SevenBridges

Cancer bioinformatics

Maj 2022

Marko Matić

Rak je velika grupa bolesti koja je okarakterisana nasledivim abnormalnim funkcionisanjem određene populacije ćelija u organizmu koja je izazvana kombinacijom faktora kao sto su nasledjeni genotip, uticaj spoljasnjih faktora na genom, epigenom i "microenviroment" ćelija, ...

Bujanje nepoželjnog tkiva u organizmu koje troši prirodan metabolizam i na taj način koristi energiju potrebnu za rad drugih organa.

Maligno stanje koje podrazumeva nekontrolisanu deobu ćelija.

Rak je stanje ćelije čija DNK mutira, vodi do gubitka sposobnosti ćelije da umre apoptozom i dovodi do abnormalnog rasta i taloženja takvih istih ćelija.

Kako shvatamo rak?

Rak je bolest životinja* koja je uslovljena takvim poremećajem DNK da se ćelije, u kojima je poremećaj prisutan, nekontrolisano dele, metastaziraju i ne vrše funkciju koju bi trebalo da vrše.
(kod biljaka ne metastazira)

Mutirane ćelije koje se neprekidno dele.

Rak je oboljenje nastalo kao posledica neobuzdanog mnozenja mutiranih celija, koje su prezivele odbrambene mehanizme ljudskog organizma koje su trebale da ih odstrane.

Rak je velika grupa bolesti koja je okarakterisana nasledivim abnormalnim funkcionisanjem određene populacije celija u organizmu koja je izazvana kombinacijom faktora kao sto su nasledjeni genotip, uticaj spoljasnjih faktora na genom, epigenom i "microenviroment" celija, virusima...

Bujanje nepoželjnog tkiva u organizmu koje troši prirodan metabolizam i na taj način koristi energiju potrebnu za rad drugih organa.

Maligno stanje koje podrazumeva nekontrolisanu deobu ćelija.

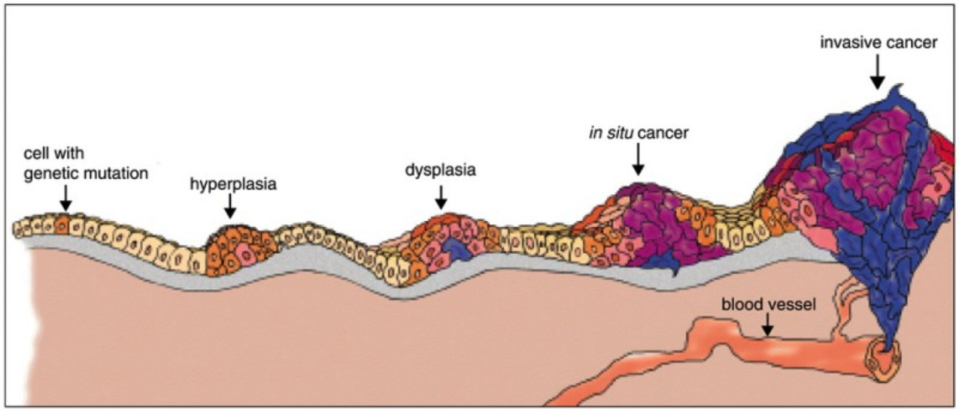
Rak je stanje ćelije čija DNK mutira, vodi do gubitka sposobnosti ćelije da umre apoptozom i dovodi do abnormalnog rasta i taloženja takvih istih ćelija.

Rak je bolest životinja* koja je uslovljena takvim poremećajem DNK da se ćelije, u kojima je poremećaj prisutan, nekontrolisano dele, metastaziraju i ne vrše funkciju koju bi trebalo da vrše.
(kod biljaka ne metastazira)

Mutirane ćelije koje se neprekidno dele.

Rak je oboljenje nastalo kao posledica neobuzdanog mnozenja mutiranih celija, koje su prezivele odbrambene mehanizme ljudskog organizma koje su trebale da ih odstrane.

Definicija(e)



Cancer Fact sheet WHO (2014):

Rak (kancer), takođe poznat kao zloćudni tumor ili maligna neoplazma, predstavlja grupu bolesti koja uključuje abnormalni rast ćelija sa potencijalnom mogućnošću da napadne ili se proširi na druge delove tela.

Karakteristike kancera

- **Abnormalnost**
  - U slucaju kancera, neke celije prestaju da se ponasaju onako kako je ocekivano, tj. prestaju da vrse bilo kakvu smislenu funkciju i tako postaju kancerogene.
- **Nekontrolisanost**
  - Fundamentalna karakteristika celija je da se umnozavaju celijskom deobom u kontrolisanom i funkcionalnom maniru. U slucaju kancera, celije se nekontrolisano dele i gomilaju, stvarajuci tumor.
- **Invazivnost**
  - Tumor ima tendenciju da se prosiri na tkiva odakle ne potice. Ova karakteristika nam omogucava da razlikujemo benigne tumore (ne prosiruju se izvan “izvornog” tkiva) i maligne tumore.

Mutacija → Besmrtnost

Normalne ćelije su smrtne i umiru u tačno određeno vreme ili u jasno definisanim situacijama

- Nakon određenog broja deoba
- Vrše samoubistvo kada im se dogode neprihvatljive greške u funkciji ili strukturi – **apoptoza**.

