

Single Cell Multi OMICs: izazovi i mogućnosti

Nadežda Bogdanović

Bioinformatics Engineering Lead,
APIS Assay Technologies Ltd.

Jelena Pejović Simeunović

Bioinformatics Engineer,
APIS Assay Technologies Ltd.



APIS Assay Technologies Ltd.



Manchester, UK

Sedište
Proizvodnja dijagnostičkih
testova i dijagnostika
biomarkera



Beograd, RS

Bioinformatika, AI/ML Generisanje
bioinformatičkih algoritama, alata,
rešenja i softvera



Bonn, DE

Clickmer Systems GmbH
Dizajn veštačkih Antitela i
Aptamera

Pregled Tema

- Single Cell sekvenciranje i analiza podataka
- Multi OMICs
- Izazovi u Single Cell multiomics polju bioinformatike
- Mogućnosti: interaktivna analiza i interpretacija podataka
 - Bioinformatička platforma
 - Bioinformatički tokovi (pipeline-i)
 - Interaktivne vizualizacije



Single Cell Seq vs. Bulk Seq analiza

Analiza podataka dobijenih sekvenciranjem pojedinačne ćelije

vs.

Analiza podataka dobijenih sekvenciranjem grupnog uzorka (grupe ćelija)

Single Cell Seq

Bulk Cell Seq

EUR 1,500 – 2,000

Obilje resursa i radne snage



Manje resursa i radne snage

EUR 200 – EUR 2,000

CENA

Nestalan uzorak

Jedna ćelija



Grupa ćelija

Robustan uzorak

VOLUMEN - PROPUSNOST

- Veoma precizno
- Moguće profilisanje ćelija i tkiva
- Nedostajući podaci
- Nedostatak reference

