

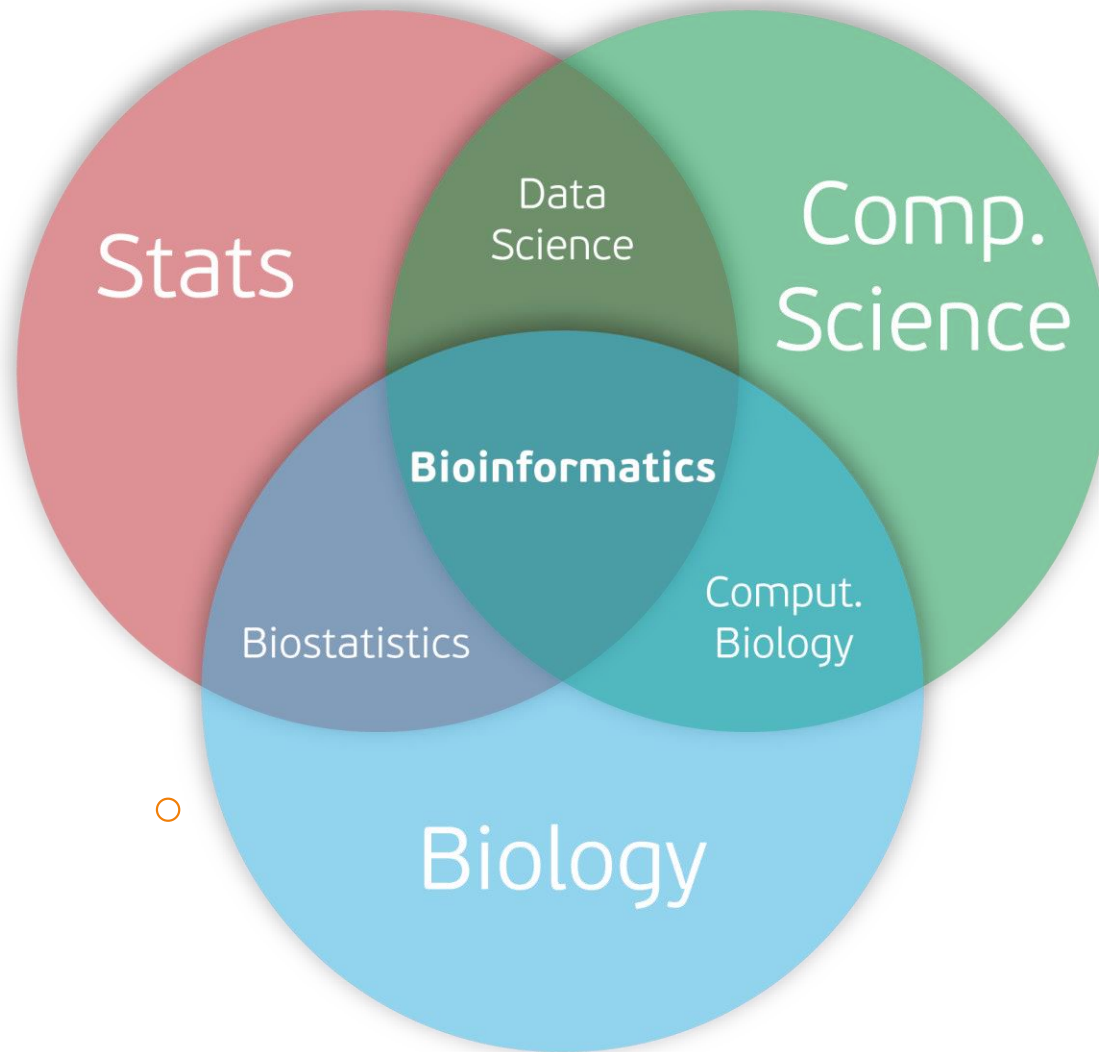
How does the “stuff” happen?

Miglè Survilaitè



Šiek tiek apie mane

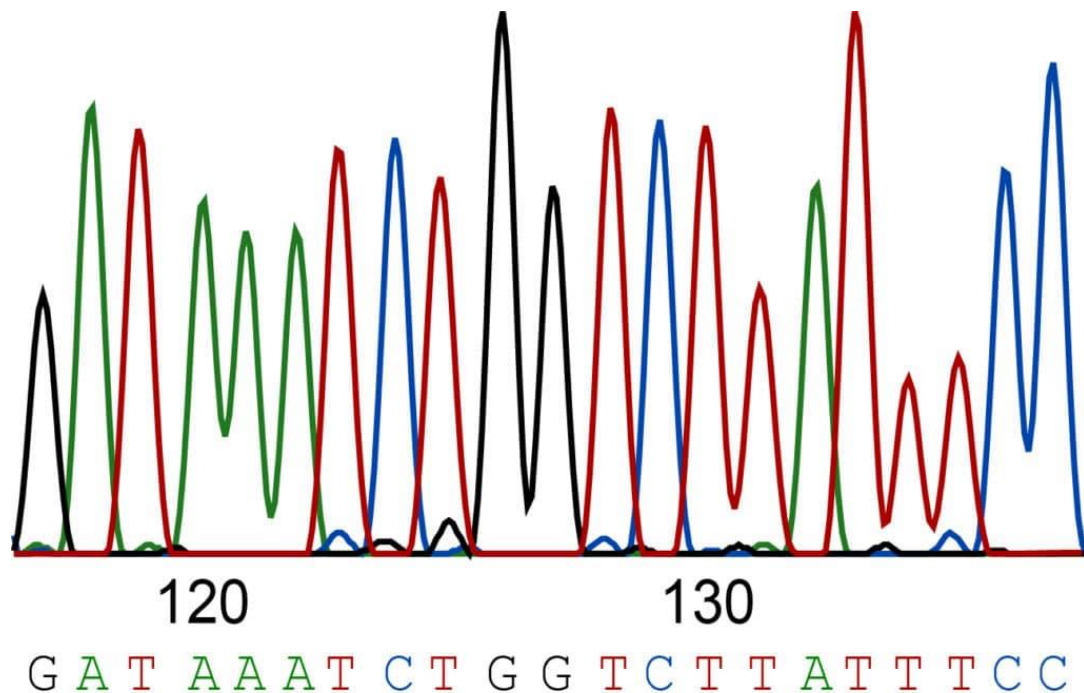
- 2012 – 2016 Bakalauro studijos Vilniaus Universitete, Biochemija
- 2016 – 2018 Magistro studijos Kopenhagos Universitete, Bioinformatika
- 2017 – ... bio.tools, bioinformatinių tools'ų kuravimas
- 2017 – ... Genominės medicinos centras, Rigshospitalet



Kas ta bioinformatika?

- Wikipedia:
 - Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data.
- Even better:
 - Bioinformatics is an interdisciplinary field that develops and applies methods and software tools for understanding biological data.
- Pagrindiniai bioinformatikų tipai?

Kaip atpažinti genus iš sekoskaitos rezultatų?



A Contig Created
From Fragments

Character
Fragments

contigs_are_contiguous_sequences_of_DNA
contigs_are_ ntigu ences_
re_conti us_se es_of_DNA
guous_sequen