Geni koji se staraju za smrt ćelije se nazivaju **tumor-supresorski geni**.

Mutacije u ovim genima mogu da izazovu besmrtnost ćelije.

Besmrtnost

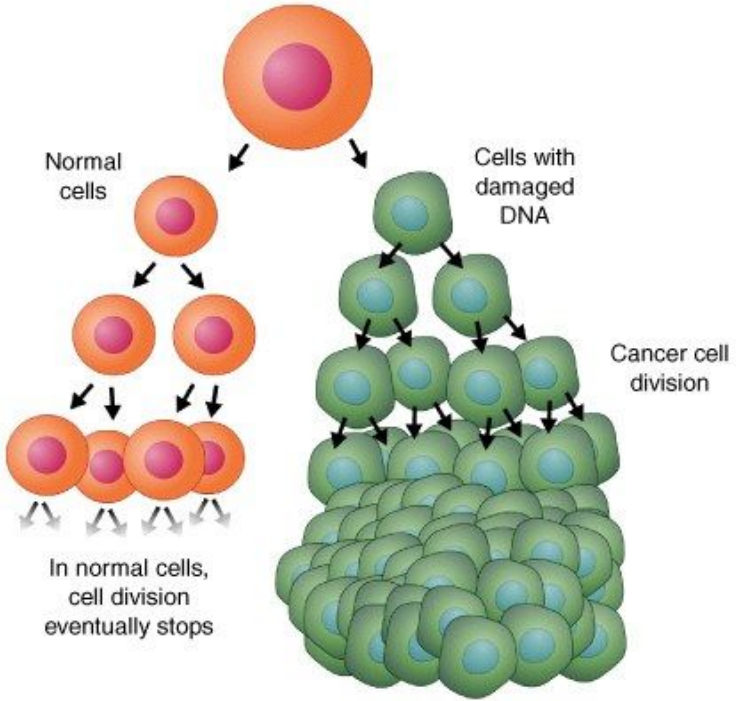
Besmrtnost, kao posledica mutacije, je centralna za razumevanje raka.

Ćelije raka se ne moraju deliti češće, samo nikada ne umiru prirodno.

Kada umiru, umiru nekontrolisano - nekroza.

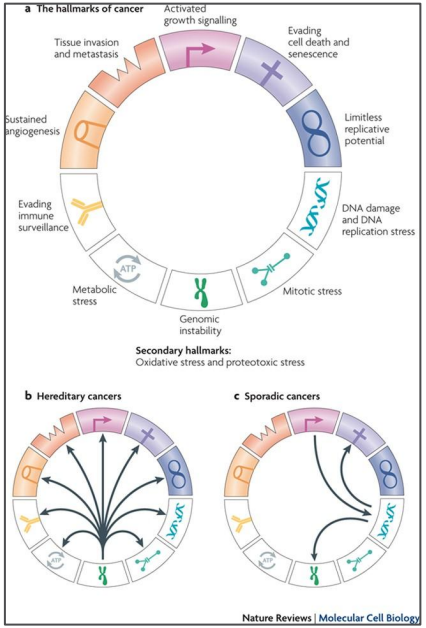
Umiru od spoljašnjih faktora, a retko od unutrašnjih.

Kroz svoj dug život nakupljaju sve više oštećenja – funkcija sve više promenjena.

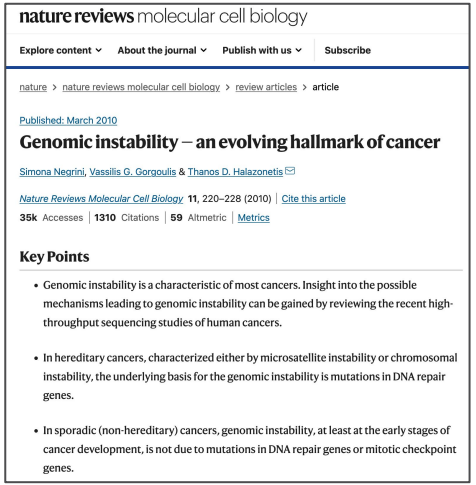


Posledice poremećaja funkcije

- Nakupljanje mutacija dovodi do **genomske nestabilnosti** - ubrzane akumulacije dodatnih mutacija.
- Kako ćelija stiče nove mutacije, njena normalna funkcija se narušava, a abnormalna funkcija sve jača.
- Narušava se fina kontrola rasta u prostoru.
- Zahtevi za energijom su sve veći – delimično zbog rasta, delimično zbog promenjene funkcije.



[Genomic instability — an evolving hallmark of cancer]



Karakteristike – revidirane

1. Izazvane poremećajem genoma (mutacija) usled raznih faktora
2. Besmrtne
   → Nekontrolisane
   → Stalno u deobi
   → Ne vrše funkciju
   → Energetski zahtevne

? Vrše invaziju i metastaziraju
? Izbegavaju odbrambene mehanizme

Tumori – benigni i maligni

Tumor se definiše kao otok tkiva, često izazvan abnormalnim rastom ćelija.