Detaljnija genomski specifikacija jedne ćelije



Genomska specifikacija populacije ćelija

- Neprecizno
- Marginalne greške

PRECIZNOST

Kompleksna analiza



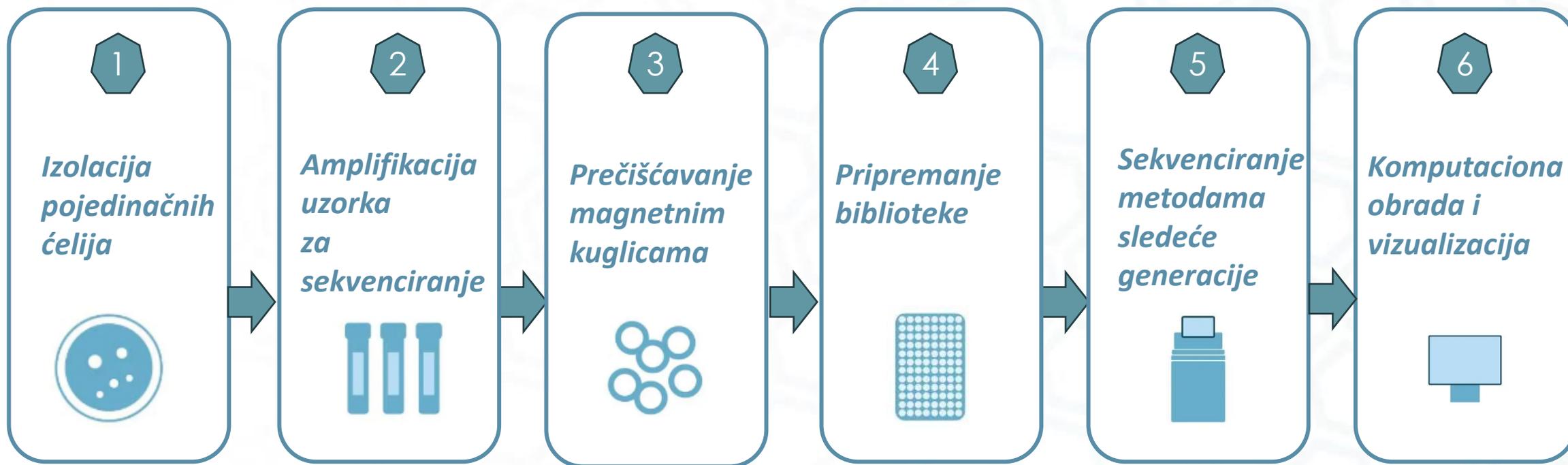
Jednostavnija analiza

- Obilje podataka
- Postojanje reference

KOMPLEKSNOŠĆ ANALIZE

Single Cell Sekvenciranje

Od izolacije ćelije do bioinformatičke analize



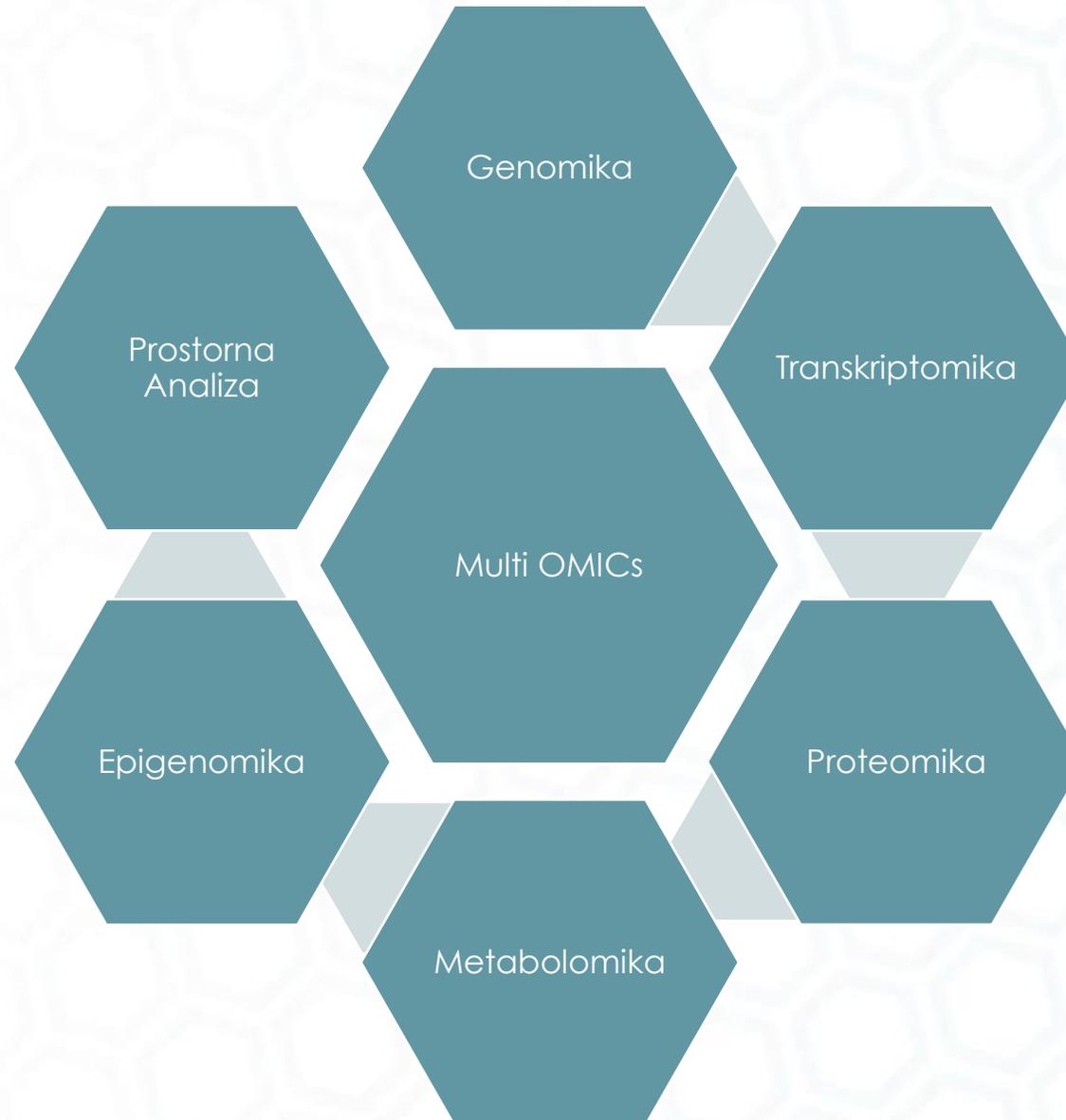
Preciznost bioinformatičke analize u velikom zavisi od laboratorijske pripreme i tehnike sekvenciranja



Proučavanje, analiza i interpretacija genomske sekvence

Proučavanje i interpretacija prostornih podataka različitih „omika“ i njihovih karakteristika

Proučava epigenetske promene: promene na nivou regulacije gena, a ne na nivou DNK sekvence



Proučavanje celog skupa transkriptnih molekula: različite vrste RNK

Proučavanje identifikacije, karakterizacije, kvantifikacije, interakcija i funkcija proteina

