičnost vezave antiga ima ključno vlogo regija HCDR3, tretja regija CDR (angl.

complementarity determining region) na težki verigi (Xu in Davis, 2000), ki nastane z rekombinacijo segmentov VDJ *de novo* (Maizels, 2005). Somatske hipermutacije po stimulaciji zrelih celic B z antigeni v germinalnih centrih, pa tudi zunaj klasi nih germinalnih centrov (William in sod., 2002), omogo ijo dodatno raznolikost protiteles. Vsi ti procesi vklju ujejo lome DNA in naknadne popravljalne mehanizme. Napake, ki se ob teh procesih zgodijo, pa lahko povzro ajo B-celi ne rakotvorbe.

Visok delež bolnikov s KLL ima podoben preferen ni repertoar dolo enih genov težkih in lahkikh verig imunoglobulinov, kot tudi vzorce somatskih hipermutacij, ki vodijo do uporabe podobnih aminokislinskih motivov v HCDR3. Ta dejstva so vodila do hipoteze, da je pri patogenezi vpletena klonalna selekcija s specifi nim skupnim antigenom (Chiorazzi in Ferrarini, 2003; Messmer in sod., 2004; Murray in sod., 2008). To hipotezo podpira tudi dejstvo, da pri patologiji KLL igrajo pomembno vlogo stimulacija preko BCR, avtoantigeni ter mehanizmi eliminacije apoptoti nih celic in patogenih bakterij (Caligaris-Cappio in Ghia, 2008).

Izvor prizadetih celic B v kroni ni levkemiji je bil sprva nekoliko kontroverzen. Raziskovalci so domnevali, da KLL predstavlja dve izvorno razli ni bolezni. Približno polovica celic KLL izraža CD5 in IgM/IgD ter ima fenotip podoben fenotipu naivnih celic, ki imajo v normalnih pogojih nemutirane gene imunoglobulinov (Pascual in sod., 1994). Med 50 - 70% pacientov s KLL pa ima somatske mutacije genov IgVH (variabilna regija težke verige imunoglobulina) (Schroeder in sod., 1994) in zgleda, da so te celice prestale proces zorenja afinitete in naj bi izvirale iz izkušenih celic B. Prisotnost somatskih mutacij je povezana z dolo enimi IgVH geni. Na primer, specifi ni aleli IgVH1-69 (Kipps in Carson, 1993) in IgVH4-39 (Chiorazzi in Ferrarini, 2003) imajo nemutiran profil. Prisotnost gena IgVH3-21 pa izjemoma pomeni slabši izid bolezni neodvisno od mutacijskih sprememb (Thorsélius in sod., 2006).

Profil mutacij genov IgVH je dobro prognosti no sredstvo – bolniki z nemutiranim profilom imajo slabšo prognozo. Celice nemutirane oblike KLL izražajo polireaktivna protitelesa, ki se vežejo nespecifi no na lastne in tuje antigene, ve ina mutiranih pa ne (Pritsch in sod., 1993; Herve in sod., 2005). Od statusa IgVH je odvisen odziv pacientov na

stimulacijo BCR, in sicer imajo bolniki z nemutiranimi geni IgVH boljši odziv (Lanham in sod., 2003). Obstaja mo na povezava nemutiranega statusa IgVH s povišano ravnjo tirozin kinaze ZAP70 (Rosenwald in sod., 2001). Protein ZAP 70 se običajno nahaja v celicah T ter celicah naravnih ubijalkah, v normalnih celicah B pa ga ni, le na stopnjah pred-B in pro-B celic. Njegova ekspresija v celicah KLL omogoča uinkovito signaliziranje preko IgM, kar je možen razlog za agresiven potek bolezni. Z nemutiranimi geni IgVH in slabšo prognozo povezujemo povišano raven še nekaterih drugih molekul, npr. CD38, ki je pomembna za medcelijsne interakcije in povišano raven citidin deaminaze (AICDA), ki ima vlogo pri somatskih mutacijah in preklopu težkih verig. Danes je znano, da vse celice KLL izražajo fenotip antigensko izkušenih in aktiviranih celic B ter imajo, ne glede na mutacijski status genov IgVH, ekspresijski profil podoben spominskim celicam B (Hamblin in sod., 1999; Marti in sod., 2005). Vse celice KLL imajo skupen celijski izvor oz. skupen mehanizem transformacije. Rawstron in sod. (2002) so v krvi 3,5% zdravih oseb, starejših od 40 let, opazili populacijo monoklonskih limfocitov s fenotipskimi lastnostmi celic KLL. Pri ožjih sorodnikih bolnikov s KLL pa je bila pogostost takšnih celic med 13,5% in 18% (Rawstron in sod., 2002; Marti in sod., 2003). Ali predstavljajo majhne monoklonske populacije celic B kritičen korak v razvoju KLL trenutno še ni znano in je to predmet intenzivnih raziskav.

Navkljub dostopnosti tumorskih celic iz krvi, je bilo o patofiziologiji bolezni do sedaj malo znanega. Rezultati *in vivo* raziskav potrjujejo hipotezo, da imajo celice KLL nizko proliferativno sposobnost in da pride do akumulacije celic B zaradi okvare v programirani celijski smrti. V nasprotju s tem pa v *in vitro* kulturah pride do apoptoze, kar namiguje na velik pomen mikrokolja za preživetje celic KLL (Caligaris-Cappio, 2003). *In vivo* je mikrokolje za proliferacijo možno v psevdo-foliklih v bezgavkah ter celijskih agregatih v kostnem mozgu (Granziero, 2001), kjer se poleg deležnih se celic B nahaja tudi veliko CD4+ celic T.

Na razvoj KLL naj bi imele vpliv tudi okužbe, saj so posamezniki z zgodovino pljučnic bolj dovetni za razvoj KLL (Hamblin, 2006; Landgren in sod., 2007). Trdnih dokazov za povezavo KLL z izpostavljenostjo sevanju ali kemičnim snovem, razen povezave s herbicidi pri delavcih v kmetijstvu, zaenkrat ni. Dokazana je povezava KLL z v

Vietnamaski vojni uporabljenim herbicidom Agent Orange (Committee to ..., 2003), novejše raziskave pa kažejo, da izklju evanje ionizirajo ega sevanja kot vzroka za KLL ni utemeljeno (Hamblin, 2008).

#### **2.4.1 Genetski in epigenetski vzroki kroni ne limfocitne levkemije**

Genetsko je KLL posebna bolezen, saj navadno vklju uje delecije, medtem ko so za ostale hematološke rakotvorbe bolj zna ilne translokacije. Najbolj pogoste so delecije na regijah 13q14 in 11q22-q23, ki naj bi imeli skupen ancestralni izvor (Auer in sod., 2007). Z dovolj ob utljivimi metodami lahko zaznamo kromosomske nepravilnosti pri veini bolnikov s KLL (Döhner in sod., 2000). Preglednica 2 prikazuje najbolj pogoste kromosomske napake in njihove pogostnosti. Običajno gre za posamezne nepravilnosti, z napredovanjem bolezni pa je možnih ve kromosomskih sprememb (Döhner in sod., 2000). Bolniki z delecijo 13q kot edino kromosomsко napako imajo najbolj ugodno prognozo. Najslabša prognoza pa je povezana z izgubo gena *TP53* pri pacientih z delecijo 17p. Delecija 11q pomeni izgubo funkcije tumorskega supresorskega gena *ATM* pri podskupini bolnikov s KLL. Izguba funkcij genov *ATM* na 11q in *TP53* na 17p lahko razloži slabši izid bolezni pri teh kariotipskih podskupinah, vendar ta gena verjetno nista odgovorna za etni razvoj bolezni. Poleg tega lahko deleciji 17p in 11q zaznamo kot sekundarna inka klonalne evolucije bolezni. Nasprotno pa drži za regijo 13q14, ki se izgubi pri ve kot 54% bolnikov s KLL. Izguba prve kopije 13q14 vedno predstavlja primarni dogodek (Stilgenbauer in sod., 2007).

Glede na Knudsonov teoreti ni model, delecije namigujejo na prisotnost tumorskih supresorskih genov. Alfred Knudson je namre leta 1971 na primeru retinoblastoma prvi predlagal obstoj tumorskih supresorskih genov ter oblikoval koncept dominantnega dedovanja za nagnjenost k raku z recessivnim mehanizmom razvoja raka na celi ni ravni, saj je za nastanek fenotipa potrebno izni enje obeh kopij alelov. Kljub intenzivnim raziskavam pa je bilo pri KLL identificiranih zelo malo vzro nih mutacij. Zato se je fokus od predlagane hipoteze, ki se je izkazala za pretirano poenostavitev, preusmeril na druge alternativne mehanizme inaktivacije genov, npr. haplo-nezadostnost (angl. haploinsufficiency) in epigenetsko utišanje.

Preglednica 2: Pogostnost kromosomskih nepravilnosti pri 325 bolnikih s kroni no limfocitno levkemijo (KLL). \* 175 bolnikov je imelo eno kromosomske napako, 67 dve in 24 bolnikov je imelo več kot dve kromosomske napaki (povzeto po Döhner in sod., 2000; 1911).

Kromosomska napaka	Število pacientov (%) *
Delecija 13q	178 (55)
Delecija 11q	58 (18)
Trisomija 12q	53 (16)
Delecija 17p	23 (7)
Delecija 6q	21 (6)
Trisomija 8q	16 (5)
t(14q32)	12 (4)
Trisomija 3q	9 (3)
Klonalne nepravilnosti	268 (82)
Normalen kariotip	57 (18)

#### 2.4.1.1 Haplo-nezadostnost (angl. haploinsufficiency)

Haplo-nezadostnost opisuje proces, kjer je posamezen funkcionalen alel nezadosten za normalno aktivnost genskega produkta. Delecija enega alela brez mutacije v drugem alelu povzroči spremembo genske doze. Mnogo genov je pleiotropnih, imajo pa inke v vseh celihih procesih, njihova funkcija pa je lahko odvisna od ravni ekspresije (oz. doze). Na primer, *TP53* pri nizki ravni ekspresije zaustavi celinski cikel, pri visokih pa inducira apoptozo. Čeprav nimajo vsi geni dvojnega odziva glede na gensko dozo, pa je možno da imajo tisti, ki so ob utljivi na ravni genske doze, ob nizki ekspresiji potencial za karcinogenezo (Cotter in Auer, 2007). Vendar pa na odvisnost tumorskih supresorjev od genske doze ne vpliva samo genetski profil organizma, temveč pa komopleksnosti prispevajo tudi razlike med tkivi. Ta koncept postaja vedno bolj privlačen tudi za razlaganje mehanizma delecij KLL (Mertens in sod., 2002; Haslinger in sod., 2004; Kienle in sod., 2005; Dickinson in sod., 2006), še posebej zaradi neuspehov detekcije mutacij alela na normalnem kromosomu bolnikov s posamezno delecijo (Mabuchi in sod., 2001; Migliazza in sod., 2001; Wolf in sod., 2001).

#### 2.4.1.2 Epigenetska supresija

Epigenetsko utišanje pomeni reverzibilno spremembo genskega izražanja brez spremembe v nukleotidnem zaporedju. To vključuje metilacijo DNA in število histonskih modifikacij, od katerih je najbolj znana acetilacija. Hipermetilacija DNA in hipoacetilacija histonov sta

znaka utišanja genov. Nasprotno, hipometilacija DNA in acetilirani histoni spodbujajo aktivnost transkripcije. Pri raku je najbolj pogosta epigenetska spremembra metilacija. Medtem ko je globalno metilacija DNA znižana, pa lokalna hipermetilacija genskih promotorjev vodi v utišanje genov. Dobro je preu ena metilacija tumorskih supresorskih genov, hipermetilacija promotorskih regij pa je bila identificirana v ve kot polovici genov, ki povzro ajo družinske oblike raka (Baylin in sod., 2001). Takšne spremembe so prisotne tudi pri nefamiliarnih rakih in so pogosto opisane pri hematopoetskih neoplazmah (Baylin in sod., 1998). Vendar pa so epigenetski mehanizmi, ki vplivajo na gensko ekspresijo kompleksni in medsebojno prepleteni. Na primer, metilacija promotorja vpliva na stopnjo acetilacije in metilacije histonov (Bachman in sod., 2003).

Posebej zanimiva je delecija regije 13q14, v kateri se nahajata zapisa za miR-15a in miR-16 in jo ima ve kot polovica bolnikov s KLL (Damle in sod., 1999; Döhner in sod., 2000).

eprav je raven izražanja ve ine genov znotraj te regije znižana (Mertens in sod., 2002; Dickinson in sod., 2006), pa vzro nih mutacij do sedaj niso odkrili. Metilacija je bila prav tako izklju ena kot potencialni mehanizem transkripcijskega utišanja (Mertens in sod., 2002). Ker pa je raven izražanja v primerih delecije ene kopije znižana za faktor ve ji kot dva, kolikor bi pri akovali z obi ajnim u inkom doze, je vseeno možen epigenetski razlog.

Tega vprašanja so se lotili Mertens in sod. (2006) ter preu ili ali so tumorski supresorski geni v tej regiji pri vzorcih zdravih oseb izraženi monoalelno in prepoznali veliko podobnosti pri celicah B in T med procesi na genih imunoglobulinov ter dogodki na regiji 13q14. Demonstrirali so, da se pri zdravih osebah kopiji regije 13q14 replicirata asinhrono, kar nakazuje razli no pakiranje kromatina. Asinhrona replikacija kopij pomeni, da se en alel pomnoži bolj zgodaj v fazi S celi nega cikla kot drugi in predstavlja eno od najbolj zgodnjih epigenetskih diskriminacij. Posledi no so pri specifi nih genih zaznali monoalelno ekspresijo, pri pacientih s KLL z monoalelno delecijo 13q pa je epigenetsko utišanje ene kopije zadoš alo za izgubo funkcije gena.

Za enkrat so pri loveku poznani trije mehanizmi monoalelnega izražanja: i) preferen na ekspresija specifi nih alelov polimorfizma posameznih nukleotidov (SNP), ii) imprinting, in iii) alelna izklju itev (angl. allelic exclusion). Mertens in sod. (2006) so našli le en SNP

s preferen no ekspresijo, mehanizem imprintinga pa je bil izklju en, ker je bila monoalelno lahko utišana tako maternalna kot paternalna kopija. Mehanizem monoalelne ekspresije je izredno podoben alelni izklju itvi genov imunoglobulinskih receptorjev, kjer je utišanje ene kopije nujno, da pote e proces rekombinacije in sestavljanja BCR le enkrat in ne pride do avtoimunosti. Rekombinacija vklju uje kompleksne procese dvooverižnih lomov za preureditev genskih segmentov imunoglobulinskih receptorjev. Glede na to, da je za bolnike s KLL zna ilna nenavadno pogosta uporaba specifi nih genov za imunoglobuline, hkrati pa so pogoste tudi delecije 13q14, bi bilo zanimivo vedeti, ali imata obe zna ilnosti KLL soroden vzrok v okvari mehanizmov, ki so sicer odgovorni za normalne rearanžmaje genov *BCR*. Dalje, ker so odkrili monoalelno izražanje regije 13q14 tako pri celicah B kot celicah T, pride o itno do utišanja zelo zgodaj v razvoju ter se to lahko stabilno ohranja preko številnih celi nih delitev. Obstaja pa tudi razlika med utišanjem 13q14 ter alelni izklju itvio. Pri utišanju regije 13q14.3 je v celotnem tkivu aktivna enaka kopija, pri alelni izklju itvi pa pride do utišanja v vsaki celici neodvisno. S tem sta obe kopiji približno enako izraženi v celotnem tkivu. Ti rezultati ponujajo potencialno razlagu znižanega izražanja kandidatnih genov in miRNA v regiji 13q14 pri KLL.

Rassenti in Kipps (1997) sta pokazala, da pri bolnikih s KLL alelna izklju itev ni popolna in se izraža ve kot ena funkcionalna težka veriga imunoglobulinov. Ker se zdi, da lahko vleemo paralele med utišanjem regije 13q14 in utišanjem lokusov imunoglobulinov, je možno domnevati o transkripcijskih in posttranskripcijskih mehanizmih pri KLL, ki so bili že dokazani kot pomembni igralci pri združevanju segmentov VDJ (Bolland in sod., 2004). Tako na genskih kot intergenskih delih lokusa IgVH so raziskave odkrile znatno protismiselno (angl. antisense) transkripcijo. Na splošno prihaja v zadnjem asu do odkritij pri razli nih organizmih, ki povezujejo utišanje s kromatinom s prisotnostjo ali odsotnostjo nekodirajo ih RNA. To vklju uje tako dolge nekodirajo e RNA, ki imajo dokazano vlogo pri kompenzaciji genske doze in imprintingu, kot tudi majhne nekodirajo e RNA, ki sodelujejo pri združevanju heterokromatina po poti RNA interference (Bühler, 2009). To zastavlja nova vprašanja o vlogi nekodirajo ih RNA pri kontroli genskega izražanja na ravni kromatina. Prav tako za enkrat še ni poznana vloga nekodirajo ih RNA pri monoalelni ekspresiji. Danes je znano, da je pri loveku in miški med 1 in 5% avtosomalnih genov izraženih monoalelno (Gimelbrant in sod., 2007) ter da je transkripcija

evkariontskih genomov veliko bolj obsežna kot so sprva mislili. Dejstvo, da so Mertens in sod. (2006) odkrili veliko spojitenih (angl. splicing) variant genov nekodirajo ih RNA v regiji 13q14, hkrati pa je znižanje ekspresije miR-15 in miR-16, ki sta locirani v tej regiji, močno povezano s KLL, ponuja možnost, da imajo miRNA in drugi razredi nekodirajo ih RNA pomembno funkcijo pri razvoju in napredovanju kroni ne limfocitne levkemije.

Pri pregledu literature za ostale kromosomske napake pri KLL nismo zasledili podobnih objav o monoalelnem utišanju, vendar pa so druge kromosomske spremembe nasprotno slabše raziskane. Morda je vreden omembe gen *ATM*, ki se nahaja znotraj minimalne regije delecije 11q (Stilgenbauer in sod., 1996), mutacije pa so prisotne pri 20% pacientov s KLL (Schaffner in sod., 1999; Stankovic in sod., 1999). Poleg poznane funkcije pri popravljalnih mehanizmih DNA in s tem pri razvoju raka, ima pomembno vlogo pri alelni izključitvi ter preprevanju bialelnih rekombinacij (Hewitt in sod., 2009). Zanimivo je tudi, da je delecija 11q močno povezana z nemutiranim statusom genov IgVH.

## 2.5 ŠTUDIJSKI PRISTOPI DOSEDANJIH GENETSKIH ANALIZ KRONI NE LIMFOCITNE LEVKEMIJE

Eprav genetski vzroki KLL še vedno niso pozani, pa je uporaba različnih tipov eksperimentalnih postopkov in analiz pokazala širok razpon sprememb na ravni DNA.

### 2.5.1 Živalski modeli

Večina celičnih linij, ki so bile pripisane KLL, izvira iz pacientov s plastičnimi limfomoma, ki posnema KLL, ali pa so to limfoblastoidne linije, kontaminirane z normalnimi limfociti. Sveže vpogledi v KLL so omogočili živalski modeli za KLL, ki zaenkrat dokazujejo deregulacijo treh bioloških poti: poti *TCL1-AKT*, poti *TNF-NF-kB* in z *BCL2* posredovane antiapoptozne poti. Bichi in sod. (2002) so objavili opis transgenih miških z onkogenom *TCL1* pod kontrolo za celice B specifičnega promotorja ter ojačevalca. *TCL1* je aktivator signalne poti serin/treonin kinaze AKT, ki je vpletena v celično proliferacijo, preživetje in smrt. Povišana raven izražanja *TCL1* je povezana z agresivnim fenotipom KLL. Leta 2004 so Zapata in sod. (2004) naredili dvojno transgeno miško iz modela, ki v limfoidnem tkivu miške prekomerno izraža *BCL2* (Katsumata in sod., 1992)

in transgene miške, ki v celicah B in T prekomerno izraža mutirano izoformo TRAF2 (Lee in sod., 1997). TRAF2 je lan družine adaptorskih proteinov, ki se vežejo na receptorje družine TNF ter aktivirajo NF- $\kappa$ B in JNK. Ve ina dvojnih transgenih mišk razvije B-celi no levkemijo, ki je podobna KLL. Planelles in sod. (2004) so opisali model, ki podobno kaže na vlogo aktivacije TNF. Naredili so transgeno mišjo linijo, ki pod za celice T specifi nim promotorjem izraža *APRIL* (angl. a proliferation-inducing ligand). Skupaj z BAFF (angl. B cell-activating factor of tumor necrosis factor family ) predstavljata novejše lane družine citokinov TNF. Izražata se v skoraj vseh hematopoetskih celicah (Haiat in sod., 2006). Miške s transgenom *APRIL* imajo manj agresiven fenotip kot zgoraj opisana modela. Tako je možno, da *APRIL* sicer predisponira za maligno transformacijo, vendar je za le-to potreben še drug dogodek, npr. prekomerno izražanje *BCL2* ali *TCL1*. Pri miškah NZB (New Zealand Black) z avtoimunimi in limfoproliferativnimi obolenji, ki so model za KLL, so Raveche in sod. (2007) dokazali nepravilnosti v lokusu miR-16, nedavno pa so Brugge in sod. (2009) opisali nov mišji model KLL z insercijo gena za *SV40 T* antigen v zrele celice B.

### 2.5.2 Citogenetske in molekularno-citogenetske študije

Sprva so bile citogenetske analize KLL omejene, saj s konvencionalnimi metodami proganja v levkemi nih celicah, pri katerih je mitoza skoraj odsotna, ni možno inducirati metafaze (Winkler in sod., 2006). Kasneje so s tehnikami komparativne genomske hibridizacije (CGH) in iskanjem mikrosatelitov postale možne bolj natan ne analize kromosomskih napak. Fluorescen na in situ hibridizacija (FISH) omogo a ob utljivo detekcijo specifi nih kromosomskih aberacij tako metafaznih kot tudi interfaznih celi nih jeder. Obširna analiza s FISH na 325 primerih KLL je detektirala genomske aberacije pri ve kot 80% bolnikov (Döhner in sod., 2000; preglednica 2). Genomske napake imajo tudi prognosti ni pomen. Delecija 13q14 je povezana z daljšim preživetjem, deleciji 11q22-q23 in 17p13 predstavljalja slabšo prognozo, primeri brez aberacij ali s trisomijo 12 pa so glede na prognozo in preživetje nekje vmes (Döhner in sod., 2000). V želji, da bi mehanizem KLL razumeli še bolje, so raziskovalci razvili klasi no komparativno genomsko hibridizacijo (cCGH) za analizo genomske sprememb v številu kopij genov. Tehnološko še bolj napredna je tehnologija komparativne genomske hibridizacije na mreži (angl. array

comparative genomic hybridization, aCGH), ki predstavlja ob utljiv in robusten genomski diagnosti ni test (Schwaenen in sod., 2004). V tem postopku se celotno genomsko DNA celic KLL kompetitivno hibridizira z razli no ozna enim kontrolnimi DNA na mikromrežo s tar nimi fragmenti DNA. S tem pristopom so bile poleg že poznanih genomskejih regij identificirane dodatne aberacije pri KLL (Schwaenen in sod., 2004), saj omogo a detekcijo tar nih genov ter submikroskopskejih sprememb, ki so še posebej informativni pri kariotipsko normalnih osebah s KLL. Pomanjkljivost aCGH je ob utljivost detekcije v primeru majhnega števila malignih celic v vzorcu (Tyybäkinoja in sod., 2007). V takšnih primerih je potrebna analiza s FISH. Grubor in sod. (2009) so za preu evanje sprememb števila kopij pri KLL uporabili analizo ROMA (angl. representational oligonucleotide microarray analysis), ki je oblika CGH z zaenkrat najboljšo lo ljivostjo.

### **2.5.3 Študije genske povezanosti**

Študije genske povezanosti (angl. linkage analysis) lokalizirajo gene glede na hkratno dedovanje genskih markerjev in bolezni v družinah. Temeljna ideja je, da si družinski lani s KLL delijo gensko tveganje. S poznavanjem družinskih odnosov in statusov bolezni lahko preiš emo genom za skupne gene, ki si jih delijo bolj pogosto kot po naklju ju. Ker je o itno, da dedovanje predisponira za KLL, naj bi raziskave družin same po sebi omogo ale identifikacijo pomembnih genov za razvoj KLL. Do danes so bile za KLL izvedene štiri ve je genomske študije genske povezanosti. Goldin in sod. (2003) so genotipizirali 359 mikrosatelitskih markerjev pri 94 osebah (38 s KLL) iz 18 družin, ki so imele vsaj dva ali ve družinskih lanov s KLL. Po parametri ni kot tudi neparametri ni analizi genske povezanosti, so pri prizadetih sorodnikih izpadle kromosomske regije 1q, 3q, 6q in 13q. Vendar pa nadaljnje študije niso identificirale zarodnih mutacij v teh regijah (Ng in sod., 2007). Sellick in sod. (2005) so genotipizirali 11.560 polimorfizmov posameznih nukleotidov (SNP) pri 228 osebah iz 115 družin. Povezanost je bila dokazana na kromosomski lokaciji 11p11. Druge zanimive regije so bile 5q22-23, 6p22, 10q25 in 14q32. Nobena od regij, identificiranih s tema študijama, ni bila statisti no zna ilna. Navkljub podobnim metodam tudi prekrivanja identificiranih regij med študijama ni bilo. Najverjetnejši razlog za to je genska raznolikost pri KLL ter dejstvo, da je ve ina družin

imela samo dva lana s KLL. Kasneje so Sellick in sod. (2007) raziskavo razširili na 206 družin in z recessivnim modelom dedovanja dokazali povezanost z regijama 2q21 in 6p22 ter povezanost 18q21 z dominantnim modelom. Fuller in sod. (2008) pri najveji do sedaj poznani družini z 11 lani treh generacij s KLL niso odkrili dokazov za povezanost posameznih genov s KLL. Za identifikacijo lokusov, ki vplivajo na nagnjenost h KLL, so potrebne statistične bolje podprtne študije, ki pa so možne le s sodelovanjem večjih centrov, kot sta npr. Genetic Epidemiology of CLL Consortium in International CLL Consortium.

#### **2.5.4 Asociacijske in mutacijske študije**

SNP-ji (polimorfizmi posameznih nukleotidov) so najbolj pogosti tip polimorfizma v loveškem genomu (Kruglyak in sod., 2001) in mnogi geni so že bilo preučeni kot možni kandidati za tveganje KLL. Sprva so te študije preučevale manj SNP-jev iz manjšega števila genov, ki so bili izbrani za raziskavo zaradi domneve o njihovem biološkem pomenu. Kljub temu pa je bilo statistično znano ilno s KLL povezanih zelo malo genov in skoraj noben konsistentno potrjen z drugimi neodvisnimi študijami. Wiley in sod. (2002) so prvi našli povezavo med tveganjem za KLL in receptorjem P2RX7. Thunberg in sod. (2002) so prav tako odkrili znano ilno povezavo, vendar v obratni smeri. Pet nadaljnih študij ni odkrilo nikakršne povezave med SNP-ji in KLL (Starczynski in sod., 2003; Zhang in sod., 2003; Nückel in sod., 2004; Sellick in sod., 2004; Cabrini in sod., 2005). Podobno so Demeter in sod. (1997) poročali o povezavi KLL s SNP-jem v promotorju faktorja tumorske nekroze alfa (TNFa), kasnejše raziskave pa tega niso potrdile (Wihlborg in sod., 1999; Mainou-Fowler in sod., 2000; Au in sod., 2006; Bogunia-Kubik in sod., 2006).