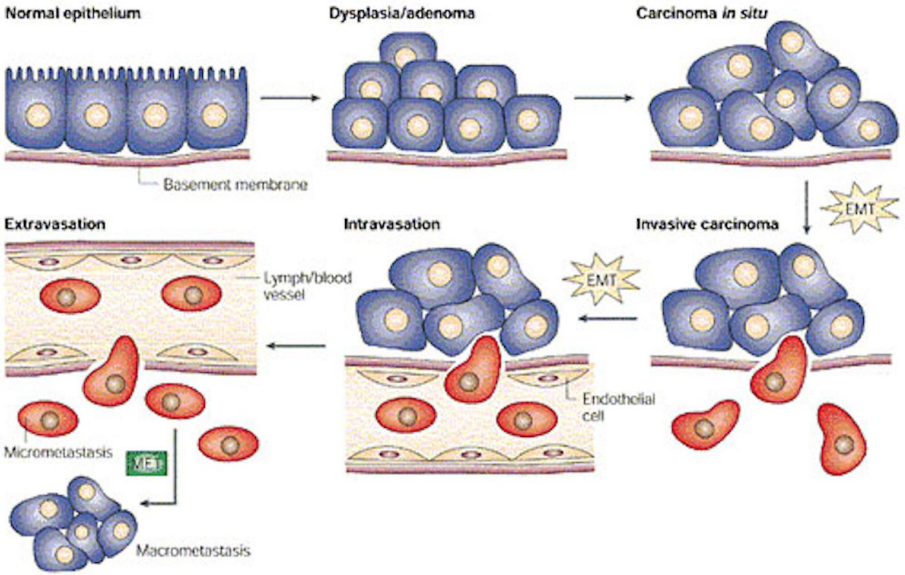
Benigni tumori:

- Rastu nekontrolisano
- Ćelije benignih tumora su besmrtne
- Mogu imati poremećenu funkciju
- Energetski zahtevni
- Pritiskaju susedno tkivo - lokalno kompresivni efekat

Maligni tumori: ?

Maligni tumori

- Invazivnost i metastaziranje su glavne odrednice maligniteta.
- Benigni tumori su lokalno **kompresivni**, maligni su lokalno **infiltrativni/invazivni**.
- Metastaziranje je mogućnost korišćenja limfotoka i krvotoka za migraciju u udaljene delove tela.



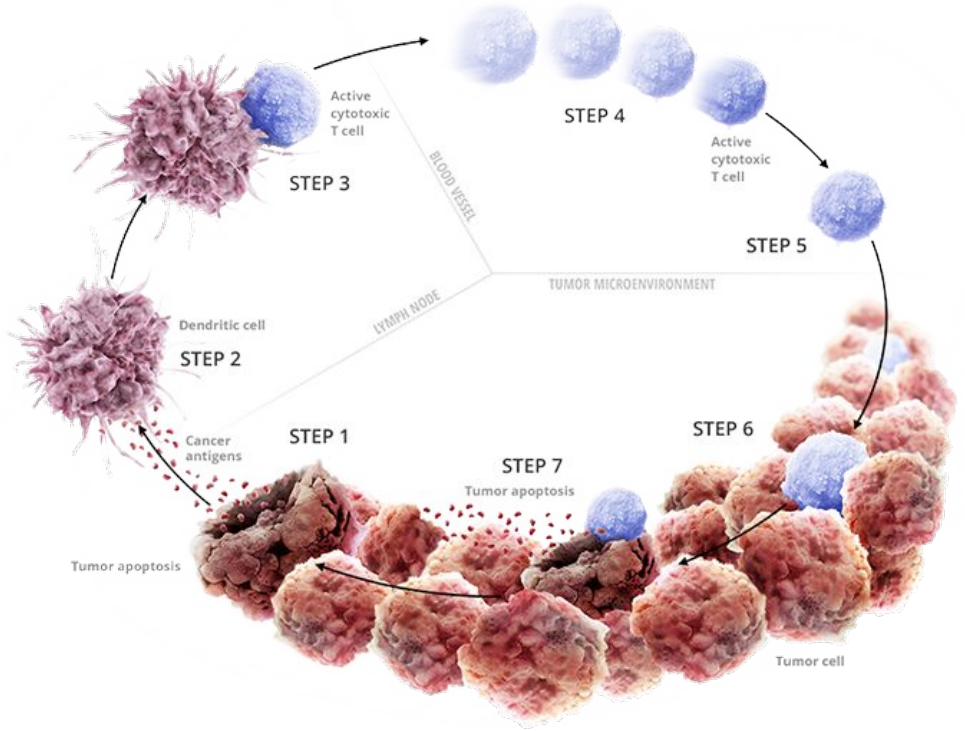
Imuni odgovor na ćelije raka

Antigeni su molekuli koji izazivaju imuni odgovor.

Promenjeni proteini usled mutacija u genomu su imunogeni – **neoantigeni**.

U teoriji, više mutacija znači jači imuni odgovor.

Više mutacija znači i agresivnije ponašanje raka.