Identifikacija, kvantifikacija, karakterizacija i interpretacija metabolita

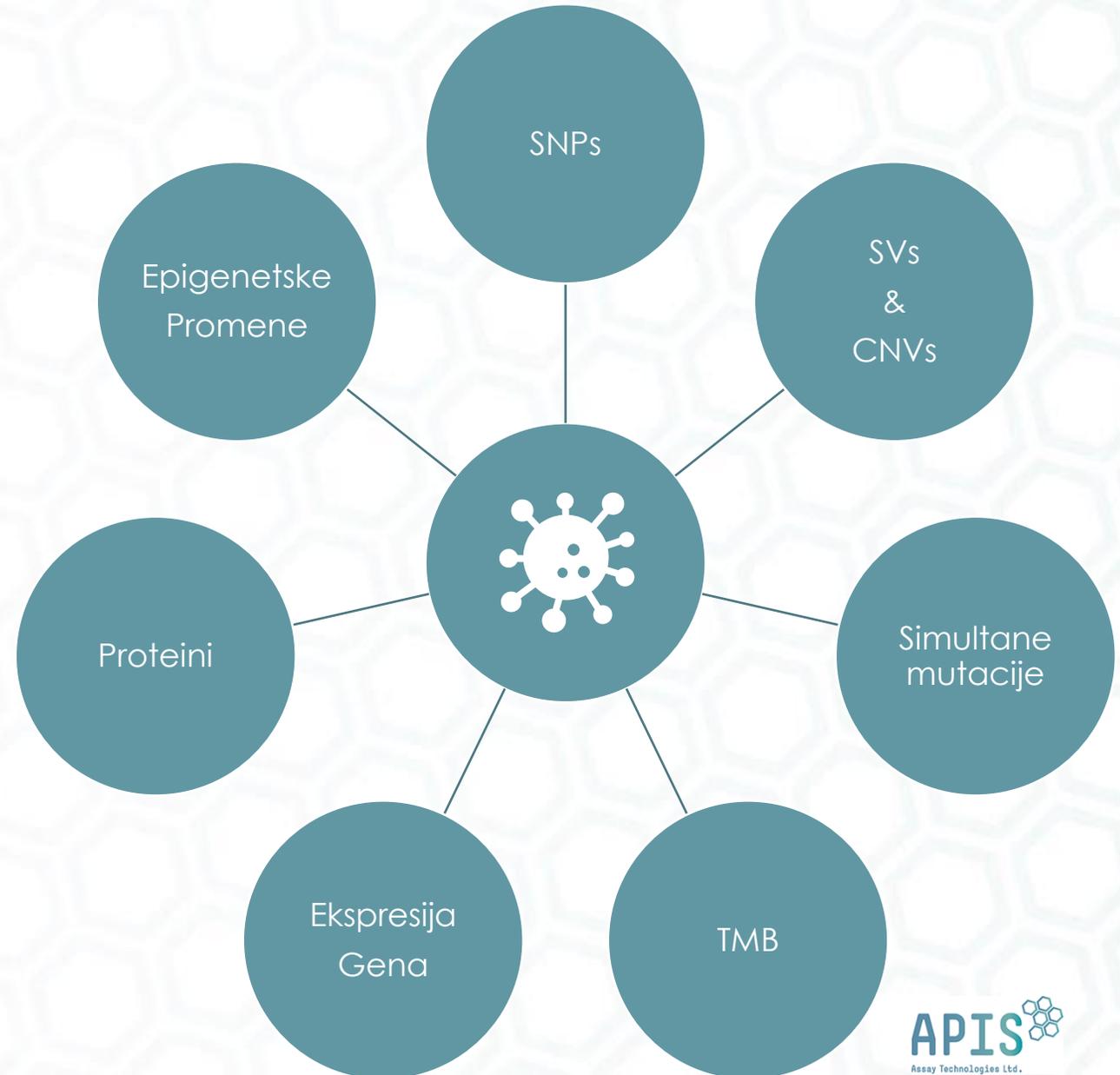


Izazovi

Šta utiče na sudbinu jedne ćelije?

Industrijski izazovi:

- Podaci:
 - Veličina
 - Kompleksnost
 - Brzina obrade
 - Čuvanje i pristup
 - Integracija
 - Nedostatak
- Standardizacija
- Tehnološki napredak
- Biološka interpretacija
- Nedostatak kvalifikovanih bioinformatičara



Mogućnosti

BaseJumper: Platforma za multiomics analizu i vizualizaciju rezultata

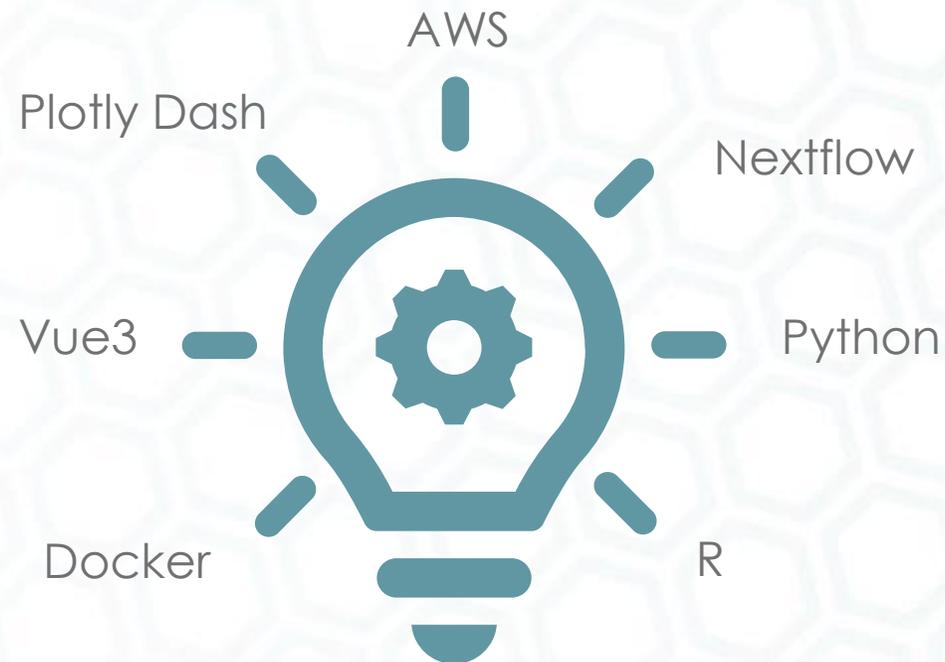


Izazovi i Mogućnosti

- Čuvanje podataka?
- Integrativna Analiza?
- Konkurentnost?
- Interpretacija?