B

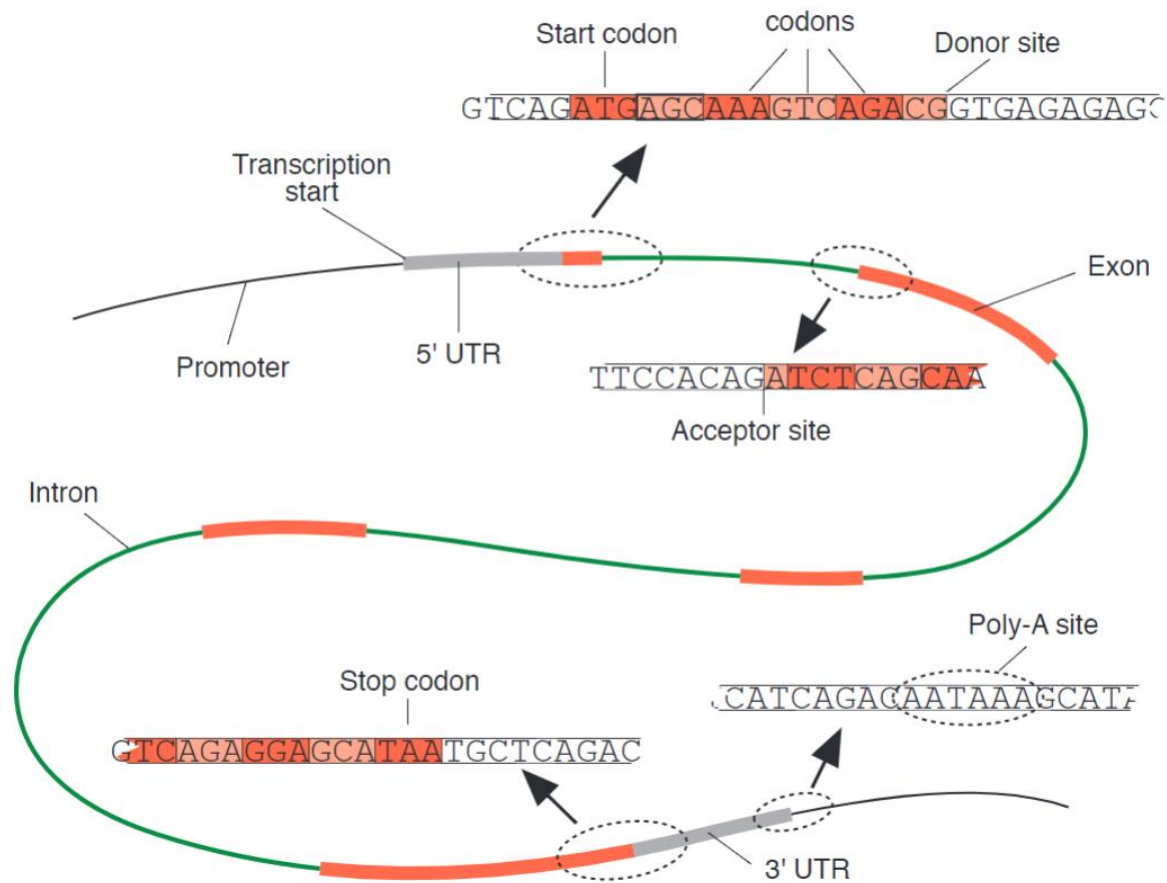
Contig

Sequenced DNA
Fragments

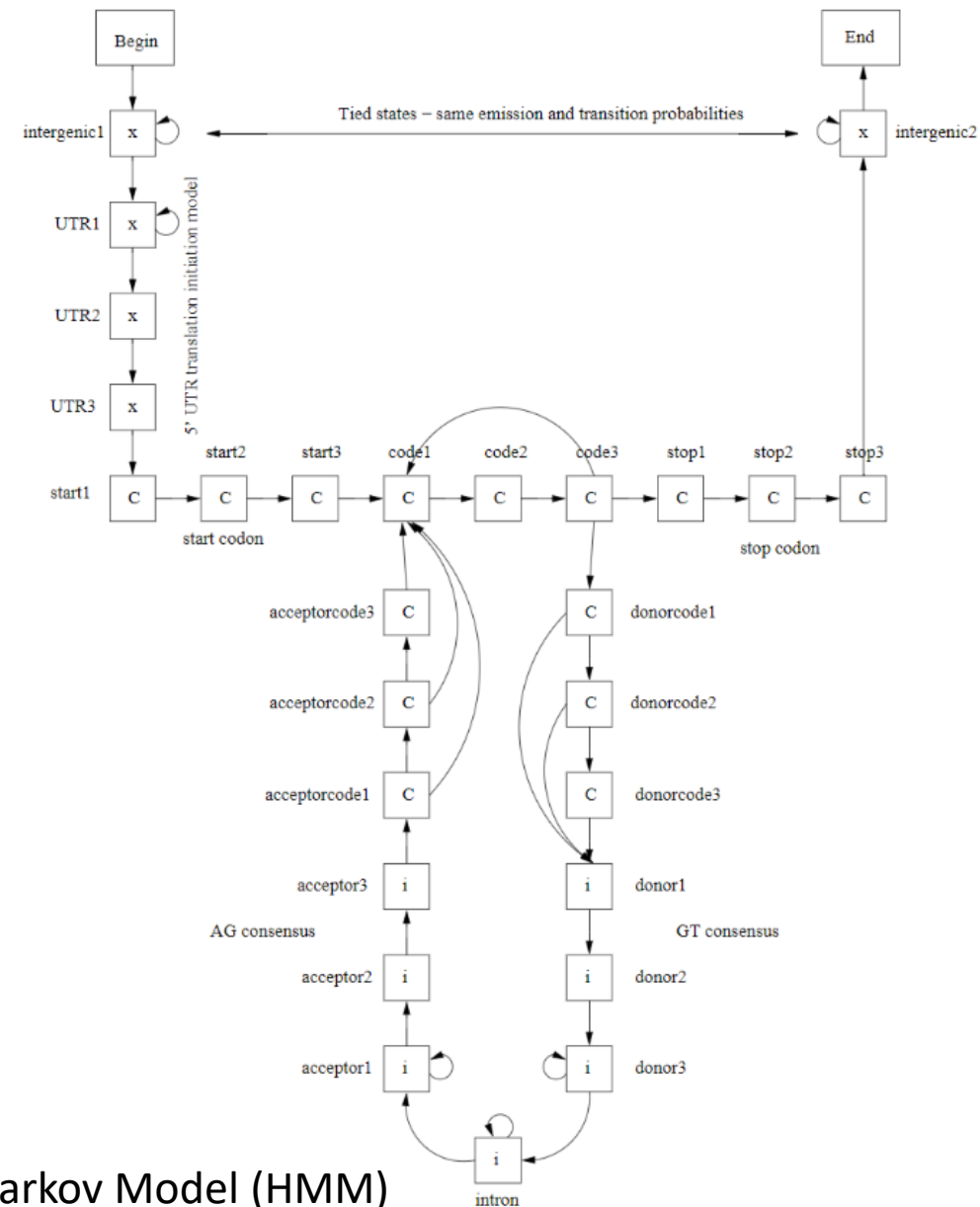
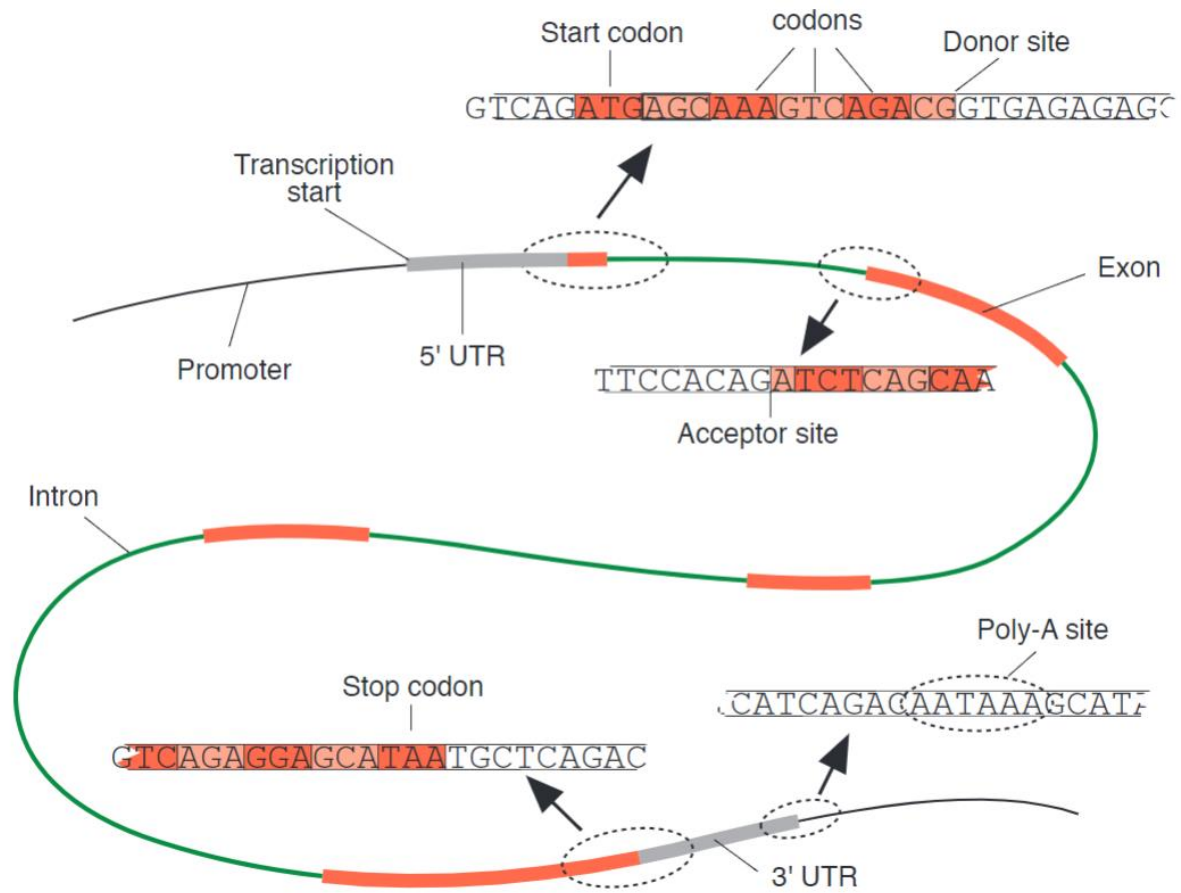


Kas yra sekoskaita?

Kaip atpažinti genus iš sekoskaitos rezultatų?