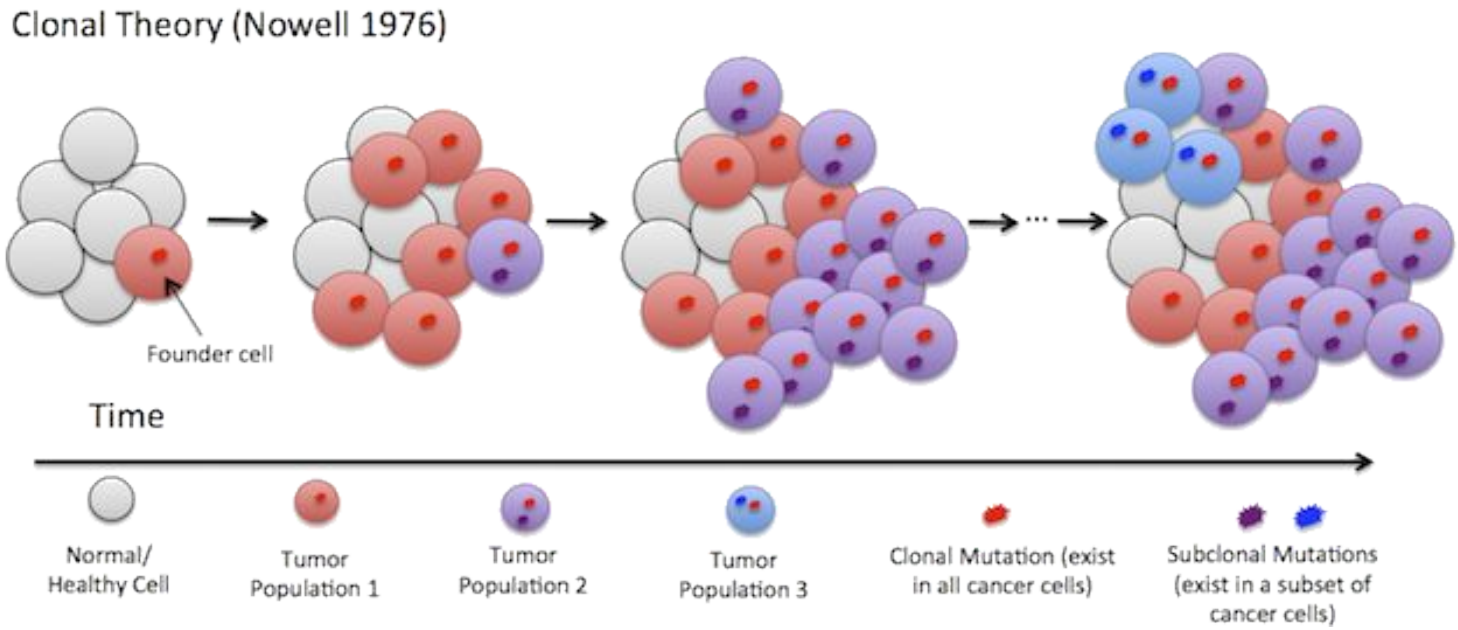
Karakteristike – kompletirane

Rak je izazvan poremećajem genoma (mutacija) usled raznih faktora.
Ćelije tumora su besmrtne i:

- Nekontrolisane
- Stalno u deobi
- Ne vrše funkciju
- Energetski zahtevne

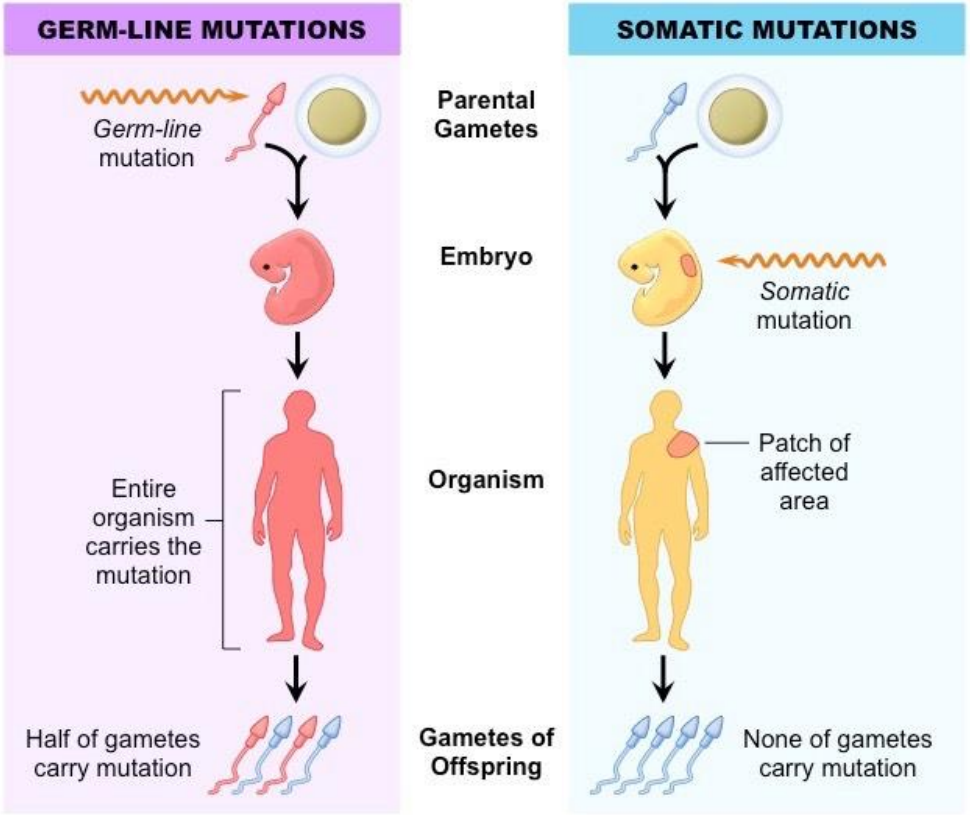
Malignitet – vrše invaziju i metastaziraju
Da bi se manifestovali – izbegavaju odbrambene mehanizme

