Amplicon Architect

A tool for detecting and reconstructing extrachromosomal circular DNAs from NGS data.

Maciej Hryc, MSc. student in Bioinformatics

AMU, Faculty of Biology

November 25, 2024

Purpose of this presentation

- Provide a (**very**) brief introduction to eccDNA research and its methods
- Introduce basic concepts and showcase one of the computational problems
- Broadly explain one of the available solutions (*Amplicon Architect*)

Intro to eccDNA biology

Definition

EccDNA (Extrachromosomal circular DNA) is broad class of covalently closed ssDNA and dsDNA of genomic origin, varying in size from 100 bp to 5 Mbp, and is ubiquitous in eukaryotes.

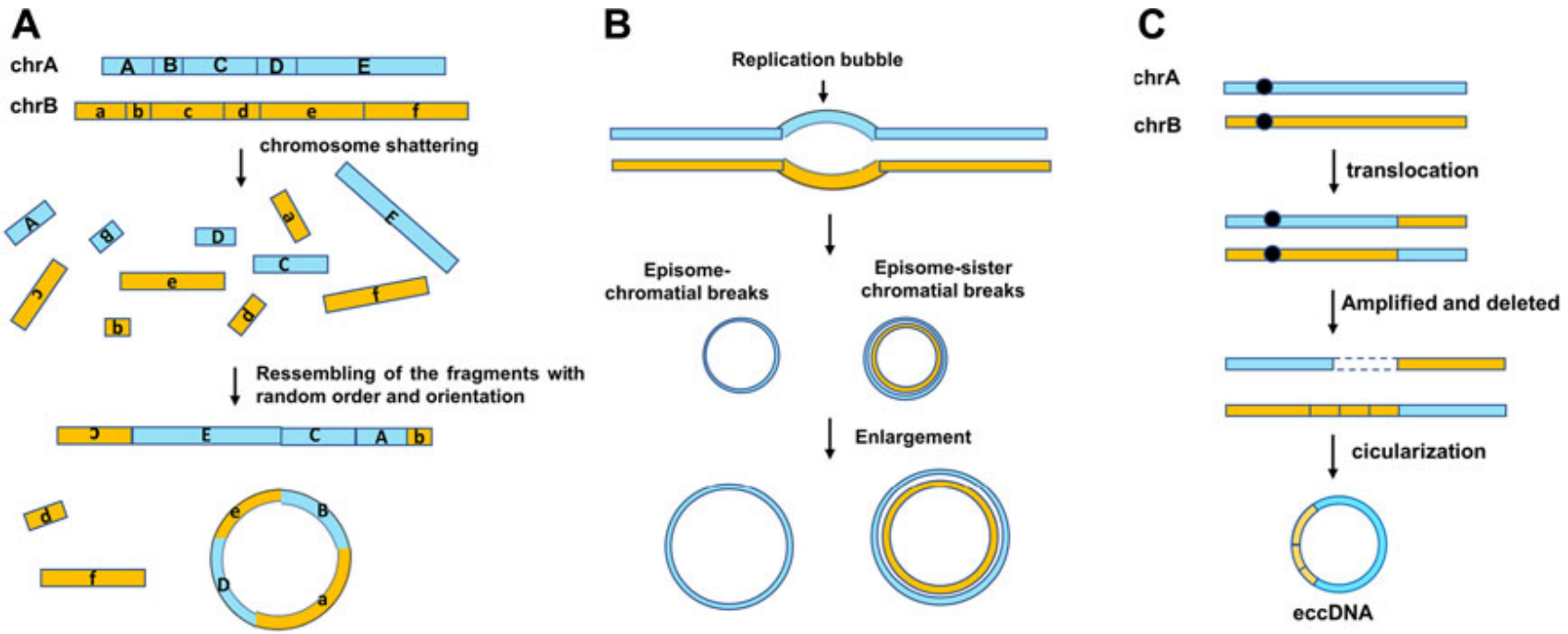
Key traits

- highly accessible chromatin
- independant replication
- uneven segregation during mitosis
- can carry genomic sequences, eg. genes, enhancers
- can be a "collage" of distant genomic sequences
- one of the main causes of focal amplification in cancer