Lako dostupno

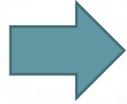
(Cognito, Amplify, AppSync, DynamoDB, S3, EventBridge, Lambda Funkcije, Batch...)



Proces kreiranja platforme

Od ideje do realizacije

Istraživanje, dizajn,
planiranje



Agilno razvijanje



Dizajn i
zahtevi



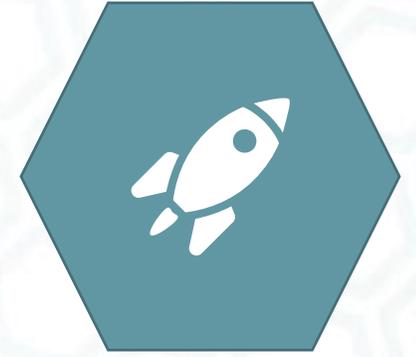
Razvijanje



Validacija i
testiranje



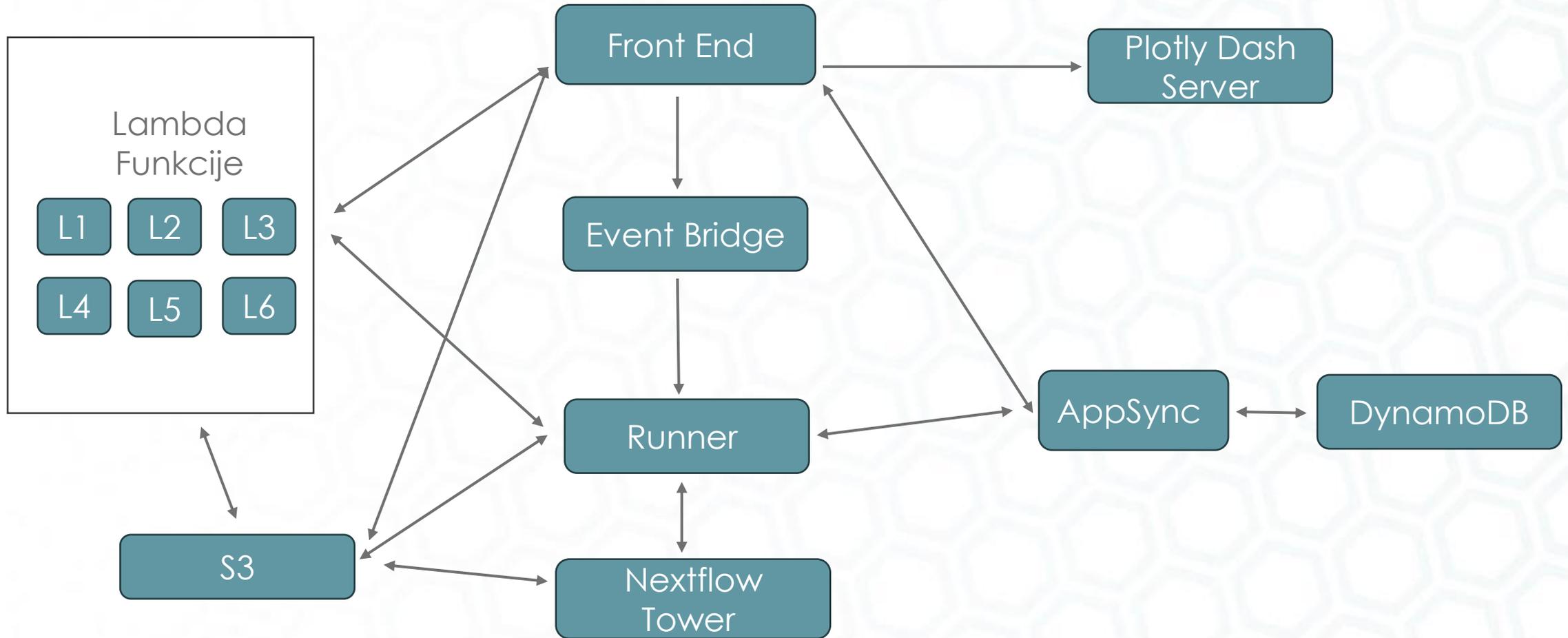
Lansiranje



Povratna informacija



BaseJumper Platforma: Komponente

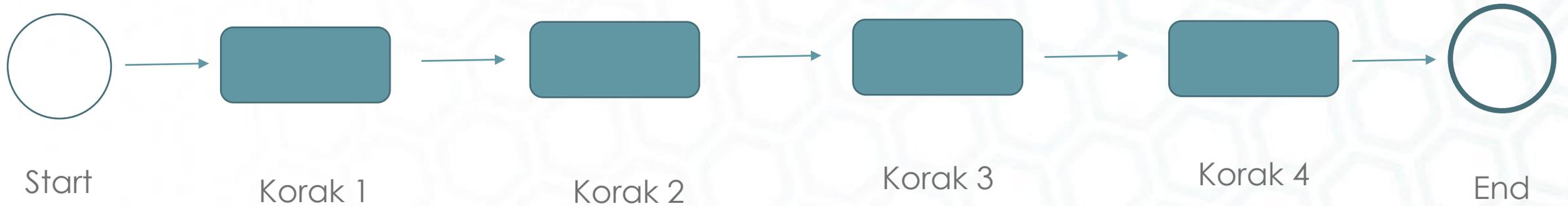


Bioinformatička Analiza



Bioinformatički Tokovi (Pipeline-i)

- Serije komputacionih koraka: ulaz u jedan korak je izlaz iz prethodnog koraka
- Komunikacija preko rezultujućih fajlova



- Koraci su crne kutije u komputacionom smislu
- Nextflow + Docker = Omotač oko postojećih alata