Genomska nestabilnost



Variant Calling

Varijante (mutacije)



Variant calling koncept

- Checking if all reads at a tested position support the reference
- Reference supporting reads – REF
- Variant (Alternative) supporting reads – ALT
- Depth/Coverage = REF + ALT (number of reads covering that position)
- Variant Allele Frequency = ALT / (REF + ALT)
  - Coverage 30 – 20 REF reads, 10 ALT reads
    - VAF = 0,33 or 33%
- Genotypes
  - 0/0 = Both alleles match the reference (homozygous)
  - 0/1 = One allele matches reference and one does not (heterozygous)
  - 1/1 = Both alleles do not match reference (homozygous)
  - 1/2 = One allele contains one variant and the other another one (heterozygous)

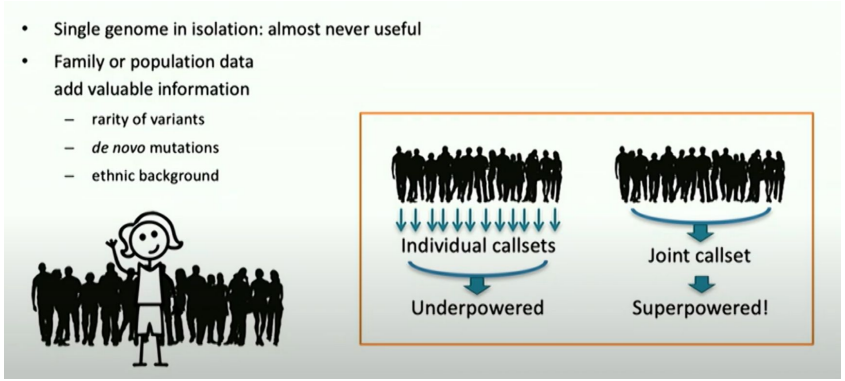
Variant calling rezultati

- Example of VCF format
- Each row represents one mutation

CHR	POS	REF	ALT	FORMAT	NA12878
1	14300	A	G	GT, VAF	0/1, 0.4
2	15367	A	C	GT, VAF	1/1, 0.9
3	25612	C	G,A	GT, VAF	1/2, ?
5	5632	TA	T	GT, VAF	0/1, 0.5
7	7824	T	TA	GT, VAF	1/1, 0.8

Variant calling tipovi

- **Multisample calling**
  - Multisample does not take other samples into consideration when deciding if there is a variant at the position in the currently tested sample.
  - **Somatic calling** – a type of multisample calling
    - Two samples are compared for differences
    - Parameters are adjusted to accommodate cancer properties
- **Joint calling** takes into consideration variants at the position in question in other samples than the one currently being tested.