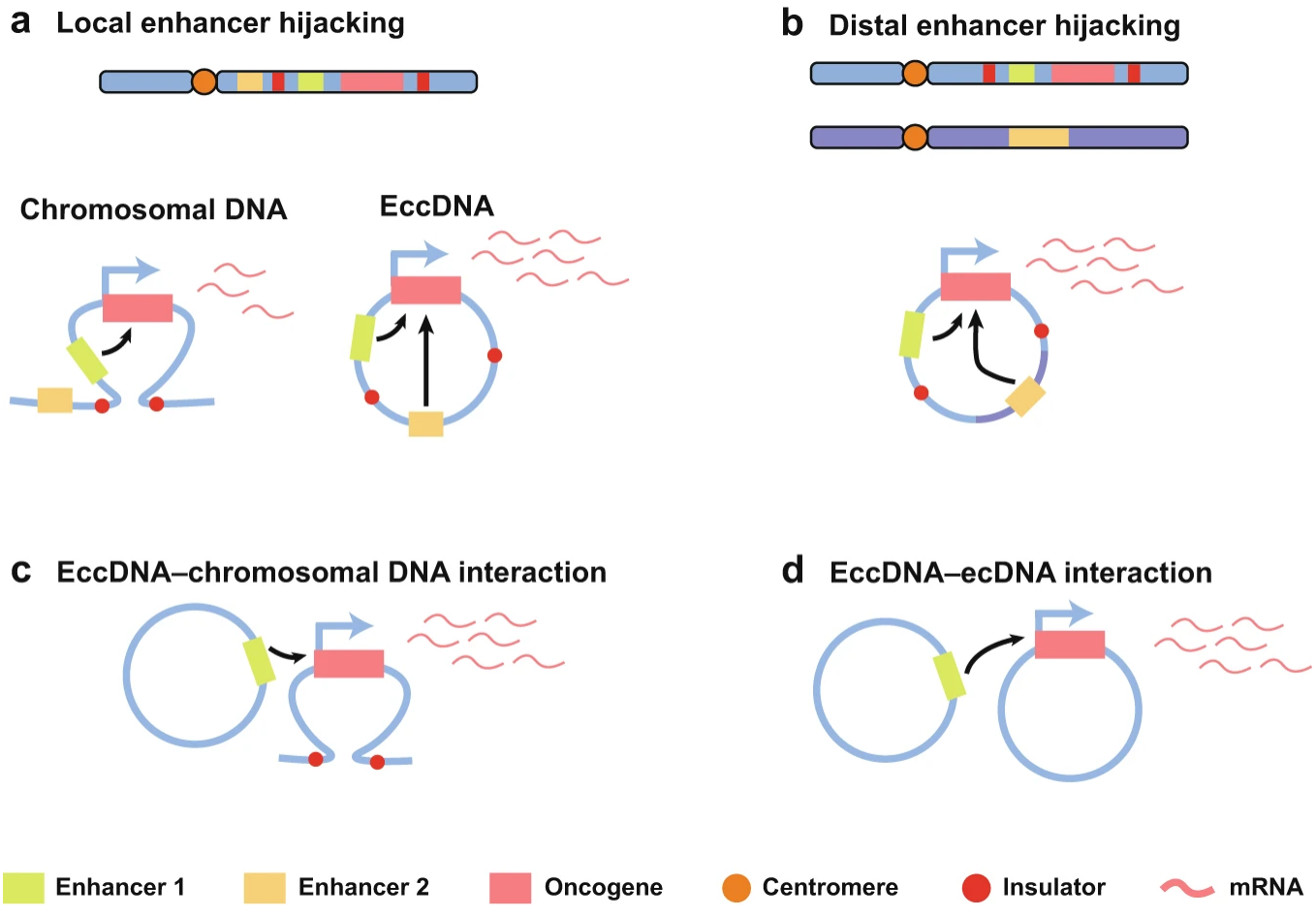
Biogenesis

Known pathways

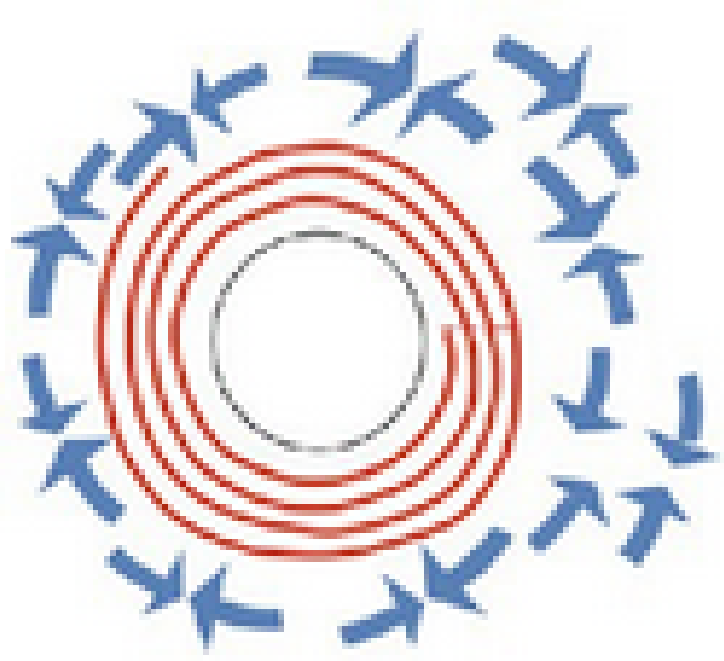
- Chromothripsis (Chromosome shattering)
- Episome model (polymerase slippage)
- Translocation-Deletion-Amplification mechanism
- Breakage-Fusion-Bridge cycle