- 5'UTR
- Start kodonas
- Kodonai
- Donorinė seka
- Intronas
- Akceptorinė seka
- Stop kodonas
- Poly-A seka



Hidden Markov Model (HMM)

Kaip atpažinti genus iš sekoskaitos rezultatų?

Realybėje viskas daug sudėtingiau 😊

Ką gali bioinformatika?

- Sequence analysis
- Structural bioinformatics
- Phylogenetic analysis/population genetics
- Bioinformatics of high-throughput analysis
- Statistics, visualizations

FASTA formatas

- Daugumos sequence analysis tools'ų input formatas yra FASTA.
- FASTA failai susideda iš sekos pavadinimo (1 eilutė prasidedanti > ženklu) ir nukleotidų/aminorūgščių sekos.
- Pavyzdys:
 - > Some fancy name
 - TACAANAAAATTAGCCAGGCATGATNGTGCATGNCTATGGTCCCANNTNCNNGGNAGGCTGAGGCAGGAGNATNGNTTGAACCTGN
 - GANGNNNNGGNTNCAGTNNNNCANGTNNCNCNNNNNGCNNTCCCGNNNNNGNANNNNGNGCNNNNC
- FASTQ failai susideda iš 4 eilučių (@ prasidedantis *sequence identifier*; nukleotidų/aminorūgščių seka; + (ir *optional sequence identifier*); nukleotidų/aminorūgščių *quality scores*)
 - @NS500314:279:HFT5JBGX5:1:11101:4769:1044 1:N:0:AGTCAA
 - AGGAANTAAAATTAAATATGTCTTCNTCCACCTNAGCCAGGGATGNNANCNNCTNCTGTGAGCCCTGGACNCTNANCATGCAGCCNG
 - CNCNNNNNANGNTCGNNNNNGCNCNCNNCNGNNNNCCNNTCACANNNNNNCNANNNCNTANNN
 - +
 - 6AAAA#EEEEEEEEEEEEEEEEEEEE#EEEEEEEE#EEEEEEEEEEEE##E#E##/E#EEEEEEAEEA/EEEE#AE#/#/EEE</EEE#EE#E#####E#/#E
 - EE#####A/#E#A##/#/####6E##EE//A#####E#E###/#E/###

Sequence alignment – sekų palyginimas

Scarites	C	T	T	A	G	A	T	C	G	T	A	C	C	A	A	-	-	-	A	A	T	A	T	T	A	C
Carenum	C	T	T	A	G	A	T	C	G	T	A	C	C	A	C	A	-	T	A	C	-	T	T	T	A	C
Pasimachus	A	T	T	A	G	A	T	C	G	T	A	C	C	A	C	T	A	T	A	A	G	T	T	T	A	C
Pheropsophus	C	T	T	A	G	A	T	C	G	T	T	C	C	A	C	-	-	-	A	C	A	T	A	T	A	C
Brachinus armiger	A	T	T	A	G	A	T	C	G	T	A	C	C	A	C	-	-	-	A	T	A	T	A	T	T	C
Brachinus hirsutus	A	T	T	A	G	A	T	C	G	T	A	C	C	A	C	-	-	-	A	T	A	T	A	T	A	C
Aptinus	C	T	T	A	G	A	T	C	G	T	A	C	C	A	C	-	-	-	A	C	A	A	T	T	A	C
Pseudomorpha	C	T	T	A	G	A	T	C	G	T	A	C	C	-	-	-	-	-	A	C	A	A	A	T	A	C

Pairwise alignment – dynamic programming

	C	T	A	A	G
C					
A					
G					

Pairwise alignment – dynamic programming

		C	T	A	A	G
	-	-	-	-	-	-
C	-	1	1	1	1	1
A	-	1	1	2	2	2
G	-	1	1	2	2	3

Match (+1)

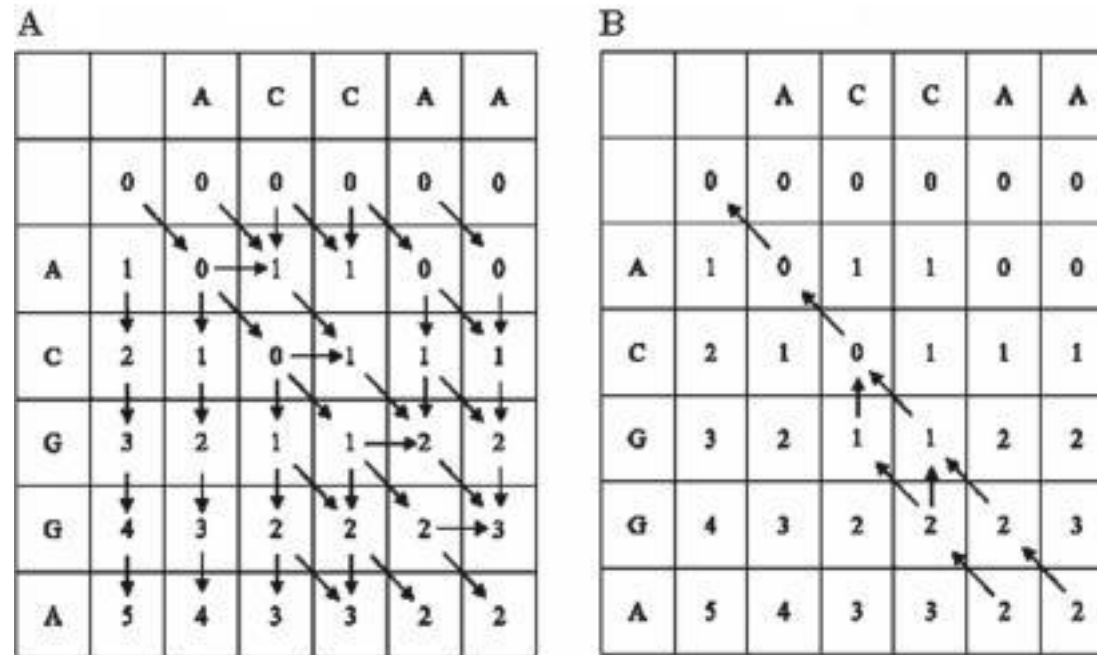
Pairwise alignment – dynamic programming

		C	T	A	A	G
	-	-	-	-	-	-
C	-	1	0	-1	-2	-3
A	-	0	0	-1	0	-1
G	-	-1	-1	-1	-2	1

Gap (-1)

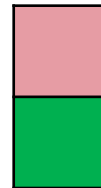
Match (+1)/Mismatch (-1)

Pairwise alignment – dynamic programming



Pairwise alignment – dynamic programming

		C	T	A	A	G
	-	-	-	-	-	-
C	-	1	0	-1	-2	-3
A	-	0	0	-1	0	-1
G	-	-1	-1	-1	-1	1



Gap (-1)

Match (+1)/Mismatch(-1)

Pairwise alignment – dynamic programming

	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
A	5	-1	-2	-1	-3	0	-2	-1	-1	-2	-1	-1	-1	-1	-2	1	0	0	-3	-2
C	-1	13	-4	-3	-2	-3	-3	-2	-3	-2	-2	-2	-4	-3	-4	-1	-1	-1	-5	-3
D	-2	-4	8	2	-5	-1	-1	-4	-1	-4	-4	2	-1	0	-2	0	-1	-4	-5	-3
E	-1	-3	2	6	-3	-3	0	-4	1	-3	-2	0	-1	2	0	-1	-1	-3	-3	-2
F	-3	-2	-5	-3	8	-4	-1	0	-4	1	0	-4	-4	-4	-3	-3	-2	-1	1	4
G	0	-3	-1	-3	-4	8	-2	-4	-2	-4	-3	0	-2	-2	-3	0	-2	-4	-3	-3
H	-2	-3	-1	0	-1	-2	10	-4	0	-3	-1	1	-2	1	0	-1	-2	-4	-3	2
I	-1	-2	-4	-4	0	-4	-4	5	-3	2	2	-3	-3	-3	-4	-3	-1	4	-3	-1
K	-1	-3	-1	1	-4	-2	0	-3	6	-3	-2	0	-1	2	3	0	-1	-3	-3	-2
L	-2	-2	-4	-3	1	-4	-3	2	-3	5	3	-4	-4	-2	-3	-3	-1	1	-2	-1
M	-1	-2	-4	-2	0	-3	-1	2	-2	3	7	-2	-3	0	-2	-2	-1	1	-1	0
N	-1	-2	2	0	-4	0	1	-3	0	-4	-2	7	-2	0	-1	1	0	-3	-4	-2
P	-1	-4	-1	-1	-4	-2	-2	-3	-1	-4	-3	-2	10	-1	-3	-1	-1	-3	-4	-3
Q	-1	-3	0	2	-4	-2	1	-3	2	-2	0	0	-1	7	1	0	-1	-3	-1	-1
R	-2	-4	-2	0	-3	-3	0	-4	3	-3	-2	-1	-3	1	7	-1	-1	-3	-3	-1
S	1	-1	0	-1	-3	0	-1	-3	0	-3	-2	1	-1	0	-1	5	2	-2	-4	-2
T	0	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	0	-1	-1	-1	2	5	0	-3	-2
V	0	-1	-4	-3	-1	-4	-4	4	-3	1	1	-3	-3	-3	-3	-2	0	5	-3	-1
W	-3	-5	-5	-3	1	-3	-3	-3	-3	-2	-1	-4	-4	-1	-3	-4	-3	-3	15	2
Y	-2	-3	-3	-2	4	-3	2	-1	-2	-1	0	-2	-3	-1	-1	-2	-2	-1	2	8