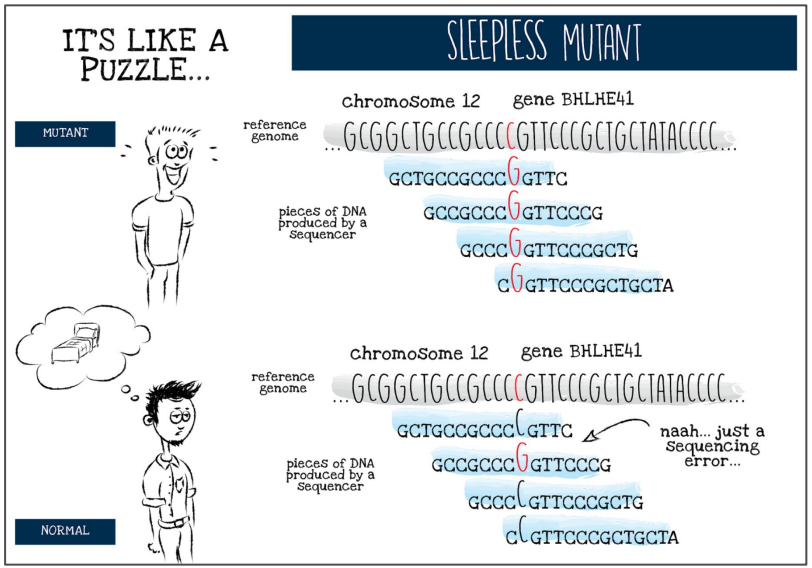


Multisample VCF primer

CHR	POS	REF	ALT	FORMAT	NA12877	NA12878
1	14125	T	A	GT, VAF	0/1, 0.43	0/0, 0.1
5	14125	A	G	GT, VAF	0/0, 0.0	0/1, 0.4

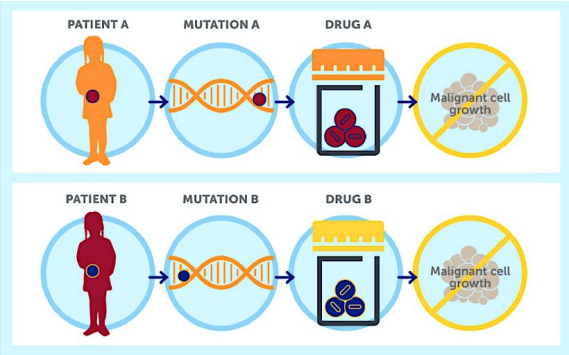
Germline vs Somatic calling

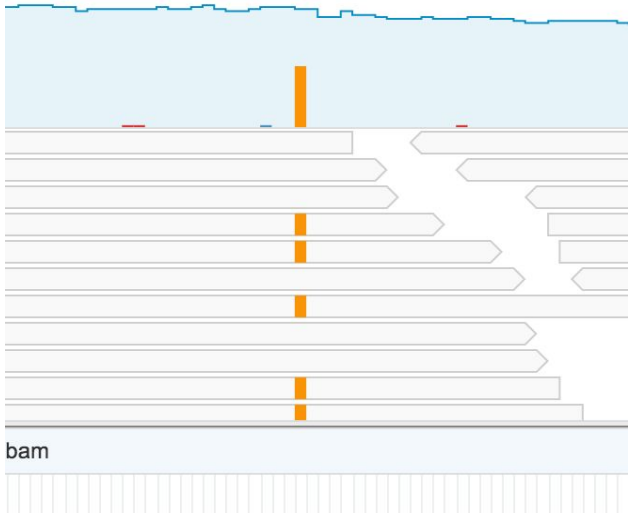
- **Germline Variant Calling**
  - Difference of a tested sample to a reference genome
  - Each sample compared to a reference genome (single sample or joint calling)
- **Somatic Variant Calling**
  - Difference of both Normal tissue sample (Normal) and Tumor tissue sample (Tumor) to a reference genome
  - Finding mutations (usually by applying statistical methods) that occur in Tumor but are not present in Normal



Somatic calling ciljevi

- Finding variations that occur in tumor tissue but do not occur in normal tissue
  - Possible cause of cancer occurrence
  - Variations frequently occurring in the same tumor type (early diagnosis)
  - Possible targets for cancer treatment
- Engineering vaccines and medications that interact with the tumor tissue but do not damage the normal tissue





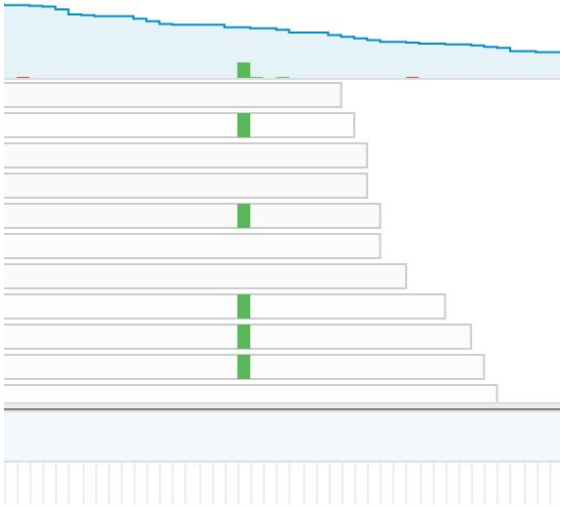
GERMLINE MUTATION

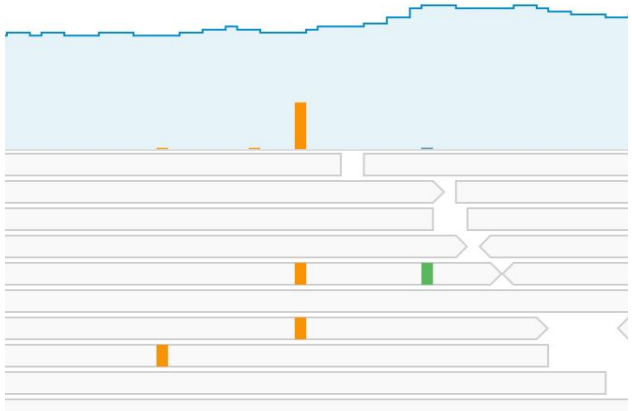
TUMOR SAMPLE

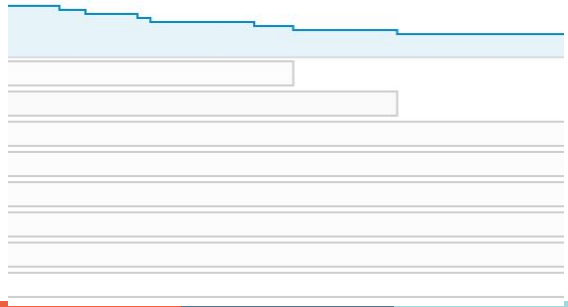
Somatic Calling

NORMAL SAMPLE

SOMATIC MUTATION

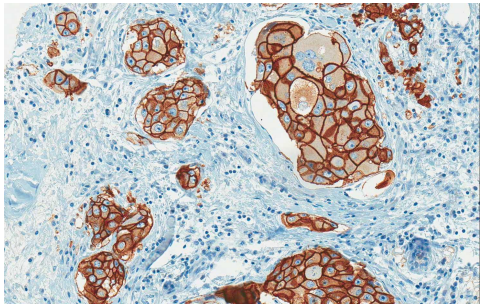




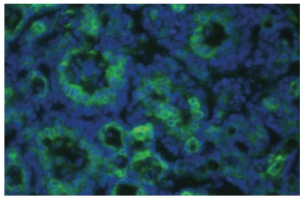


Somatic variant calling izazovi

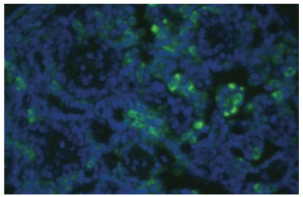
- Tumor heterogeneity
  - Different biopsy locations yield different variations and/or different variant allele frequencies
- Tumor purity
  - Hard to isolate tumor cells from healthy cells when preparing for sequencing
- Low variant allele frequency (directly related to above mentioned) in combination with sequencing errors
  - Hard to conclude that something is a somatic variant
- No datasets for benchmarking (Truth VCF)
  - Hard to benchmark available tools and analysis methods