Alternativno lahko SNP-je brez predhodnih predpostavk o biološki vlogi kandidatnih genov iščemo z genomske raziskavami. Trenutno je zaradi visokih stroškov genotipiziranja in potrebe po velikem številu vzorcev objavljena le ena takšna študija pri KLL (Di Bernardo in sod., 2008), ki je izmed 299.983 SNP-jev identificirala šest novih lokusov za dovzetnost za KLL. Pa vendarle je bilo opravljenih tudi kar nekaj asociacijskih študij v večjem merilu. Rudd in sod. (2006) so določili 1467 nesintonimnih SNP-jev v 865 genih, saj so predpostavili, da je biološki vpliv nesintonimnih SNP-jev na dovzetnost za KLL večji ter odkrili za KLL pomembne variante v genih celi nega cikla in DNA

popravljalnih mehanizmov. Podobno so Sellick in sod. (2008) in Enjuanes in sod. (2008) odkrili nesinonimne SNP-je, povezane z napredovanjem bolezni v genih, ki uravnavajo imunski sistem, odziv na poškodbe DNA in apoptozo.

Za asociacijske študije je lahko problemati en vir DNA. V nekaterih primerih je ta namre izolirana iz vzorcev periferne krvi brez separacije tumorske in normalne DNA. iš enje celic B pa je za velike asociacijske študije nepraktično. Alternativni vir DNA so celice bukalne sluznice. Veliko genov potrebuje validacijo v neodvisnih vzorcih primerne velikosti in definiran fenotip bolezni.

Napredek genotipiziranja je omogočil hkratno določanje velikega števila SNP-jev, s tem pa je lahko veliko število SNP-jev signifikantnih po naključju. Zaradi pogosto premajhnega števila vzorcev so možni tako lažno pozitivni kot lažno negativni rezultati asociacijskih študij. Običajni pristopi k nadzoru števila testiranih SNP vključujejo Bonferronijevo korekcijo, FDR (angl. false discovery rate), FPRP (angl. false positive report probability), permutacije in TS (angl. tail strength).

Dalje, mnogo študij preučuje Ne-Hodgkinov limfom (NHL), ki glede na klasifikacijo WHO (World Health Organisation) vključuje tudi KLL, vendar je število vzorcev KLL v takšnih raziskavah običajno manjše in s tem tudi zanesljivost podatkov. Poleg tega so vzorci KLL in SLL (angl. small lymphocytic lymphoma) zaradi enakega imunofenotipa običajno v takšnih študijah združeni, delež KLL vzorcev pa večinoma ni podan. Tako je eden od pomembnih izzivov asociacijskih študij že sama definicija obolenja, saj v primeru ko sta KLL in SLL genetsko različni, to povzroča šum v rezultatih in manjšo statistično moč testov. Potrebno bo raziskati ustreznike združevanja vzorcev KLL in drugih podtipov NHL. To pa je zaradi relativno redke pojavnosti KLL možno le preko večjih konzorcijev kot je InterLymph.

### **2.5.5 Ekspresijske študije**

Mikromreže so postale popularno orodje za primerjalne analize genskega izražanja v večjem merilu. Najbolj običajne so cDNA mikromreže in oligonukleotidne mikromreže.

Za cDNA mrežo robot to kovno nanese klonirane cDNA na nosilec, obi ajno na stekleno plošico ali najlonsko membrano. Oligonukleotidne mreže pa vsebujejo veliko število oligonukleotidnih sond, ki se sintetizirajo direktno na silicijevem ipu, v primeru daljših oligonukleotidov pa se to kovno nanesajo ali natisnejo s tehnologijo »ink-jet«. Hibridizacijo vzorca lahko, odvisno od platforme, zaznamo z radioaktivnimi ali fluorescenčnimi sondami. Za na najlonski membrani to kovno naneseno cDNA je tarata obi ajno radioaktivno ozna ena s  $^{33}\text{P}$ . Če pa je nanesena na steklenem nosilcu, hibridiziramo po označevanju z različnimi fluoroforoma, tipi no s cianinom 5 (Cy5) in s cianinom 3 (Cy3), hkrati preučevani in referenčni vzorec. Barvili se med vzorci tudi zamenja (angl. dye swap). Na oligonukleotidne mikromreže na silicijevih ipih se hibridizira posamezna tarata, tipi no fluorescenčno označena s streptavidinom-fikoeritrinom. Tehnične razlike med platformami lahko vodijo do razlik v rezultatih. Mreže, ki hibridizirajo postopno posamezne vzorce (cDNA mikromreže z radioaktivno označenimi taratami ali silicijeve oligonukleotidne mreže), merijo gensko izražanje na osnovi intenzitete signala vsakega gena. cDNA mikromreže s fluorescenčno označenimi taratami pa merijo razmerje signala testnega vzorca proti signalu referenčnega vzorca. Zaradi teh tehničnih razlik je primerjava med različnimi platformami otežena. Prav tako je primerljivost rezultatov študij z različnimi platformami problematična zaradi bioloških razlik med vzorci, npr. zaradi druge ne izbire referenčnega vzorca. Do razlik pa pride tudi pri študijah z enakimi platformami. To vključuje razlike v pripravi vzorcev, kvaliteti RNA, označevanje sond, izbiro klona, hibridizaciji in kakovost same mikromreže. Biološka variabilnost pa vključuje genske razlike, genomske nestabilnosti in normalna vsakodnevna nihanja ravni izražanja genov. Zaradi naštetih možnih virov variabilnosti ter zaradi pomanjkanja standardizacije postopkov ostaja, kljub obdelavi podatkov z normalizacijo, primerjava ekspresijskih profilov velik izziv.

Vendarle ponujajo ekspresijske študije veliko prednosti. Ekspresijske profile lahko testiramo v različnih kontekstih. Na primer, ekspresijski profil lahko prevzame vlogo biološkega fenotipa. Biološko stanje v obliki vzorca genske ekspresije predstavlja koncept »podpisa« genskega izražanja (angl. gene expression signature). Omogoča pa nam na primer tudi ocenjevanje primerljivosti genskega izražanja v celičnih linijah z *in vivo* vzorci. V kontekstu raka omogoča ekspresijsko profiliranje določanje novih podtipov in

predikcijo klini nega izida. Obstaja mnogo primerov u inkovite raz lenitve heterogenih fenotipov. Prvi takšen dokaz je bila lo itev akutne mieloidne in akutne limfoblastne levkemije (Golub in sod., 1999), sledile pa so študije, ki so identificirale prej nepoznane razrede difuznega velikoceli nega limfoma B (Alizadeh in sod., 2000). Z mikromrežami lahko identificiramo tudi specifi ne molekularne poti in potencialne diagnosti ne in prognosti ne markerje, terapevtske tar e, ter odkrivamo mehanizme delovanja razli nih zdravil.

Že leta 2001 sta dve študiji genske ekspresije raziskali ali predstavljata obliki KLL z mutiranimi in nemutiranimi geni IgVH razli ni bolezni (Rosenwald in sod., 2001; Klein in sod., 2001). Rezultati obeh študij so kljub razli nim platformam potrdili, da gre za eno bolezen, saj je bila, ne glede na status IgVH, za KLL karakteristi na ekspresija 300 genov (Rosenwald in sod., 2001).

Ekspresijsko profiliranje je za KLL še posebej primerno, saj je na in pridobivanja ve jih koli in istih, razvojno homogenih populacij levkemi nih celic relativno enostaven. Vendar pa je zaradi pogosto majhnega števila preu evanih bolnikov potrebno oceniti reprezentativnost genske ekspresije. Dva glavna pristopa k analizi genske ekspresije sta nadzorovana in nenadzorovana analiza.

Nenadzorovana analiza razvrš a vzorce v skupine brez predhodnega znanja ali predvidevanj o vzro nih mehanizmih danega ekspresijskega profila. S to metodo lahko klasificiramo biološke vzorce v kategorije, za katere prej nismo vedeli, da obstajajo. Primera algoritmov za nenadzorovano analizo sta samoorganizacijske mreže (angl. self-organizing maps, SOM) in hierarhi no razvrš anje.

Pri nadzorovani analizi upoštevamo obstoje e informacije o razredih fenotipov in kontrol. Namen je odkriti gene, ki so neposredno povezani s klini no relevantnim fenotipom. Ker je pri KLL pogosto problemati no majhno število preu evanih pacientov, hkrati pa želja po ugotavljanju spremenjenega izražanja im ve jega števila genov, je potrebno nameniti nekaj pozornosti tudi napovedni sposobnosti podpisa. Ali je podpis resni no povezan s fenotipom, preverimo s statisti nimi meritvami. Pogosto se uporablja navzkrižno

preverjanje tipa »izpusti enega« (angl. leave-one-out crossvalidation), kjer enega vzorca ne uporabimo za ustvarjanje podpisa, temve z njim preverimo e bi podpis pravilno napovedal njegovo biološko stanje. Z nadzorovano analizo so Rosenwald in sod. (2001) odkril ve kot 100 različnih izraženih genov med mutiranimi in nemutiranimi primeri KLL. Najbolje je skupini razlikoval protein ZAP70 in drugi proteini povezani s stimulacijo BCR. Temu lahko sledi še nadaljnja validacija z neodvisnimi vzorci, ki niso bili uporabljeni pri razvoju podpisa, izražajo pa dani fenotip. To so običajno tumorski vzorci oz. vzorci, ki jim želimo določiti profil. Pri tem merimo verjetnost, da rezultati testiranja z mikromrežo izražajo vzorec, ki je bil definiran s podpisom. Te verjetnosti prikažemo s topotnim grafom (angl. heatmap). Če želimo vzorce preučiti za več podpisov, jih lahko razvrstimo v skupine (angl. clustering). Algoritmi nadzorovane analize vključujejo ponderirano glasovanje, k-najbližjih sosedov, metodo podpornih vektorjev (support vector machine, SVM) in umetne nevronske mreže.

Pomembne metode za validacijo rezultatov mikromrež pa tudi neodvisno spremljanje genskega izražanja so še PCR v realnem času, reverzna transkripcija s PCR, prenos po Northernu...

Mnoge druge nadaljnje študije so potrdile diferencialno izražanje genov med mutiranimi in nemutiranimi primeri KLL (Klein in sod., 2001; Ferrer in sod., 2004; Abruzzo in sod., 2005) ter odkrile razlike v ravni izražanja tudi glede na prognostične markerje, ki so neposredno povezani z nemutiranim statusom genov IgVH, kot npr.: ZAP70 (Stamatopoulos in sod., 2009a), *LPL* (Bilban in sod., 2006; Lewintre in sod., 2009) in *CD38* (Düring in sod., 2003; Joshi in sod., 2007). Prav tako je bila predmet mnogih raziskav spremenjena ekspresija glede na razlike med celicami KLL in normalnimi celicami B (Zheng in sod., 2002; Jelinek in sod., 2003), spremenjeno izražanje glede na genomske aberacije (Haslinger in sod., 2004; Dickinson in sod., 2005, 2006; Hernandez in sod., 2009) ali glede na druga merila napredovanja bolezni, npr. sistem razvrščanja bolezni po Raiju in Binetu (Aalto in sod., 2001; Jelinek in sod., 2003; Van't Veer in sod., 2006), OS (angl. overall survival), TFS (angl. treatment free survival), PFS (angl. progression free survival) in LDT (angl. lymphocyte doubling time) (Dickinson in sod., 2005; Stamatopoulos in sod., 2009a).

### 2.5.6 Epigenetske študije

Prve indikacije vpletjenosti sprememb metilacije DNA v razvoj in napredovanje KLL so odkrili Wahlfors in sod. (1992). Odkrili so globalno hipometilacijo genomske DNA pri skoraj vseh analiziranih vzorcih bolnikov. Večinski delež izgube globalne metilacije se pripisuje daljšim repetitivnim zaporedjem DNA, vendar nekateri poročajo tudi o hipometilaciji genov *MYC* (Deguchi in sod., 1987), *BCL2* (Hanada in sod., 1993) in *TCL1A* (Yuille in sod., 2001).

Več skupin je poročalo o epigenetskem utišanju tumorskih supresorskih genov pri KLL. To vključuje gene vpletene v apoptozo, npr.: *DAPK1*, *WIF1*, *SFRP* (Chim in sod., 2006, 2008; Liu in sod., 2006), gene regulatorjev celičnega cikla *CDKN2A* (*p16INK4*) in *CDKN2B* (*p15INK4*) (Martel in sod., 1997; Chim in sod., 2006; Tsirigotis in sod., 2006), tumorski supresorski gen *DLEU7* (Hammersund in sod., 2003). Transkripcijski faktor in regulator poti p53, *TWIST2*, je selektivno utišan pri pacientih z mutiranimi geni IgVH, ki imajo bolj ugodno prognozo kot tisti pacienti s KLL, ki imajo nemutirane IgVH (Raval in sod., 2005). Podobno je bila metilacija dinukleotida CG, 344 bp navzdol od mesta za etika transkripcije *ZAP70*, povezana z nižjo ravnjo izražanja tega znanega prognostičnega markerja (Corcoran in sod., 2005).

Rush in sod. (2004) so za iskanje novih metiliranih genov pri KLL uporabili RLGS (angl. Restriction Landmark Genomic Scanning). RLGS je dvodimensionalna gelska elektroforeza, ki loči radioaktivno označene restrikcijske fragmente DNA. Restrikcijska mesta encimov so preferenčno locirana v CpG otokih. V raziskavi je bilo od 3000 analiziranih CpG otokov v povprečju 4,8% metiliranih. *GRM7*, enega od lokusov s spremenjeno metilacijo, so validirali z bisulfitnim sekvenciranjem ter dokazali izražanje gena po aplikaciji inhibitorja metiltransferaz.

### 2.5.7 Študije miRNA

miRNA ne predstavljajo samo diagnostičnih markerjev, ampak lahko ekspresijska raven miRNA napoveduje tudi klinični potek. Calin in sod. (2005) so prvi identificirali miRNA podpis, ki je bil diferencialno izražen pri pacientih glede na status IgVH in *ZAP70* ter

predlagali, da igrata miR-15a in miR-16 pomembno vlogo pri KLL. Tej študiji sta sledili še dve pomembnejši študiji (Fulci in sod., 2007; Marton in sod., 2008), vendar je bilo med rezultati razli nih raziskav zelo malo prekrivanja. Vzrok je verjetno v uporabi razli nih metod pri profiliranju miRNA. Kljub temu so vse študije identificirale znižano izražanje miR-223 in miRNA družine miR-29, kar je bilo povezano z agresivno obliko KLL. Stamatopoulos in sod. (2009b) so prvi sistematično analizirali ekspresijo miR-223 in miR-29c glede na več prognostičnih faktorjev pri KLL in odkrili, da imata obe miRNA znižano izražanje pri bolnikih s slabšo prognozo. Calin in sod. (2008) so v levkemični celični liniji in primarnih vzorcih KLL proučili profil genov, ki ga regulirata miR-15a in miR-16 in identificirali 60 njunih takih genov. V prihodnosti lahko pri akujemo odkritje novih miRNA vpletenih v KLL ter več eksperimentalno dokazanih takih.

Oligonukleotidna mikromreža je najbolj pogosto uporabljana tehnika za ocenjevanje za rak specifičnih ravni izražanja velikega števila miRNA. To je tehnika ekspresijskega profiliranja in tako predstavlja podobne izzive kot so opisani zgoraj. Poleg tega je zaradi kratke dolžine možna napaka na hibridizacija miRNA. Problematične so tudi talilne temperature različnih miRNA, ki so lahko zelo različne in tako lahko na intenzitetu signalov na mikromreži v veliki meri vplivajo intrinzične lastnosti molekul miRNA kot pa je to njihova koncentracija.

Druge metode za določanje ravni izražanja miRNA so še pretečna citometrija, ki temelji na kroglicah s sondami, komplementarnimi z miRNA (angl. bead-based flow-cytometric technique), kvantitativna PCR v realnem času za prekurozorske ali aktivne miRNA, miRAGE (zaporedna analiza genskega izražanja (SAGE) za genomsko analizo miRNA), RAKE test (angl. RNA-primed Array-based Klenow Extension) (Calin in Croce, 2006).

### 3 MATERIALI IN METODE

#### 3.1 ZBIRANJE LITERATURE

Vse obstoječe podatke o genetskih vzrokih kroni ne limfocitne levkemije smo zbrali tako, da smo pregledali podatkovne zbirke in znanstveno literaturo, dostopno v angleškem jeziku, objavljeno do julija 2009. Iskali smo po podatkovnih zbirkah PubMed (<http://www.ncbi.nlm.nih.gov/>) in Web of Science (<http://isiknowledge.com>) z iskalnimi pojmi: »chronic lymphocytic leukaemia, genetics, candidate genes, cytogenetics, epigenetics, DNA methylation, susceptibility, progression, cancer, miRNA, non-coding RNA, linkage, expression, microarray, association, SNP«. S temi iskalnimi nizi smo pridobili 253 relevantnih znanstvenih lankov. Upoštevali smo lanke, ki so vsebovali podatke o KLL na osnovi primarnih vzorcev ljudi, celi nih linij ali živalskih modelov. Vključene so bile tako študije posameznih genov, kot tudi genomske študije. Raziskave, ki niso natančno definirale metodologije, ali pa so bili rezultati prikazani dvoumno, niso bile upoštevane.

Informacije smo iskali še z orodji podatkovnega rudarjenja (angl. data mining) PubMeth (<http://www.pubmeth.org/>) in PolySearch (<http://wishart.biology.ualberta.ca/polysearch/>). Dodatno smo iskali po podatkovnih zbirkah: Nucleic Acids Research Database (<http://www.oxfordjournals.org/nar/database/c/>), Cancer Genome Anatomy Project (<http://cgap.nci.nih.gov/>), GoGene (<http://www.gopubmed.org/gogene/>), Atlas of Genetics and Cytogenetics in Oncology and Haematology (<http://atlasgeneticsoncology.org/index.html>), Genetic Association Database (<http://geneticassociationdb.nih.gov/cgi-bin/index.cgi>), GeneCards (<http://www.genecards.org/>), COSMIC (<http://www.sanger.ac.uk/genetics/CGP/cosmic/>), MethyCancer (<http://methycancer.genomics.org.cn/>), miR2disease Base (<http://www.mir2disease.org/>), The human microRNA disease database (HMDD, <http://cmbi.bjmu.edu.cn/hmdd>). Podatke o živalskih modelih smo dodatno preverili še v podatkovnih zbirkah MGI (<http://www.informatics.jax.org>) in RGD (<http://rgd.mcw.edu>).

### 3.2 IZDELAVA PODATKOVNE ZBIRKE GENSKIH LOKUSOV, POVEZANIH S KRONI NO LIMFOCITNO LEVKEMIJO

Podatkovno zbirko smo organizirali glede na sedem sklopov različnih študijskih pristopov: raziskave na i) živalskih modelih, ii) citogenetske študije, iii) analize genske povezanosti (angl. linkage analysis), iv) asociacijske in mutacijske študije, v) epigenetske študije, vi) ekspresijske študije, ter ločeno – vii) raziskave izražanja nekodirajočih RNA.

Ker rezultati raziskav pogosto niso podani kvantitativno, so informacije v podatkovni zbirki izključno kvalitativne ali semikvantitativne narave. Vrednosti in razmerja izražanja ter vključitvenega praga nismo definirali. V primeru ekspresijskih študij smo znižano gensko izražanje označili z -, zvišano pa z +.

### 3.3 IZBOR MED NEJŠIH KANDIDATNIH GENOV ZA KRONI NO LIMFOCITNO LEVKEMIJO

Podatkovno zbirko smo pregledali za gene, ki so bili v povezavi s KLL opisani v več študijah (z enakim ali drugačnim študijskim pristopom). Med najše kandidatne gene, povezane s KLL, smo izbrali s presekom podatkov mutacijskih, asociacijskih, ekspresijskih, epigenetskih študij in živalskih modelov. Za vključitev gena v izbor med najših kandidatov je moral ta zadostiti sledečim kriterijem:

- gen je bil povezan s KLL v več neodvisnih študijah z različnimi študijskimi pristopi, ali
- gen je bil povezan s KLL v več različnih neodvisnih študijah z enakim študijskim pristopom, ali
- gen je bil povezan s KLL v eni raziskavi z različnimi študijskimi pristopi, ali
- gen je bil povezan z več različnimi prognostičnimi markerji ali kliničnimi stanji KLL v eni raziskavi z enakim študijskim pristopom.

Za vsako neodvisno pojavnost gena v literaturi, ki je ustrezala zgornjim kriterijem, smo v preglednici med najših kandidatnih genov pripisali en znak +.

Ve SNP-jem enega gena, odkritih v posamezni študiji, smo pripisali le en znak +.

Ker je bil namen tega postopka poiskati pomembne gene, povezane s KLL, njihove funkcije pa smo preu ili v kasnejšem koraku raziskave, smo v izbor mo nejših kandidatnih genov prav tako vklju ili gene, ki so se pojavili v ve razli nih študijah, vendar so se ugotovljene funkcije genov ali ugotovljena raven izražanja med razli nimi študijami razlikovali ali celo izklju evali.

### 3.4 FUNKCIJSKA ANALIZA GENOV

Najbolj zastopane funkcije izbranih mo nejših kandidatnih genov smo ugotovili z orodji funkcijalne anotacije podatkovne zbirke DAVID (The Database for Annotation, Visualization and Integrated Discovery, verzija 2008; <http://david.abcc.ncifcrf.gov/home.jsp>; Dennis in sod., 2003; Huang in sod., 2009). Anotacije genov smo najprej preverili po skupinah. Uporabili smo funkcijo »Functional Annotation Clustering« z nastavljivo najbolj strogih kriterijev, ki temelji na hipotezi, da imajo podobne anotacije podobne gene. Nato smo v razdelku splošnih anotacij s funkcijo Cytoband preverili najbolj zastopane kromosomske lokacije genov, v razdelku bolezni termine povezane z obolenji (Disease Association Database), v razdelku poti zastopanost bioloških poti iz podatkovne zbirke KEGG (Kyoto Encyclopedia of Genes and Genomes), v razdelku funkcionalnih kategorij s funkcijo »Uniprot Sequence Feature« pa globalne sekvene motive. Nato smo z orodjem miRNAPath (<http://lgmb.fmrp.usp.br/mirnopath/index.php>; Chiromatzo in sod., 2007) preu ili odnose med geni za KLL, miRNA za KLL in metabolnimi potmi. S funkcijo analize genov smo vnesli imena genov za KLL, s funkcijo analize miRNA pa molekule miRNA, povezane s KLL. Orodje po vrsti razvrsti miRNA, ki lahko regulirajo najve je število vnesenih genov oz. domnevnih tar vnesenih miRNA. Za predikcije tar uporablja program miRNAPath algoritem miRBase Targets.

### 3.5 BIOINFORMACIJSKE ANALIZE miRNA V POVEZAVI S KRONI NO LIMFOCITNO LEVKEMIJO

Z bioinformacijskimi analizami smo raziskali polimorfizem ter poiskali še neodkrite povezave na treh ravneh z miRNA posredovane regulacije.

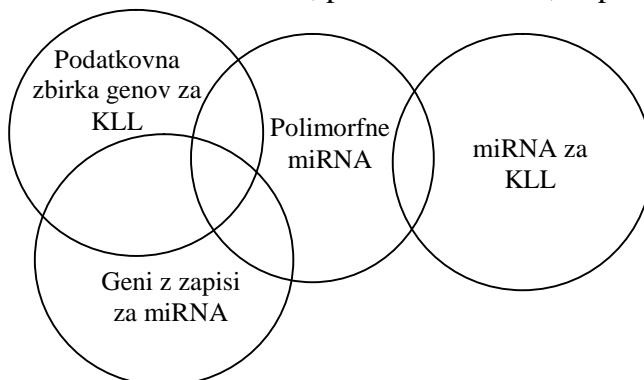
#### 3.5.1 miRNA geni

Informacije o lokacijah prekrivajo ih genov s transkripti miRNA (angl. miRNA host genes) smo pridobili preko podatkovne zbirke miRBase (<http://microrna.sanger.ac.uk>). Preverili smo, e se v katerem od kandidatnih genov iz preseka podatkovne zbirke za KLL nahaja miRNA, za katero še ni bila dokazana vpletjenost pri KLL. Funkcije genov z zapisimi za miRNA smo preverili v podatkovni zbirki DAVID z anotacijo funkcionalnih kategorij (»PIR keyword«).

S podatkovno bazo Patrocles (<http://www.patrocles.org/>) smo za gene povezane s KLL, v katerih se nahajajo geni za miRNA, še dodatno preverili, e so miRNA molekule polimorfne. Prav tako pa smo preverili, e se nahaja SNP v katerem od s KLL povezanih genov za miRNA.

Naredili smo:

- prese no množico podatkovne zbirke genov za KLL in genov, v katerih se nahajajo geni za miRNA
- prese no množico genov, povezanih s KLL, v katerih se nahajajo miRNA, in geni polimorfnih miRNA;
- prese no množico miRNA, povezanih s KLL, in polimorfnih miRNA (slika 3).



Slika 3: Vennov diagram prese nih množic za gene miRNA pri kroni ni limfocitni levkemiji

### 3.5.2 miRNA tar e

Eksperimentalno potrjene tar e miRNA, ki so geni za KLL, smo pridobili preko podatkovne baze miRecords (<http://mirecords.umn.edu/miRecords/>).

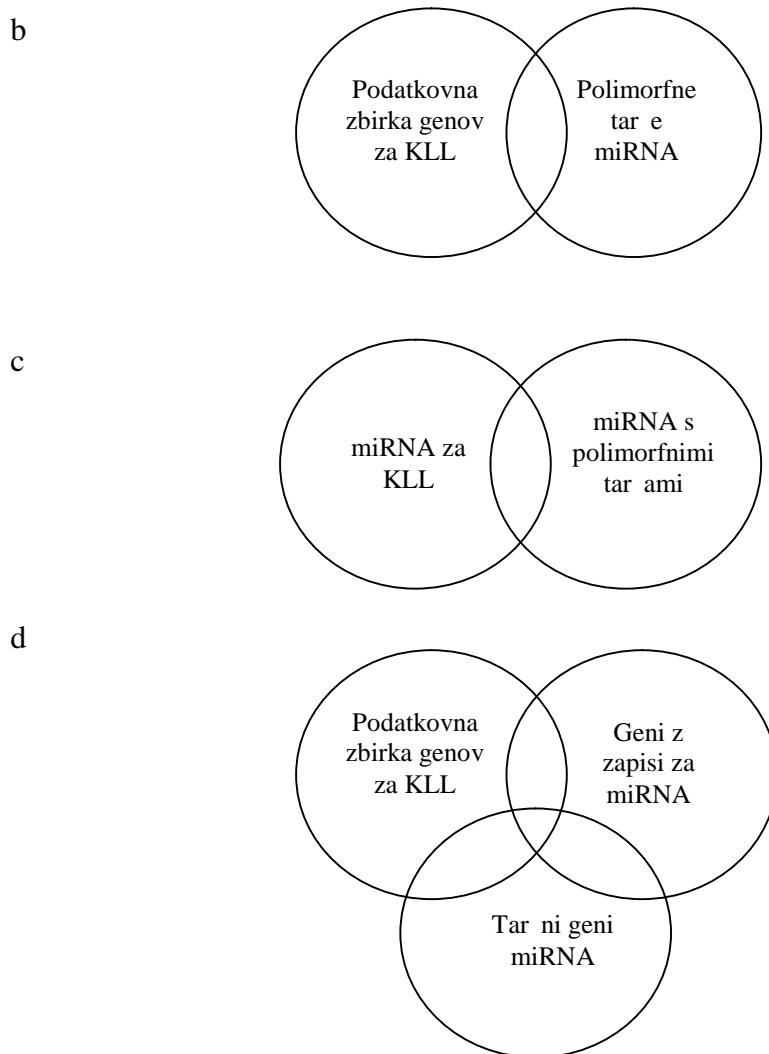
S podatkovno zbirko Patrocles smo preverili, e obstaja še nedokazana povezava med polimorfizmom znotraj kandidatnih genov za KLL in miRNA ali pa povezava med kandidatnimi miRNA in polimorfizmi tar nih genov. Prav tako smo analizirali, e so s KLL povezane tar e miRNA, ki ležijo znotraj genov povezanih s KLL.