		H	E	A	G	A	W	G	H	E	E
P			-2	-2	-2	-2	-2	-2	-2	-2	-2
A			-2	-2	1	-2	1	-2	-2	-2	-2
W			-2	-2	-2	-2	-2	1	-2	-2	-2
H			1	-2	-2	-2	-2	-2	-2	1	-2
E			-2	1	-2	-2	-2	-2	-2	-2	1
A			-2	-2	1	-2	1	-2	-2	-2	-2
E			-2	1	-2	-2	-2	-2	-2	-2	1

Gap penalty -2

Pairwise alignment – dynamic programming

	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
A	5	-1	-2	-1	-3	0	-2	-1	-1	-2	-1	-1	-1	-1	-2	1	0	0	-3	-2
C	-1	13	-4	-3	-2	-3	-3	-2	-3	-2	-2	-2	-4	-3	-4	-1	-1	-1	-5	-3
D	-2	-4	8	2	-5	-1	-1	-4	-1	-4	-4	2	-1	0	-2	0	-1	-4	-5	-3
E	-1	-3	2	6	-3	-3	0	-4	1	-3	-2	0	-1	2	0	-1	-1	-3	-3	-2
F	-3	-2	-5	-3	8	-4	-1	0	-4	1	0	-4	-4	-4	-3	-3	-2	-1	1	4
G	0	-3	-1	-3	-4	8	-2	-4	-2	-4	-3	0	-2	-2	-3	0	-2	-4	-3	-3
H	-2	-3	-1	0	-1	-2	10	-4	0	-3	-1	1	-2	1	0	-1	-2	-4	-3	2
I	-1	-2	-4	-4	0	-4	-4	5	-3	2	2	-3	-3	-3	-4	-3	-1	4	-3	-1
K	-1	-3	-1	1	-4	-2	0	-3	6	-3	-2	0	-1	2	3	0	-1	-3	-3	-2
L	-2	-2	-4	-3	1	-4	-3	2	-3	5	3	-4	-4	-2	-3	-3	-1	1	-2	-1
M	-1	-2	-4	-2	0	-3	-1	2	-2	3	7	-2	-3	0	-2	-2	-1	1	-1	0
N	-1	-2	2	0	-4	0	1	-3	0	-4	-2	7	-2	0	-1	1	0	-3	-4	-2
P	-1	-4	-1	-1	-4	-2	-2	-3	-1	-4	-3	-2	10	-1	-3	-1	-1	-3	-4	-3
Q	-1	-3	0	2	-4	-2	1	-3	2	-2	0	0	-1	7	1	0	-1	-3	-1	-1
R	-2	-4	-2	0	-3	-3	0	-4	3	-3	-2	-1	-3	1	7	-1	-1	-3	-3	-1
S	1	-1	0	-1	-3	0	-1	-3	0	-3	-2	1	-1	0	-1	5	2	-2	-4	-2
T	0	-1	-1	-1	-2	-2	-2	-1	-1	-1	-1	0	-1	-1	-1	2	5	0	-3	-2
V	0	-1	-4	-3	-1	-4	-4	4	-3	1	1	-3	-3	-3	-3	-2	0	5	-3	-1
W	-3	-5	-5	-3	1	-3	-3	-3	-3	-2	-1	-4	-4	-1	-3	-4	-3	-3	15	2
Y	-2	-3	-3	-2	4	-3	2	-1	-2	-1	0	-2	-3	-1	-1	-2	-2	-1	2	8

		H	E	A	G	A	W	G	H	E	E	
		0	-2	-4	-6	-8	-10	-12	-14	-16	-18	-20
P		-2	-2	-4	-6	-8	-10	-12	-14	-16	-18	-20
A		-4	-4	-4	-3	-5	-7	-9	-11	-13	-15	-17
W		-6	-6	-6	-5	-5	-7	-6	-8	-10	-12	-14
H		-8	-5	-7	-7	-7	-7	-8	-8	-7	-9	-11
E		-10	-7	-4	-6	-8	-9	-9	-10	-9	-6	-8
A		-12	-9	-6	-3	-5	-7	-9	-11	-11	-8	-8
E		-14	-11	-8	-5	-5	-7	-9	-11	-13	-10	-7

Gap penalty -2

Pairwise alignment – gap penalty

- Scoring matrix:
 - PAM
 - BLOSUM
- Linear – netenkamas vienodas kiekis taškų už kiekvieną *gap'ą*
 - $\gamma(g) = -gd$
 - g – gap length
 - d – gap penalty
- Affine – netenkama daugiau taškų už *gap'o* iniciavimą ir mažiau už *gap'o* *extendinimą*
 - $\gamma(g) = -d - (g-1)e$
 - g – gap length
 - d – gap-open penalty
 - e – gap extension penalty

Local vs global alignment

Local Alignment

Target Sequence

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

|||| ||||| |||||

Query Sequence

5' TACTCACGGATGAGGTACTTTAGAGGC 3'

Global Alignment

Target Sequence

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

||||| ||||| |||||

5' ACTACTAGATT---ACGGATC--GTACTTTAGAGGCTAGCAACCA 3'

Query Sequence

Pairwise alignment with linear gap score

GLOBAL (NEEDLEMAN-WUNCSH)

$$F_{(i,j)} = \max \begin{cases} F_{(i-1,j-1)} + s(x_i, y_i) \\ F_{(i,j-1)} - d \\ F_{(i-1,j)} - d \end{cases}$$

$$F_{(i,0)} = -id$$

$$F_{(0,j)} = -jd$$

$F_{(i,j)}$ - matricos pozicija

d - gap penalty

s - score

LOCAL (SMITH-WATERMAN)

$$F_{(i,j)} = \max \begin{cases} F_{(i-1,j-1)} + s(x_i, y_i) \\ F_{(i,j-1)} - d \\ F_{(i-1,j)} - d \\ 0 \end{cases}$$

BLAST – Basic Local Alignment Search Tool

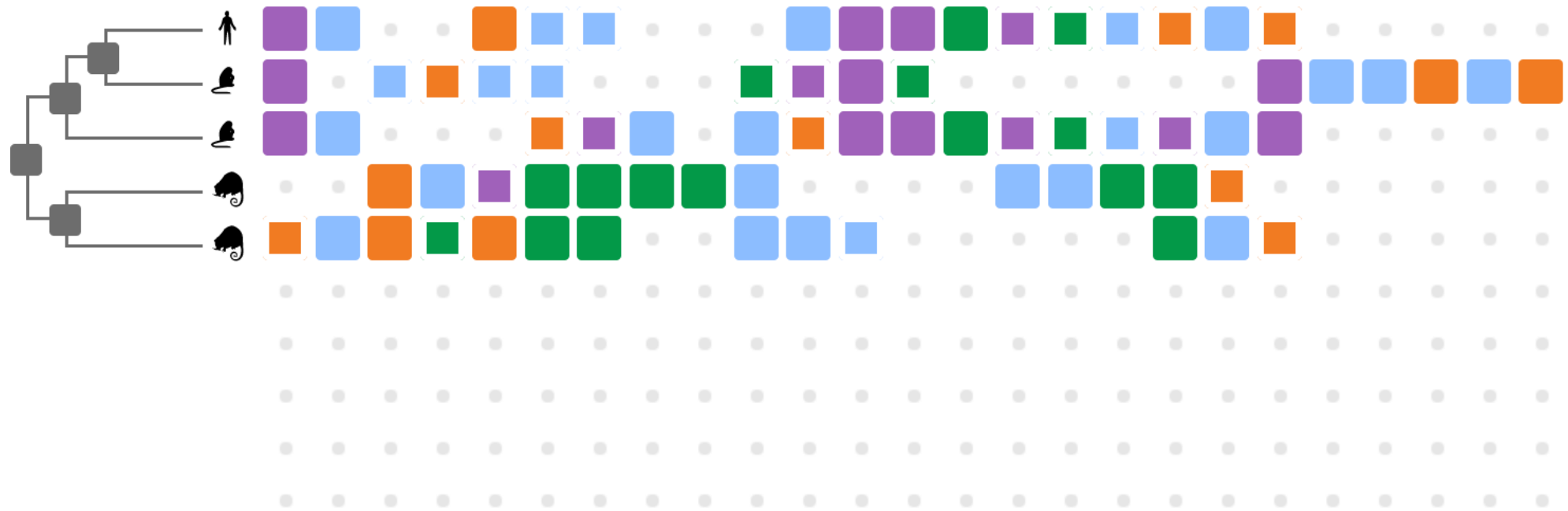
- Heuristinis metodas – mažesnis jautrumas/tikslumas, sutaupoma laiko ir vietos (*storage*)
- Identifikuojami trumpi sutampantys regionai (žodžiai)
- Jei žodžio *score* yra didesnis nei riba (*threshold*), žodis yra plečiamas (*extended*) į abi puses be *gap'ų*
- Algoritmas sustoja plėsti *alignment'ą*, kai žodis pasiekia maksimalią taškų sumą
- Naujos BLAST versijos bando sujungti kelis *alignment'us* tarp kurių yra *gap'as*

BLAST – praktinė užduotis

- Nueikite į UNIPROT duomenų bazę (<http://www.uniprot.org/>) ir raskite P38398 baltymą.
- Kairėje pasirinkite *Sequences* ir atsidarykite *Isoform1* FASTA formatu
- Nukopijuokite seką į BLAST webserver'į (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), pasirinkite blastp (baltymų sekoms) ir paleiskite su *default* parametrais
- Kokių organizmų sekas gaunate? Ar tai tik žmogaus genomo sekos? Kokias išvadas galima daryti iš pateiktų BLAST rezultatų?