Tumor (brown) interlaced with normal (blue)



SMA

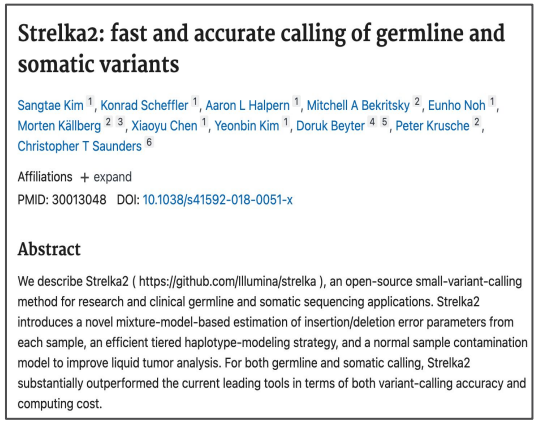


Keratin 8

Different markers shown by different cells

Somatic Variant Callers

- A vast number of callers
- There is no “star” among somatic variant callers, they perform based on the issue at hand
- Types of Somatic Variant callers
  - Position based e.g. **Varscan**
  - Context aware (perform realignment / reassembly) e.g. **Strelka**
- Different callers have different pros and cons
  - Quality of calls
  - Runtime
  - Memory consumption
  - Possibility to adjust parameters
  - Type of analysis they are suitable for (WGS, WES)



Strelka2: fast and accurate calling of germline and somatic variants

Sangtae Kim [1], Konrad Scheffler [1], Aaron L Halpern [1], Mitchell A Bekritsky [2], Eunho Noh [1], Morten Källberg [2 3], Xiaoyu Chen [1], Yeonbin Kim [1], Doruk Beyter [4 5], Peter Krusche [2], Christopher T Saunders [6]