Naredili smo:

- prese no množico podatkovne zbirke genov za KLL in validiranih tar miRNA (slika 4a);
- prese no množico podatkovne zbirke genov za KLL, ki so hkrati tar e miRNA, in miRNA za KLL (slika 4a);
- prese no množico podatkovne zbirke genov za KLL in vseh polimorfnih tar nih genov (slika 4b);
- prese no množico zbirke miRNA v povezavi s KLL in vseh miRNA s polimorfnimi tar ami (slika 4c);
- prese no množico podatkovne zbirke genov za KLL in tar nih genov miRNA, ki se nahajajo v genih, povezanih s KLL (slika 4d).

a





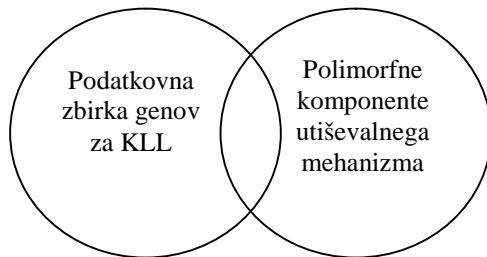
Slika 4: Vennovi diagrami: a) prese ne množice podatkovne zbirke genov za kroni no limfocitno levkemijo (KLL), validiranih tar miRNA in miRNA za KLL, b) prese na množica podatkovne zbirke genov za KLL in polimorfnih tar miRNA, c) prese na množica miRNA, povezanih s KLL in miRNA s polimorfnimi tar ami, d) prese na množica podatkovne zbirke genov za KLL, in tar nih genov miRNA, ki se nahajajo v genih, povezanih s KLL.

Funkcije genov, ki so predstavljali validirane in polimorfne tar e, smo preverili z orodjem za združevanje v skupine podatkovne zbirke DAVID.

### 3.5.3 Komponente utiševalnega mehanizma miRNA

S podatkovno zbirko Patrocles smo preverili, e se kateri od genov polimorfnih komponent utiševalnega kompleksa nahaja v podatkovni zbirki za KLL in tako vpliva na razvoj KLL

(slika 5). Funkcijo genov smo preverili še z orodji za anotacijo v podatkovni zbirkì DAVID.



Slika 5: Prese na množica podatkovne zbirke genov za KLL in polimorfnih komponent utiševalnega mehanizma

### 3.6 GENSKE MREŽE Z miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI

Da bi spoznali funkcionalno povezane gene in nove povezave med geni pri KLL smo gene in miRNA povezane s KLL narisali v genske mreže. Preizkusili smo dva programa za risanje genskih mrež: Ingenuity Pathway Analysis in MetaCore (GeneGo).

#### 3.6.1 Ingenuity Pathway Analysis (IPA 7.5 software, Ingenuity Systems®, [www.ingenuity.com](http://www.ingenuity.com) )

Zbirki 241 kandidatnih genov in 197 miRNA, ki so deregulirane pri KLL, smo lo eno prenesli v datoteko »My Lists« ter nato na zbirki kandidatnih genov izvedli »Core analysis«. Izbrali smo vse tipe identifikatorjev, saj je v zbirki ve tipov identifikatorjev genov ter analizo omejili na loveka.

Po analizi so mreže hierarhi no urejene po vrsti glede na rezultat. To je števil na vrednost, ki meri relevantnost mreže glede na gene v vhodni podatkovni zbirki. Pri tem je upoštevano število fokusnih genov (geni iz vhodne podatkovne zbirke, ki imajo direktne povezave z drugimi geni iz vhodne podatkovne zbirke) v mreži in velikost mreže. Identificirane mreže so predstavljene v obliki grafov, ki kažejo odnose med geni oz. genskimi produkti. Geni so predstavljeni kot vozliš a (angl. nodes), biološki odnosi med dvema vozliš ema pa z robovi (angl. edges) oz. rtami. Vozliš a so razli nih oblik glede na funkcionalni razred genskega produkta. Geni, ki niso obarvani, niso bili prisotni v izvorni podatkovni zbirki, temve so integrirani v mrežo na osnovi podatkov, shranjenih v bazi

znanja IPA. Predstavljajo nove kandidatne gene za diagnosti ne ali prognosti ne ozna evalce ali terapevtske tar e, ki smo jih dodatno analizirali z orodji podatkovne zbirke DAVID. Vozliš a, ki imajo veliko povezav z drugimi vozliš i imenujemo povezovalniki (angl. hubs).

Nato smo za gene miRNA uporabili funkcijo »IPA-Biomarker<sup>TM</sup> Analysis«. S filtri smo analizo omejili na limfocitne celice in bolezni raka pri loveku. miRNA, ki so preše selekcijo smo integrirali v mrežo genov, povezanih s KLL.

Podatkovna baza interakcij IPA je s strani raziskovalcev manualno vzdrževana in posodobljena štirikrat letno. Rezultati naše raziskave so bili pridobljeni julija 2009.

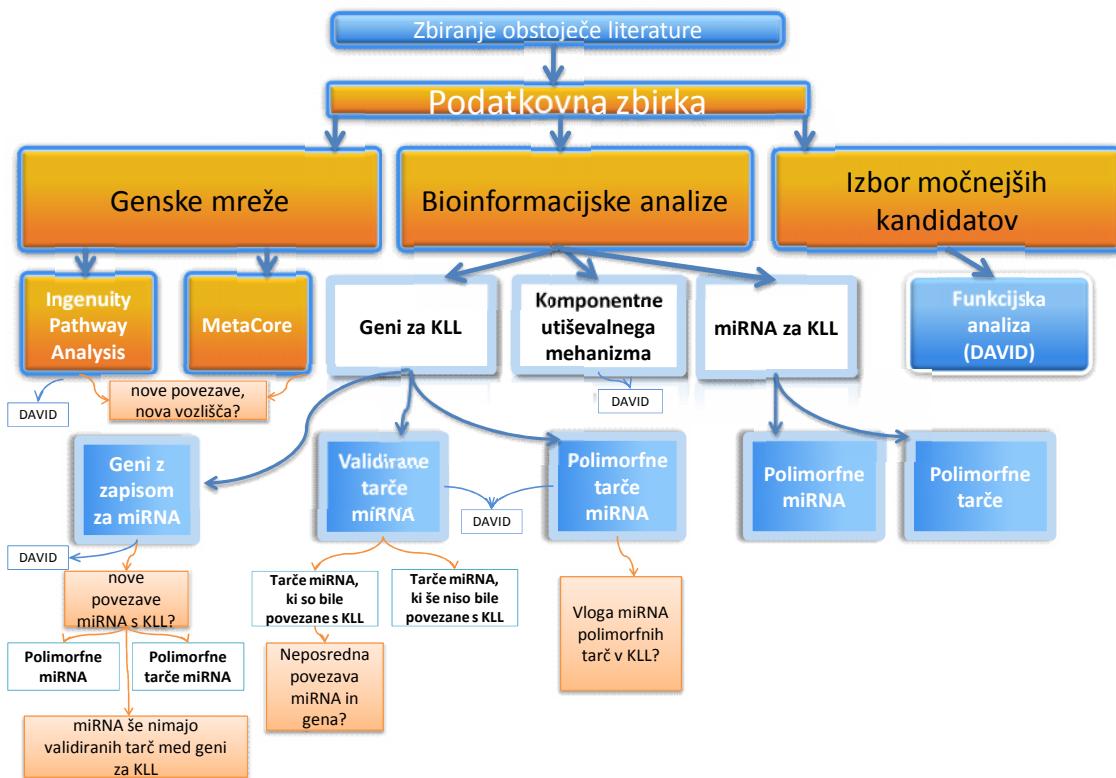
### 3.6.2 MetaCore<sup>TM</sup> (GeneGo, St Joseph, MI, [www.genego.com](http://www.genego.com) )

MetaCore<sup>TM</sup> prav tako omogo a analizo eksperimentalnih podatkov in primerjavo z manualno vzdrževano bazo znanja. V tem programu smo analizirali miRNA, povezane s KLL, in njihove interakcije z geni, ki so bili povezani s KLL, in hkrati predstavljajo njihove eksperimentalno validirane tar e, ki smo jih pridobili iz podatkovne baze miRecords (<http://mirecords.umn.edu/miRecords/>). Z orodjem MetaLink<sup>TM</sup> smo vnesli podatkovno zbirko 98 interakcij miRNA in njihovih tar , ki so bili eksperimentalno povezani s KLL. MetaCore ponuja možnost uporabe štirih razli nih algoritmov. Preizkusili smo algoritem »Analyze networks«, ki omogo a analizo ve jega števila genov in mrežo razdeli v podmreže ter jih razvrsti glede na obogatenost s podatki iz vhodne datoteke in glede na prisotnost kanoni nih poti. Podmreže, ki so rangirale više ter so bile posredno povezane s KLL, smo združili v eno mrežo.

Prav tako smo uporabili Dijkstrov algoritem »Shortest paths«, ki poiš e im bolj neposredne povezave med geni iz vhodne podatkovne zbirke ter dodatnimi elementi iz baze programa.

### 3.7 SHEMA INTEGRATIVNEGA GENOMSKEGA PRISTOPA

Na sliki 6 je shematsko prikazan potek dela s sistemskim študijskim pristopom.

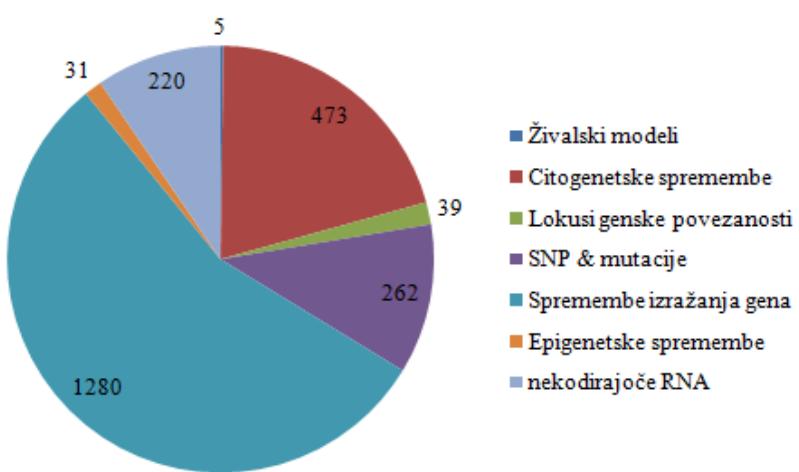


Slika 6: Shema integrativnega genomskega pristopa

## 4 REZULTATI

### 4.1 IZDELAVA PODATKOVNE ZBIRKE GENSKIH LOKUSOV, POVEZANIH S KRONI NO LIMFOCITNO LEVKEMIJO

S sistematskim in intenzivnim pregledovanjem obstoječe literature smo odkrili 253 študij in več kot 2000 genskih lokusov, povezanih z genetiko KLL. V teh študijah so bili uporabljeni različni pristopi za identifikacijo genov, ki smo jih razvrstili v sedem sklopov: živalski modeli, citogenetske študije, analize genske povezanosti, asociacijske in mutacijske študije, ekspresijske analize, epigenetske študije ter letno analize nekodirajočih RNA pri KLL (slika 7). Tako smo organizirali tudi podatkovno zbirkovo vseh znanih genskih lokusov v povezavi s KLL, ki je prosto dostopna na spletni strani <http://www.bf.uni-lj.si/fileadmin/groups/2726/ChronicLymphocyticLeukemia.pdf> in vsebuje informacije o imenih genov, citogenetskih lokacijah, SNP-jih, genski ekspresiji ali metilaciji, reference ustreznih študij ter nekatere druge relevantne informacije v povezavi z bolezni, npr. povezanost s specifičnimi prognostičnimi markerji ali ocenami napredovanja bolezni. Podatkovna zbirkova, na osnovi katere smo opravili nadaljnje analize, je vključevala literaturo, objavljeno do julija 2009. V prihodnosti bomo podatkovno zbirkovo dopolnjevali.

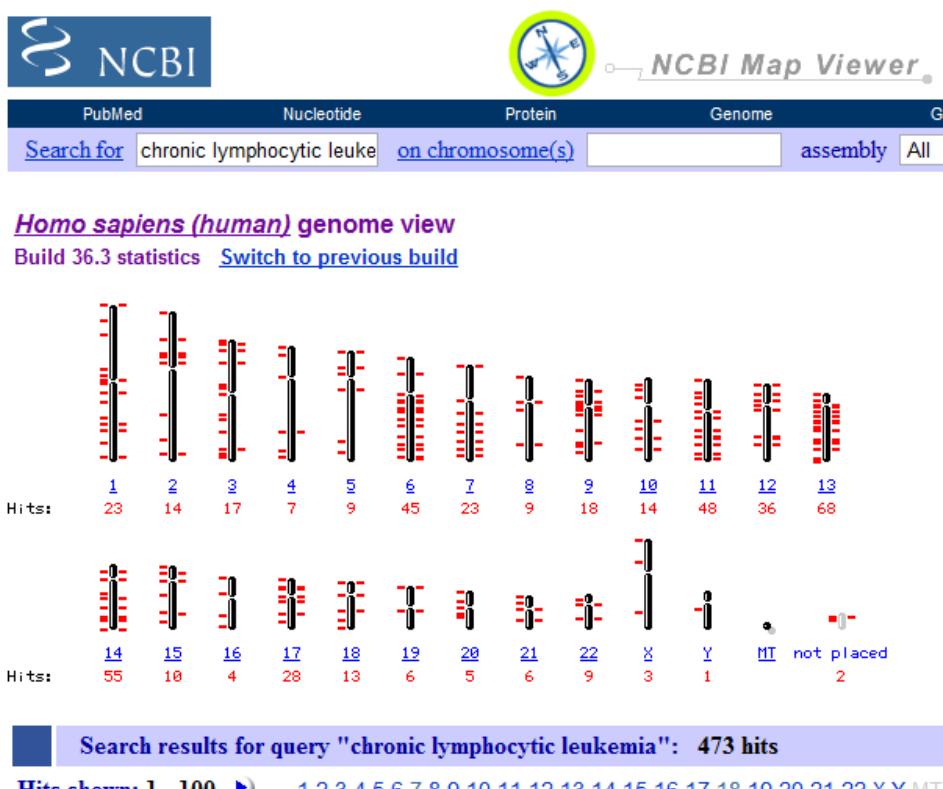


Slika 7: Število lokusov, povezanih s kroni no limfocitno levkemijo, odkritih z različnimi študijskimi pristopi<sup>1</sup>

<sup>1</sup> Prekrivanje genov odkritih z različnimi študijskimi pristopi ni upoštevano.

V literaturi smo odkrili pet živalskih modelov povezanih s KLL. Od teh je v podatkovni zbirki Mouse Genome Informatics (MGI) le miška s transgenom za *TCL1*. Pregledali smo tudi podatkovno zbirko Rat Genome Database (RGD) in nismo našli nobenih informacij o živalskih modelih za KLL.

V citogenetskih študijah smo zasledili 152, z orodjem NCBI Map Viewer (<http://www.ncbi.nlm.nih.gov/projects/mapview/>; slika 8) pa 473 kromosomskih mutacij v povezavi s KLL.



Slika 8: NCBI Map Viewer je brskalno orodje, ki omogoča iskanje in grafični prikaz genomskega informacijskega sklopa na citogenetski lokacijski ravni. Za kroni ni limfocitno levkemijo je posebej velik zadelek na kromosomih 6 in 13.

Z analizami genske povezanosti je bilo odkritih 49 citogenetskih lokacij na 19 kromosomih, ki vplivajo na dedno nagnjenost na KLL, vendar so bile signifikantne vrednosti LOD ali NPL (angl. nonparametric linkage) večkrat neodvisno odkrite le pri nekaterih regijah.

Zbrali smo podatke iz 34 asociacijskih raziskav, v katerih poro ajo o 262 SNP-jih, povezanih s KLL. Povezanost SNP-jev in mutacij s KLL je bila v razli nih raziskavah dokazana ve kot enkrat za gene *HLA-DRB1*, *POLB*, *P2X7* in *TNFA*, *ATM*, *CD79b*, *p16(INK4a)* in *TP53*.

S profiliranjem genskega izražanja je bilo skupno odkritih 1280 diferencialno izraženih genov pri KLL. Od tega je bilo 33 genov diferencialno izraženih v štirih ali ve neodvisnih študijah. Zaradi razlik med uporabljenimi platformami, sprememb ravni ekspresije v okviru diplomske naloge nismo upoštevali oz. te pogosto v poro ilih znanstvenih raziskav sploh niso bile podane. Zato je možno, da so razlike v izražanju nekaterih genov v podatkovni zbirki majhne.

Odkrili smo 21 študij o povezavi epigenetske regulacije pri KLL za 31 genov, od tega za šest genov (*DAPK1*, *SFRP1*, *SFRP2*, *SFRP4*, *SFRP5*, *p16*) v dveh ali ve neodvisnih študijah.

S pregledom literature smo pridobili 14 lankov o nekodirajo ih RNA, vpletenih v KLL, in identificirali 197 miRNA (vklju no s prekurzorskimi miRNA) in 23 T-UCR (angl. transcribed ultra-conserved regions). Ve ina raziskav je identificirala spremembe izražanja miRNA, obstaja pa tudi nekaj dokazanih mutacij v genih miRNA (Calin in sod., 2005; Raveche in sod., 2007).

#### 4.2 IZBOR MO NEJŠIH KANDIDATNIH GENOV ZA KRONI NO LIMFOCITNO LEVKEMIJO

Za nadaljnje bioinformacijske analize smo želeli izbrati mo nejše kandidatne gene, ki so bili s KLL povezani v ve študijah oz. z razli nimi študijskimi pristopi ali pa povezani z razli nimi prognosti nimi ozna evalci. V izboru mo nejših kandidatov (priloga A) smo združili 241 genov iz mutacijskih, asociacijskih, ekspresijskih ali epigenetksih študij in živalskih modelov, ki predstavlja 15% vseh genov teh pristopov. Štiriintrideset genov je bilo povezanih s KLL z ve kot enim tipom študijskega pristopa. To vklju uje tako prognosti ne kot diagnosti ne kandidatne gene. V okviru naše naloge nismo upoštevali

terapevtskih markerjev, saj bi še ve ja kompleksnost seznama genov ovirala nadaljnjo interpretacijo informacij. Diagnosti ne ali prognosti ne implikacije so razvidne iz podatkovne zbirke. S stališ a zdravljenja so informacije o diagnosti nih, prognosti nih ali terapevtskih markerjih bolj uporabne, e so predstavljene lo eno, kot pa skupaj. Vendar, ker smo za nadaljnjo analizo sklepali, da pride tekom bolezni do deregulacije interakcij genov, povezanih tako z inicacijo kot napredovanjem bolezni, poleg tega pa to delo predstavlja preliminarno študijo, smo analizo zbranih genov poenostavili. Vse gene smo obravnavali enakovredno, ne glede na povezanost s specifi nimi zna ilnostmi fenotipa KLL, spremenjeno raven izražanja, epigenetsko utišanje ali aktivacijo gena ter ne glede na asovno specifiko izražanja.

#### 4.3 FUNKCIJSKA ANALIZA GENOV

Funkcije izbora mo nejših kandidatnih genov smo preverili z bioinformacijskim orodjem za anotacijo skupin genov podatkovne zbirke DAVID ter potrdili, da so geni za KLL vpleteni v številne biološke funkcije. Glavne anotacijske skupine zbranih genov za KLL so bile povezane z imunoglobulini in B-celi nim receptorjem, regulacijo fosforilacije, negativno regulacijo celi ne rasti in regulacijo aktivacije celic B (slika 9a). S funkcijo Cytoband smo odkrili, da so najbolj zastopane kromosomske lokacije 17p13.1, 11q23, 4q34 in 12p12-p13 (slika 9b). V anotaciji s pojmi bolezni je bilo 25,5% genov povezanih s pojmom imunski, 24,5% s pojmom rak, 7,1% pa s pojmom hematološki (slika 9c). Geni so povezani tudi s pojmi infekcija, staranje, sr no-žilni in razvoj. Za kandidatne gene so iz baze KEGG najbolj zna ilne poti: signalna pot p53, hematopoetska celi na linija, apoptoza, signalna pot BCR in celi ni cikel (slika 9d). Najpomembnejše funkcijске kategorije proteinov so sekven ne variante, mesto mutageneze in pa proteini z jedrnim lokalizacijskim signalom (slika 9e).

a

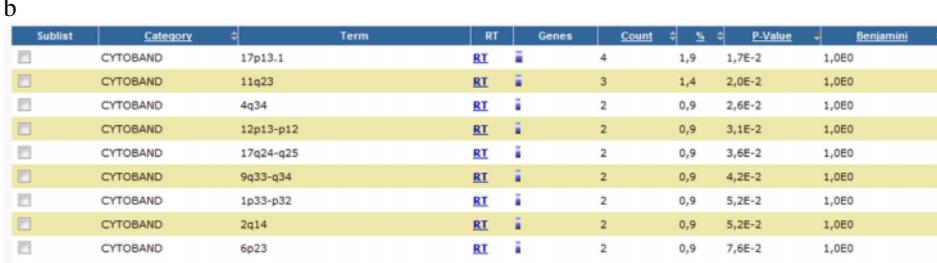
<http://david.abcc.ncifcrf.gov/termTerm.jsp>



The screenshot shows the DAVID Bioinformatics Resources 2008 interface with the title "Functional Annotation Clustering". It displays six annotation clusters with their enrichment scores and associated biological processes. The columns include Annotation Cluster, Enrichment Score, RT, Genes, Count, P-Value, Fold Enrichment, Bonferroni, Benjamini, and FDR.

Annotation Cluster	Enrichment Score	RT	Genes	Count	P-Value	Fold Enrichment	Bonferroni	Benjamini	FDR
Annotation Cluster 1	4.76	RT	1	6.9E-6	82.2	6.0E-3	2.0E-3	0	
Annotation Cluster 2	4.21	RT	1	6.9E-6	82.2	6.0E-3	2.0E-3	0	
Annotation Cluster 3	3.99	RT	1	1.1E-1	38.6	4.5E-1	5.8E-3	0.2	
Annotation Cluster 4	3.81	RT	1	4.6E-5	11.1	2.1E-1	2.1E-1	0.1	
Annotation Cluster 5	3.69	RT	1	6.5E-5	7.0	2.0E-1	3.7E-3	0.1	
Annotation Cluster 6	3.59	RT	1	6.5E-5	7.0	2.0E-1	3.7E-3	0.1	

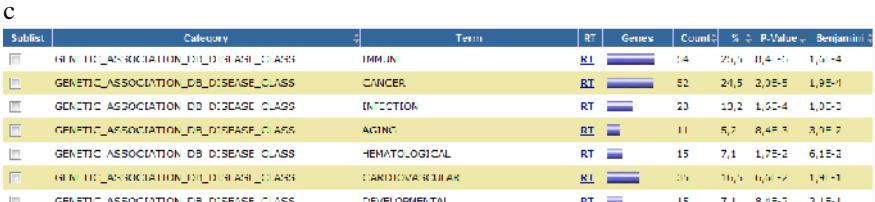
b



A table listing the top 10 chromosomal locations (Sublist, Category, Term, RT, Genes, Count, %, %, P-Value, Benjamini) for KLL genes. The data shows various chromosomes and bands, such as 17p13.1, 11q23, 4q34, etc.

Sublist	Category	Term	RT	Genes	Count	%	%	P-Value	Benjamini
CYTOBAND	17p13.1	RT	1	4	1,9	1,7E-2	1,0E0		
CYTOBAND	11q23	RT	1	3	1,4	2,0E-2	1,0E0		
CYTOBAND	4q34	RT	1	2	0,9	2,6E-2	1,0E0		
CYTOBAND	12p13-p12	RT	1	2	0,9	3,1E-2	1,0E0		
CYTOBAND	17q24-q25	RT	1	2	0,9	3,6E-2	1,0E0		
CYTOBAND	9q33-q34	RT	1	2	0,9	4,2E-2	1,0E0		
CYTOBAND	1p33-p32	RT	1	2	0,9	5,2E-2	1,0E0		
CYTOBAND	2q14	RT	1	2	0,9	5,2E-2	1,0E0		
CYTOBAND	6p23	RT	1	2	0,9	7,6E-2	1,0E0		

c



A table listing genetic associations (Sublist, Category, Term, RT, Genes, Count, %, %, P-Value, Benjamini) for KLL genes. Categories include Genetic Association DB Disease Class, Infection, Aging, Hematological, Cardiovascular, and Developmental.

Sublist	Category	Term	RT	Genes	Count	%	%	P-Value	Benjamini
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	IMMUN	KL	1	14	21,4	0,4	0	1,1E-4	
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	CANCER	RT	1	52	21,5	2,0E-5	1,9E-1		
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	INFECTION	RT	1	23	10,2	1,0E-4	1,0E0		
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	AGING	RT	1	11	5,2	8,4E-5	3,7E-2		
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	HEMATOLOGICAL	RT	1	15	7,1	1,7E-2	6,1E-2		
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	CARDIOVASCULAR	KL	1	16	10,6	0,6	0	1,4E-1	
GENETIC_ASSOCIATION_DB_DISEASE_CLASS	DEVELOPMENTAL	RT	1	15	7,1	6,1E-2	2,1E-1		

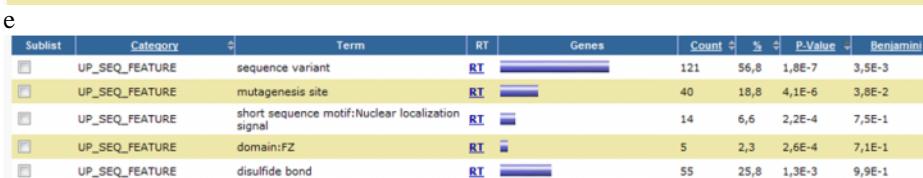
d



A table listing KEGG pathway annotations (Sublist, Category, Term, RT, Genes, Count, %, %, P-Value, Benjamini) for KLL genes. Categories include KEGG PATHWAY, p53 signaling pathway, Hematopoietic cell lineage, Apoptosis, B cell receptor signaling pathway, and Cell cycle.

Sublist	Category	Term	RT	Genes	Count	%	%	P-Value	Benjamini
KEGG PATHWAY	p53 signaling pathway	RT	1	13	0,7	3,1E-7	6,2E-5		
KEGG PATHWAY	Hematopoietic cell lineage	RT	1	14	0,8	5,5E-7	5,5E-5		
KEGG PATHWAY	Apoptosis	RT	1	13	0,7	3,3E-6	2,0E-4		
KEGG PATHWAY	B cell receptor signaling pathway	RT	1	10	0,5	7,6E-5	3,8E-3		
KEGG PATHWAY	Cell cycle	RT	1	13	0,7	7,7E-5	3,1E-3		

e



A table listing UP\_SEQ\_FEATURE annotations (Sublist, Category, Term, RT, Genes, Count, %, %, P-Value, Benjamini) for KLL genes. Categories include UP\_SEQ\_FEATURE, sequence variant, mutagenesis site, short sequence motif:Nuclear localization signal, domain:FZ, and disulfide bond.