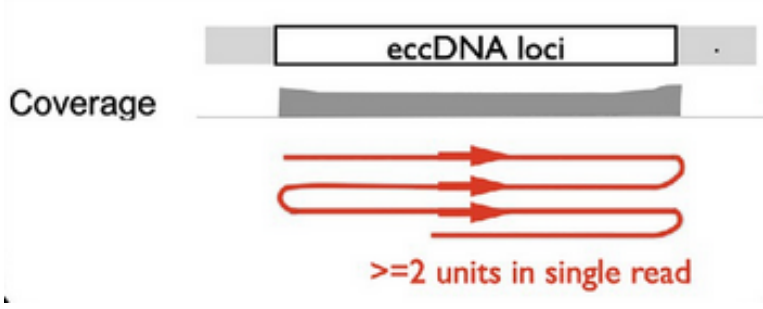


Impact on gene expression



Sequencing strategies — Φ29 polymerase RCA

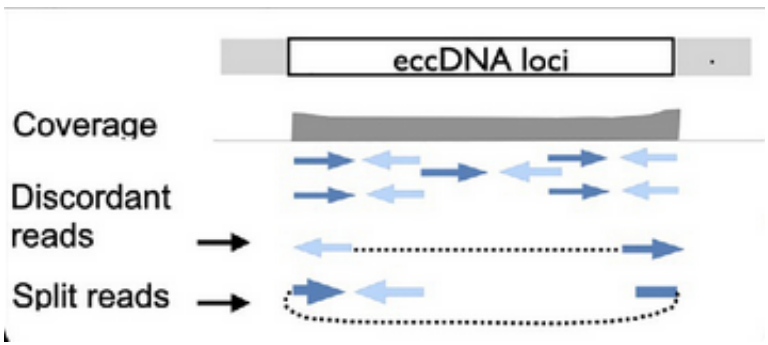


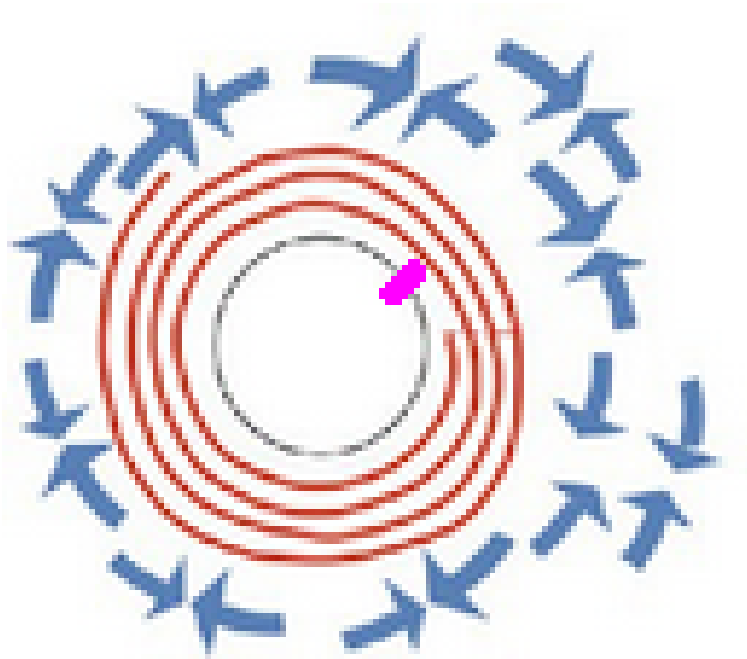


Basic concepts

Simple solution?