nextflow



Jednostavna
konfiguracija



Poliglota



Konkurentnost

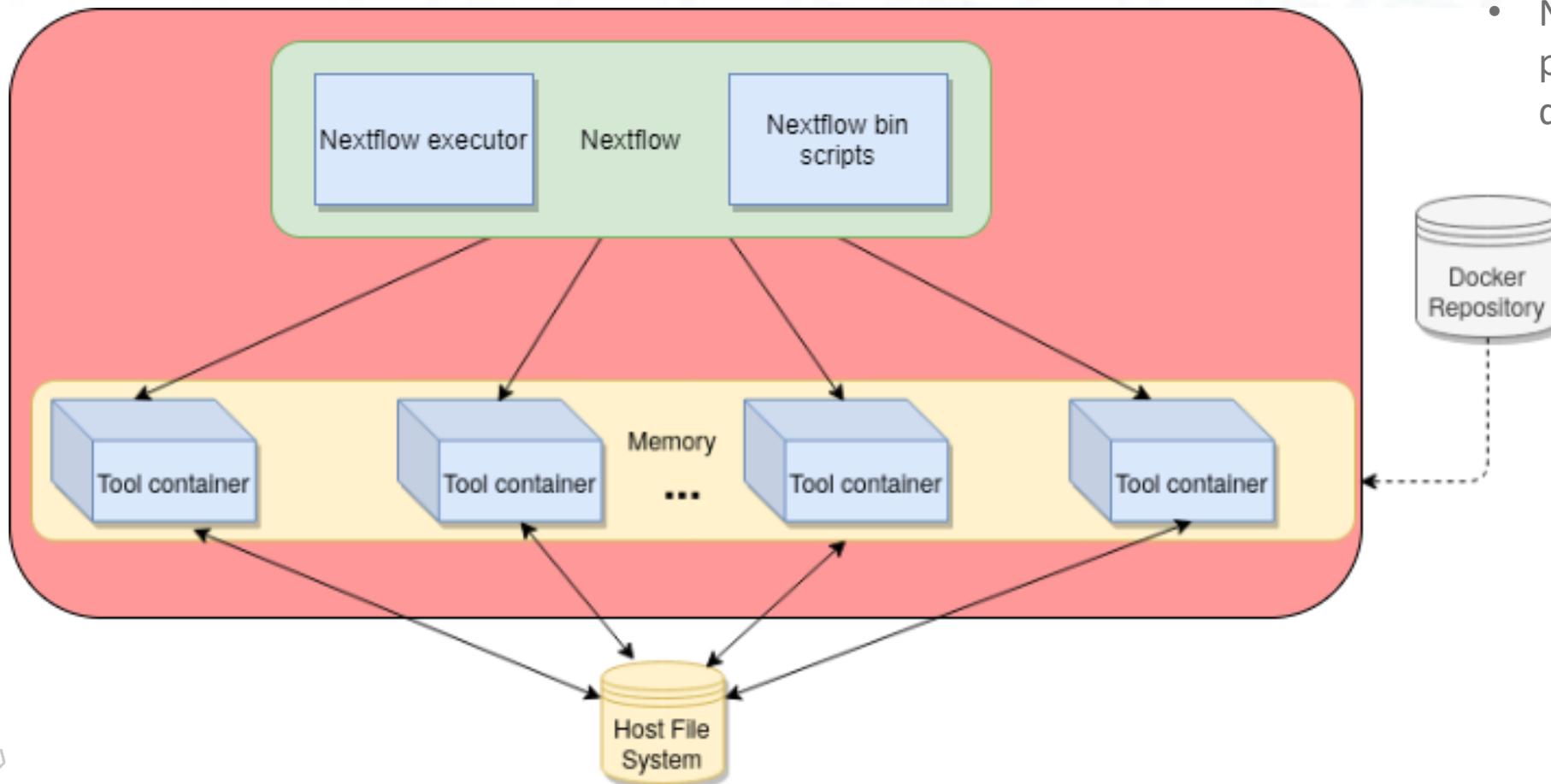


Skalabilnost



Kontejnerizacija („Pakovanje“) Alata

- Svaki alat je upakovan u kontejner i pokreće se u okviru njega
- Kontejner živi sve dokle traje pokretanje alata
- Nextflow orkestrira pokretanje kontejnera i dodeljuje resurse



Modularnost i Ponovna Iskoristivost

Moguće je pokretati više različitih tokova sekundarne analize od jednog toka primarne analize i tako redom

Primer: Preprocesiranje očitavanja

Koraci: trimovanje adaptera, filtriranje po kvalitetu, računanje metrika

Primer: Poravnanje očitavanja (WGS/WES, RNA...)

Koraci: poravnanje, deduplikacija poravnatih očitavanja, rekalkulacija

Primer: SNP poziv, SV poziv, CNV poziv, TMB estimacija, Detekcija Fuzije gena...