Sublist	Category	Term	RT	Genes	Count	%	%	P-Value	Benjamini
UP_SEQ_FEATURE	sequence variant	RT	1	121	56,8	1,8E-7	3,5E-3		
UP_SEQ_FEATURE	mutagenesis site	RT	1	40	18,8	4,1E-6	3,8E-2		
UP_SEQ_FEATURE	short sequence motif:Nuclear localization signal	RT	1	14	6,6	2,2E-4	7,5E-1		
UP_SEQ_FEATURE	domain:FZ	RT	1	5	2,3	2,6E-4	7,1E-1		
UP_SEQ_FEATURE	disulfide bond	RT	1	55	25,8	1,3E-3	9,9E-1		

Slika 9: Rezultati analize funkcij mo nejših kandidatnih genov za kroni no limfocitno levkemijo (KLL) z orodji podatkovne zbirke DAVID. a) Prve štiri anotirane skupine genov, b) Najbolj zastopane kromosomske lokacije genov za KLL, c) Anotacija s pojmi bolezni, d) Anotacija z glavnimi biološkimi potmi iz baze KEGG, e) Glavni sekveni ni motivi genov za KLL.

Z orodjem miRNPath smo ugotovili, da imajo največ tar nih genov izmed izbranih kandidatnih genov molekule miRNA družine let-7 in miR-181c (preglednica 3). Ko smo v program miRNPath vnesli seznam dereguliranih miRNA pri KLL, smo pridobili informacijo, da imajo veliko predvidenih tar nih mest družine miR-181, miR-30 in miR-15 (preglednica 4).

Preglednica 3: Rezultati analize izbranih močnejših kandidatnih genov za kroni no limfocitno levkemijo (KLL) glede na miRNA z orodjem miRNPath<sup>2</sup>

miRNA	Število tar nih genov iz izbora	Število poti
hsa-let-7c	38	68
hsa-let-7f	37	63
hsa-let-7g	36	63
hsa-let-7b	35	62
hsa-miR-181c	35	44
hsa-let-7e	34	63
hsa-let-7i	34	74
hsa-miR-29c	32	31
hsa-miR-29b	32	32
hsa-let-7d	32	56

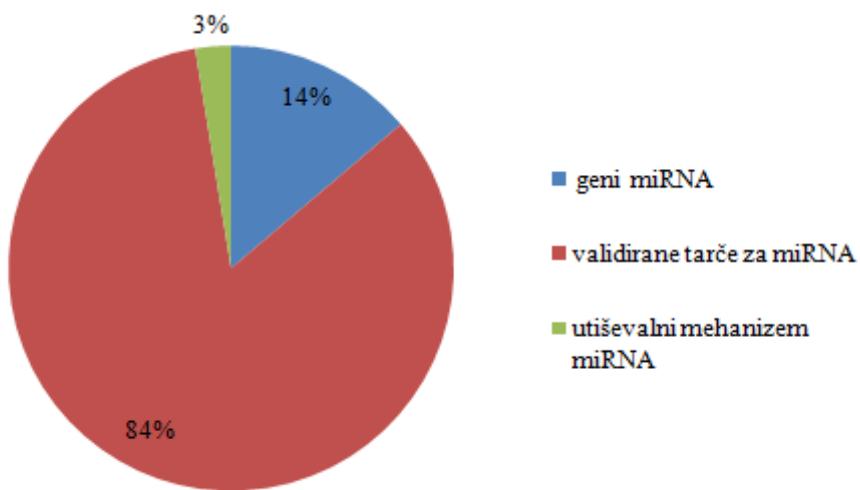
Preglednica 4: Rezultati analize miRNA pri kroni ni limfocitni levkemiji (KLL) glede na število genov, ki jih regulirajo, z orodjem miRNPath<sup>3</sup>

miRNA	Število genov	Število poti
hsa-miR-181b	1758	162
hsa-miR-30c	1752	161
hsa-miR-30d	1731	164
hsa-miR-181c	1721	167
hsa-miR-181a	1719	168
hsa-miR-15b	1718	164
hsa-miR-30b	1717	164
hsa-miR-27b	1682	157
hsa-miR-107	1663	161
hsa-miR-15a	1663	165

<sup>2,3</sup> Prikazanih je prvih deset zadetkov.

#### 4.4 BIOINFORMACIJSKE ANALIZE miRNA

Med geni za KLL, pri katerih smo našli povezavo z miRNA, je bilo 3% genov (3 geni, preglednica 6) vpletenih v utiševalni mehanizem miRNA, 14% genov (16 genov, preglednica 5) je vklju evalo zapis za gene miRNA, 84% genov (97 genov) pa so eksperimentalno potrjene tar e molekul miRNA (slika 10).



Slika 10: Deleži genov za kroni no limfocitno levkemijo (KLL), glede na vlogo, povezano z miRNA

Z izdelano podatkovno zbirko in s spletno bazo miRBase smo odkrili, da se v 16 od 1578 genov (vsote genov ekspresijskih, epigenetskih in asociacijskih študij ter živalskih modelov<sup>4</sup>; preglednica 5), povezanih s KLL, nahajajo že odkrite miRNA. Od 17 genov za miRNA, ki smo jih odkrili v 16 genih za KLL, je bila za dve miRNA (hsa-miR-32, hsa-miR-1200) predhodno že dokazana povezava s KLL (preglednica 5). Preostalih 15 miRNA pa bi veljalo preizkusiti, e so tako kot njihovi gostiteljski geni deregulirane v KLL. Ve ina miRNA je bila intronskih, razen miR-220b, ki se nahaja v eksonu gena *TUBB4* in je poleg tega polimorfna. Izmed gostiteljskih genov za miRNA je posebej zanimiv gen *TOP3B*, v katerem se nahajata gena za dve miRNA (hsa-mir-301b, hsa-mir-649). miRNA, ki se nahajajo v genih za KLL, trenutno še nimajo eksperimentalno potrjenih tar . Preverili smo njihovo genetsko variabilnost in sicer se je izkazalo, da je pet od 16 miRNA

<sup>4</sup> Prekrivanje genov ni upoštevano.

polimorfnih, 11 pa jih ima polimorfne tar e. Funkcijska analiza genov je pokazala, da je pri genih za KLL z zapisi za miRNA pogosto alternativno izklju evanje eksonov.

Preglednica 5: Geni, povezani s kroni no limfocitno levkemijo (KLL), v katerih se nahajajo zapisi za miRNA ter odgovarjajo i polimorfizmi danih miRNA in polimorfizmi njihovih tar glede na podatkovno zbirko Patrocles

Prekrivajo i kodirajo i gen, povezan s KLL	Gen za miRNA	Že povezava s KLL	ugotovljena miRNA s polimorfizem	Ugotovljen polimorfizem miRNA	Polimorfne miRNA	tar e
<i>C9orf5</i>	hsa-mir-32	✓ (Calin in sod., 2004; Wang in sod., 2008b)	✗		✓	
<i>GPC1</i>	hsa-mir-149	✗	rs2292832; prekurzor		✓	
<i>TUBB4</i>	hsa-mir-220b	✗	rs10532262; prekurzor		✓	
<i>TOP3B</i>	hsa-mir-301b	✗	✗		✓	
<i>TOP3B</i>	hsa-mir-649	✗	✗		✗	
<i>SGCZ</i>	hsa-mir-383	✗	✗		✗	
<i>DMD</i>	hsa-mir-548f-5	✗	✗		✗	
<i>MRE11</i>	hsa-mir-548l	✗	rs11020790; prekurzor rs13447640; zrela miRNA		✓	
<i>PDE4D</i>	hsa-mir-582	✗	✗		✓	
<i>FUT8</i>	hsa-mir-625	✗	rs12894182; prekurzor		✓	
<i>PRKCA</i>	hsa-mir-634	✗	✗		✓	
<i>DAPK3</i>	hsa-mir-637	✗	✗		✗	
<i>RUNX1</i>	hsa-mir-802	✗	✗		✓	
<i>ATF2</i>	hsa-mir-933	✗	✗		✗	
<i>ELMO1</i>	hsa-mir-1200	✓ (Marton in sod., 2008)	✗		✗	
<i>PP3CA</i>	hsa-mir-1255a	✗	rs28664200; prekurzor		✗	
<i>CSMD3</i>	hsa-mir-2053	✗	✗		✓	

✓ povezava s KLL je že odkrita/ tar a dane miRNA je polimorfna;

✗ povezava s KLL še ni odkrita/ polimorfizem v dani miRNA še ni odkrit/ polimorfizem v dani tar i miRNA še ni odkrit.

S primerjavo podatkovne zbirke genov za KLL in informacij o polimorfnih komponentah miRNA utiševalnega mehanizma smo odkrili, da trije geni, povezani s fenotipom KLL, nosijo zapis za proteine z vlogo v utiševalnem mehanizmu miRNA (preglednica 6). Vpletjenost v metabolne procese RNA in jedrno lokalizacijo potrjujejo tudi orodja podatkovne baze DAVID.

Preglednica 6: Geni povezani s KLL, ki imajo funkcijo pri utiševalnem mehanizmu miRNA

Oznaka gena	Ime gena
<i>GEMIN4</i>	gem (nuclear organelle) associated protein 4
<i>CNOT8</i>	CCR4-NOT transcription complex, subunit 8
<i>SMAD3</i>	SMAD family member 3

Z bazo miRecords smo ugotovili, da 97 genov za KLL predstavlja že validirane tar e miRNA (preglednica 7). Od teh je bilo 75 (77%) takšnih, ki so bile tar e molekul miRNA, ki so že bile povezane s KLL. Dejansko število genov, ki jih miRNA regulirajo, je predvidoma še veje, eprav pri 1457 genih (93%) povezanih s KLL še nismo odkrili povezave z miRNA.

Med geni za KLL je 58 genov s spremenjenim izražanjem, 9 genov s SNP-ji, pet genov z epigenetskimi spremembami, dva gena s spremenjeno ekspresijo in SNP-ji ter en gen z ekspresijsko in epigenetsko spremembo ter SNP-jem, ki predstavljajo potrjene tar e za miRNA, ki so bile prav tako predhodno že povezane s KLL (preglednica 7). Vendar pa je bila povezava tar in odgovarjajo ih miRNA pri KLL do sedaj dokazana le za osem genov (*BCL2*, *MCL1*, *TCL1*, *ASXL2*, *CENPJ*, *GTF2H1*, *JUN*, *UGDH*). Tako za preostalih 67 povezav miRNA in genov, neodvisno povezanih s KLL, predpostavljamo, da bi lahko bile pomembne pri KLL. Interakcije z miRNA bi lahko bile motene zaradi polimorfizmov posameznih nukleotidov, odkritih v genih za KLL, ali vzroki ne za spremenjeno ekspresijo in epigenetske spremembe tar nih genov za KLL. Posebej zanimiv je gen *RUNX1*, ki predstavlja gostiteljski gen za miR-802, hkrati pa je tar a molekul miR-17, miR-20a, miR-27a, miR-106a (preglednici 5, 7). Najbolj tipično za validirane tar e miRNA povezane s KLL je, da so njihovi proteinski produkti vpleteni v regulacijo fosforilacije, predstavljajo kinaze in se nahajajo v lumnu jedra oz. drugih organelov.

Preglednica 7: Geni, ki so predhodno že bili povezani s KLL, ter predstavljajo validirane tar e miRNA, za katere smo odkrili, da so prav tako že bile povezane s KLL

Geni, povezani s KLL, ki predstavljajo validirane tar e miRNA	E	Genetska sprememba	miRNA, ki so povezane s KLL, in imajo validirane tar e gene, ki so povezani s KLL
<i>ACPL2</i>	E	x	
<i>ACTA2</i>	E	hsa-miR-21	se nadaljuje

## nadaljevanje

<i>ACVR1B</i>	E	hsa-miR-24	
<i>AHR</i>	E	hsa-miR-124	
<i>ANXA1</i>	E	✗	
<i>APAF1</i>	E	hsa-miR-21	
<i>APC</i>	S	✗	
<i>APP</i>	E	hsa-miR-106a	
<i>ASXL2</i>	E	hsa-miR-15a, hsa-miR-16	
<i>ATXN1</i>	E	hsa-miR-19a, hsa-miR-101, hsa-miR-130a	
<i>BAK1</i>	S	hsa-miR-125b	
<i>BCL2</i>	E, S, EP	hsa-miR-15, hsa-miR-16, hsa-miR-34,	
<i>BCL6</i>	E, S	hsa-miR-127-3p	
<i>BCL7A</i>	E	hsa-let-7b	
<i>BRCA1</i>	E	hsa-miR-146a	
<i>BTG3</i>	E	hsa-miR-124	
<i>C2orf18</i>	E	✗	
<i>CASP3</i>	S	hsa-let-7a	
<i>CASP6</i>	E	hsa-miR-125b	
<i>CBX7</i>	E	hsa-miR-125b	
<i>CCND1</i>	E, EP	hsa-miR-20a	
<i>CCND3</i>	E	hsa-miR-16	
<i>CD44</i>	E	✗	
<i>CD83</i>	E	✗	
<i>CDK4</i>	E	hsa-miR-124	
<i>CDKN1B</i>	E	hsa-miR-181a	
<i>CDKN2A</i>	E, EP	hsa-miR-24	
<i>CENPF</i>	E	✗	
<i>CENPJ</i>	E	hsa-miR-15a, hsa-miR-16	
<i>CLCN3</i>	E	✗	
<i>CLU</i>	E	hsa-miR-125b	
<i>COL1A1</i>	E	hsa-miR-29c	
<i>COL4A2</i>	E	hsa-miR-29c	
<i>CREB1</i>	S	✗	
<i>CXCL12</i>	S	hsa-miR-23a	
<i>CXCR4</i>	E, S	hsa-miR-146a	
<i>CYP1B1</i>	S	hsa-miR-124	
<i>DHFR</i>	E	hsa-miR-192	
<i>E2F1</i>	E	hsa-miR-20a	
<i>ESRI</i>	E	hsa-miR-22	
<i>EZH2</i>	E, S	✗	
<i>FAS</i>	E	hsa-miR-21	
<i>GAS2L1</i>	E	hsa-miR-124	se nadaljuje

## nadaljevanje

<i>GMCL1</i>	E	hsa-miR-124
<i>GSN</i>	E	hsa-miR-124
<i>GTF2H1</i>	E	hsa-miR-15a, hsa-miR-16
<i>HIF1A</i>	E	✗
<i>HMGA1</i>	E	✗
<i>ICOS</i>	S	hsa-miR-101
<i>ID2</i>	E	hsa-miR-125b
<i>ID3</i>	E	hsa-miR-125b
<i>JUN</i>	E	hsa-miR-15a, hsa-miR-16
<i>LDOC1</i>	E	hsa-miR-155
<i>LITAF</i>	E	hsa-miR-124
<i>LMNB1</i>	E	hsa-miR-124
<i>LMO2</i>	E	hsa-miR-223
<i>MCL1</i>	E	hsa-miR-15a, hsa-miR-16, hsa-miR-29b, hsa-miR-101, hsa-miR-193a-3p, hsa-miR-302
<i>MYC</i>	E	hsa-miR-145
<i>OSBPL8</i>	E	hsa-miR-124
<i>P2RX7</i>	S	hsa-miR-150
<i>PAK1</i>	E	✗
<i>PFTK1</i>	E	✗
<i>PHF19</i>	E	hsa-miR-124
<i>PIM1</i>	E	✗
<i>PKM2</i>	E	hsa-miR-133a
<i>PRKCE</i>	E	hsa-miR-205
<i>PRKD2</i>	S	hsa-miR-17
<i>PTBP1</i>	E	hsa-miR-124
<i>PTPN12</i>	E	hsa-miR-124
<i>PTPRO</i>	EP	✗
<i>RAF1</i>	E	✗
<i>RB1</i>	E	hsa-miR-106a
<i>RUNX1</i>	E	hsa-miR-17, hsa-miR-20a, hsa-miR-27a, hsa-miR-106a
<i>SERP1</i>	E	hsa-miR-124
<i>SFPQ</i>	E	hsa-miR-29b, hsa-miR-141
<i>SFRS9</i>	E	✗
<i>SLC7A6</i>	E	hsa-miR-30a
<i>SOCS1</i>	E	hsa-miR-19a
<i>STK4</i>	E	✗
<i>TGFB2</i>	E	hsa-miR-141
<i>TCL1</i>	E, EP	hsa-miR-29b, hsa-miR-181b
<i>TGFBR2</i>	E	hsa-miR-21
<i>TGFBR3</i>	E	hsa-miR-21
<i>THBS1</i>	S	hsa-miR-30a
		se nadaljuje

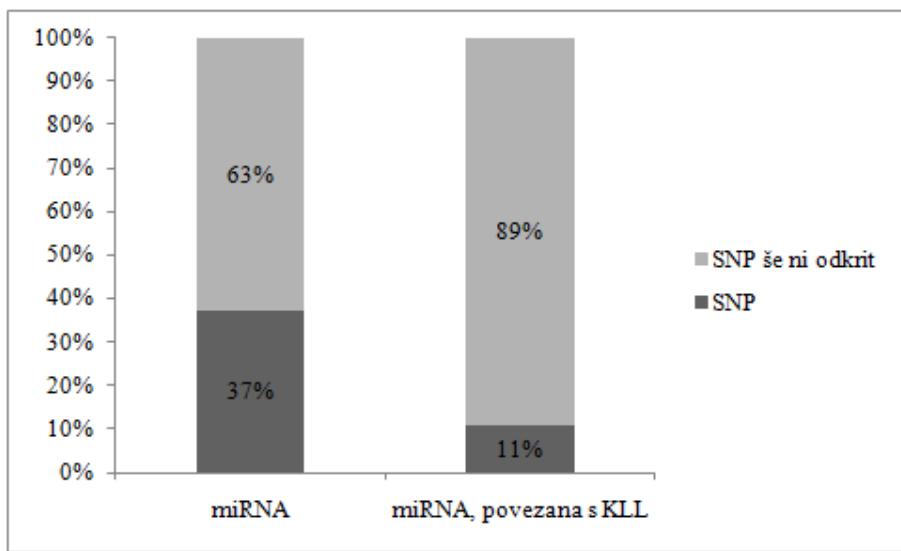
nadaljevanje

<i>TPII</i>	E	hsa-miR-15a, hsa-miR-16
<i>TRIM2</i>	E	×
<i>TRPS1</i>	E	×
<i>TTK</i>	E	×
<i>TUBA1A</i>	E	hsa-miR-30a
<i>TUBB6</i>	E	hsa-miR-124
<i>TWIST2</i>	EP	hsa-miR-124
<i>UGDH</i>	E	hsa-miR-15a, hsa-miR-16
<i>UHMK1</i>	E	hsa-miR-124
<i>VCAM1</i>	S	hsa-miR-126
	E	hsa-miR-15a, hsa-miR-16, hsa-miR-17, hsa-miR-205, hsa-miR-20a, hsa-miR-34a, hsa-miR-106a, hsa-miR-126
<i>VEGFA</i>		
<i>ZBED3</i>	E	hsa-miR-124

E: spremembra ekspresije gena, S: SNP, EP: epigenetska spremembra  
× miRNA, ki ima vezna mesta za dani gen, še ni bila povezana s KLL

V diplomskem delu smo z *in silico* analizo žeeli odkriti kakšen je pomen polimorfizma v miRNA pri KLL, eprav so možni tudi drugi mehanizmi deregulacije miRNA, ki vklju ujejo mutacije, delecije in spremembe ravni izražanja.

Izmed 197 s KLL povezanih miRNA smo identificirali 24 (11%) genov za miRNA, ki so vsebovali vsaj en SNP (priloga B). Pri tem se SNP-ji nahajajo v prekurzorju 18 miRNA, v šestih zrelih oblikah miRNA in od tega se nahaja en SNP v podro ju, odgovornem za vezavo na mRNA (angl. seed region). V primerjavi z vsemi 264 miRNA s SNP izmed 706 trenutno registriranih miRNA, to predstavlja manjšo variabilnost (slika 11). Z vlogo miRNA v širokem spektru razli nih celi nih procesov bi odkriti polimofizmi znali imeti pomembno vlogo pri razvoju bolezni.



Slika 11: Primerjava deleža polimorfnih miRNA glede na 706 miRNA v registru podatkovne zbirke miRBase ter 197 miRNA, za katere smo odkrili povezanost s KLL

Izmed ca. 2000 genov za KLL smo identificirali šest polimorfnih tar nih genov, povezanih s KLL, ki so imeli skupno 20 SNP-jev (preglednica 8), ki bi lahko predstavljali potencialne ozna evalce za KLL. Ali so odgovarjajo e miRNA povezane s KLL bi bilo potrebno preveriti eksperimentalno, saj imajo miR-19a, miR-19b in miRNA družine hsa-let-7 že dokazano vlogo pri KLL (preglednica 8). Polimorfne tar e miRNA, povezane s KLL, imajo glede na funkcionalno analizo vlogo v metabolnih in celi nih procesih.

Preglednica 8 : Polimorfni tar ni geni v KLL, z odgovarjajo imi SNP-ji

Geni povezani s KLL, ki predstavljajo polimorfne tar e za miRNA	SNP	miRNA polimorfnih tar
<i>AKR7A3</i>	rs1065660 rs1061670	/ /
<i>E2F2</i>	rs2811970	hsa-miR-483-3p
<i>HSPB7</i>	rs11552854 rs11556 rs11552855 rs1763600 rs1048238 rs1048237	hsa-miR-532-3p hsa-miR-765 hsa-miR-330-5p hsa-miR-326 hsa-miR-637 hsa-miR-611 hsa-miR-1224-5p hsa-miR-532-3p hsa-miR-214* hsa-miR-632 hsa-miR-346
<i>MTHFR</i>	rs35134728 rs34762557	hsa-miR-363* hsa-miR-363* se nadaljuje

nadaljevanje		
<i>RPL11</i>	rs3737966	hsa-miR-1228*
	rs1061313	miR-1261
	rs2279865	/
<i>TNFRSF1B</i>	rs5746065	hsa-miR-486-3p
	rs5746067	hsa-miR-1308
	rs6668561	hsa-miR-1306 hsa-miR-616
	ENSSNP11488865	(hsa-miR-19b) (hsa-miR-19a)
	rs11807940	hsa-let-7e; hsa-let-7d, hsa-let-7g, hsa-let-7f, hsa-miR-98, hsa-let-7i, hsa-let-7c, hsa-let-7b, hsa-let-7a, hsa-miR-1248
	rs5746075	hsa-miR-1248

/: V podatkovni bazi Patrocles ni podatka

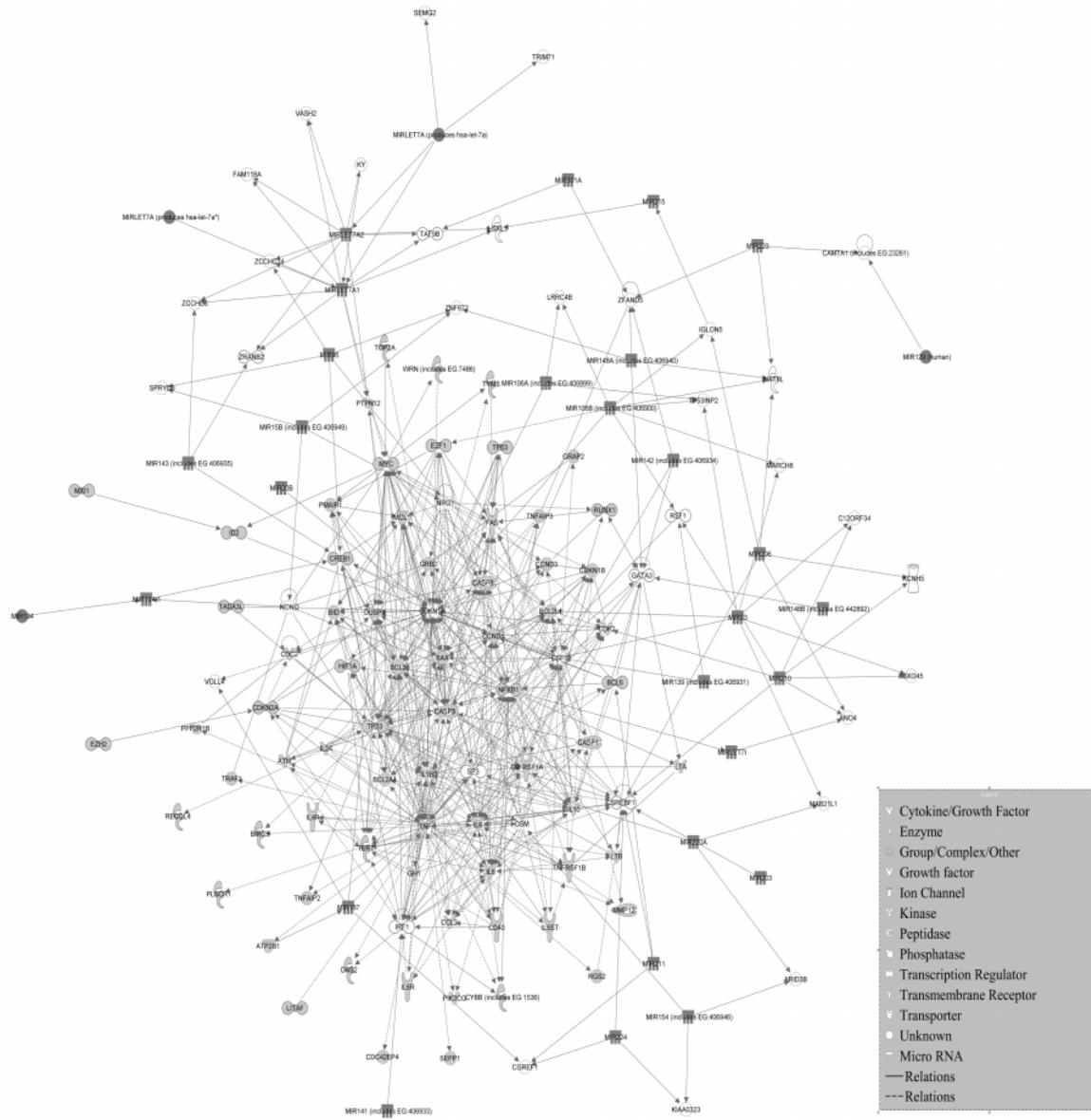
V prilogi C so predstavljene miRNA, ki so povezane s KLL in imajo polimorfne tar e. Njihovo vlogo pri KLL je potrebno preveriti eksperimentalno.

#### 4.5 GENSKE MREŽE Z miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI

Ker nas je zanimalo, kako so prepletene interakcije genov za KLL in miRNA ter smo želeli identificirati povezovalnike kot kandidatne gene, smo na osnovi zbirk genov za KLL in miRNA narisali genske mreže. Za vejo raznolikost pridobljenih informacij in možnost različnih interpretacij smo uporabili testni dostop dveh komercialno dostopnih programov: Ingenuity Pathway Analysis in MetaCore (GeneGo).

V mrežo, narisano s programom Ingenuity Pathway Analysis, ki uporablja algoritem najkrajših poti, je vpletenih 31 miRNA, ki vplivajo na KLL (slika 12). Geni oz. miRNA, ki predstavljajo vozlišča in imajo več interakcij z drugimi geni, so vpleteni v več bioloških procesov ter vplivajo na patologijo tudi drugih tipov raka. To so npr. miRNA družine let-7, miRNA-21, miRNA-146, miRNA-15b in geni *TP53*, *MYC*, *SMAD4*, *STAT3*, *BAX*, *BCL2*, *NFKB1*, *CASP3*, *CASP8*, *FAS*, *MCL*, *BCL6*, *TNF*, *IL6*, *IL8*, *IL10*, *BCL2A*, *IL1B*, *CD40*, *TERT*, *DUSP*... Glavne povezovalnike predstavljajo geni, kot so onkogeni *BCL2*, *TNF* in *MYC* ter tumorski supresorji, kot sta *P53* in *SMAD4*.