- concordant read pairs,
- discordant read pairs (unexpected mapping),
- split reads

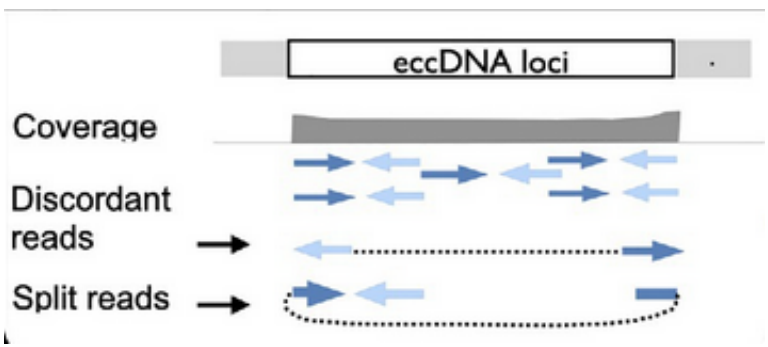


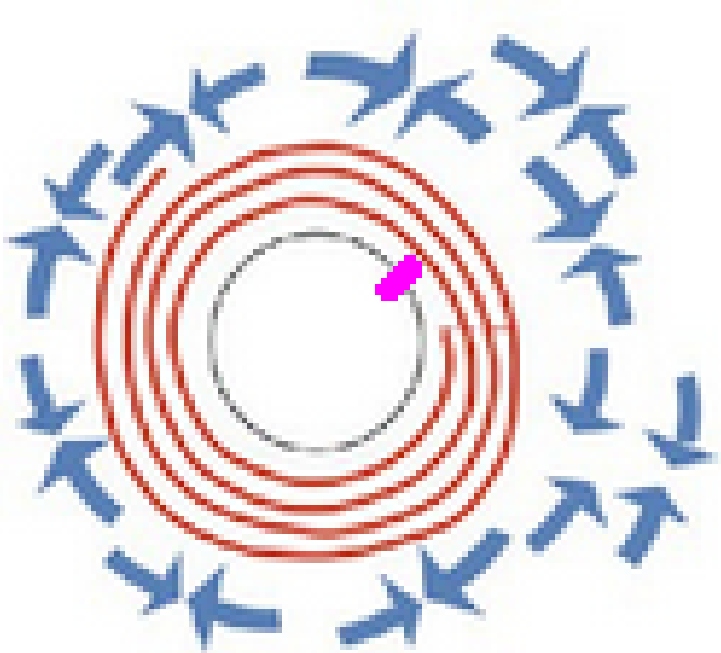


Basic concepts

~~Simple solution?~~ This is not enough!