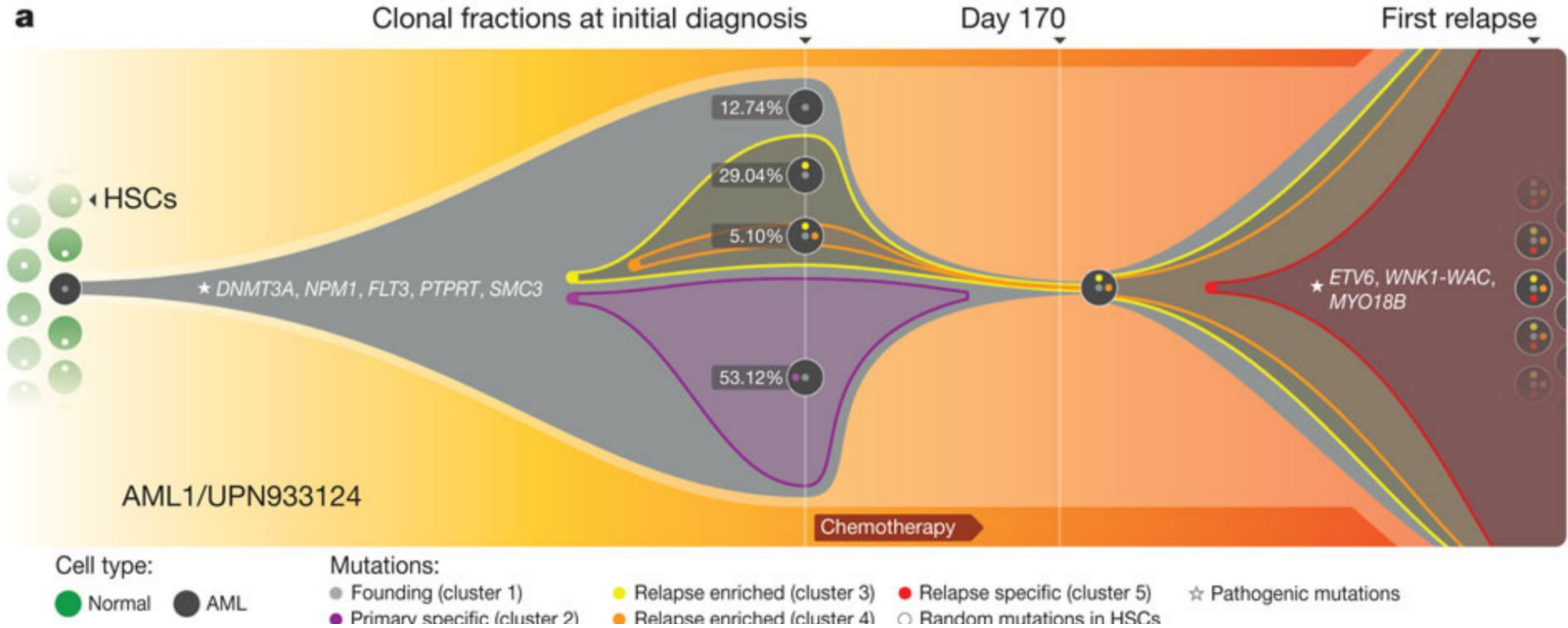
Affiliations + expand

PMID: 30013048 DOI: 10.1038/s41592-018-0051-x

Abstract

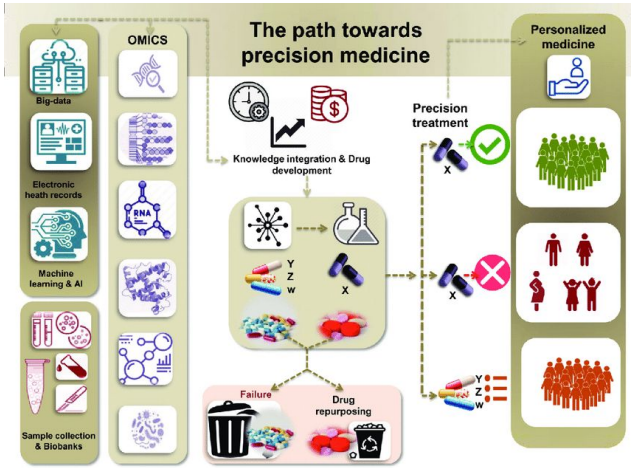
We describe Strelka2 (https://github.com/Illumina/strelka), an open-source small-variant-calling method for research and clinical germline and somatic sequencing applications. Strelka2 introduces a novel mixture-model-based estimation of insertion/deletion error parameters from each sample, an efficient tiered haplotype-modeling strategy, and a normal sample contamination model to improve liquid tumor analysis. For both germline and somatic calling, Strelka2 substantially outperformed the current leading tools in terms of both variant-calling accuracy and computing cost.

Fishplot

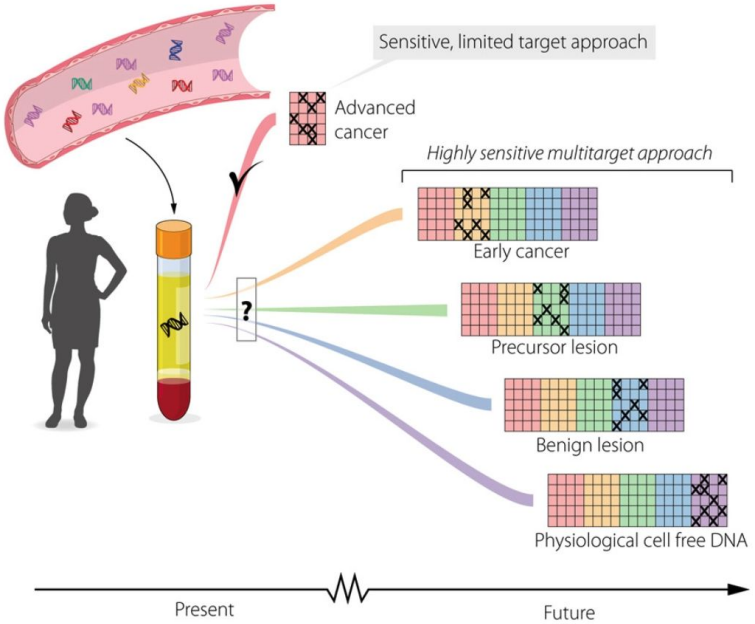

a
Clonal fractions at initial diagnosis
Day 170
First relapse
HSCs
12.74%
29.04%
5.10%
53.12%
★ DNMT3A, NPM1, FLT3, PTPRT, SMC3
★ ETV6, WNK1-WAC, MYO18B
AML1/UPN933124
Chemotherapy
Cell type:
Normal
AML
Mutations:
Founding (cluster 1)
Primary specific (cluster 2)
Relapse enriched (cluster 3)
Relapse enriched (cluster 4)
Relapse specific (cluster 5)
Random mutations in HSCs
☆ Pathogenic mutations

Zaključak

- Personalized cancer therapy is coming
- Understanding the genomic cancer landscape is becoming practically useful
- Somatic analysis is not a direct extension of germline analysis
- Many concepts are purely abstract
- Computational processing limited by sequencing technology



Razvoj personalizovane medicine - primer



- Cell-free methylated DNA immunoprecipitation-sequencing (cfMeDIP-seq) identifies genomic regions with DNA methylation, using a protocol adapted to work with low-input DNA samples and with cell-free DNA (cfDNA). This method allows for DNA methylation profiling of circulating tumour DNA in cancer patients' blood samples. Such epigenetic profiling of circulating tumour DNA provides information about in which tissues tumour DNA originates, a key requirement of any test for early cancer detection. In addition, DNA methylation signatures provide prognostic information and can detect relapse. For robust quantitative comparisons between samples, immunoprecipitation enrichment methods like cfMeDIP-seq require normalization against common reference controls.
- Paper: Sensitive and reproducible cell-free methylome quantification with synthetic spike-in controls

Pitanja / Ispit

milan.kovacevic@sbgenomics.com
marko.matic@sbgenomics.com