Mreže so rezultat vnosa vhodnih podatkov ter informacij iz baz znanj programov, kar omogočajo identifikacijo novih kandidatnih genov, ki so na mrežah neobarvani (preglednica 9, slika 12). To so geni za proteine z vezavnimi domenami cinkov prst, DNA-vezavnimi domenami in vezavnimi domenami za kovine, katione, ione, ki imajo pomembno vlogo pri transkripcijski regulaciji.



Slika 12: Genska mreža pri KLL z integriranimi miRNA, izrisana s programom Ingenuity Pathway Analysis

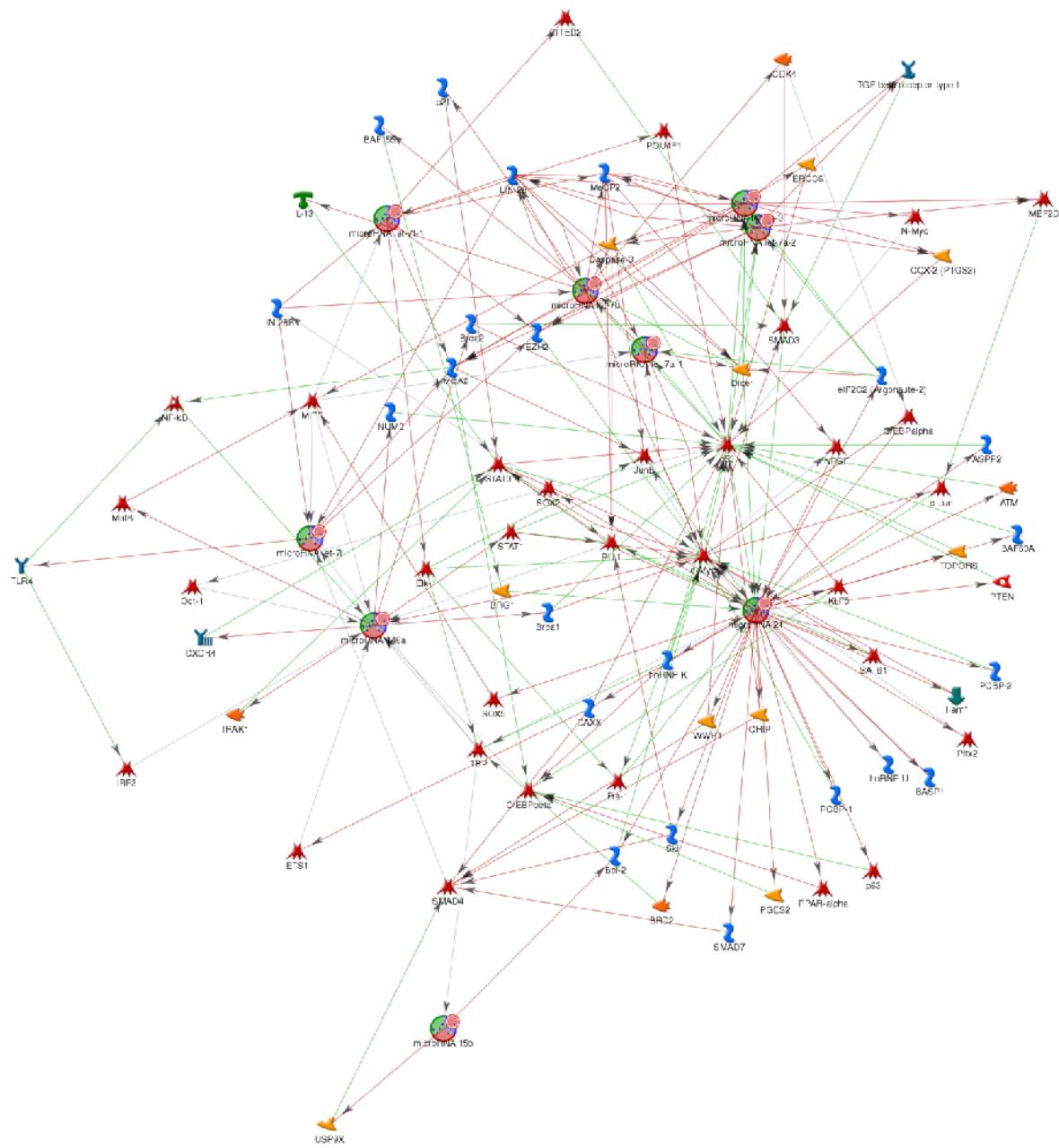
Preglednica 9: Novih 37 kandidatnih genov glede na izrisano gensko mrežo z miRNA s programom IPA

---

<i>ANO4</i>	anoctamin 4
<i>ARID3B</i>	AT rich interactive domain 3B (BRIGHT-like)
<i>C12ORF34</i>	chromosome 12 open reading frame 34
<i>CAMTA1</i>	calmodulin binding transcription activator 1
<i>CDKN1A</i>	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
<i>CGREF1</i>	cell growth regulator with EF-hand domain 1
<i>CSF</i>	colony stimulating factor 1 (macrophage)
<i>FAM118A</i>	family with sequence similarity 118, member A
<i>FBXO45</i>	F-box protein 45
<i>GATA3</i>	GATA binding protein 3
<i>IGLON5</i>	IgLON family member 5
<i>IRF1</i>	interferon regulatory factor 1
<i>KCNH5</i>	potassium voltage-gated channel, subfamily H (eag-related), member 5
<i>KIAA0323</i>	KIAA0323
<i>KY</i>	kyphoscoliosis peptidase
<i>LRRC4B</i>	leucine rich repeat containing 4B
<i>MAB21L1</i>	mab-21-like 1 ( <i>C. elegans</i> )
<i>MARCH8</i>	membrane-associated ring finger (C3HC4) 8
<i>NAT8L</i>	N-acetyltransferase 8-like (GCN5-related, putative)
<i>NONO</i>	non-POU domain containing, octamer-binding
<i>NRG1</i>	neuregulin 1
<i>OSM</i>	oncostatin M
<i>RSF1</i>	remodeling and spacing factor 1
<i>SEMG2</i>	semenogelin II
<i>SP3</i>	Sp3 transcription factor
<i>SPRYD3</i>	SPRY domain containing 3
<i>TAE9B</i>	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor
<i>TP53INP2</i>	tumor protein p53 inducible nuclear protein 2
<i>TRIM71</i>	tripartite motif-containing 71
<i>VASH2</i>	tripartite motif-containing 71
<i>VGLL4</i>	vestigial like 4 ( <i>Drosophila</i> )
<i>ZCCHC14</i>	zinc finger, CCHC domain containing 14
<i>ZCCHC3</i>	zinc finger, CCHC domain containing 3
<i>ZFAND5</i>	zinc finger, AN1-type domain 5
<i>ZNF672</i>	zinc finger protein 672
<i>ZRANB2</i>	zinc finger, RAN-binding domain containing 2

---

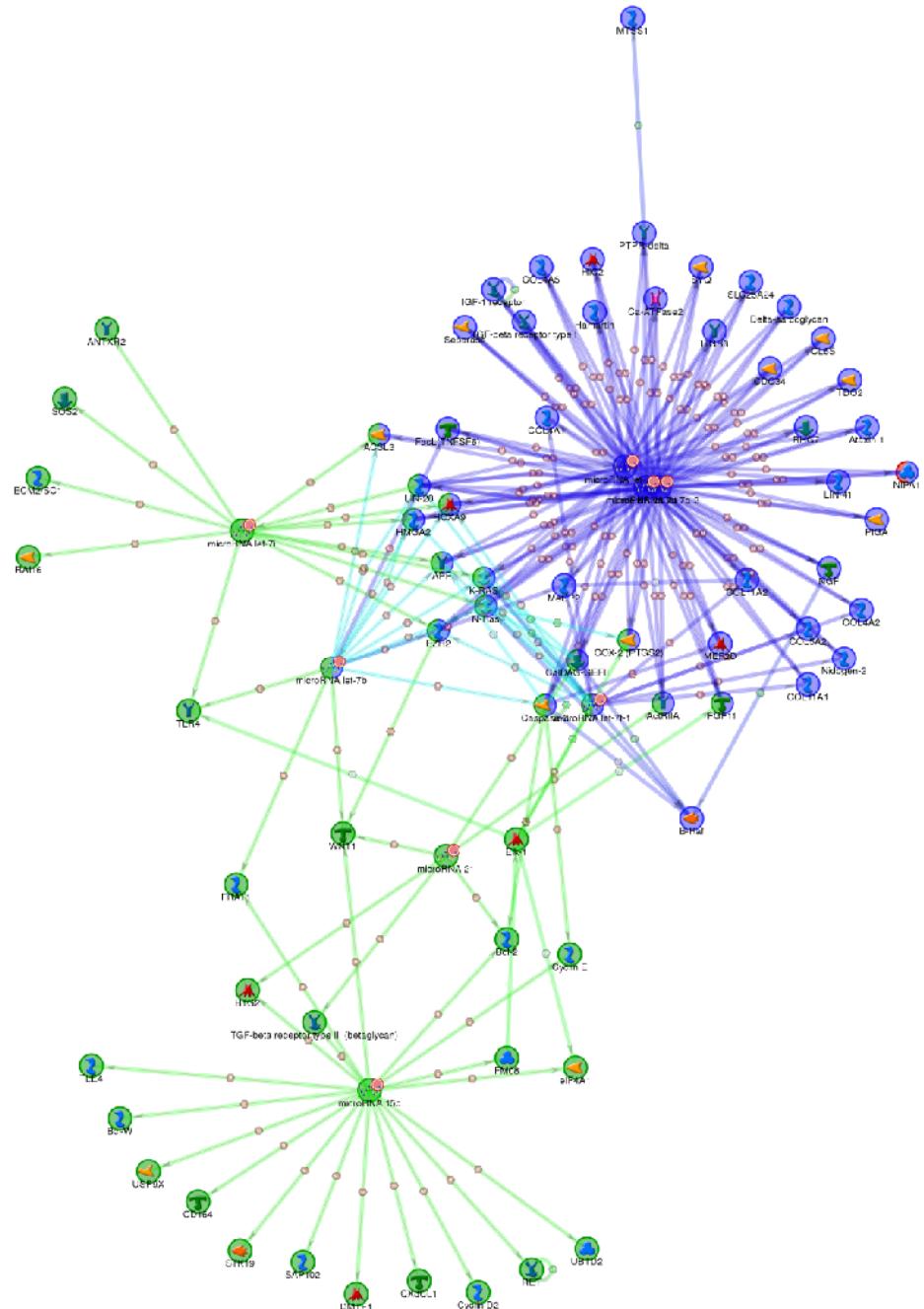
Mrežo na sliki 13 smo vizualizirali s funkcijo najkrajših poti (Dijkstrov algoritmom) s programom Metacore ter orodjem Metalink. Na sliki 14 pa je izrisana mreža, ki je rezultat združevanja podmrež.



Slika 13: Genska mreža pri KLL z integriranimi miRNA, izrisana z algoritmom »Shortest path« v programu MetaCore, z uporabo orodja Metalink

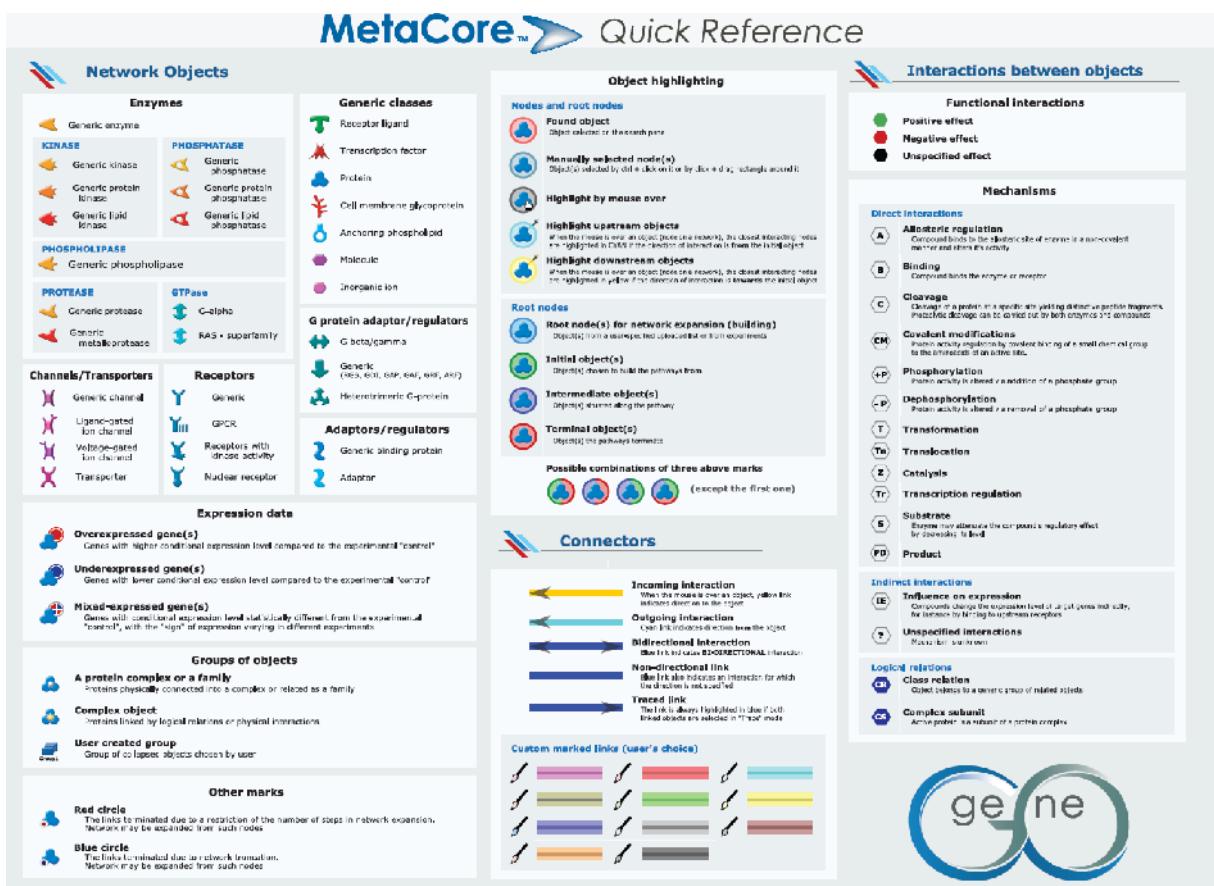
V genskih mrežah lahko poleg zanimivih komponent prepoznamo nekatere regulatorne mehanizme. Na vseh treh mrežah so narisane miRNA družine let-7, ki imajo na splošno pogosto znižano izražanje pri raku. Na obeh mrežah, narisanih s programom MetaCore (slike 13, 14, 15) se obakrat pojavi miR-21, ki je prav tako prekomerno izražena pri veinih tipov raka in predstavlja znano oncomiR. Iz slike 13 je razvidno, da ta miRNA vpliva tudi na EIF2C2, lana družine Argonavtov, miR-7a-1 pa na Dicer. Opazili smo, da imajo

nekatere miRNA homologne tar e, npr. miR-146a se veže na *BRCA1* in *BRCA2*, miRNA let-7a-1 in let-7a-2 pa na kolagene *COL4A2*, *COL5A2*, *COL11A1*, *COL11A2*. Prav tako drži obratno, mnoge tar e so regulirane z molekulami miRNA iste družine. Družina let-7 ima na primer veliko skupnih tar . Nadalje lahko opazimo poleg interakcij miRNA s transkripcijskimi faktorji tudi interakcije z vezavnimi proteini, npr. poleg Dicerja in Argonavta EIF2C2 vplivajo na miRNA RNA-vezavni proteini LIN28 in LIN28B.



Slika 14: Genska mreža pri KLL, izrisana z združevanjem podmrež s programom MetaCore ter orodjem MetaLink

Mreža na sliki 14 kaže povezanost različnih podmrež miRNA in nakazuje, da so te prepletene preko nekaterih znanih onkogenov, kot sta *NRAS* in *KRAS* in proteinov pomembnih za apoptozo kot so kaspaze in Fas ligand. V mrežah so pogoste interakcije miRNA s transkripcijskimi faktorji ter geni, ki imajo pomembno vlogo pri kromatinski strukturi, npr. *HOXA9*, *HMGA1*, *EZH2*, *BRCA2*, *MeCP2*, *SMAD4*, *TOP2A*, *MEF2*, *STAT1*, *SOX2*, *SOX5*, *HIC2*...



Slika 15: Legenda k sliki 13 in sliki 14

## 5 RAZPRAVA IN SKLEPI

Diplomsko delo je z izdelano podatkovno zbirkovo, uporabo bioinformacijskih orodij in rekonstrukcijo genskih mrež ponudilo osnovo za sklepanja o uporabljenih pristopih in o možnih vzrokih ter posledicah motenih interakcij miRNA pri KLL.

### 5.1 UPORABLJENI INTEGRATIVNI PRISTOPI

Sprva smo kot model za študij integrativnih genomskega pristopov pri iskanju vzročnih genov in njihovih povezav v genskih mrežah z miRNA nameravali izbrati rak dojke. Nato pa smo tekom raziskovalnega procesa ugotovili, da nam na področju tako intenzivnih raziskav raka, ki je svetovno gledano druga najbolj pogosta vrsta raka, v mejah diplomskega dela ne bi uspelo zbrati vseh znanih in predvidenih genetskih vzrokov ter povezav z miRNA. Zato smo se odločili, da za prvi model izberemo bolezen z bolje dokazanimi povezavami z miRNA in na ta način za enostavnejši študirati kroni no limfocitno levkemijo v sodelovanju z vodilnim raziskovalcem na področju miRNA in levkemij, prof. Calinom iz MD Anderson Centra (Houston, Texas). Vendar pa smo kasneje ugotovili, da je bilo pri levkemijah, limfomih in sarkomih odkritih mnogo več genov za raka kot pa pri drugih tipih rakov, eprav naštete maligne transformacije predstavljajo le 10% vseh vrst raka pri loveku (Futreal in sod., 2005).

#### 5.1.1 Izdelava podatkovne zbirke genskih lokusov za kroni no limfocitno levkemijo

V študiji smo manualno pregledali 253 lankov, zbrali informacije o ca. 2000 lokusih za KLL in na ta način odkrili 241 možno nejših kandidatnih genov za KLL. Uporabili smo integrativni pristop, združili informacije o živalskih modelih, citogenetskih spremembah, mutacijah in polimorfizmih posameznih nukleotidov, genskem izražanju, epigenetiki in miRNA mehanizmih, ter predstavili v stopenjski proces do identifikacije pomembnejših vzročnih lokusov danega fenotipa. Odkrite lokuse smo združili v podatkovno zbirko, ki je prosto dostopna na spletni strani <http://www.bf.uni-lj.si/fileadmin/groups/2726/ChronicLymphocyticLeukemia.pdf>. V izdelani podatkovni zbirki so dostopne informacije o ca. 2000 lokusih povezanih s KLL, ki so zaradi nadaljnjih

analiz in boljšega pregleda razdeljene po različnih študijskih pristopih. To omogoča lažjo oceno zanesljivosti vpletjenosti danega lokusa. Na primer, medtem ko je zanesljivost ekspresijskih in asociacijskih študij relativno nizka, pa so citogenetske študije kljub nižji resoluciji definirane regije, bolj zanesljive, saj je lažje odkriti povezavo med fenotipom pacienta in citogenetsko spremembo kot pa posameznim genom, poleg tega da so v kliniki že daljša uveljavljene.

KLL je heterogena bolezen in tako obstaja veliko različnih študij, katerih namen je bil odkriti molekularne mehanizme in razvoj označevalcev te bolezni, vendar so študije posameznih genov pogosto pristranske, saj preverjajo v naprej postavljeno hipotezo. Sistemska biologija ima enake cilje kot klasična molekularna biologija, vendar pa upošteva dejstvo, da je biologija integriran sistem in ne zbirka posameznih gradbenih elementov. Takšni pogledi so bili najprej implementirani v raziskavah genomskega izražanja z mikromrežami in s proteomskim profiliranjem. Tako je danes ta tip informacij že dostopen preko podatkovnih zbirk, kot so GEO (Gene Expression Omnibus), BIND (Biomolecular Interaction Network Database), Oncomine, IntAct, DIP (Database of Interacting Proteins), MINT (Molecular Interactions Database). Posledi no vključuje večina raziskav genskih in proteinskih mrež izključno ta tip informacij, ki pa lahko vodi do napaknih zaključkov o biologiji bolezni. Od prvih transkriptomskih analiz pri KLL v letu 2001 (Klein in sod., 2001; Rosenwald in sod., 2001), so se genomske raziskave KLL razmahnile, ustvarile veliko kopico podatkov, a s tem zaradi tehničnih in bioloških razlik tudi veliko šuma v informacijah. Zato je prikazan pristop integracije informacij različnih komplementarnih študij pomemben korak naprej, ki poveča možnost detekcije zanesljivih in relevantnih informacij ter zasnova novih eksperimentov. Repozitorij genov, povezanih s KLL, je prva podatkovna zbirka, ki zajema vse poznane genetske podatke o KLL vseh študijskih pristopov. Do sedaj so namreč bile narejene integracije le nekaj študijskih pristopov ali pa so se raziskave osredotočile na določen manjši del kromosoma. Zato bo repozitorij omogočil validacijo genov za nove specifične diagnostike, prognostične ali terapevtske markerje oz. set markerjev, saj je uporaba več označevalcev bolj zanesljiva. Zbirka bo omogočila raziskovalcem na področju KLL pridobiti širok pregled podatkov, ki so trenutno dostopni.

Ob zbiranju genov v podatkovno zbirko pa smo bili soo eni ne samo z veliko koli ino podatkov, temve tudi veliko raznolikostjo informacij, ki smo jih želeli združiti v smiselno celoto. Zbrali smo informacije razli nih raziskav, pri emer je bila signifikantnost rezultatov razli na, prekrivanje odkritih genov pa relativno nizko. To je lahko posledica razli nih izvorov in istosti celic B, ki se tudi slabo delijo in odzivajo na mitogene, premajhnega vzorca pacientov, oteženega pridobivanja vzorcev bolnikov z boleznijo, ki se pojavi v kasnejši fazi življenja, razlik v eksperimentalnih postopkih, preu evanih klini nih parametrih, specifi nih vprašanj, ki so si jih raziskovalci v študijah zastavili itd. Pri pregledovanju lankov smo imeli opravka z rezultati študij, ki so jim avtorji pripisovali razli en pomen. Kljub dolo enim zahtevam dokazovanja hipotez ter standardom pisanja znanstvenih lankov, pa ti še vedno dopuš ajo prostor za izražanje lastnih idej avtorjev. Pregledovanje utemeljenosti subjektivnih interpretacij pa zahteva za izdelavo podatkovne zbirke veliko truda. Pogosto nam je problem predstavljal ravno nasprotno – pomanjkanje interpretacije rezultatov študij in le navajanje suhoparnih podatkov, pri emer je bilo bralcu preprič eno, da si rezultate razloži po svoje. Pri iskanju vzro nih genov za KLL smo imeli tudi popolnoma nevsebinske ovire. Izpolnitev ambiciozne želje, da zberemo vse genetske povezave s KLL, je ovirala omejena dostopnost znanstvenih lankov. V nekaterih primerih smo iskali informacije iz dopolnilnega materiala lankov, ki naj bi bile dostopne na ustreznih spletnih straneh revij, vendar pa se nam je zgodilo, da citirani podatki niso bili dostopni (tak primer je lanek Stamatopoulos in sod., 2009a). Podoben banalen primer je slaba lo ljestvica slik topotnih grafov (angl heatmap), iz katerih v nekaterih primerih ni bilo možno razbrati imen genov s spremenjeno ravnjo izražanja (primer lanka - Joshi in sod., 2007). Nadaljnji izziv nam je predstavljal uporaba razli nih oznak genov. Tako je zaenkrat v izdelani podatkovni zbirki še vedno možno, da se isti gen ponavlja z razli nimi imeni. To pomanjkljivost nameravamo v podatkovni zbirki v prihodnje odpraviti. Problem dodajanja novih imen za obstoje e gene je bil mo no prisoten v preteklosti. Trenutna strategija ve ine revij je, da morajo avtorji imena genov uskladiti s potrjenimi imeni na internetni strani genenames.org ter uporabljati simbole genov odobrene s strani komiteja za nomenklaturo genov HUGO (Human Genome Organisation). Uporaba uradnih imen genov nam bo omogo ila, da jih združujemo v genske družine in tako lažje prepoznamo pomenljive vzorce v podatkovnih zbirkah in skupne funkcije družin genov.

Zbiranje informacij je potekalo ro no, saj ima zaenkrat avtomatizirana ekstrakcija vzro nih genov iz literature še veliko pomanjkljivosti. Kljub temu je množi na uporaba opisanega pristopa malo verjetna, saj je z eksponentnim generiranjem novih informacij ro no zbiranje podatkov naporno, asovno zahtevno in dopuš a možnost napak ter nenamernega izpušanja vzro nih genov. Poleg tega pa zaradi pospešenih raziskav takšna podatkovna zbirka, e ni redno vzdrževana, hitro zastara. Podatkovno zbirko bi morali vklju iti v strežnik, ki bi omogo al lažje upravljanje s podatki in redno posodabljanje, takoj ko bi bile objavljene nove informacije o vzro nih genih.

### 5.1.2 Izbor mo nejših kandidatnih genov

Ker ima vsak od sedmih študijskih pristopov svoje prednosti in slabosti, smo mo nejše kandidatne gene izbrali z integracijo ve pristopov, ki skupaj omogo ajo ve jo verjetnost vzro nosti gena. Pri izboru pomembnejših kandidatnih genov smo glede na kriterije, navedene v metodah združili informacije asociacijskih, mutacijskih, ekspresijskih, epigenetskih študij in živalskih modelov.

Rezultati funkcijskih analiz potrjujejo, da smo z opisanim pristopom združili gene, ki so relevantni za KLL (slika 9). Prikazan integrativni pristop je torej možno uporabiti pri analizi drugih poligenskih bolezni in fenotipov.

Prikazan izbor mo nejših kandidatov pa kljub temu še vedno ni popolnoma objektiven in je primernost uporabljenih kriterijev lahko predmet razprave. Obstaja namre možnost, da preu evanje dolo enega gena v razli nih študijah ni neodvisno, temve je posledica pristranskega osredoto anja raziskav na dani gen.

V diplomskem delu smo združili samo pet tipov informacij, saj analiza dodatnih informacij drugih študijskih pristopov iz podatkovne zbirke, kot so kromosomske aberacije in lokusi genske povezanosti, v genskih mrežah ali z orodji anotacije genov, zaenkrat ni možna. V prihodnosti pa nameravamo vklju iti tudi te podatke. Poleg uporabe prese nih množic in risanja genskih mrež na rtujemo uporabo tretjega na ina iskanja mo nejših kandidatnih genov. Narisali bomo gensko karto, ki bi omogo ila vizualizacijo prekrivanja citogenetskih

sprememb z drugimi spremembami na ravni genoma ter iskanje novih genov kandidatov, podobno kot je bilo to narejeno na primeru genske mape za mle no žlezo in mastitis (Ogorevc in sod.; 2009). Vendar pa nameravamo za vizualizacijo uporabiti avtomatiziran pristop, npr. z uporabo orodij zbirke programov GMOD (Generic Model Organism Database project).

### **5.1.3 Funkcijska analiza genov**

Na osnovi 241 izbranih genskih povezav s KLL, smo funkcije mo nejših kandidatnih genov preverili z bioinformacijskimi orodji za anotacijo (DAVID). Ta prosto dostopna podatkovna zbirka z eksponentno rasto o citiranostjo je enostavna za uporabo in omogo a hitro analizo ve jih zbirk genov.