- structural variants can lead to similar read mappings
- no information about the number of eccDNAs
- no structure reconstruction

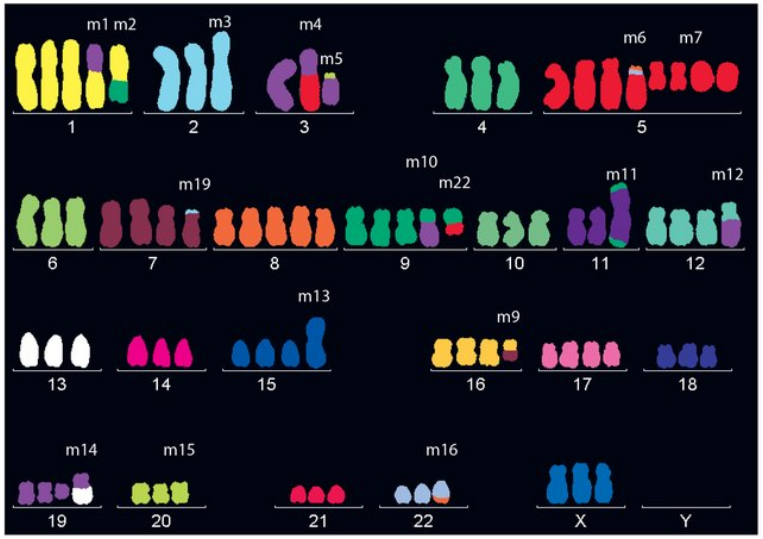




Amplicon Architecht: NGS data from tumor cells

Input

- WGS paired-end reads mapped to reference
- a set of seed intervals



Seed interval selection

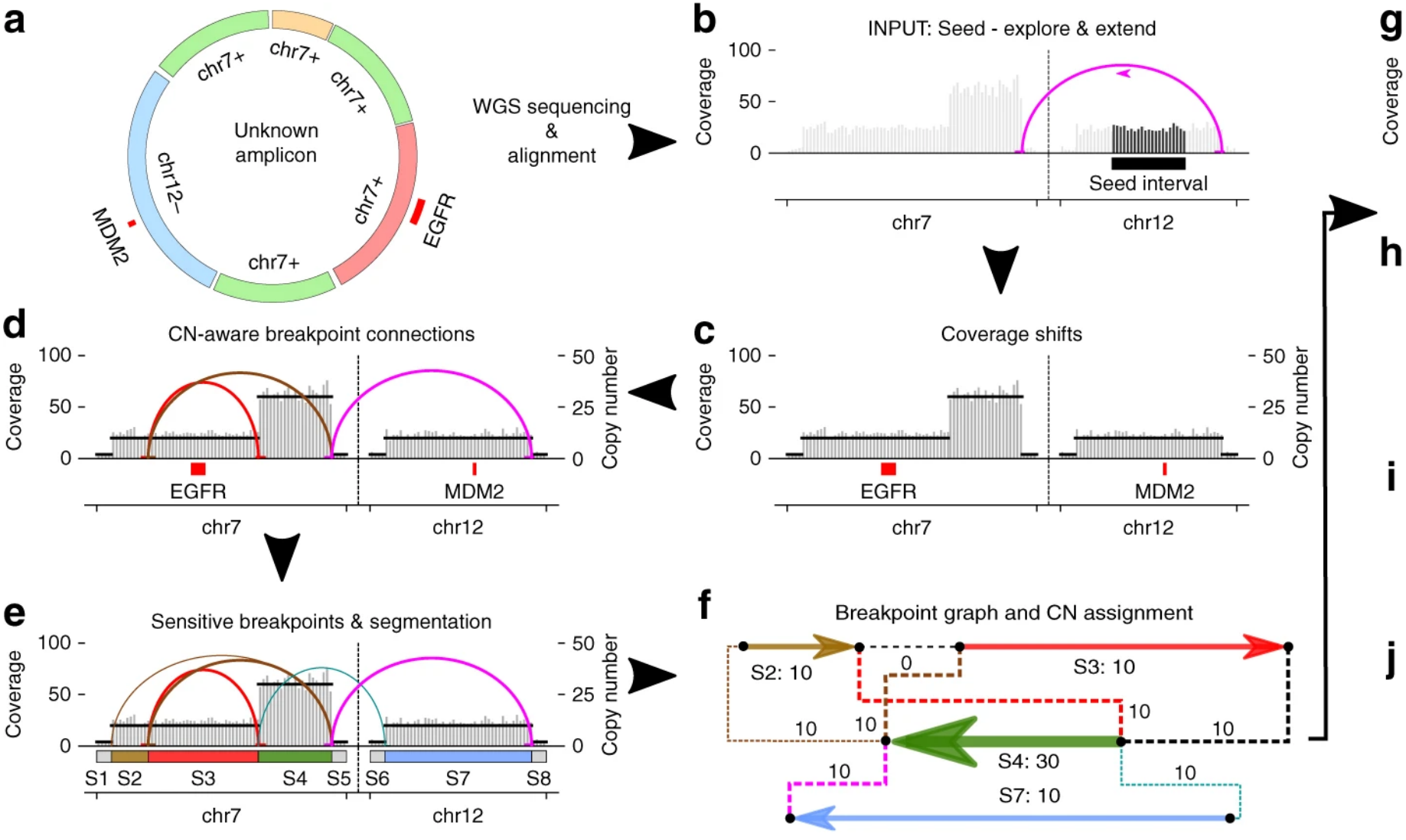
- Find CNV calls (*Read Depth*)
- filter out:
  - blacklisted regions (CNV calls from non-cancerous cells)
  - regions with high average repetitiveness (*Duke Mapability Uniqueness*)
  - known segmental duplications
- correct for aneuploidy