Tokovi
Primarne
Analize



Tokovi
Sekundarne
Analize



Tokovi
Tercijarne
Analize

Ulazi: neprerađeni FASTQ fajlovi

Izlazi: prerađeni FASTQ fajlovi sa metrikama

Ulazi: prerađeni FASTQ fajlovi

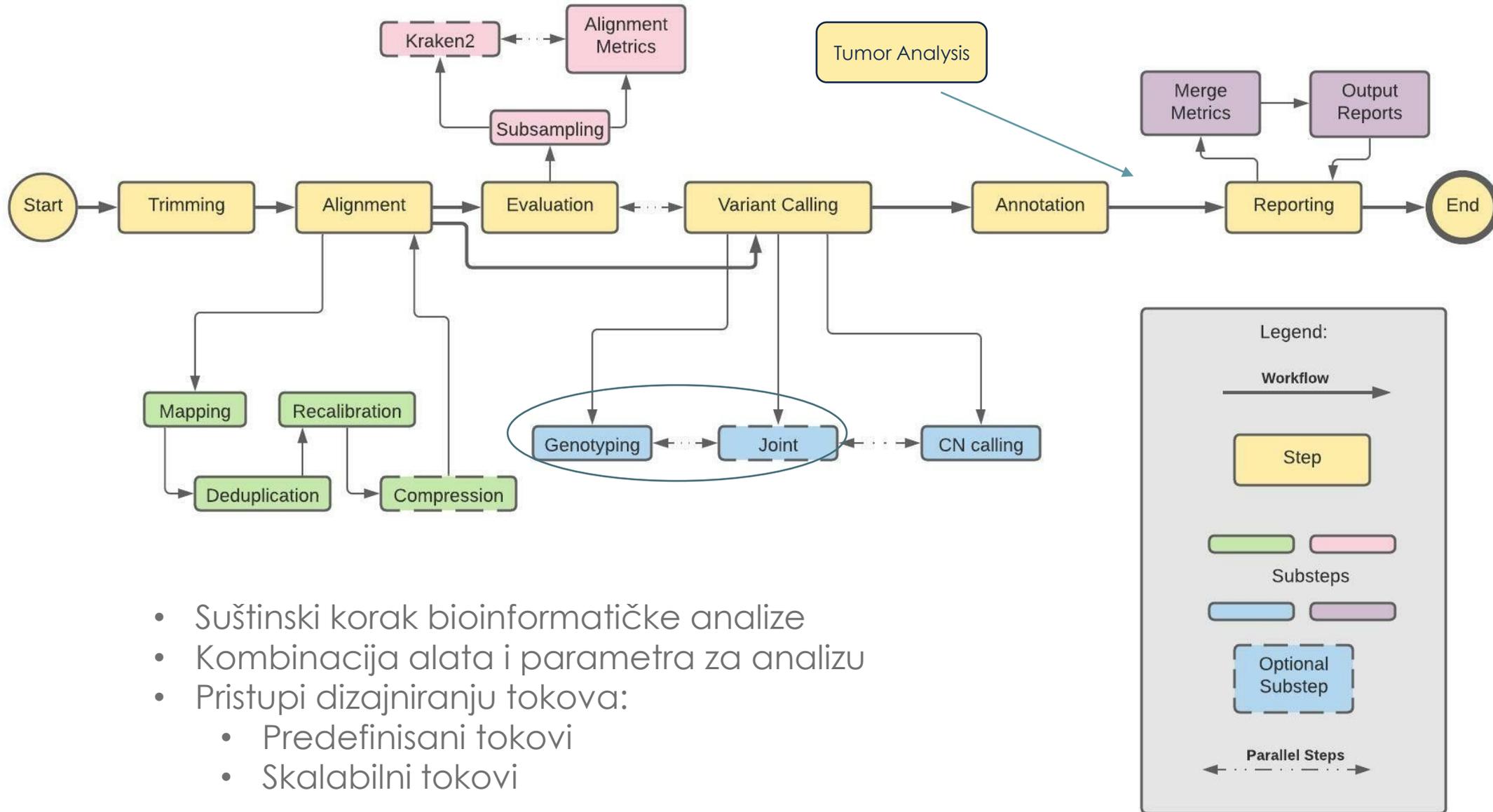
Izlazi: deduplicirani BAM fajlovi

Ulazi: deduplicirani BAM fajlovi