Vendar je analiza funkcij genov z orodjem DAVID koli insko omejena. Tako za analizo bioloških poti nismo uporabili vseh genov iz podatkovne zbirke, temve le izbor mo nejših kandidatov. Slabost pri tem je, da na ta na in izpustimo mnoge paralogne gene iste genske družine, ki bi lahko imeli podobne funkcije in tako ne prepoznamo skupin genov, ki bi v bolezni lahko igrali skupno vlogo. Prav tako bi bilo uporabno, e bi s tem orodjem lahko analizirali tudi gene za miRNA, vendar pa je to težavno že zaradi do sedaj še vedno pomanjkljivega znanja o funkcijah miRNA. To pomanjkanje smo poskusili omiliti z orodjem miRNAPath, ki lahko poiš e glavne biološke poti, tar ne gene ali miRNA povezane z vhodnim seznamom genov oz. miRNA.

### **5.1.4 Bioinformacijske analize miRNA**

S presekom izdelane podatkovne zbirke ter spletnih podatkovnih zbirk miRBase, miRecords in Patrocles smo med geni za KLL odkrili eksperimentalno potrjene tar e miRNA, gene, ki nosijo zapis za miRNA ter gene, katerih proteinski produkti so vpleteni v utiševalne mehanizme.

S pristopi prese nih množic informacij razli nih podatkovnih zbirk smo odkrili veliko novih domnevnih povezav z miRNA ter genetsko variabilnosti na ve ravneh miRNA regulacije, ki bodo omogo ile nove hipoteze in zasnovo eksperimentov. Identificirali smo 15 miRNA, za katere na podlagi njihove lokacije v genih za KLL, predpostavljamo vpletenost v KLL. Med geni za KLL smo odkrili tri polimorfne komponente z vlogo v utiševalnem mehanizmu. Za 67 povezav tar nih genov in miRNA, obojih neodvisno povezanih s KLL, predlagamo, da se eksperimentalno preveri signifikantnost teh interakcij pri KLL. Za miRNA 22 drugih tar nih genov za KLL pa predlagamo, da se prav tako natna no preu i njihova funkcija v KLL. Nadalje smo odkrili 24 polimorfnih genov za miRNA (18 s SNP v prekursorju, šest s SNP v zreli obliki miRNA in od tega en SNP v podro ju, odgovornem za vezavo na mRNA (angl. seed region)) in šest polimorfnih tar nih genov, povezanih s KLL s skupno 20 SNP-ji, ki predstavljajo potencialne ozna evalce za KLL.

Manualno pregledovanje prekrivajo ih se podatkov je zamudno, zato bi bila dobrodošla baza, ki bi omogo ala integrirano uporabo vseh podatkovnih zbirk in programov o regulaciji genskega izražanja z miRNA.

### 5.1.5 Genske mreže

Razumljivo je, da zaradi omejitev eksperimentalnih postopkov, tudi nadaljnje sklepanje o vzrokih KLL lahko le deloma nadgradi dosedanje pojasnitve. Vendar pa smo v diplomskem delu združili znanje, vsebovano v podatkovni zbirki s funkcionalno anotacijo genov (gensko ontologijo) ter znanjem o interakcijah tar nih genov z miRNA ter tako prispevali k odkrivanju strukturne in funkcijalne organizacije genov znotraj genskih mrež, ki pri KLL še ni dobro poznana. Genska mreža prikazuje stati ne in kvalitativne odnose, na osnovi katerih lahko sklepamo o strukturi sistema. Pokazali smo vlogo 37 novih vozliš v genskih mrežah ter potrdili že poznano vpletenost nekaterih genov. Za glavna vozliš a v genskih mrežah so se izkazali klju ni regulatorji signalnih kaskad, ki so bili odkriti tudi z analizo funkcijalne anotacije bioloških poti. V novih vozliš ih pa se pogosto nahajajo geni za proteine s specifi nimi DNA-vezavnimi domenami (npr. domene cinkov prst).

Genske mreže smo narisali samo na osnovi seznama genov, numeri nih parametrov in meritev študij pa nismo upoštevali. Tako je bil večinski delež genov, ki smo jih vnesli v program za risanje mrež, produkt ekspresijskih študij, ki pa so zaradi tehničnih omejitev nagnjene k odkrivanju genov, ki imajo močno spremenjeno ekspresijo. Možno je, da genov z manjšimi, a vseeno pomembnimi spremembami izražanja na mrežah ni. Poleg tega iz genske ekspresije ni vedno možno odkriti vzroke nih genskih faktorjev, ki bi predisponirali k razvoju obolenja. Na sploh pa ne moremo sklepati o kvantitativnih prispevkih posameznih genov, saj smo vnesli le seznam genov. Prav tako se moramo zavedati, da so različne interakcije genov aktivne pod različnimi specifičnimi pogojimi in druga ne v različnih celih nih kompartimentih ter avsonno specifične. Pri KLL bi bilo npr. še posebej zanimivo opazovati razlike med genske interakcije tokom celih nega cikla. Prav tako je zaželeno, da orodja za izris genskih mrež z miRNA v podatkovnih bazah vključujejo znanja o tkivni in avsonni specifičnosti izražanja genov kot tudi miRNA in to upoštevajo glede na tip tkiva ali vrsto celic, katerih gensko mrežo preučimo. V prihodnosti bi bilo zaželeno, da bi bila prisotnost podatkov različnih študijskih pristopov v mreži sorazmerna njihovi statistiki in značilnosti, a za to bo verjetno potreben globalen dogovor o ponderiranih merilih za vključevanje podatkov in standardizacija eksperimentalnih ter analitskih postopkov.

Za izris mrež smo uporabili dva programa, kakršne ponujajo mnoga podjetja. Ti programi združujejo orodja za risanje mrež z informacijami iz vseh podatkovnih zbirk interakcij, ki jih manualno pregledujejo in redno dopolnjejo z objavljenimi eksperimentalnimi podatki o interakcijah. Takšni programi so Ingenuity Pathway Analysis, MetaCore, Pathway Analysis, Pathway Studio itd. Nepraktično pri tem je, da gre za komercialne programe. Testni dostopi pa so v nekaterih primerih možni krajevi, saj kot bi ga uporabnik potreboval, da bi lahko preizkusil vse funkcije programa. Pomanjkljivosti podatkovnih baz teh programov je, da še vedno vsebujejo pretežni delež interakcij protein-protein, informacije o miRNA interakcijah pa so zaenkrat pomanjkljive. Nadaljnja komplikacija je, da so interakcije med geni močno odvisne od pogojev, v katerih se celice nahajajo ter tako prikazani robovi med geni morda niso vzrok ničesar. Možno je, da program kljub vneseni miRNA, te na mrežo ne nariše. Le eden od takšnih primerov je transkripcijski faktor

*RUNX1*, ki je bil povezan s KLL neodvisno v dveh študijah, z *in silico* analizo pa smo odkrili, da je validirana tudi štiri miRNA, poleg tega, da se v tem genu nahaja tudi zapis za miR-802. Vendar pa iz mrež (slika 12) nobena od teh informacij ni razvidna. miR-17, miR-27a in miR-20a, ki imajo validirana vezavna mesta v tem genu, na mreži sploh niso prisotne. miR-106a je narisana, vendar pa njene interakcije z *RUNX1* ni, eprav je bila ta interakcija eksperimentalno potrjena (Fontana in sod., 2007). Zanimivo, te interakcije ne nariše niti program Metacore, eprav smo v tem programu s funkcijo Metalink vnesli direktne podatke o interakcijah genov z miRNA. Prav tako zaenkrat v mreži ni možno označiti, v katerih genih se nahajajo zapisi za gene miRNA ali v program vnesti gene za nekatere druge nekodirajoče RNA, npr. T-UCR. Na tej toki se postavlja tudi vprašanje, ali so geni z veliko interakcijami objektivno predstavljeni kot povezovalniki, ali pa je morda možen vpliv večjega števila podatkov o interakcijah genov, za katere že vemo, da igrajo ključno vlogo in predstavljajo fokus raziskav; ter obratno, ali imajo geni z majhnim številom interakcij teh resni no malo, ali pa jih še nismo odkrili? Za reševanje takšnih težav so potrebni ustrezni algoritmi, ki bi povezovalnikom z veliko interakcijami, a majhnim številom povezav s kandidatnimi geni, pripisali ustrezne uteži. Pomanjkljivosti izrisanih genskih mrež so tako posledica nepopolnih informacij o post-transkripcijski regulaciji v podatkovnih zbirkah programov za risanje genskih mrež, informacijskih in matematičnih izzikov modeliranja kot tudi pomanjkljivega znanja povprečnega biologa o ustreznih algoritmih za vizualizacijo želene mreže. Mreže, predstavljene v diplomskem delu predstavljajo le prve korake pri rekonstrukciji pomenljivih genskih mrež.

Potencial genskih mrež je velik, vendar še ne popolno izkoristi en. Ker je znanje o post-transkripcijski regulaciji še pomanjkljivo, bi morda veljalo le-to združiti z obstoječim znanjem o interakcijah proteinov, metabolitov in bioloških poteh, ki je že dostopno, a fragmentirano v različnih podatkovnih zbirkah (KEGG, BIND...). V prihodnosti pa bi vključili še informacije o post-translacijskih modifikacijah ter drugih mehanizmih, ki lahko vodijo do raka. Zanimiva, a zaenkrat še nemožna, bi bilo primerjava biološke mreže genotipa bolezni in zdravega genotipa, kot tudi primerjati mreže med različnimi živalskimi vrstami ter celimi nimi tipi, kar bi omogočilo odkritje interakcij, ki so evolucijsko ohranjene ali specifične za dan fenotip.

## 5.2 INTERPRETACIJA POVEZAV miRNA PRI KRONI NI LIMFOCITNI LEVKEMIJI

S funkcionalno analizo smo potrdili, da so v izboru kandidatnih genov res vzročni geni za rak imunskega sistema, kjer pride do motenega razvoja in delitev celic. Najpomembnejše funkcionalne kategorije proteinov so sekvenčne variante, mestna mutageneze in pa proteini z jedrnim lokalacijskim signalom. Na podlagi tega lahko sklepamo, da so za KLL pomembni procesi regulacije genskega izražanja na ravni DNA in RNA. Alternativno vključevanje eksonov je še posebej pogosto pri genih za KLL, ki nosijo zapis za gene miRNA. Možno je, da so za patologijo KLL pomembne napake istoasnega procesiranja mRNA in miRNA, v primeru da so te regulirane z istim promotorjem. Večina kandidatnih genov se nahaja na lokusih 17p in 11q, kar ustreza tudi ugotovljenim delečijam. Ker vzročni geni na regiji 13q ne poznamo, genov za miRNA pa s tem orodjem ne moremo analizirati, te regije v rezultatu analize ni bilo.

Z orodjem miRNAPath smo ugotovili, da imajo največ tarčnih genov izmed izbranih kandidatnih genov molekule miRNA družine let-7 in miR-181c. Prve predstavljajo znane antitumorske molekule, slednje miRNA pa se preferentno izražajo v celicah B, kar je v kontekstu KLL smiselno. Z vnosom miRNA, ki so povezane s KLL, smo odkrili, da imajo veliko predvidenih tarčnih mest tudi miRNA družine miR-30 in miR-15. Za miR-15a je že dokazano, da v KLL cilja gene, ki regulirajo apoptozo, (npr. *BCL2*, *MCL1*, *VEGFA*), gene, ki regulirajo gensko ekspresijo, (npr. *JUN* in *ASXL2*), gene, ki vplivajo na celične delitve (npr. *CENPJ*), ter gene za popravljalne mehanizme poškodb DNA (npr. *GTF2H1*). Patološke spremembe vseh tarčnih miRNA pa bodo pri KLL verjetno razjasnjene v prihodnosti.

Med geni, ki so bili predhodno povezani s KLL, smo odkrili gene za proteine, ki sodelujejo v utiševalnem mehanizmu miRNA, gostiteljske gene za miRNA in tarčne miRNA. Med slednjimi so bili tudi tarčni geni z vezavnimi mestami za molekule miRNA, ki so bile predhodno že povezane s KLL (preglednica 7), vendar pa domnevna povezava med danimi tarčnimi in odgovarjajočimi miRNA pri KLL do sedaj še ni bila preučena. Gostiteljski geni za miRNA imajo tipične spojiteljne (angl. splicing) variant, zato bi bilo potrebno

raziskati ali alternativno vključevanje eksonov teh genov vpliva na biogenezo miRNA oz. ali obstaja obratna povezava. Poleg tega predlagamo, da se vlogo miRNA, ki imajo zapise znotraj genov za KLL ter domnevno povezavo med miRNA z ugotovljeno vlogo pri KLL in geni za KLL preveri eksperimentalno.

Opazili smo polimorfizem na vseh treh ravneh z miRNA posredovane regulacije. Med miRNA za KLL je bil delež polimorfnih molekul manjši kot delež polimorfnih miRNA med vsemi do sedaj registriranimi miRNA. To namiguje, da so v KLL vpletene miRNA s pomembnejšimi funkcijami, ki so torej podvržene selekcijskemu pritisku. Zato ima polimorfizem, ki ga odkrijemo pri teh molekulah, bolj verjetno vpliv na fenotip. Sklepamo, da podobno drži za tarene gene in gene za utiševalne komplekse. Polimorfizmi v teh genih bi bili lahko vzroki za spremenjeno ekspresijo in epigenetske spremembe genov za KLL.

Z risanjem genskih mrež smo odkrili, da so v KLL pomembne interakcije genov, ki imajo vlogo tudi v drugih tipih rakov, v razvoju ali hematopoezi. Poleg tega smo kot povezovalnike v genskih mrežah identificirali nove kandidatne gene za KLL, ki v večji meri predstavljajo regulatorne faktorje. Eprav sprva nepričakovano, so mreže nasprotno bogate z interakcijami med miRNA in transkripcijskimi faktorji ter geni, ki vplivajo na strukturo kromatina. Ta opažanja se skladajo z rezultati analiz genske ontologije. Interakcije so dvostranske; miRNA aktivno regulirajo izražanje teh genov ter so tudi same regulirane s strani transkripcijskih faktorjev in epigenetskih sprememb. Vse to nas je spodbudilo k razmišljanju, da ima sodelovanje nekodirajočih RNA in epigenetskih mehanizmov verjetno vpliv na genomsko integriteto, ter morda na delecije, ki so pri KLL bolj pogoste. Še vedno, z vizualizacijo miRNA v genskih mrežah smo odkrili, da funkcije miRNA niso regulirane samo na ravni transkripcije, temveč imajo povezave z miRNA tudi vezavni proteini, kar pomeni, da gre za post-transkripcijsko uravnavanje miRNA.

Odkrili smo, da imajo nekatere miRNA homologne tare, npr. miR-146a se veže na *BRCA1* in *BRCA2*, miRNA let-7a-1 in let-7a-2 pa na kolagene *COL4A2*, *COL5A2*, *COL11A1*, *COL11A2*. Tudi obratno, mnoge tare so regulirane z molekulami miRNA iste družine. To nakazuje, da bi lahko bilo v patologijo KLL morda vpletene tare miRNA dane družine. Zanimivo bi bilo vedeti, ali je podobnost posledica konvergentne ali divergentne

evolucije. Za podobne gene miRNA, ki ležijo blizu skupaj, zaenkrat velja, da so nastali z duplikacijo (Zhang in sod., 2009), vendar pa bi bila pri homolognih miRNA, ki so locirane na različnih kromosomih zaradi kratkega področja za vezavo na mRNA (angl. seed region) možna tudi konvergentna evolucija.

Lu in sod. (2007) so pokazali, da imajo povezovalniki oz. vozlišča z večjim številom interakcij manjšo spremembo genskega izražanja kot pa vozlišča z manj povezavami. To se ujema tudi z odkritji našega dela, saj v centru mrež ni bilo genov, ki so bili pri KLL povezani z najbolj spremenjeno ekspresijo (npr. ZAP70). Tako centralna vozlišča ne predstavljajo markerjev, specifičnih za KLL, lahko pa predstavljajo nove terapevtske tarče. Geni z večjo spremembo izražanja se nahajajo na periferiji mreže ali pa jih na mreži sploh ni. So manj esencialni in predstavljajo kandidatne gene za bolezen. Pridobljene informacije bodo olajšale razumevanje bolezni in razvoj diagnostike in prognostičnih markerjev, kot tudi izboljšanje zdravljenja KLL.

### 5.3 SKLEPI IN PRIHODNJE DELO

V diplomskem delu smo prikazali sistemski pristop k potrjevanju obstoječih in odkrivanju novih kandidatnih genov na modelu kronične limfocitne levkemije, ki ga lahko apliciramo pri drugih kompleksnih fenotipih. Po naših podatkih je to prva študija, ki je uporabila takšen integrativni pristop pri KLL. Odkrili smo nove povezave in interakcije miRNA pri KLL ter variabilnost na vseh poznanih ravneh z RNA posredovane regulacije, ki bi lahko imele vlogo pri razvoju obolenja. Ugotovitve in rezultati pri ujemanju del se vklapljam v širšo teorijo, da je KLL kompleksna bolezen, za razumevanje katere je potrebno ukoreninjen linearni miselni vzorec centralne dogme: DNA-RNA-protein preoblikovati v več dimenzionalen pogled, saj so možne napake v interakcijah med različnimi komponentami, ki lahko vodijo v fenotip bolezni. Za več plastno razumevanje danih interakcij pa je priporočljivo uporabiti komplementarne, sistemske in integrativne pristope.

V prihodnosti bo potrebno ločeno preučiti genske mreže, ki so specifične za podtipa mutirane in nemutirane oblike KLL ali pa genske mreže za specifične klinične parametre.

S pristopi sistemске biologije lahko v prihodnosti pri akujemo tudi boljše razumevanje razli nih mehanizmov iniciacije KLL ter napredovanja bolezni. Prav tako lahko pri akujemo razjasnitev interakcij genov pri KLL v celicah T, v katerih pride do spremenjenega izražanja genov pod vplivom celic KLL (Görgün in sod., 2005). V prihodnosti bodo genske mreže omogo ale tudi vpogled v dinamiko bolezni, zato bo potrebno v analizo z genskimi mrežami vklju iti tudi kvantitativne parametre. Nadalje, še vedno ni razjasnjen natan en pomen interakcij miRNA s tar nimi geni in vloga drugih nekodirajo ih RNA. Ker prihaja v zadnjem asu do odkritij, da miRNA nimajo samo inhibitorne narave, bi bilo potrebno raziskati vlogo aktivacije miRNA v genskih mrežah KLL. Zelo uporabna bi bila tudi mreža, iz katere bi bile razvidne razlike interakcij med razli nimi variantami polimorfnih miRNA ter tar kot tudi razli nih izoform alternativnega spajanja. Širše gledano bi bila uporabna genska mreža, na kateri bi bili prisotni podatki o tem, ali je bila za dan gen odkrita spremenjena ekspresija, polimorfizem ali epigenetska sprememba.

## 6 POVZETEK

Razumevanje genske regulacije je v biologiji raka še vedno velik izviv. eprav je v uporabi veliko razli nih metodoloških pristopov k identifikaciji sprememb na ravni genov, pa z nobenim od njih ne pridobimo zadovoljivega vpogleda v medsebojne odnose razli nih molekulskih mehanizmov. Na primeru kroni ne limfocitne levkemije (KLL) je bilo pred kratkim prvi dokazano, da igrajo pri razvoju in napredovanju raka pomembno vlogo interakcije z regulatornimi molekulami mikroRNA (miRNA).

Namen diplomskega dela je bil združiti vsa obstoje a znanja o genskih lokusih in miRNA, povezanih s KLL, prikazati integrativni genomski pristop ter na tej osnovi analizirati domnevne vzroke in posledice deregulacije miRNA pri KLL.

Izdelali smo podatkovno zbirk, ki vsebuje ca. 2000 genskih lokusov za KLL, na osnovi katere smo identificirali 241 mo nejših kandidatnih genov in analizirali njihove glavne biološke funkcije. Na podlagi genske lokacije v genih za KLL smo predpostavili novo vlogo pri KLL za 15 miRNA. Odkrili smo 67 novih domnevnih povezav med geni za KLL in miRNA za KLL. Pokazali smo genetsko variabilnost 24 genov za miRNA, šestih tar miRNA s skupno 20 SNP-ji in treh komponent z vlogo v utiševalnem mehanizmu, ki predstavljajo potencialne ozna evalce za KLL. Vizualizirali smo genske mreže z molekulami miRNA pri KLL, pokazali vlogo 37 novih vozliš v genskih mrežah za KLL ter potrdili že poznano vpletjenost nekaterih genov.

Ugotovitve diplomskega dela bodo prispevale k boljšemu razumevanju KLL, prikazan sistemski pristop pa bo mogo e aplicirati pri analizi drugih poligenskih bolezni in kompleksnih lastnosti.

## 7 VIRI

- Aalto Y., El-Rifa W., Vilpo L., Ollila J., Nagy B., Vihinen M., Vilpo J., Knuutila S. 2001. Distinct gene expression profiling in chronic lymphocytic leukemia with 11q23 deletion. Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K., 15: 1721-8
- Abruzzo L.V., Wang J., Kapoor M., Medeiros L.J., Keating M.J., Edward Highsmith W., Barron L.L., Cromwell C.C., Coombes K.R. 2005. Biological validation of differentially expressed genes in chronic lymphocytic leukemia identified by applying multiple statistical methods to oligonucleotide microarrays. The Journal of Molecular Diagnostics, 7: 337-45
- Alfarano A., Indraccolo S., Circosta P., Minuzzo S., Vallario A., Zamarchi R., Fregonese A., Calderazzo F., Faldella A., Aragno M. in sod. 1999. An alternatively spliced form of CD79b gene may account for altered B-cell receptor expression in B-chronic lymphocytic leukemia. Blood, 93: 2327-35
- Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X. in sod. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature, 403: 503-11
- Alon U. 2006. An introduction to systems biology. CH/CRC Mathematical and Computational Biology Series, Volume 10. **Boca Raton**, CRC Press: 301 str.
- Andreeff M., Darzynkiewicz Z., Sharpless T.K., Clarkson B.D., Melamed M.R.. 1980. Discrimination of human leukemia subtypes by flow cytometric analysis of cellular DNA and RNA. Blood, 55: 282-93
- Aqeilan R.I., Calin G.A., Croce C.M. 2009. miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. Cell death and differentiation (v tisku)
- Aragues R., Sali A., Bonet J., Marti-Renom M.A., Oliva B. 2007. Characterization of protein hubs by inferring interacting motifs from protein interactions. PLoS Computational Biology 3: 1761-71
- Au W.Y., Fung A., Wong K.F., Chan C.H., Liang R. 2006. Tumor necrosis factor alpha promoter polymorphism and the risk of chronic lymphocytic leukemia and myeloma in the Chinese population. Leukemia & lymphoma, 47: 2189-93
- Auer R.L., Riaz S., Cotter F.E. 2007. The 13q and 11q B-cell chronic lymphocytic leukaemia associated regions derive from a common ancestral region in the zebrafish. British Journal of Haematology, 137: 443-453

- Awan A., Bari H., Yan F., Moksong S., Yang S., Chowdhury S., Cui Q., Yu Z., Purisima E.O., Wang E. 2007. Regulatory network motifs and hotspots of cancer genes in a mammalian cellular signalling network. *IET systems biology*, 1: 292-7
- Babu M.M. Aravind L. 2006. Adaptive evolution by optimizing expression levels in different environments. *Trends in microbiology*, 14: 11-4
- Bachman K.E., Park B.H., Rhee I., Rajagopalan H., Herman J.G., Baylin S.B., Kinzler K.W., Vogelstein B. 2003. Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell*, 3: 89-95
- Barabási A., Oltvai Z.N. 2004. Network biology: understanding the cell's functional organization. *Nature reviews. Genetics*, 5: 101-13
- Barrett C.L., Palsson B.O. 2006. Iterative reconstruction of transcriptional regulatory networks: an algorithmic approach. *PLoS Computational Biology*, 2: e52
- Bartel D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116: 281-97
- Batada N.N., Hurst L.D., Tyers M. 2006. Evolutionary and physiological importance of hub proteins. *PLoS Computational Biology*, 2: e88
- Baylin S.B., Esteller M., Rountree M.R., Bachman K.E., Schuebel K., Herman J.G. 2001. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Human Molecular Genetics*, 10: 687-692
- Baylin S.B., Herman J.G., Graff J.R., Vertino P.M., Issa J.P. 1998. Alterations in DNA methylation: A fundamental aspect of neoplasia. *Advances in Cancer Research*, 72: 141-196.
- Bichi R., Shinton S.A., Martin E.S., Koval A., Calin G.A., Cesari R., Russo G., Hardy R.R. Croce CM. 2002. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 6955-60
- Bilban M., Heintel D., Scharl T., Woelfel T., Auer M.M., Porpaczy E., Kainz B., Kröber A., Carey V.J., Shehata M., Zielinski C., Pickl W., Stilgenbauer S., Gaiger A., Wagner O., Jäger U., German CLL Study Group. 2006. Deregulated expression of fat and muscle genes in B-cell chronic lymphocytic leukemia with high lipoprotein lipase expression. *Leukemia: official journal of the Leukemia Society of America*, Leukemia Research Fund, U.K., 20: 1080-8

- Binet J., Caligaris-Cappio F., Catovsky D., Cheson B., Davis T., Dighiero G., Döhner H., Hallek M., Hillmen P., Keating M.,Montserrat E., Kipps T.J., Rai K., International Workshop on Chronic Lymphocytic Leukemia (IWCLL). 2006. Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. *Blood*, 107: 859-61
- Bjørnstad O.N. Harvill E.T. 2005. Evolution and emergence of *Bordetella* in humans. *Trends in microbiology*, 13: 355-9
- Bogunia-Kubik K., Mazur G., Urbanowicz I., Wróbel T., Kuliczkowski K., Wo niak M., Lange A. 2006. Lack of association between the TNF-alpha promoter gene polymorphism and susceptibility to B-cell chronic lymphocytic leukaemia. *International Journal of Immunogenetics*, 33: 21-4
- Bolland D.J., Wood A.L., Johnston C.M., Bunting S.F., Morgan G., Chakalova L. Fraser P.J., Corcoran A.E. 2004. Antisense intergenic transcription in V(D)J recombination. *Nature Immunology*, 5: 630-637
- Buhler M. 2009. RNA turnover and chromatin-dependent gene silencing. *Chromosoma*, 118: 141-151
- Cabrini G., Falzoni S., Forchap S.L., Pellegatti P., Balboni A., Agostini P., Cuneo A., Castoldi G., Baricordi O.R., Di Virgilio F. in sod. 2005. A His-155 to Tyr polymorphism confers gain-of-function to the human P2X7 receptor of human leukemic lymphocytes. *Journal of Immunology*, 175: 82-9
- Caligaris-Cappio F., Ghia P. 2008. Novel insights in chronic lymphocytic leukemia: are we getting closer to understanding the pathogenesis of the disease? *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 26: 4497-503
- Caligaris-Cappio F. 1996. B-chronic lymphocytic leukemia: a malignancy of anti-self B cells. *Blood*, 87: 2615-20
- Caligaris-Cappio F. 2003. Role of the microenvironment in chronic lymphocytic leukaemia. *British Journal of Haematology*, 123: 380-8
- Calin G.A., Croce C.M. 2006. MicroRNA signatures in human cancers. *Nature reviews. Cancer*, 6: 857-66
- Calin G.A., Cimmino A., Fabbri M., Ferracin M., Wojcik S.E., Shimizu M., Taccioli C., Zanesi N., Garzon R., Aqeilan R.I., Alder H., Volinia S., Rassenti L., Liu X., Liu C.G., Kipps T.J., Negrini M., Croce C.M. 2008. MiR-15a and miR-16-1 cluster functions in human leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 5166-71