Multiple sequence alignment – 5 minutes žaidimo

<http://phylo.cs.mcgill.ca/tutorial/index.html#EN>

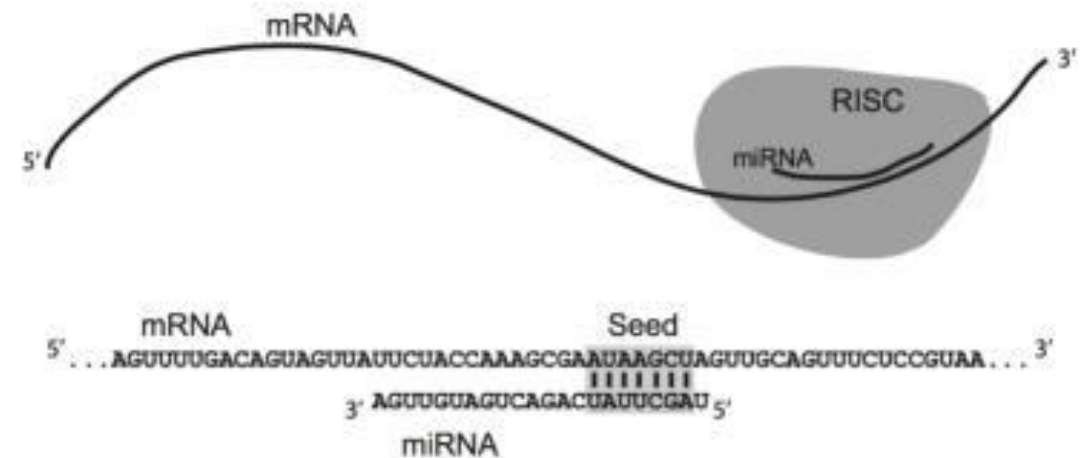


Motif discovery - užduoties įvadas

A microRNA (miRNA) is a small RNA molecule, around 22 bases long, which regulates gene expression by inhibiting the translation of mRNAs. A miRNA is incorporated into the RNA-induced silencing complex (RISC), which facilitates the base-pairing of the miRNA with the target single stranded mRNA. This yields a partially double stranded mRNA which cannot be translated by the ribosome. Most often the mRNA target site is complementary to the sequence going from positions 2-7 or 2-8 of the miRNA. The complementary region in the mRNA is called a **seed site**. Normally the seed sites are located in the 3' untranslated part of the mRNA (the 3'UTR)

An example for the human miRNA 21 (hsa-mir-21) is in the right.

How can one identify the target sites of a particular miRNA? The idea is simply to perturb the expression levels of the miRNA and then do a **motif finding analysis** on the set of up or down regulated genes. The perturbation is most commonly done in two ways. In the first one you transfect into the cell multiple copies of the miRNA of interest, thereby increasing the cellular levels of the corresponding miRNA, which in turn enhances the repression of the miRNA target genes. In the second one you introduce in the cell a sequence which is complementary to the miRNA, which will bind the miRNA and thus it will block its binding to the seed site. Thus, in contrast to the first approach, in this case the function of the miRNA will be inhibited.

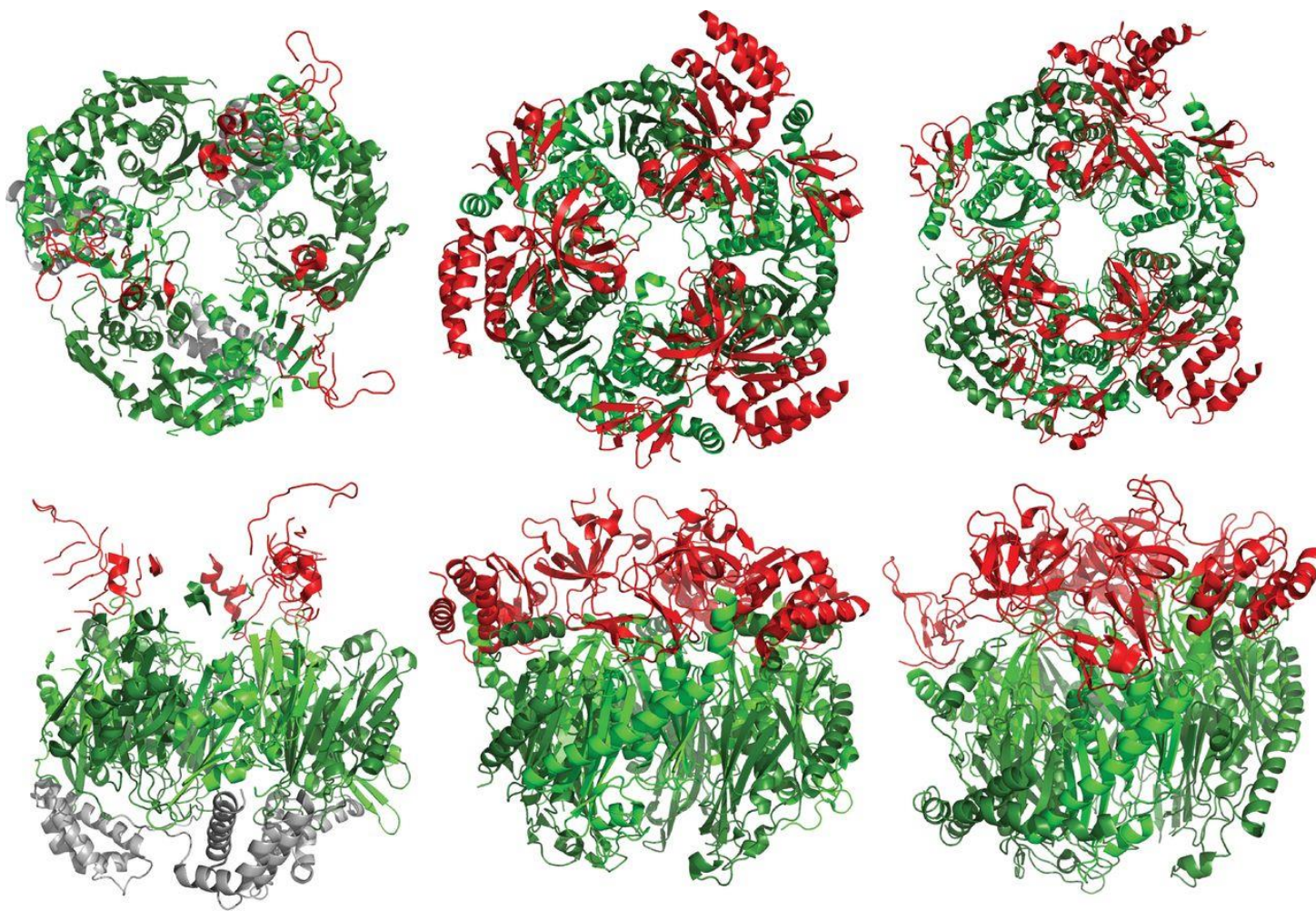


Motif discovery - užduotis

- Eksperimento metu žmogaus ląstelių linija buvo transfekuota miR-16, 500 labiausiai *up-regulated* genų buvo naudojami tolimesnėje analizėje
- Nueikite į MEME webserver'į (<http://meme-suite.org/tools/meme>)
- Įkelkite pateiktą FASTA failą (PUM2.top500.fa) į MEME webserver'į ir paleiskite su *default* parametrais
- Pasirinkite HTML output'ą, naudodamiesi pateiktais LOGO atsakykite į pateiktus klausimus
- Ar visi identifikuoti motyvai yra vienodai tikėtini būti *true positives*?
- Ar galite pasakyti kokį motyvą didžiausia tikimybė rasti pagal MEME? Jei motyvas nėra visiškai konservatyvus, kokia yra labiausiai specifinė seka?
- Kokią išvadą apie miR-16 *seed site* galima daryti iš gautų rezultatų?