Izlazi: VCF fajlovi, tabularni fajlovi...



Dizajn Bioinformatičkih Tokova

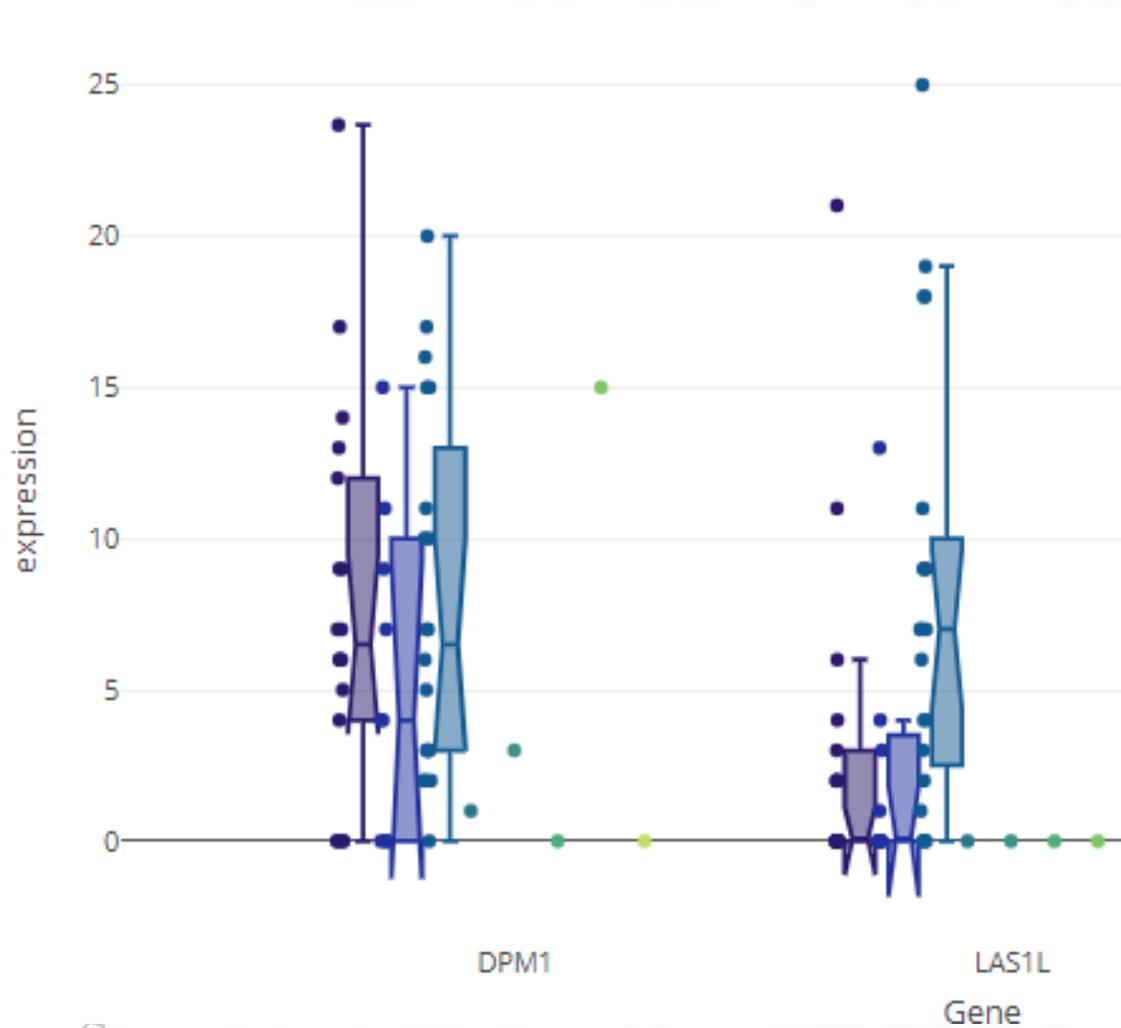


- Suštinski korak bioinformatičke analize
- Kombinacija alata i parametra za analizu
- Pristupi dizajniranju tokova:
 - Predefinisani tokovi
 - Skalabilni tokovi