- Calin G.A., Dumitru C.D., Shimizu M., Bichi R., Zupo S. Noch E., Aldler H., Rattan S., Keating M., Rai K., Rassenti L., Kipps T., Negrini M., Bullrich F., Croce C.M. 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences*, 99: 13-18
- Calin G.A., Ferracin M., Cimmino A., Leva G.D., Shimizu M., Wojcik S.E., Iorio M.V., Visone R., Sever N.I., Fabbri M., Iuliano R., Palumbo T., Pichiorri F., Roldo C., Garzon R., Sevignani C., Rassenti L., Alder H., Volinia S., Liu C., Kipps T.J., Negrini M., Croce C.M. 2005. A MicroRNA Signature Associated with Prognosis and Progression in Chronic Lymphocytic Leukemia. *The New England Journal of Medicine*, 353: 1793-1802
- Calin G.A., Liu C., Ferracin M., Hyslop T., Spizzo R., Sevignani C., Fabbri M., Cimmino A., Lee E.J., Wojcik S.E. Shimizu M., Tili E., Rossi S., Taccioli C., Pichiorri F., Liu X., Zupo S., Herlea V., Gramantieri L., Lanza G., Alder H., Rassenti L., Volinia S., Schmittgen T.D., Kipps T.J., Negrini M., Croce C.M. 2007. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell*, 12: 215-29
- Calin G.A., Liu C., Sevignani C., Ferracin M., Felli N., Dumitru C.D., Shimizu M., Cimmino A., Zupo S., Dono M., Dell'Aquila M.L., Alder H., Rassenti L., Kipps T.J., Bullrich F., Negrini M., Croce C.M. 2004a. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 11755-60
- Calin G.A., Sevignani C., Dumitru C.D., Hyslop T., Noh E., Yendamuri S., Shimizu M., Rattan S., Bullrich F., Negrini M. Croce C.M. 2004b. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 2999-3004
- Catovsky D. 1982. Symposium: classification of leukemia. 1. The classification of acute leukemia. *Pathology*, 14: 277-81
- Chen B. Wu W. 2007. Underlying principles of natural selection in network evolution: systems biology approach. *Evolutionary Bioinformatics Online*, 3: 245-62
- Chen C., Li L., Lodish H.F., Bartel D.P. 2004. MicroRNAs modulate hematopoietic lineage differentiation. *Science*, 303: 83-6
- Chim C.S., Fung T.K., Wong K.F., Lau J.S., Law M., Liang R. 2006. Methylation of INK4 and CIP/KIP families of cyclin-dependent kinase inhibitor in chronic lymphocytic leukaemia in Chinese patients. *Journal of Clinical Pathology*, 59: 921-6
- Chim C.S., Pang R., Liang R. 2008. Epigenetic dysregulation of the Wnt signalling pathway in chronic lymphocytic leukaemia. *Journal of Clinical Pathology*, 61: 1214-9

- Chiorazzi N., Ferrarini M. 2003. B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. *Annual review of Immunology*, 21: 841-94
- Chiromatzo A.O., Oliveira T.Y., Pereira G., Costa A.Y., Montesco C.A., Gras D.E., Yosetake F., Vilar J.B., Cervato M., Prado P.R., in sod. 2007. miRNAPath: a database of miRNAs, target genes and metabolic pathways. *Genetics and Molecular Research*, 6: 859-65
- Chuang H., Lee E., Liu Y., Lee D., Ideker T. 2007. Network-based classification of breast cancer metastasis. *Molecular Systems Biology*, 3: 140
- Cimmino A., Calin G.A., Fabbri M., Iorio M.V., Ferracin M., Shimizu M., Wojcik S.E., Aqeilan R.I., Zupo S., Dono M., in sod. 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America*, 102: 13944-9
- Committee to review the health effects in Vietnam veterans of exposure to herbicides (fourth biennial update). *Veterans and Agent Orange*: 2002. 2003. Washington, DC, National Academic Press: 373-77
- Corcoran M., Parker A., Orchard J., Davis Z., Wirtz M., Schmitz O.J., Oscier D. 2005. ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. *Haematologica*, 90: 1078-88
- Cotter F.E., Auer R.L. 2007. Genetic alteration associated with chronic lymphocytic leukemia. *Cytogenetic and Genome Research*, 118: 310-319
- Cui Q., Ma Y., Jaramillo M., Bari H., Awan A., Yang S., Zhang S., Liu L., Lu M., O'Connor-McCourt M., Purisima E.O., Wang E. 2007. A map of human cancer signaling. *Molecular Systems Biology*, 3: 152
- Cui Q., Yu Z., Purisima E.O., Wang E. 2006. Principles of microRNA regulation of a human cellular signaling network. *Molecular Systems Biology*, 2: 46
- Damle R.N., Wasil T., Fais F., Ghiootto F., Valetto A., Allen S.L., Buchbinder A., Budman D., Dittmar K., Kolitz J., Lichtman S.M., Schulman P., Vinciguerra V.P., Rai K.R., Ferrarini M., Chiorazzi N. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*, 94: 1840-1847
- Dazzi F., D'Andrea E., Biasi G., De Silvestro G., Gaidano G., Schena M., Tison T., Vianello F., Girolami A., Caligaris-Cappio F. 1995. Failure of B cells of chronic lymphocytic leukemia in presenting soluble and alloantigens. *Clinical Immunology and Immunopathology*, 75: 26-32
- Deguchi Y., Negoro S., Kishimoto S. 1987. Methylation of c-myc gene changes in human lymphoproliferative diseases. *Bioscience Reports*, 7: 637-43

- Dekel E., Alon U. 2005. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436: 588-92
- Demeter J., Porzsolt F., Rämisch S., Schmid M., Messer G. 1997. Polymorphism of the tumour necrosis factor-alpha and lymphotoxin-alpha genes in hairy cell leukaemia. *British Journal of Haematology*, 97: 132-4
- Dennis G. Jr., Sherman B.T., Hosack D.A., Yang J., Gao W., Lane H.C., Lempicki R.A. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, 4: P3
- Di Bernardo M.C., Crowther-Swanepoel D., Broderick P., Webb E., Sellick G., Wild R., Sullivan K., Vijayakrishnan J., Wang Y., Pittman A.M., Sunter N.J., Hall A.G., Dyer M.J., Matutes E., Dearden C., Mainou-Fowler T., Jackson G.H., Summerfield G., Harris R.J., Pettit A.R., Hillmen P., Allsup D.J., Bailey J.R., Pratt G., Pepper C., Fegan C., Allan J.M., Catovsky D., Houlston R.S. 2008. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nature Genetics*, 40: 1204-10
- Dickinson J.D., Joshi A., Iqbal J., Sanger W., Bierman P.J., Joshi S.S. 2006. Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. *International Journal of Molecular Medicine*, 17: 769-778
- Dickinson J.D., Smith L.M., Sanger W.G., Zhou G., Townley P., Lynch J.C., Pavletic Z.S., Bierman P.J., Joshi S.S. 2005. Unique gene expression and clinical characteristics are associated with the 11q23 deletion in chronic lymphocytic leukaemia. *British Journal of Haematology*, 128: 460-71
- Döhner H., Stilgenbauer S., Benner A., Leupolt E., Krober A., Bullinger L., Dohner K., Bentz M., Lichter P. 2000. Genomic aberrations and survival in chronic lymphocytic leukemia. *New England Journal of Medicine*, 343: 1910-1916
- Dürig J., Nückel H., Hüttmann A., Kruse E., Hölter T., Halfmeyer K., Führer A., Rudolph R., Kalhori N., Nusch A., Deaglio S., Malavasi F., Möröy T., Klein-Hitpass L., Dührsen U. 2003. Expression of ribosomal and translation-associated genes is correlated with a favorable clinical course in chronic lymphocytic leukemia. *Blood*, 101: 2748-55
- Ekman D., Light S., Björklund A.K., Elofsson A. 2006. What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome Biology*, 7: R45
- Enjuanes A., Benavente Y., Bosch F., Martín-Guerrero I., Colomer D., Pérez-Alvarez S., Reina O., Ardanaz M.T., Jares P., García-Orad A., Pujana M.A., Montserrat E., de-Sanjose S., Campo E. 2008. Genetic variants in apoptosis and immunoregulation-related genes are associated with risk of chronic lymphocytic leukemia. *Cancer Research*, 68: 10178-86

- Ferrer A., Ollila J., Tobin G., Nagy B., Thunberg U., Aalto Y., Vihinen M., Vilpo J., Rosenquist R., Knuutila S. 2004. Different gene expression in immunoglobulin-mutated and immunoglobulin-unmutated forms of chronic lymphocytic leukemia. *Cancer Genetics and Cytogenetics*, 153: 69-72
- Filipowicz W., Bhattacharyya S.N., Sonenberg N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews. Genetics*, 9: 102-14
- Fong S.S., Joyce A.R., Palsson B.Ø. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Research*, 15: 1365-72
- Fontana L., Pelosi E., Greco P., Racanicchi S., Testa U., Liuzzi F., Croce C.M., Brunetti E., Grignani F., Peschle C. 2007. MicroRNAs 17-5p-20a-106a control monocytopoiesis through AML1 targeting and M-CSF receptor upregulation. *Nature Cell Biology*, 9: 775-87
- Fulci V., Chiaretti S., Goldoni M., Azzalin G., Carucci N., Castellano L., Magrelli A., Citarella F., Messina M., Maggio R., Peragine N., Santangelo S., Mauro F.R., Landgraf P., Tuschl T., Weir D.B., Chien M., Russo J.J., Ju J., Sheridan R., Sander C., Zavolan M., Guarini A., Foà R., Macino G. 2009. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood*, 109: 4944-4951
- Fuller S.J., Papaemmanuil E., McKinnon L., Webb E., Sellick G.S., Dao-Ung L., Skarratt K.K., Crowther D., Houlston R.S., Wiley J.S. 2008. Analysis of a large multi-generational family provides insight into the genetics of chronic lymphocytic leukemia. *British Journal of Haematology*, 142: 238-245.
- Futreal P.A., Coin L., Marshall M., Down T., Hubbard T., Wooster R., Rahman N., Stratton M.R. 2004. A census of human cancer genes. *Nature Reviews. Cancer*, 4: 177-83
- Gimelbrant A., Hutchinson J.N., Thompson B.R., Chess A. 2007. Widespread monoallelic expression on human autosomes. *Science*, 318: 1136-1140
- Goh K., Cusick M.E., Valle D., Childs B., Vidal M. 2007. The human disease network. *Proceedings of the National Academy of Sciences*, 104: 8685-8690
- Goldin L.R., Ishibe N., Sgambati M., Marti G.E., Fontaine L., Lee M.P., Kelley J.M., Scherbier T., Buetow K.H., Caporaso N.E. 2003. A genome scan of 18 families with chronic lymphocytic leukaemia. *British Journal of Haematology*, 121: 866-73
- Golub T.R., Slonim D.K., Tamayo P., Huard C., Gaasenbeek M., Mesirov J.P., Coller H., Loh M.L., Downing J.R., Caligiuri M.A., Bloomfield C.D., Lander E.S. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286: 531-7

- Görgün G., Holderried T.A., Zahrieh D., Neuberg D., Gribben J.G. 2005. Chronic lymphocytic leukemia cells induce changes in gene expression of CD4 and CD8 T cells. *The Journal of Clinical Investigation*, 115: 1797-805
- Granziero L., Ghia P. Circosta P., Gottardi D., Strola G., Geuna M., Montagna L., Piccoli P., Chilosì M., Caligaris-Cappio F. 2001. Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. *Blood*, 97: 2777-83
- Grubor V., Krasnitz A., Troge J.E., Meth J.L., Lakshmi B., Kendall J.T., Yamrom B., Alex G., Pai D., Navin N., Hufnagel L.A., Lee Y.H., Cook K., Allen S.L., Rai K.R., Damle R.N., Calissino C., Chiorazzi N., Wigler M., Esposito D. 2009. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). *Blood*, 113: 1294-303
- Gunsalus K.C., Ge H., Schetter A.J., Goldberg D.S., Han J.J., Hao T., Berriz G.F., Bertin N., Huang J., Chuang L., Li N., ,ani R., Hyman A.A., Sönnichsen B., Echeverri C.J., Roth F.P., Vidal M., Piano F. 2005. Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature*, 436: 861-5
- Haiat S., Billard C., Quiney C., Ajchenbaum-Cymbalista F., Kolb J. 2006. Role of BAFF and APRIL in human B-cell chronic lymphocytic leukaemia. *Immunology*, 118: 281-92
- Hamblin T. 2006. Is chronic lymphocytic leukemia a response to infectious agents? *Leukemia research*, 30: 1063-4
- Hamblin T.J., Davis Z., Gardiner A., Oscier D.G., Stevenson F.K. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*, 94: 1848-54
- Hammarsund M., Corcoran M.M., Wilson W., Zhu C., Einhorn S., Sangfelt O., Grande D. 2004. Characterization of a novel B-CLL candidate gene - DLEU7 - located in the 13q14 tumor suppressor locus. *FEBS Letters*, 556: 1-6
- Han J.J., Bertin N., Hao T., Goldberg D.S., Berriz G.F., Zhang L.V., Dupuy D., Walhout A.J., Cusick M.E., Roth F.P., Vidal M. 2004. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature*, 430: 88-93
- Han J.J. 2008. Understanding biological functions through molecular networks. *Cell Research*, 18: 224-37
- Hanada M., Delia D., Aiello A., Stadtmauer E., Reed J.C. 1993. bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. *Blood*, 82: 1820-8

- Haslinger C., Schweifer N., Stilgenbauer S., Dohner H., Lichter P., Kraut N., Stratowa C., Abseher R. 2004. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. *Journal of Clinical Oncology*, 22: 3937-3949
- Hernández J.Á., Rodríguez A.E., González M., Benito R., Fontanillo C., Sandoval V., Romero M., Martín-núñez G., Coca A.G., Fisac R., Galende J., Recio I., Ortuño F., García J.L., de las Rivas J., Gutiérrez N.C., San Miguel J.F., Hernández J.M. 2009. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica*, 94: 364-371
- Hervé M., Xu K., Ng Y., Wardemann H., Albesiano E., Messmer B.T., Chiorazzi N., Meffre E. 2005. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *The Journal of Clinical Investigation*, 115: 1636-43
- Hewitt S.L., Yin B., Ji Y.H., Chaumeil J.K., Marszalek K., Tenthorey J., Salvagiotto G., Steinle N., Ramsey L.B., Ghysdael J., Farrar M.A., Sleckman B.P., Schatz D.G., Busslinger M., Bassing C.H. Skok J.A. 2009. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nature Immunology*, 10: 655-U131
- Huang D.W., Sherman B.T., Lempicki R.A. 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols*, 4: 44-57
- Hwang H., Wentzel E.A., Mendell J.T. 2007. A hexanucleotide element directs microRNA nuclear import. *Science*, 315: 97-100
- Jansen R., Yu H., Greenbaum D., Kluger Y., Krogan N.J., Chung S., Emili A., Snyder M., Greenblatt J.F., Gerstein M. 2003. A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science*, 302: 449-53
- Jelinek D.F., Tschumper R.C., Stolovitzky G.A., Iturria S.J., Tu Y., Lepre J., Shah N., Kay N.E. 2003. Identification of a global gene expression signature of B-chronic lymphocytic leukemia. *Molecular Cancer Research*, 1: 346-61
- Jemal A., Murray T., Ward E., Samuels A., Tiwari R.C., Ghafoor A., Feuer E.J., Thun M.J. 2005. Cancer Statistics. *CA: a Cancer Journal for Clinicians*, 55: 10-30
- Jeong H., Mason S.P., Barabási A.L., Oltvai Z.N. 2001. Lethality and centrality in protein networks. *Nature*, 411: 41-2
- Jonsson P.F., Bates P.A. 2006. Global topological features of cancer proteins in the human interactome. *Bioinformatics*, 22: 2291-7

- Jordan I.K., Wolf Y.I., Koonin E.V. 2003. No simple dependence between protein evolution rate and the number of protein-protein interactions: only the most prolific interactors tend to evolve slowly. *BMC Evolutionary Biology*, 3: 1
- Joshi A.D., Hegde G.V., Dickinson J.D., Mittal A.K., Lynch J.C., Eudy J.D., Armitage J.O., Bierman P.J., Bociek R.G., Devetten M.P., Vose J.M., Joshi S.S. 2007. ATM, CTLA4, MNDA, and HEM1 in high versus low CD38 expressing B-cell chronic lymphocytic leukemia. *Clinical Cancer Research : an official journal of the American Association for Cancer Research*, 13: 5295-304
- Katsumata M., Siegel R.M., Louie D.C., Miyashita T., Tsujimoto Y., Nowell P.C., Greene M.I., Reed J.C. 1992. Differential effects of Bcl-2 on T and B cells in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 89: 11376-80
- Kienle D.L., Korz C., Hosch B., Benner A., Mertens D., Habermann A., Kröber A., Jäger U., Lichter P., Döhner H. Stigenbauer S. 2005. Evidence for distinct pathomechanisms in genetic subgroups of chronic lymphocytic leukemia revealed by quantitative expression analysis of cell cycle, activation, and apoptosis-associated genes. *Journal of Clinical Oncology*, 23: 3780-3792
- Kipps T.J., Carson D.A. 1993. Autoantibodies in chronic lymphocytic leukemia and related systemic autoimmune diseases. *Blood*, 81: 2475-87
- Klein U., Tu Y., Stolovitzky G.A., Mattioli M., Cattoretti G., Husson H., Freedman A., Inghirami G., Cro L., Baldini L., Neri A., Califano A., Dalla-Favera R. 2001. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *The Journal of Experimental Medicine*, 194: 1625-38
- Knudson A.G. 1971. Mutation and cancer - statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, 68: 820-3
- Krek A., Grün D., Poy M.N., Wolf R., Rosenberg L., Epstein E.J., MacMenamin P., da Piedade I., Gunsalus K.C., Stoffel M., Rajewsky N. 2005. Combinatorial microRNA target predictions. *Nature Genetics*, 37: 495-500
- Kruglyak L., Nickerson D.A. 2001. Variation is the spice of life. *Nature Genetics*, 27: 234-6
- Landgren O., Rapkin J.S., Caporaso N.E., Mellemkjaer L., Gridley G., Goldin L.R., Engels E.A. 2007. Respiratory tract infections and subsequent risk of chronic lymphocytic leukemia. *Blood*, 109: 2198-201
- Lanham S., Hamblin T., Oscier D., Ibbotson R., Stevenson F., and Packham G. 2003. Differential signaling via surface IgM is associated with VH gene mutational status and CD38 expression in chronic lymphocytic leukemia. *Blood*, 101: 1087-93

- Lanzavecchia A. 1985. Antigen-specific interaction between T and B cells. *Nature*, 314: 537-9
- Lee I., Date S.V., Adai A.T., Marcotte E.M. 2004. A probabilistic functional network of yeast genes. *Science*, 306: 1555-8
- Lee S.Y., Reichlin A., Santana A., Sokol K.A., Nussenzweig M.C., Choi Y. 1997. TRAF2 is essential for JNK but not NF-kappaB activation and regulates lymphocyte proliferation and survival. *Immunity*, 7: 703-13
- Lewintre E. J., Martín C. R., Montaner D., Marín M., Terol M.J., Farrás R. Benet I., Calvete, J.J., Dopazo, J. García-Conde, J. 2009. Analysis of chronic lymphocytic leukemia transcriptomic profile: differences between molecular subgroups. *Leukemia and Lymphoma*, 50: 68 -79
- Lewis B.P., Burge C.B., Bartel D.P. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 120: 15-20
- Liu T., Raval A., Chen S., Matkovic J.J., Byrd J.C., Plass C. 2006. CpG island methylation and expression of the secreted frizzled-related protein gene family in chronic lymphocytic leukemia. *Cancer research*, 66: 653-8
- Lu M., Zhang Q., Deng M., Miao J., Guo Y., Gao W., Cui Q. 2008. An analysis of human microRNA and disease associations. *PloS One*, 3: e3420
- Lujambio A., Calin G.A., Villanueva A., Ropero S., Sánchez-Céspedes M., Blanco D., Montuenga L.M., Rossi S., Nicoloso M.S., Faller W.J., Gallagher W.M., Eccles S.A., Croce C.M., Esteller M. 2008. A microRNA DNA methylation signature for human cancer metastasis. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 13556-61
- Mabuchi H., Fujii H., Calin G., Alder H., Negrini M., Rassenti L., Kipps T.J., Bullrich F., Croce C.M. 2001. Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. *Cancer Research*, 61: 2870-2877
- Madan Babu M., Teichmann S.A., Aravind L. 2006. Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *Journal of Molecular Biology*, 358: 614-33
- Mainou-Fowler T., Dickinson A.M., Taylor P.R., Mounter P., Jack F., Proctor S.J., Nordon J. Middleton P.G. 2000. Tumour necrosis factor gene polymorphisms in lymphoproliferative disease. *Leukemia & Lymphoma*, 38: 547-52
- Maizels N. 2005. Immunoglobulin gene diversification. *Annual Review of Genetics*, 39: 23-46

- Mani K.M., Lefebvre C., Wang K., Lim W.K., Basso K., Dalla-Favera R., Califano A. 2008. A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Molecular Systems Biology*, 4: 169
- Marti G.E., Rawstron A.C., Ghia P., Hillmen P., Houlston R.S., Kay N., Schleinitz T.A., Caporaso N. 2005. Diagnostic criteria for monoclonal B-cell lymphocytosis. *British Journal of Haematology*, 130: 325-32
- Marton S., Garcia M.R., Robello C., Persson H., Trajtenberg F., Pritsch O., Rovira C., Naya H., Dighiero G., Cayota A., 2008. Small RNAs analysis in CLL reveals a deregulation of miRNA expression and novel miRNA candidates of putative relevance in CLL pathogenesis. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K.*, 22: 330-8
- Maslov S., Sneppen K. 2002. Specificity and stability in topology of protein networks. *Science*, 296: 910-3
- Matutes E., Owusu-Ankomah K., Morilla R., Garcia Marco J., Houlihan A., Que T.H., Catovsky D. 1994. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K.*, 8: 1640-5
- Mertens D., Wolf S., Schroeter P., Schaffner C., Dohner H., Stilgenbauer S., Lichter P. 2002. Down-regulation of candidate tumor suppressor genes within chromosome band 13q14.3 is independent of the DNA methylation pattern in B-cell chronic lymphocytic leukemia. *Blood*, 99: 4116-4121
- Mertens D., Wolf S., Tschuch C., Mund C., Kienle D., Ohl S., Schroeter P., Lyko F., Dohner H., Stilgenbauer S. Lichter P. 2006. Allelic silencing at the tumor-suppressor locus 13q14.3 suggests an epigenetic tumor-suppressor mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 7741-7746
- Messmer B.T., Albesiano E., Efremov D.G., Ghiotto F., Allen S.L., Kolitz J., Foa R., Damle R.N., Fais F., Messmer D., Rai K.R., Ferrarini M., Chiorazzi N. 2004. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *The Journal of Experimental Medicine*, 200: 519-25
- Migliazza A., Bosch F., Komatsu H., Cayanis E., Martinotti S., Toniato E., Guccione E., Qu X.Y., Chien M., Murty V.V.V., Gaidano G., Inghirami G., Zhang P., Fischer S., Kalachikov S.M., Russo J., Edelman I., Efstratiadis A., Dalla-Favera R. 2001. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. *Blood*, 97: 2098-2104
- Milo R., Shen-Orr S., Itzkovitz S., Kashtan N., Chklovskii D., Alon U. 2002. Network motifs: simple building blocks of complex networks. *Science*, 298: 824-7

- Muller-Hermelink H.K., Montserrat E., Catovsky D., Harris N.L. 2001. Chronic lymphocytic leukemia/small lymphocytic lymphoma. V: World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Jaffe E.S., Harris N.L., Stein H., Vardiman J.W. (eds.) Lyon, France, IARC Press: 127–130
- Murray F., Darzentas N., Hadzidimitriou A., Tobin G., Boudjogra M., Scielzo C., Laoutaris N., Karlsson K., Baran-Marzsak F., Tsafaris A., in sod. 2008. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*, 111: 1524-33
- National Cancer Institute. 2008. SEER cancer statistics review 1975–2006.  
[http://seer.cancer.gov/csr/1975\\_2006/results\\_merged/sect\\_13\\_leukemia.pdf](http://seer.cancer.gov/csr/1975_2006/results_merged/sect_13_leukemia.pdf) (5. avgust 2009)
- Nilsson K. 1992. Human B-lymphoid cell lines. *Human Cell: official journal of Human Cell Research Society*, 5: 25-41
- Nückel H., Frey U.H., Dürig J., Dührsen U., Siffert W. 2004. 1513A/C polymorphism in the P2X7 receptor gene in chronic lymphocytic leukemia: absence of correlation with clinical outcome. *European Journal of Haematology*, 72: 259-63
- Ogorevc J., Kunej T., Razpet A., Dov P. 2009. Database of cattle candidate genes and genetic markers for milk production and mastitis. *Animal Genetics* (v tisku)
- Ørom U.A., Nielsen F.C., Lund A.H. 2008. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular Cell*, 30: 460-71
- Pascual V., Liu Y.J., Magalski A., de Bouteiller O., Banchereau J., Capra J.D. 1994. Analysis of somatic mutation in five B cell subsets of human tonsil. *The Journal of Experimental Medicine*, 180: 329-39
- Payelle-Brogard B., Dumas G., Magnac C., Lalanne A.I., Dighiero G., Vuillier F. 2006. Abnormal levels of the alpha chain of the CD22 adhesion molecule may account for low CD22 surface expression in chronic lymphocytic leukemia. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K.*, 20: 877-8
- Payelle-Brogard B., Magnac C., Alcover A., Roux P., Dighiero G. 2002. Defective assembly of the B-cell receptor chains accounts for its low expression in B-chronic lymphocytic leukaemia. *British Journal of Haematology*, 118: 976-85
- Place R.F., Li L., Pookot D., Noonan E.J., Dahiya R. 2008. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 1608-13