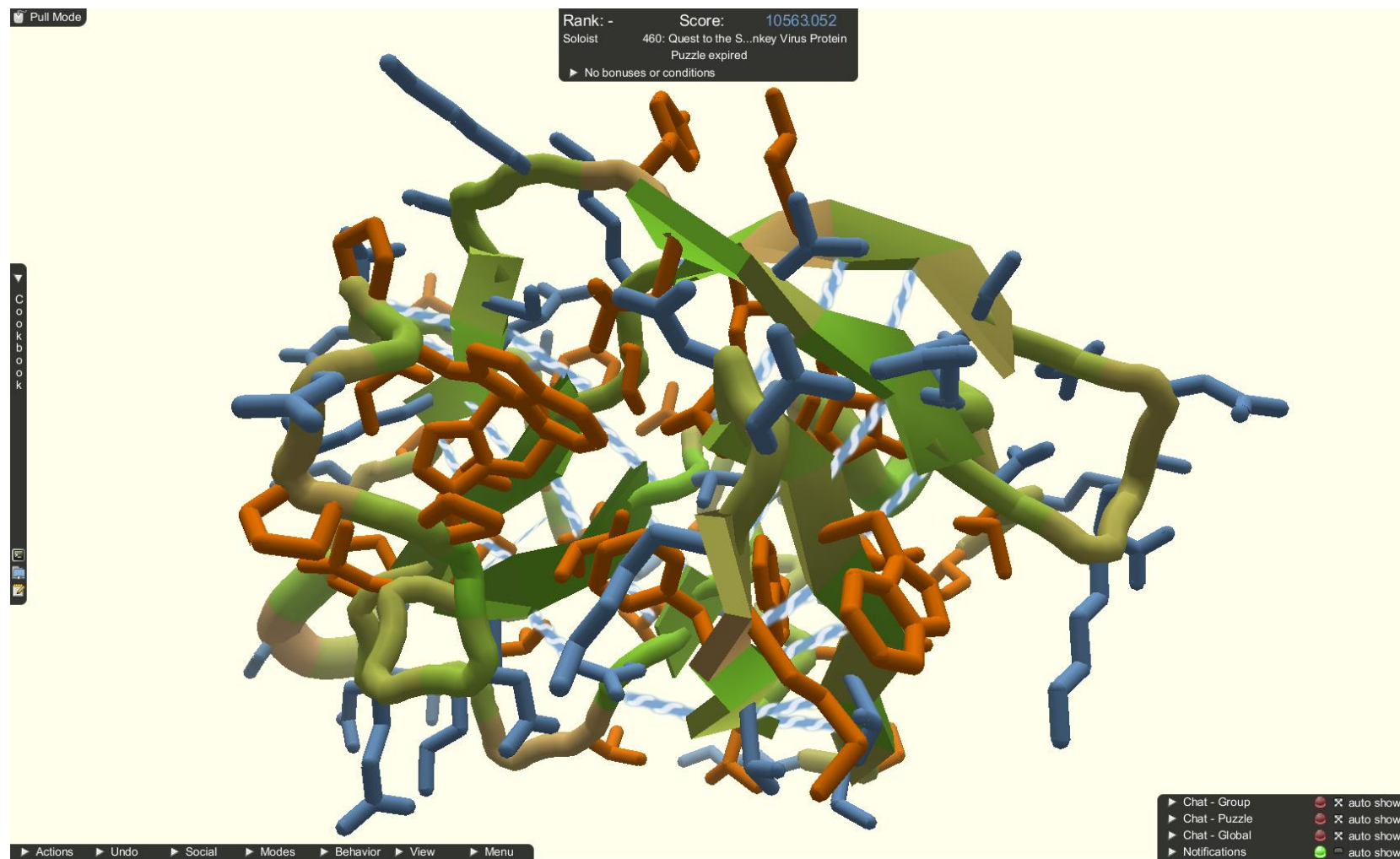
Tools'ai sekų analizei

- Pairwise alignment:
 - BLAST - <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
 - Needle (global) PROTEIN - https://www.ebi.ac.uk/Tools/psa/emboss_needle/
 - Needle (global) NUCLEOTIDE - https://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html
 - Water (local) PROTEIN – https://www.ebi.ac.uk/Tools/psa/emboss_water/
 - Water (local) NUCLEOTIDE - https://www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html
- Multiple sequence alignment:
 - Clustal W/Clustal Ω - <https://www.ebi.ac.uk/Tools/msa/clustalo/>
 - MUSCLE (protein) - <https://www.ebi.ac.uk/Tools/msa/muscle/>
 - T-Coffee - <https://www.ebi.ac.uk/Tools/msa/tcoffee/>
 - Visualization Jalview - <http://www.jalview.org/>
- Motif discovery:
 - MEME – <http://meme-suite.org/tools/meme>