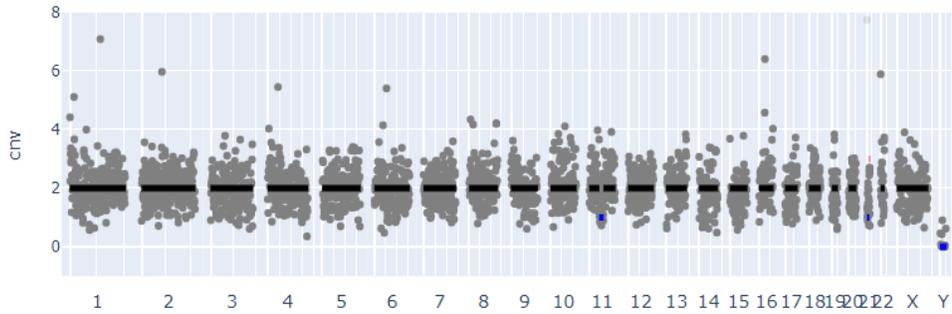
Vizualizacija i Interpretacija



Promena u broju kopija naslednog materijala i diferencijalna ekspresija gena



6500-JW-19_dots
Predicted Ploidy=2.03



Number of genes to display:

chr	start	end	gene_id	ploidy
chr21	13610228	13610749	POTED_1	1
chr21	13615122	13615237	POTED_2	1
chr21	13615396	13615570	POTED_3	1
chr21	13618348	13618455	POTED_4	1
chr21	13620157	13620295	POTED_5	1
chr21	13622804	13622875	POTED_6	1

Strukturne varijante i Fuzije gena

Multi-Omics Visualization (SV & Fusions)

Select all samples CLEAR ALL SAMPLES

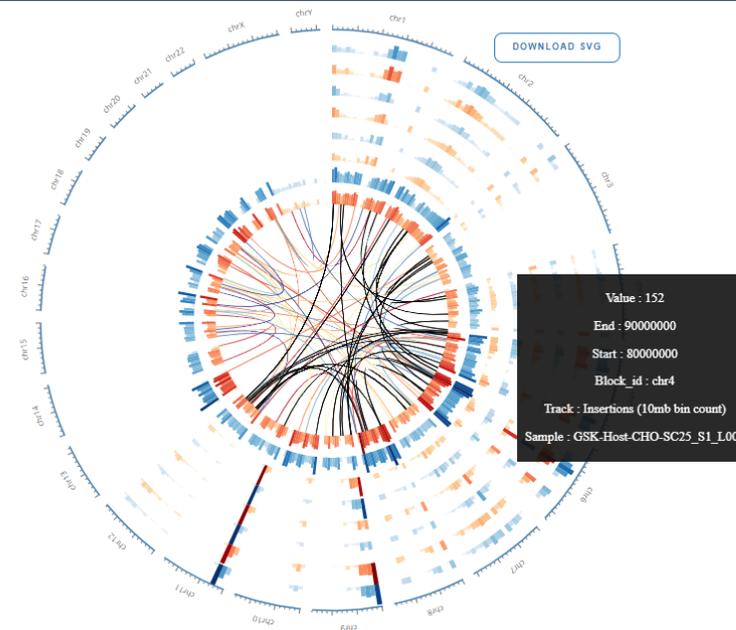
Samples: AAC7WWCM5-HCC1395MECC-SC12 | GSK-Host-CHO-SC25_S1_L001 | GSK-Host-CHO-SC33_S1_L001 | GSK-Host-CHO-SC31_S1_L001

Tracks: Indels Repeated elements Fusions Translocations

GENERATE CIRCOUS PLOT

Select all chromosomes CLEAR ALL CHROMOSOMES

chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9
 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18
 chr19 chr20 chr21 chr22 chrX chrY



Pregled i Diskusija



Pregled najvažnijih pojmova

- Single Cell Multi OMICs = Integrativna analiza podataka o pojedinačnoj ćeliji iz različitih perspektiva (naučnih polja)
- Trenutno jedno od najperspektivnijih polja u bioinformatički
- Obilje podataka koji se generišu ne mogu biti lako interpretirani bez upotrebe komputacionih metoda i bioinformatičkih rešenja
- Bioinformatičke platforme olakšavaju izučavanje ovog polja:
 - Jednostavno čuvanje i obrada podataka
 - Jednostavno pokretanja različitih bioinformatičkih analiza
 - Jednostavan prikaz i vizuelizacija podataka
 - Umanjuju potrošnju resursa
 - Ublažavaju nedostatak bioinformatičke ekspertize





Pitanja



Hvala!

Za više informacija kontaktirajte naš **EduMe** tim za edukaciju bioinformatičara: **edume.apis.belgrade@apisassay.com**



Nadežda Bogdanović
Bioinformatics Engineering Lead
APIS Assay Technologies Ltd.
nadezda.bogdanovic@apisassay.com



Jelena Pejović Simeunović
Bioinformatics Engineer
APIS Assay Technologies Ltd.
jelena.pejovic@apisassay.com



Second Floor, Citylabs 1.0
Nelson Street
Manchester
M13 9NQ



info@apisassay.com