Main pipeline

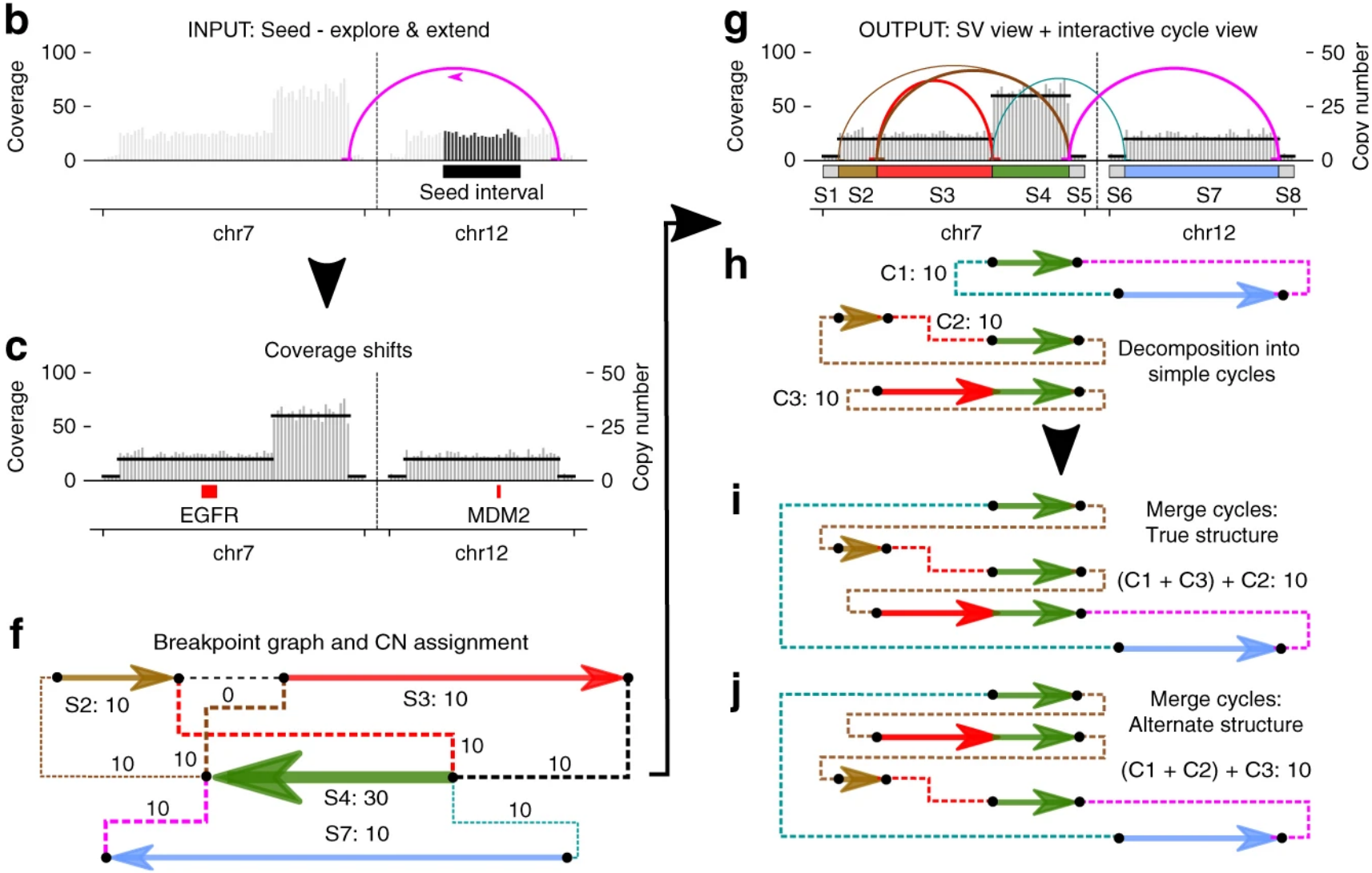
AA pipeline

- Iteratively identify intervals belonging to the same amplicon (discordant read pairs)
- Identify rearangements within the interval (based on discordant read pairs and CNV boundary detection)
- create a breakpoint graph (a graph showing how segments of the amplicon are connected) and estimate copy numbers for each segment
- Breakpoint graph decomposition into simple cycles (based on CN using a heuristic method)
- Merge simple cycles into possible amplicon structures

Schematic of Amplicon Architect pipeline



Schematic of Amplicon Architect pipeline, cont.



Additional information

Other available tools

NGS data

- Circle-Map,
- Circle_finder,
- ecc_finder,
- ECCsplorer

Long-read sequencing

- NanoCircle,
- eccDNA_RCA_nanopore,
- CReSIL,
- FLED

Better diagram of eccDNA formation pathways that was too big for the slide.

doi: 10.3389/fonc.2022.814504. PMID: 35494014; PMCID: PMC9046939.

Literature

- Zuo S, Yi Y, Wang C, Li X, Zhou M, Peng Q, Zhou J, Yang Y and He Q (2022) Extrachromosomal Circular DNA (eccDNA): From Chaos to Function. Front. Cell Dev. Biol. 9:792555. doi: 10.3389/fcell.2021.792555