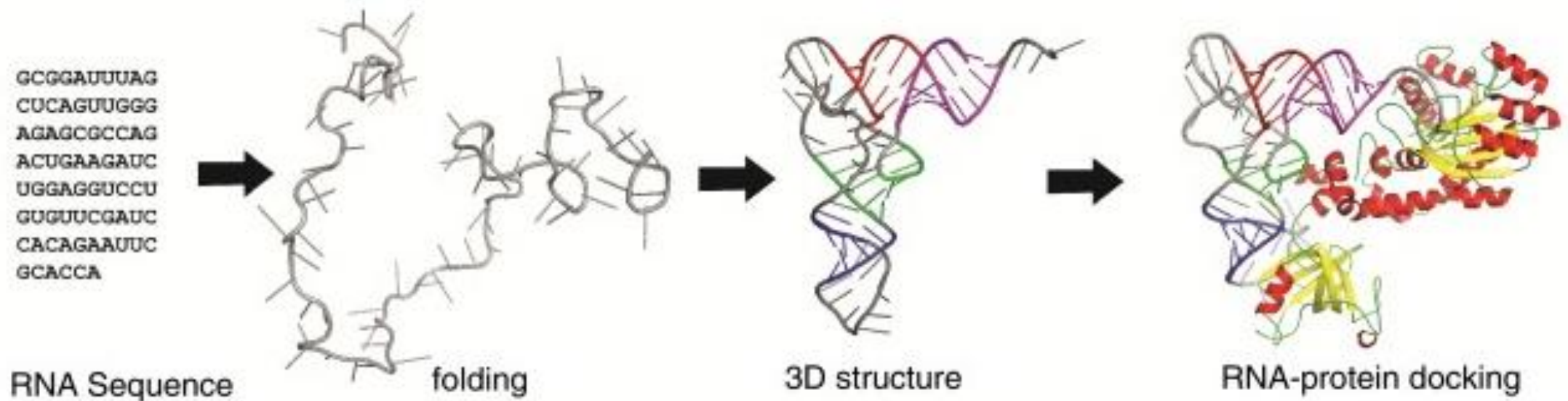
Struktūrinė bioinformatika

- Baltymų struktūrų analizė
- RNR struktūrų analizė



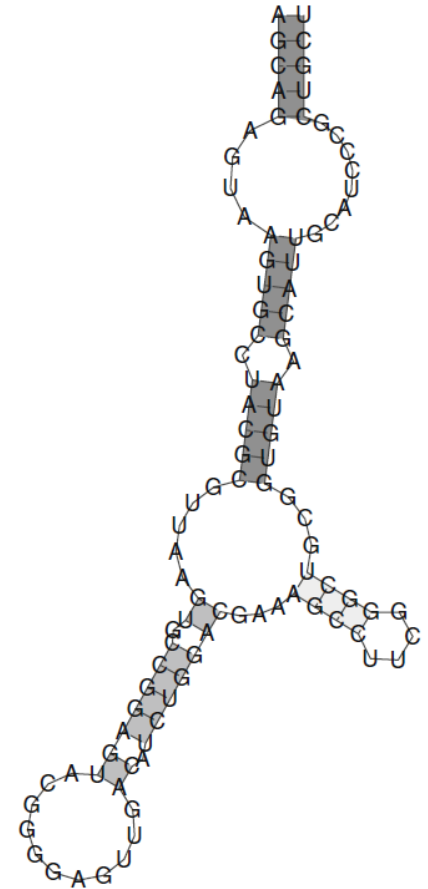
Fold.it

RNR erdvinės struktūros

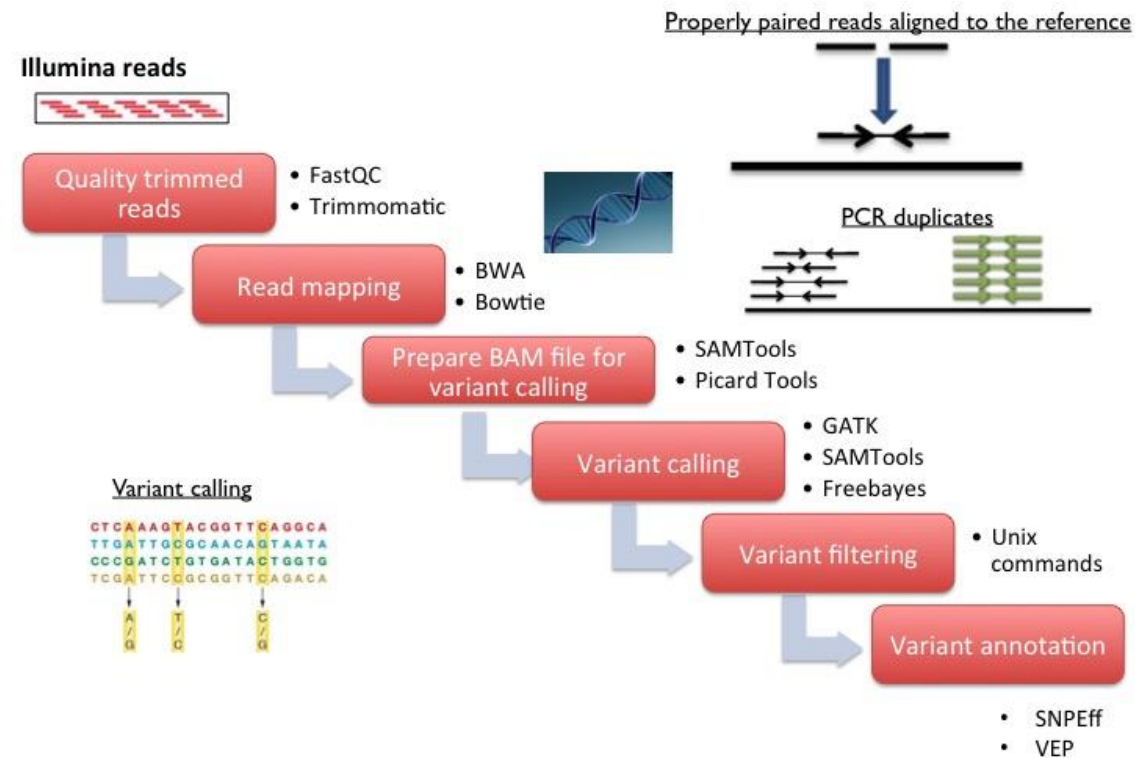


RNR struktūrų analizės užduotis

- Nukopijuokite RNR seką į RNAfold webserver'į ir paleiskite su default parametrais (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RFold.cgi>)
 - AGCAGAGUAAGUGCCUACGCGUUAAGUGCCGGAGUACGGGGAGUUGACAUCUGGA
CGAAAGCCUUCGGGCUGCGGUGUAAGCAUUGCAUCCCGCUGCU
- Palyginkite gautą struktūrą (MFE) su struktūra, pavaizduota skaidrės dešinėje
- Nueikite į RNR anotacijų duomenų bazę Rfam (<http://rfam.xfam.org/>) ir raskite **RF01831** RNR šeimą (*Jump to*), pasirinkite pirmas 10 sekų multiple alignmen'tus (Eikite į *Alignment*, tuomet *download/view* (pasirinkite *View*) *FASTA alignment in gapped format*)
- Paleiskite rastas 10 sekų su RNAalifold naudojant default parametrus (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAalifold.cgi>)
- Palyginkite gautą struktūrą su struktūra, pavaizduota skaidrės dešinėje, ir struktūra iš RNAfold webserver'io? Kodėl matote jų skirtumus? (Hint: konservatyvios sekos)



Bioinformatikos panaudojimas medicinoje

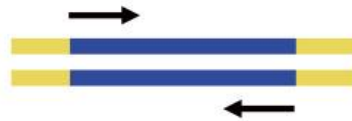


Bioinformatikos panaudojimas medicinoje

READ MAPPING

A

Paired-end
Illumina sequencing



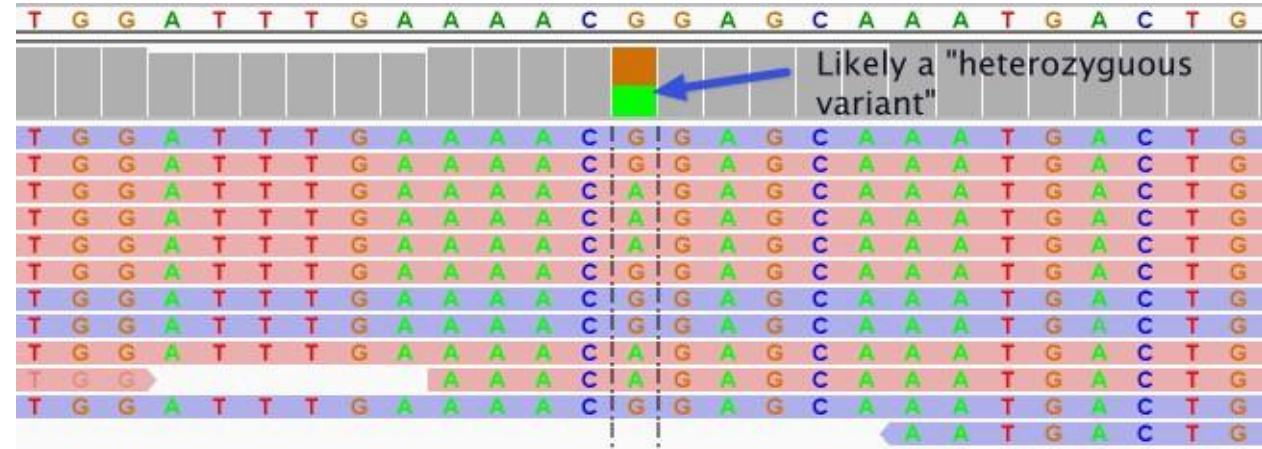
Paired-end
FASTQ sequences



Mapping to the
reference genome



VARIANT CALLING




Bioinformatikos panaudojimas medicinoje

VARIANT IDENTIFICATION

Secure | https://variants.ingenuity.com/va/hgmd.jsp?id=CM002750

HGMD Professional subscriber? Click [here](#) to view this report in HGMD Professional



HGMD accession	Reported disease/phenotype	Variant class	Gene symbol	Codon change	Amino acid change	Codon number
CM002750	Breast cancer, association with	DFP	BRCA2	AAT-CAT	Asn-His	372
The N372H substitution exhibits a shift in polarity from polar to positively charged and displays an increase in Kyte-Doolittle hydrophobicity from -3.5 to -3.2. Approximately 0.21% of missense mutations in HGMD are Asn-His. The mutation occurs 3047 amino acids from the end of the protein.						
Literature citation		Citation type	Support BETA	Comments/notes		
1. Healey (2000) <i>Nat Genet</i> 26 : 362 PubMed: 11062481		Primary literature report		No comments		
2. Bayrakli (2011) <i>J Neurosurg Pediatr</i> 8 : 476 PubMed: 22044372		Additional phenotype		Medulloblastoma; in homozygous form.		
3. Bodian (2014) <i>PLoS One</i> 9 : e94554 PubMed: 24728327		Additional literature report		found in healthy ancestrally diverse cohort. Table S1.		
4. Feliubadaló (2013) <i>Eur J Hum Genet</i> 21 : 864 PubMed: 23249857		Additional literature report		Descr. as polymorphism in Suppl. Table 1 (online).		
5. Guidugli (2014) <i>Hum Mutat</i> 35 : 151 PubMed: 24323938		Additional literature report		IARC Class 1 - not pathogenic		
6. Hondow (2011) <i>BMC Cancer</i> 11 : 265 PubMed: 21702807		Additional literature report		Descr. as polymorphism (in Table S3, online).		
7. Hu (2008) <i>Sichuan Da Xue Xue Bao Yi Xue Ban</i> 39 : 973 PubMed: 19253839		Additional phenotype		Breast cancer, association with		
8. Johnston (2012) <i>Am J Hum Genet</i> 91 : 97 PubMed: 22703879		Additional literature report	—	found in 298/572 individuals		
9. Kote-Jarai (2011) <i>Br J Cancer</i> 105 : 1230 PubMed: 21952622		Additional phenotype	—	Prostate cancer ?; descr as unclassified variant		
10. Maxwell (2016) <i>Am J Hum Genet</i> 98 : 801 PubMed: 27153395		Additional literature report		Table S5. Final call Benign.		
11. Minucci (2015) <i>Expert Rev Mol Diagn</i> 15 : 1383 PubMed: 26306726		Additional literature report		no clinical significance. Table 8.		
12. Palli (2007) <i>BMC Cancer</i> 7 : 170 PubMed: 17267207		Additional phenotype		Male BC risk		
13. Pilato (2011) <i>Breast Cancer Res Treat</i> 125 : 651 PubMed: 20352487		Additional literature report	?	?		
14. Qiu (2010) <i>Breast Cancer Res Treat</i> 123 : 487 PubMed: 20135345		Additional phenotype		Breast cancer, susceptibility to, association with; low penetrance allele.		
15. Seymour (2008) <i>Oncol Rep</i> 19 : 783 PubMed: 18788416		Additional literature report		None		
16. Wenham (2003) <i>Clin Cancer Res</i> 9 : 4396 PubMed: 14555511		Additional literature report		no association with ovarian cancer		
17. Wu (2005) <i>Cancer Res</i> 65 : 417 PubMed: 15695382		Functional characterisation		None		