- Pritsch O., Magnac C., Dumas G., Egile C., Dighiero G. 1993. V gene usage by seven hybrids derived from CD5+ B-cell chronic lymphocytic leukemia and displaying autoantibody activity. *Blood*, 82: 3103-12
- Rajewsky N. 2006. microRNA target predictions in animals. *Nature Genetics*, 38 Suppl: S8-13
- Ranheim E.A., Kipps T.J. 1993. Activated T cells induce expression of B7/BB1 on normal or leukemic B cells through a CD40-dependent signal. *The Journal of Experimental Medicine*, 177: 925-35
- Rassenti L.Z., Kipps T.J. 1997. Lack of allelic exclusion in B cell chronic lymphocytic leukemia. *The Journal of Experimental Medicine*, 185: 1435-45
- Raval A., Lucas D.M., Matkovic J.J., Bennett K.L., Liyanarachchi S., Young D.C., Rassenti L., Kipps T.J., Grever M.R., Byrd J.C., Plass C. 2005. TWIST2 demonstrates differential methylation in immunoglobulin variable heavy chain mutated and unmutated chronic lymphocytic leukemia. *Journal of Clinical Oncology: official journal of the American Society of Clinical Oncology*, 23: 3877-85
- Raveche E.S., Salerno E., Scaglione B.J., Manohar V., Abbasi F., Lin Y., Fredrickson T., Landgraf P., Ramachandra S., Huppi K., Toro J.R., Zenger V.E., Metcalf R.A., Marti G.E. 2007. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood*, 109: 5079-86
- Rawstron A.C., Green M.J., Kuzmicki A., Kennedy B., Fenton J.A., Evans P.A., O'Connor S.J., Richards S.J., Morgan G.J., Jack A.S., Hillmen P. 2002. Monoclonal B lymphocytes with the characteristics of "indolent" chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts. *Blood*, 100: 635-9
- Rhodes D.R., Tomlins S.A., Varambally S., Mahavisno V., Barrette T., Kalyana-Sundaram S., Ghosh D., Pandey A., Chinnaiyan A.M. 2005. Probabilistic model of the human protein-protein interaction network. *Nature Biotechnology*, 23: 951-9
- Rosenwald A., Alizadeh A.A., Widhopf G., Simon R., Davis R.E., Yu X., Yang L., Pickeral O.K., Rassenti L.Z., Powell J., Botstein D., Byrd J.C., Grever M.R., Cheson B.D., Chiorazzi N., Wilson W.H., Kipps T.J., Brown P.O., Staudt L.M. 2001. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *The Journal of Experimental Medicine*, 194: 1639-47
- Rudd M.F., Sellick G.S., Webb E.L., Catovsky D., Houlston R.S. 2006. Variants in the ATM-BRCA2-CHEK2 axis predispose to chronic lymphocytic leukemia. *Blood*, 108: 638-44
- Rush L.J., Raval A., Funchain P., Johnson A.J., Smith L., Lucas D.M., Bembea M., Liu T., Heerema N.A., Rassenti L., in sod. 2004. Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer Research*, 64: 2424-33

- Saeed R., Deane C.M. 2006. Protein protein interactions, evolutionary rate, abundance and age. *BMC Bioinformatics*, 7: 128
- Said M.R., Begley T.J., Oppenheim A.V., Lauffenburger D.A., Samson L.D. 2004. Global network analysis of phenotypic effects: protein networks and toxicity modulation in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 18006-11
- Schaffner C., Stilgenbauer S., Rappold G.A., Dohner H., Lichter P. 1999. Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. *Blood*, 94: 748-753
- Schroeder H.W., Dighiero G. 1994. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunology Today*, 15: 288-94
- Schwaenen C., Nessling M., Wessendorf S., Salvi T., Wrobel G., Radlwimmer B., Kestler H.A., Haslinger C., Stilgenbauer S., Döhner H., Döhner H., Bentz M., Lichter P. 2004. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 1039-44
- Sellick G.S., Goldin L.R., Wild R.W., Slager S.L., Ressenti L., Strom S.S., Dyer M.J., Mauro F.R., Marti G.E., Fuller S., Lyttelton M., Kipps T.J., Keating M.J., Call T.G., Catovsky D., Caporaso N., Houlston R.S. 2007. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. *Blood*, 110: 3326-33
- Sellick G.S., Rudd M., Eve P., Allinson R., Matutes E., Catovsky D., Houlston R.S. 2004. The P2X7 receptor gene A1513C polymorphism does not contribute to risk of familial or sporadic chronic lymphocytic leukemia. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 13: 1065-7
- Sellick G.S., Wade R., Richards S., Oscier D.G., Catovsky D., Houlston R.S. 2008. Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis. *Blood*, 111: 1625-33
- Sellick G.S., Webb E.L., Allinson R., Matutes E., Dyer M.J., Jonsson V., Langerak A.W., Mauro F.R., Fuller S., Wiley J., Lyttelton M., Callea V., Yuille M., Catovsky D., Houlston R.S. 2005. A high-density SNP genomewide linkage scan for chronic lymphocytic leukemia-susceptibility loci. *American Journal of Human Genetics* 77: 420-9
- Shiu S., Borevitz J.O. 2008. The next generation of microarray research: applications in evolutionary and ecological genomics. *Heredity*, 100: 141-9

- Stamatopoulos B., Haibe-Kains B., Equeter C., Meuleman N., Sorée A., De Bruyn C., Hanosset D., Bron D., Martiat P., Lagneaux L. 2009a. Gene expression profiling reveals differences in microenvironment interaction between patients with chronic lymphocytic leukemia expressing high versus low ZAP70 mRNA. *Haematologica*, 94: 790-9
- Stamatopoulos B., Meuleman N., Haibe-Kains B., Saussoy P., Van Den Neste E., Michaux L., Heimann P., Martiat P., Bron D., Lagneaux L. 2009b. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. *Blood*, 113: 5237-45
- Stankovic T., Weber P., Stewart G., Bedenham T., Murray J., Byrd P.J., Moss P.A.H., Taylor A.M.R. 1999. Inactivation of ataxia telangiectasia mutated gene in B-cell chronic lymphocytic leukaemia. *Lancet*, 353: 26-29
- Starczynski J., Pepper C., Pratt G., Hooper L., Thomas A., Hoy T., Milligan D., Bentley P., Fegan C. 2003. The P2X7 receptor gene polymorphism 1513 A-->C has no effect on clinical prognostic markers, in vitro sensitivity to fludarabine, Bcl-2 family protein expression or survival in B-cell chronic lymphocytic leukaemia. *British Journal of Haematology*, 123: 66-71
- Stilgenbauer S., Liebisch P., James M.R., Bentz M., Fischer K., Lichter P., Dohner H. 1996. Molecular cytogenetic delineation of a novel critical genomic region in chromosome bands 11q22.3-q23.1 in lymphoproliferative disorders. *Blood*, 88: 243-243
- Stilgenbauer S., Sander S., Bullinger L., Benner A., Leupolt E., Winkler D., Krober A., Kienle D., Lichter P., Dohner H. 2007. Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. *Haematologica-the Hematology Journal*, 92: 1242-1245
- Suzuki H., Forrest A.R., van Nimwegen E., Daub C.O., Balwierz P.J., Irvine K.M., Lassmann T., Ravasi T., Hasegawa Y., de Hoon M.J., in sod. 2009. The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. *Nature Genetics*, 41: 553-562
- Tay Y., Zhang J., Thomson A.M., Lim B., Rigoutsos I. 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, 455: 1124-8
- Ter Brugge P.J., Ta V.B., de Brujin M.J., Keijzers G., Maas A., van Gent D.C., Hendriks R.W. 2009. A mouse model for chronic lymphocytic leukemia based on expression of the SV40 large T antigen. *Blood*, 114: 119-127

- Ternynck T., Dighiero G., Follezou J., Binet J.L. 1974. Comparison of normal and CLL lymphocyte surface Ig determinants using peroxidase-labeled antibodies. I. Detection and quantitation of light chain determinants. *Blood*, 43: 789-95
- Thompson A.A., Talley J.A., Do H.N., Kagan H.L., Kunkel L., Berenson J., Cooper M.D., Saxon A., Wall R. 1997. Aberrations of the B-cell receptor B29 (CD79b) gene in chronic lymphocytic leukemia. *Blood*, 90: 1387-94
- Thunberg U., Tobin G., Johnson A., Söderberg O., Padyukov L., Hultdin M., Klareskog L., Enblad G., Sundström C., Roos G., Rosenquist R. 2002. Polymorphism in the P2X7 receptor gene and survival in chronic lymphocytic leukaemia. *Lancet*, 360: 1935-9
- Tsirigotis P., Pappa V., Labropoulos S., Papageorgiou S., Kontsioti F., Dervenoulas J., Papageorgiou E., Panani A., Mantzios G., Economopoulos T., Raptiatis S. 2006. Mutational and methylation analysis of the cyclin-dependent kinase 4 inhibitor (p16INK4A) gene in chronic lymphocytic leukemia. *European Journal of Haematology*, 76: 230-6
- Tyybakinoja A., Vilpo J., Knuutila S. 2007. High-resolution oligonucleotide array-CGH pinpoints genes involved in cryptic losses in chronic lymphocytic leukemia. *Cytogenetic and Genome Research*, 118: 8-12
- Valencia-Sanchez M.A., Liu J., Hannon G.J., Parker R. 2006. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & Development*, 20: 515-24
- Vallabhajosyula R.R., Chakravarti D., Lutfeali S., Ray A., Raval A. 2009. Identifying hubs in protein interaction networks. *PloS one*, 4: e5344
- van't Veer M.B., Brooijmans A.M., Langerak A.W., Verhaaf B., Goudswaard C.S., Graveland W.J., van Lom K., Valk P.J. 2006. The predictive value of lipoprotein lipase for survival in chronic lymphocytic leukemia. *Haematologica*, 91: 56-63
- Vasudevan S., Tong Y., Steitz J.A. 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science*, 318: 1931-4
- Vázquez A. 2003. Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 67: 056104
- von Mering C., Krause R., Snel B., Cornell M., Oliver S.G., Fields S., Bork P. 2002. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature*, 417: 399-403
- Vuillier F., Dumas G., Magnac C., Prevost M., Lalanne A.I., Oppezzo P., Melanitou E., Dighiero G., Payelle-Brogard B. 2005. Lower levels of surface B-cell-receptor expression in chronic lymphocytic leukemia are associated with glycosylation and folding defects of the mu and CD79a chains. *Blood*, 105: 2933-40

- Wahlfors J., Hiltunen H., Heinonen K., Hämäläinen E., Alhonen L., Jänne J. 1992. Genomic hypomethylation in human chronic lymphocytic leukemia. *Blood*, 80: 2074-80
- Wang E., Purisima E. 2005. Network motifs are enriched with transcription factors whose transcripts have short half-lives. *Trends in Genetics*, 21: 492-5
- Wang E. 2008a. 25 Computational Methods in Cancer Gene Networking. *Nature Reviews Genetics*: 367-374
- Wang L., Pal S., Sif S. 2008b. Protein arginine methyltransferase 5 suppresses the transcription of the RB family of tumor suppressors in leukemia and lymphoma cells. *Molecular and Cellular Biology*, 28: 6262-77
- Watts D.J., Strogatz S.H. 1998. Collective dynamics of 'small-world' networks. *Nature*, 393: 440-2
- Weinberg R.A. 2006. *The Biology of Cancer*. New York, London, Garland Science: 796 str.
- Wihlborg C., Sjöberg J., Intaglietta M., Axdorph U., Pisa E.K., Pisa P. 1999. Tumour necrosis factor-alpha cytokine promoter gene polymorphism in Hodgkin's disease and chronic lymphocytic leukaemia. *British Journal of Haematology*, 104: 346-9
- Wiley J.S., Dao-Ung L.P., Gu B.J., Sluyter R., Shemon A.N., Li C., Taper J., Gallo J., Manoharan A. 2002. A loss-of-function polymorphic mutation in the cytolytic P2X7 receptor gene and chronic lymphocytic leukaemia: a molecular study. *Lancet*, 359: 1114-9
- William J., Euler C., Christensen S., Shlomchik M.J. 2002. Evolution of autoantibody responses via somatic hypermutation outside of germinal centers. *Science*, 297: 2066-70
- Winkler D., Döhner H., Stilgenbauer S. 2006. Genetics, gene expression, and targeted therapies in chronic lymphocytic leukemia. *Current Drug Targets*, 7: 1313-27
- Wolf S., Mertens D., Schaffner C., Korz C., Dohner H., Stilgenbauer S., Lichter P. 2001. B-cell neoplasia associated gene with multiple splicing (BCMS): the candidate B-CLL gene on 13q14 comprises more than 560 kb covering all critical regions. *Human Molecular Genetics*, 10: 1275-1285
- Wuchty S., Oltvai Z.N., Barabási A. 2003. Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nature Genetics*, 35: 176-9
- Xu J.L., Davis M.M. 2000. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity*, 13: 37-45

- Yu H., Kim P.M., Sprecher E., Trifonov V., Gerstein M. 2007. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Computational Biology*, 3: e59
- Yuille M.R., Condie A., Stone E.M., Wilsher J., Bradshaw P.S., Brooks L., Catovsky D. 2001. TCL1 is activated by chromosomal rearrangement or by hypomethylation. *Genes, Chromosomes & Cancer*, 30: 336-41
- Zapata J.M., Krajewska M., Morse H.C., Choi Y., Reed J.C. 2004. TNF receptor-associated factor (TRAF) domain and Bcl-2 cooperate to induce small B cell lymphoma/chronic lymphocytic leukemia in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 16600-5
- Zhang L.Y., Ibbotson R.E., Orchard J.A., Gardiner A.C., Seear R.V., Chase A.J., Oscier D.G., Cross N.C. 2003. P2X7 polymorphism and chronic lymphocytic leukaemia: lack of correlation with incidence, survival and abnormalities of chromosome 12. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K.*, 17: 2097-100
- Zhang Y., Zhang R., Su B. 2009. Diversity and evolution of MicroRNA gene clusters. *Science in China. Series C, Life sciences*, 52: 261-6
- Zheng Z., Venkatapathy S., Rao G., Harrington C.A. 2002. Expression profiling of B cell chronic lymphocytic leukemia suggests deficient CD1-mediated immunity, polarized cytokine response, altered adhesion and increased intracellular protein transport and processing of leukemic cells. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, U.K.*, 16: 2429-37

## ZAHVALA

Posebej se zahvaljujem mentorici doc. dr. Tanji Kunej za izjemno pozornost in usmerjanje pri delu, prijazno pomoč pri problemih, s katerimi sem se soočila, potrpežljivost, spodbudo in ponujeno roko pri nadalnjih korakih v svet bioznanosti.

Iskrena zahvala prof. dr. Petru Dovu za hiter pregled diplomskega dela, vse komentarje in nasvete ter predloge glede boljše ponazoritve nekaterih idej, pomoč pri izbiri strokovnega izrazoslovja in dano možnost obetavnega izhodišča za vnaprej.

Hvala sošolcem za kolegialnost tekom študija.

Najbolj pa sem hvaležna staršem za vse, kar so mi omogočili, in vsem doma im za podporo pri študiju in vlivanje novega elana.

## PRILOGE

### Priloga A:

Prese na množica izbora mo nejših kandidatnih genov za kroni no limfocitno levkemijo

Ime gena	Asociacijske in mutacijske študije	Ekspresijske študije	Epigenetske študije	Živalski model
<i>ACVR1B</i>		++		
<i>APAF1</i>	+	+		
<i>ATM</i>	+++	++++++		
<i>BAX</i>	+++	+		
<i>BCL2</i>	++	++++++++	+	+
<i>BCL6</i>	+	++		
<i>BRCA2</i>	++			
<i>CASP1</i>	+	+		
<i>CASP10</i>	++	+		
<i>CASP3</i>	++			
<i>CCR7</i>	+	+++++		
<i>CD5</i>	++	++++		
<i>CD79B</i>	++	+		
<i>CTLA4</i>	+	++++		
<i>CXCR4</i>	+	+		
<i>DAPK1</i>	+	++	++	
<i>DMD</i>		++++++		
<i>DST</i>	+	+		
<i>EZH2</i>	+	+		
<i>FUT8</i>		++		
<i>GSTP1</i>	++			
<i>H2BFA</i>		++		
<i>IGFBP4</i>		++++		
<i>IL10</i>	+	+		
<i>IL10RA</i>	+	+++		
<i>IL16</i>	++	+		
<i>IL1B</i>	+	+		
<i>IL4R</i>		+++++		
<i>IL8</i>	+	+++		
<i>IRF4</i>	+	+		
<i>ITGA4</i>		++++		
<i>LBR</i>		++		
<i>LPL</i>	+	+++++.....		
<i>MAD2L1</i>		++		
<i>MCL1</i>		+++		
<i>MMP9</i>	+	+		
<i>MYC</i>		++++++		
<i>OGG1</i>	+	+		
<i>PNMA2</i>		++		
<i>SEC14L4</i>	+	+		
<i>SELL</i>	+	+++		
<i>TNFA</i>	++			
<i>TP53</i>	+++	++++	+	
<i>TRAF1</i>	+	++++		
<i>ZAK</i>	+	+		
<i>ZNF288</i>		++++		

se nadaljuje

---

nadaljevanje

<i>HLA-DRB1</i>	++			
<i>POLB</i>	++			
<i>P2X7</i>	++			
<i>P16 (INK4a)</i>	++		++	
<i>ZAP70</i>		+++++	++	
<i>TCL1</i>	+		+	+
<i>SFRP1</i>			++	
<i>SFRP2</i>			++	
<i>SFRP4</i>			++	
<i>SFRP5</i>			++	
<i>DKK3</i>	+		+	
<i>RBL2</i>		+	+	
<i>CDKN2A</i>		+	+	
<i>KPNA2</i>		++		
<i>EIF4A1</i>		+++		
<i>GABARAP</i>		++		
<i>CDKN1B</i>		++		
<i>BIRC2</i>		++		
<i>PPP2R1B</i>		++		
<i>ATP2B1</i>		+++		
<i>MONDOA</i>		++		
<i>HEM1</i>		++++		
<i>ARHGAP1</i>		++		
<i>CEACAM1</i>		++		
<i>RFP2</i>		+++		
<i>KPNA3</i>		++		
<i>NT5</i>		++		
<i>TPM2</i>		+++		
<i>BCL7A</i>		+++++		
<i>KLF4</i>		++		
<i>SLAMF1</i>		++		
<i>PCDH9</i>		++		
<i>SLAM</i>		+++		
<i>WSB2</i>		+++		
<i>NRIP1</i>		+++++		
<i>SORL1</i>		++++		
<i>TCF7</i>		++		
<i>KLK2</i>		++		
<i>EST</i>		++		
<i>ABCA6</i>		+++		
<i>COL9A3</i>		++		
<i>FMOD</i>		++++		
<i>TNFRSF7</i>		+++		
<i>TGFBR3</i>		++		
<i>TRA</i>		++		
<i>CSNK1E</i>		++		
<i>PTPN7</i>		++		
<i>RASGRF1</i>		+++		
<i>HIVEP2</i>		++		
<i>ID3</i>		++		
<i>LEF1</i>		++		
<i>GS3955</i>		++		
<i>LTB</i>		++		
<i>IGJ</i>		++		
<i>IGL</i>		++		

se nadaljuje

---

nadaljevanje	
<i>CYBB</i>	+++
<i>GLRX</i>	+++++
<i>CSDA</i>	++
<i>CLEC2B</i>	++
<i>TKT</i>	++
<i>CD22</i>	+++++
<i>CD79A</i>	++
<i>MS4A1</i>	+++
<i>JUNB</i>	++++
<i>MOX2</i>	++
<i>CDC25B</i>	+++
<i>FCGRT</i>	++
<i>MXI1</i>	+++
<i>PTPN22</i>	++
<i>FGR</i>	+++++
<i>TTN</i>	++++
<i>PTPN12</i>	++++
<i>IL24</i>	++
<i>FCER2</i>	+++++
<i>OAS2</i>	++
<i>PPP2R5C</i>	++
<i>TNFRSF1B</i>	+++
<i>ARHGEF6</i>	+++
<i>ALOX5</i>	++++
<i>PLCL2</i>	+++
<i>TOSO</i>	++++
<i>NIFU</i>	++
<i>COL9A2</i>	++
<i>FCGR2B</i>	+++
<i>ITGB7</i>	++
<i>PMAIP1</i>	++
<i>LCK</i>	++
<i>PPP3CC</i>	++
<i>CD83</i>	++
<i>CD69</i>	++
<i>CD1C</i>	+++
<i>NME1</i>	++
<i>SLC2A3</i>	++
<i>BCL2A1</i>	++
<i>NFKB1</i>	++
<i>TNF</i>	+
<i>ID2</i>	++
<i>PAICS</i>	++
<i>IGHM</i>	+++++
<i>BUB1</i>	++
<i>IFI30</i>	++
<i>CD38</i>	++++++
<i>MADH7</i>	++
<i>LCP1</i>	+++
<i>EIF4G1</i>	+++
<i>ADA</i>	+++
<i>CTPS</i>	++
<i>EGR3</i>	++
<i>LDHA</i>	++
<i>MME</i>	++

se nadaljuje

<b>nadaljevanje</b>		
<i>HIF1A</i>	++	
<i>BLR1</i>	+++	
<i>CDC2</i>	++	
<i>CENPF</i>	++	
<i>PLCG2</i>	++	
<i>RGS13</i>	++	
<i>CLLU1</i>	++	
<i>CCND1</i>	+++++	+
<i>CCND3</i>	+++++	
<i>NSF</i>	++	
<i>GH1</i>	++	
<i>Ror1</i>	+++	
<i>CD23</i>	++	
<i>PAX5</i>	++	
<i>CD21</i>	++	
<i>TOP2A</i>	++	
<i>EVI2B</i>	++	
<i>MNDA</i>	++++	
<i>ATF2</i>	++	
<i>PBX3</i>	++	
<i>TCF1</i>	++	
<i>DUSP22</i>	+++	
<i>ADAM29</i>	++++	
<i>CRY1</i>	+++	
<i>Wnt3</i>	++	
<i>AID</i>	+++++	
<i>Sept10</i>	+++++	
<i>AKAP12</i>	++++++	
<i>PEG10</i>	++++	
<i>CD82</i>	++	
<i>GAS7</i>	++	
<i>GRB2</i>	++	
<i>MAP3K1</i>	++	
<i>AICL</i>	+++	
<i>LDOC1</i>	++	
<i>APOBEC3G</i>	++	
<i>FGL2</i>	++	
<i>CCNB2</i>	++	
<i>BID</i>	++	
<i>CD18</i>	++	
<i>SELPLG</i>	++	
<i>SELP</i>	++	
<i>CD44</i>	+++	
<i>TGFBI</i>	++	
<i>ATXN1</i>	++	
<i>IGSF3</i>	++	
<i>SIAT1</i>	++	
<i>FLNB</i>	+++	
<i>TLR7</i>	++	
<i>FRY</i>	++	
<i>MCM4</i>	++	
<i>TYMS</i>	+++	
<i>HMGA1</i>	++	
<i>MKI67</i>	++	
<i>AURKB</i>	++	

se nadaljuje

nadaljevanje	
<i>BIN2</i>	+++
<i>FAIM3</i>	+++
<i>PRKCB</i>	+++
<i>CD37</i>	++
<i>DGKA</i>	++
<i>MYBL1</i>	++
<i>TTG-2</i>	++
<i>SYK</i>	++++
<i>CKS1B</i>	++
<i>MYBL2</i>	++
<i>HLA-DRA</i>	++++
<i>BTG1</i>	++++
<i>RGS2</i>	+++
<i>FOS</i>	++
<i>DUSP1</i>	++
<i>SPIB</i>	++
<i>hTERT</i>	+++ +
<i>CLECSF2</i>	++
<i>CDC42EP4</i>	++
<i>IGHG1</i>	++
<i>PIP5K1B</i>	++
<i>RUNX1</i>	++
<i>LRCH1</i>	++
<i>RPS9</i>	++
<i>RECQL</i>	++
<i>PI3K</i>	++
<i>TCL</i>	+++
<i>E2F</i>	++

Priloga B:

Polimorfne miRNA, povezane s kroni no limfocitno levkemijo (KLL)

Polimorfne miRNA, povezane s KLL	Lokacija SNP-ja
miR-140	prekurzor
miR-223	prekurzor
miR-187	prekurzor
miR-122a	prekurzor
miR-146	prekurzor
miR-92a-1	prekurzor
miR-34a	zrela oblika miRNA
miR-146a	prekurzor
miR-25	zrela oblika miRNA
miR-27a	prekurzor
miR-9-2	prekurzor
miR-96	prekurzor
miR-141	prekurzor
miR-154	zrela oblika miRNA
miR-184	prekurzor
miR-196a-2	prekurzor
miR-217	prekurzor
miR-92b	zrela oblika miRNA
miR-125a	zrela oblika miRNA, podro je, odgovorno za vezavo na mRNA (angl. seed region)
miR-211	prekurzor
miR-18a	prekurzor
miR-140	prekurzor
miR-208b	zrela oblika miRNA
miR-202	prekurzor

Priloga C:

miRNA, ki so predhodno že bile povezane s KLL, in imajo glede na podatkovno zbirk  
Patrocles polimorfne tar e

---

miRNA, ki so povezane s  
KLL, in imajo polimorfne  
tar e

---

hsa-let-7a  
hsa-let-7b  
hsa-let-7d  
hsa-let-7f-1  
hsa-let-7g  
hsa-let-7i  
hsa-miR-100  
hsa-miR-101  
hsa-miR-102  
hsa-miR-103  
hsa-miR-105  
hsa-miR-106  
hsa-miR-106a  
hsa-miR-107  
hsa-miR-10a  
hsa-miR-10b  
hsa-miR-1200  
hsa-miR-1201  
hsa-miR-1202  
hsa-miR-1203  
hsa-miR-125  
hsa-miR-126  
hsa-miR-129  
hsa-miR-130  
hsa-miR-132  
hsa-miR-133a-1  
hsa-miR-134  
hsa-miR-135b  
hsa-miR-136  
hsa-miR-137  
hsa-miR-138-1  
hsa-miR-139  
hsa-miR-140  
hsa-miR-141

se nadaljuje

nadaljevanje

hsa-miR-142  
hsa-miR-143  
hsa-miR-144  
hsa-miR-146  
hsa-miR-146a  
hsa-miR-148  
hsa-miR-150  
hsa-miR-152  
hsa-miR-153  
hsa-miR-154  
hsa-miR-155  
hsa-miR-15a  
hsa-miR-15b  
hsa-miR-16  
hsa-miR-17  
hsa-miR-181a  
hsa-miR-181b  
hsa-miR-181c  
hsa-miR-182  
hsa-miR-183  
hsa-miR-184  
hsa-miR-185  
hsa-miR-186  
hsa-miR-187  
hsa-miR-188  
hsa-miR-190  
hsa-miR-191  
hsa-miR-192  
hsa-miR-193  
hsa-miR-194  
hsa-miR-195  
hsa-miR-197  
hsa-miR-199a  
hsa-miR-199b  
hsa-miR-19a  
hsa-miR-19b-1  
hsa-miR-19b-2  
hsa-miR-200a  
hsa-miR-202  
hsa-miR-203  
hsa-miR-204

se nadaljuje

nadaljevanje

hsa-miR-205  
hsa-miR-206  
hsa-miR-208  
hsa-miR-20a  
hsa-miR-21  
hsa-miR-210  
hsa-miR-211  
hsa-miR-212  
hsa-miR-215  
hsa-miR-217  
hsa-miR-218-2  
hsa-miR-218-2  
hsa-miR-22  
hsa-miR-220  
hsa-miR-221  
hsa-miR-222  
hsa-miR-223  
hsa-miR-224  
hsa-miR-23a  
hsa-miR-23b  
hsa-miR-24-1  
hsa-miR-24-2  
hsa-miR-25  
hsa-miR-26  
hsa-miR-27a  
hsa-miR-27b  
hsa-miR-28  
hsa-miR-296  
hsa-miR-29a  
hsa-miR-29b  
hsa-miR-29c  
hsa-miR-301  
hsa-miR-30a  
hsa-miR-30b  
hsa-miR-30c  
hsa-miR-30d  
hsa-miR-32  
hsa-miR-320  
hsa-miR-324-5p  
hsa-miR-331  
hsa-miR-335

se nadaljuje

nadaljevanje

hsa-miR-33b

hsa-miR-340

hsa-miR-342

hsa-miR-34a

hsa-miR-363

hsa-miR-367

hsa-miR-7-1

hsa-miR-92

hsa-miR-93

hsa-miR-96

hsa-miR-98

hsa-miR-99a

---

UNIVERZA V LJUBLJANI  
BIOTEHNIŠKA FAKULTETA  
ŠTUDIJ BIOTEHNOLOGIJE

Eva EH

**INTEGRACIJA mikroRNA (miRNA) V GENSKE  
MREŽE PRI KRONI NI LIMFOCITNI LEVKEMIJI**

DIPLOMSKO DELO

Univerzitetni študij

Ljubljana, 2009