- Yang, L., Jia, R., Ge, T. et al. Extrachromosomal circular DNA: biogenesis, structure, functions and diseases. Sig Transduct Target Ther 7, 342 (2022). https://doi.org/10.1038/s41392-022-01176-8
- https://github.com/chrisamiller/readDepth
- https://github.com/virajbdeshpande/AmpliconArchitect
- https://genome.ucsc.edu/cgi-bin/hgc?hgsid=1932599340_RfoA4w7uC0f1moRCQOvdK55A36Y0&db=hg19&c=chr12&l=0&r=305275&o=0&t=305275&g=wgEncodeDukeMapabilityUniqueness35bp&i=wgEncodeDukeMapabilityUniqueness35bp
- Deshpande, V. et al., Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. Nat. Commun. 10, 392 (2019). PMID: 30674876. (Article)
- Advancements in Focal Amplification Detection in Tumor/ Liquid Biopsies and Emerging Clinical Applications; Aram Arshadi and Doron Tolomeo and Santina Venuto and Clelia Tiziana Storlazzi; 2023; 10.3390/genes14061304
- Zhang P, Peng H, Llauro C, Bucher E and Mirouze M (2021) ecc_finder: A Robust and Accurate Tool for Detecting Extrachromosomal Circular DNA From Sequencing Data. Front. Plant Sci. 12:743742. doi: 10.3389/fpls.2021.743742
- Smirnov E, Kalmárová M, Koberna K, Zemanová Z, Malínský J, Masata M, Cvacková Z, Michalová K, Raska I. NORs and their transcription competence during the cell cycle. Folia Biol (Praha). 2006;52(3):59-70. PMID: 17089916;