DIAGNOSIS



UCSC genome browser

- Nueikite į UCSC genome browser'į (<https://genome.ucsc.edu/>) ir pasirinkite *Genomes*, Human hg19
- Nueikite į chr17:41,155,717-41,318,094
- Žemiau galima pasirinkti skirtingus track'us priklausomai nuo to, ko ieškote
- Spustelėjus ant track'o pateikiama informacija apie jį
- Galimybė kurti custom track'us

UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

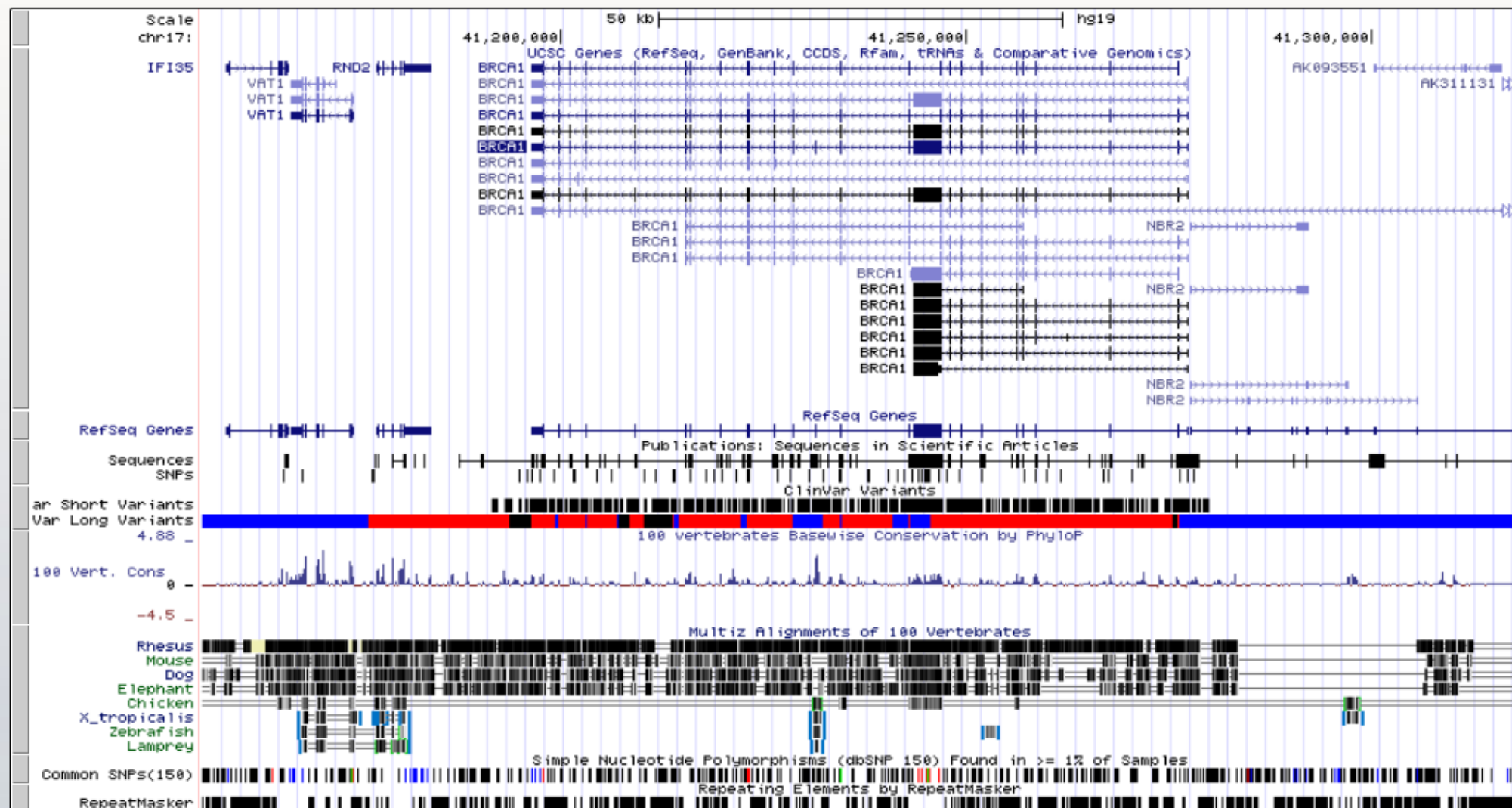
move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

chr17:41,155,717-41,318,094 162,378 bp. enter position, gene symbol, HGVS or search terms



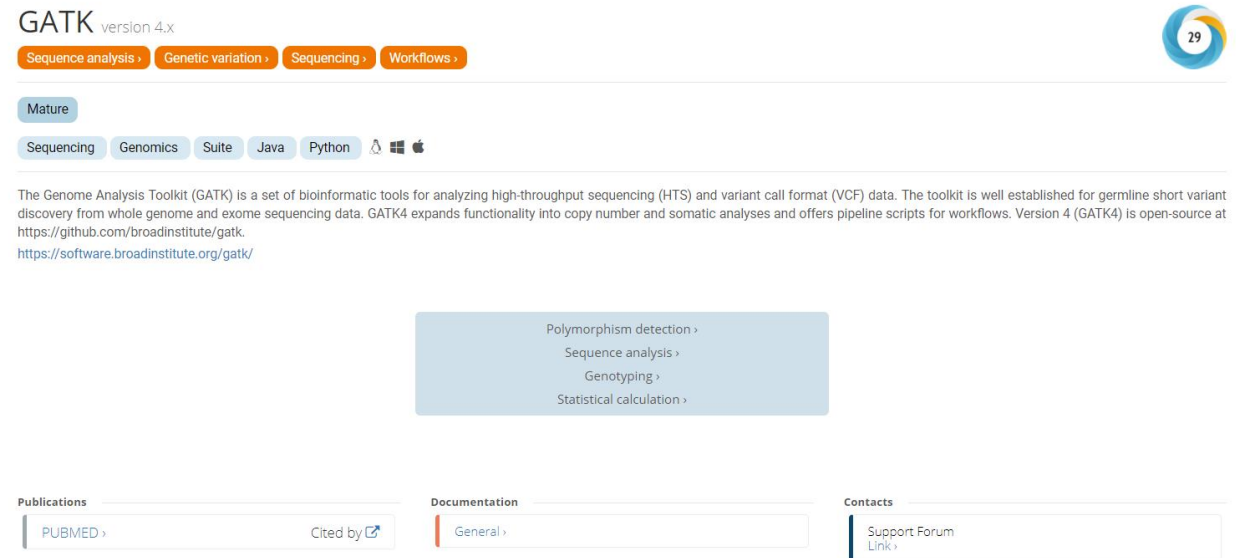
go

chr17 (q21.31) 13.3 13.2 p13.1 17p12 17p11.2 17q11.2 17q12 17q21 17q22 24.2q24.3q25.1 17q25.3



Kur rasti specifinių bioinformatinių tools'ų?

- Google
- bio.tools
- omictools.com
- mybiosoftware.com
- Kolegos, draugai, forumai



Galaxy - usegalaxy.org/ arba galaxy.pasteur.fr/

The screenshot displays the Galaxy web interface. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Help', and 'Login or Register', along with a 'Using 0%' status indicator. On the left, the 'Tools' panel lists various categories such as 'Get Data', 'Text Manipulation', 'Datamash', and 'NGS: QC and manipulation'. The central area features a large announcement for 'BOSC' (Bioinformatics Open Source Conference) with the text 'Got data?' and 'June 25-30 Portland, Oregon'. To the right of this announcement is a 'Tweets' section showing a tweet from @galaxyproject. On the far right, the 'History' panel shows an 'Unnamed history' which is currently empty, with a message indicating that users can load their own data or get data from an external source. The bottom of the image shows a Windows taskbar with various application icons and a system clock indicating 12:14 on 3/28/2018.

Galaxy is an open source, web-based platform for data intensive biomedical research. If you are new to Galaxy [start here](#) or consult our [help resources](#). You can install your own Galaxy by following the [tutorial](#) and choose from thousands of tools from the [Tool Shed](#).

Got data?

BOSC

June 25-30 Portland, Oregon

Tweets by @galaxyproject

Galaxy Project Retweeted

Ntino Krampis @bioitx
Data munching day is today... Thank you @galaxyproject for making it easier #metagenome #wgs #Bioinformatics from our MGS-Fast dlvr.it/NjyqDD with @nyulangone soon to be on @GigaScience

History

search datasets

Unnamed history
(empty)

This history is empty. You can [load your own data](#) or [get data from an external source](#)

Išvados

