



CEITEC

Central European Institute of Technology
BRNO | CZECH REPUBLIC

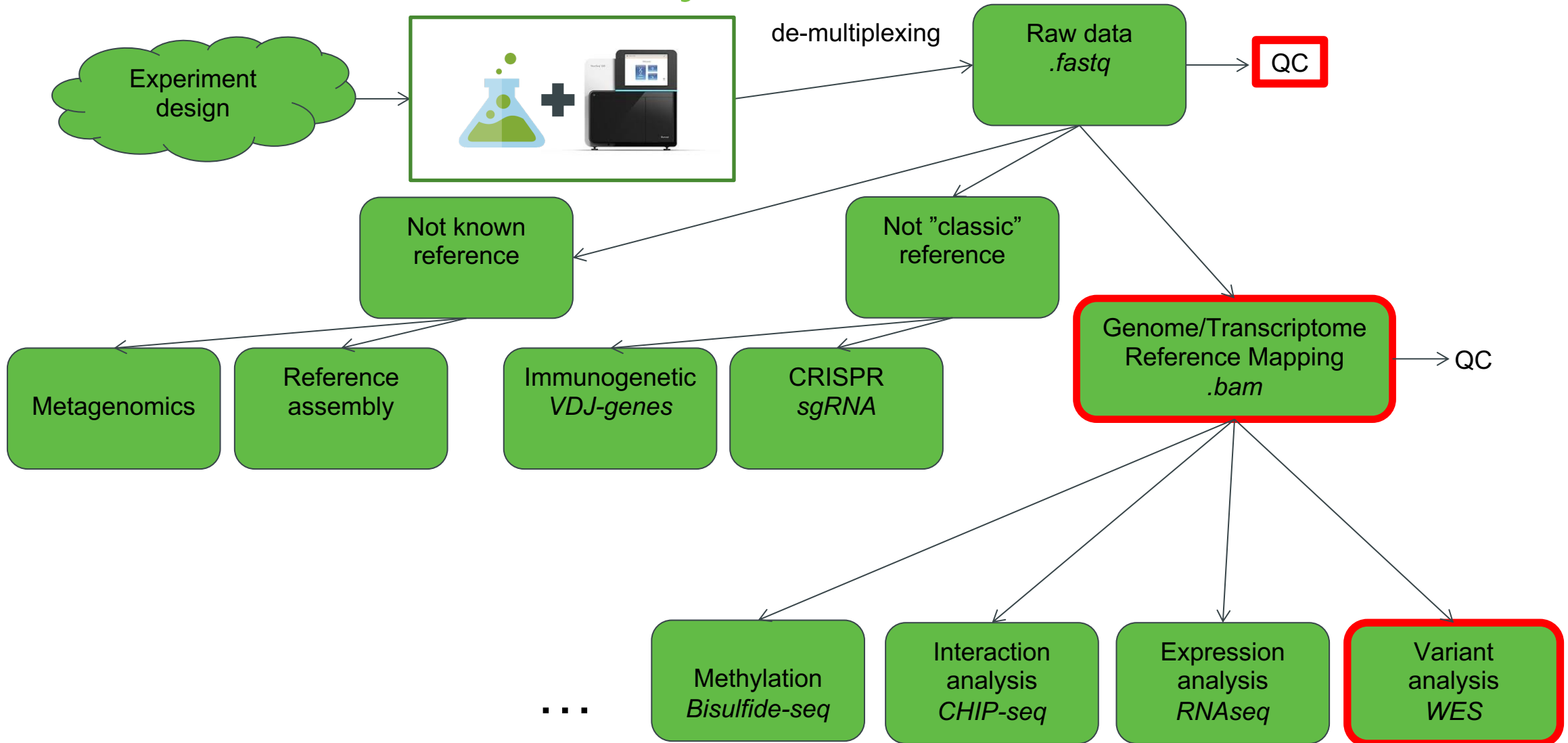


**Modern methods for genome analysis
(PřF:Bi7420)**

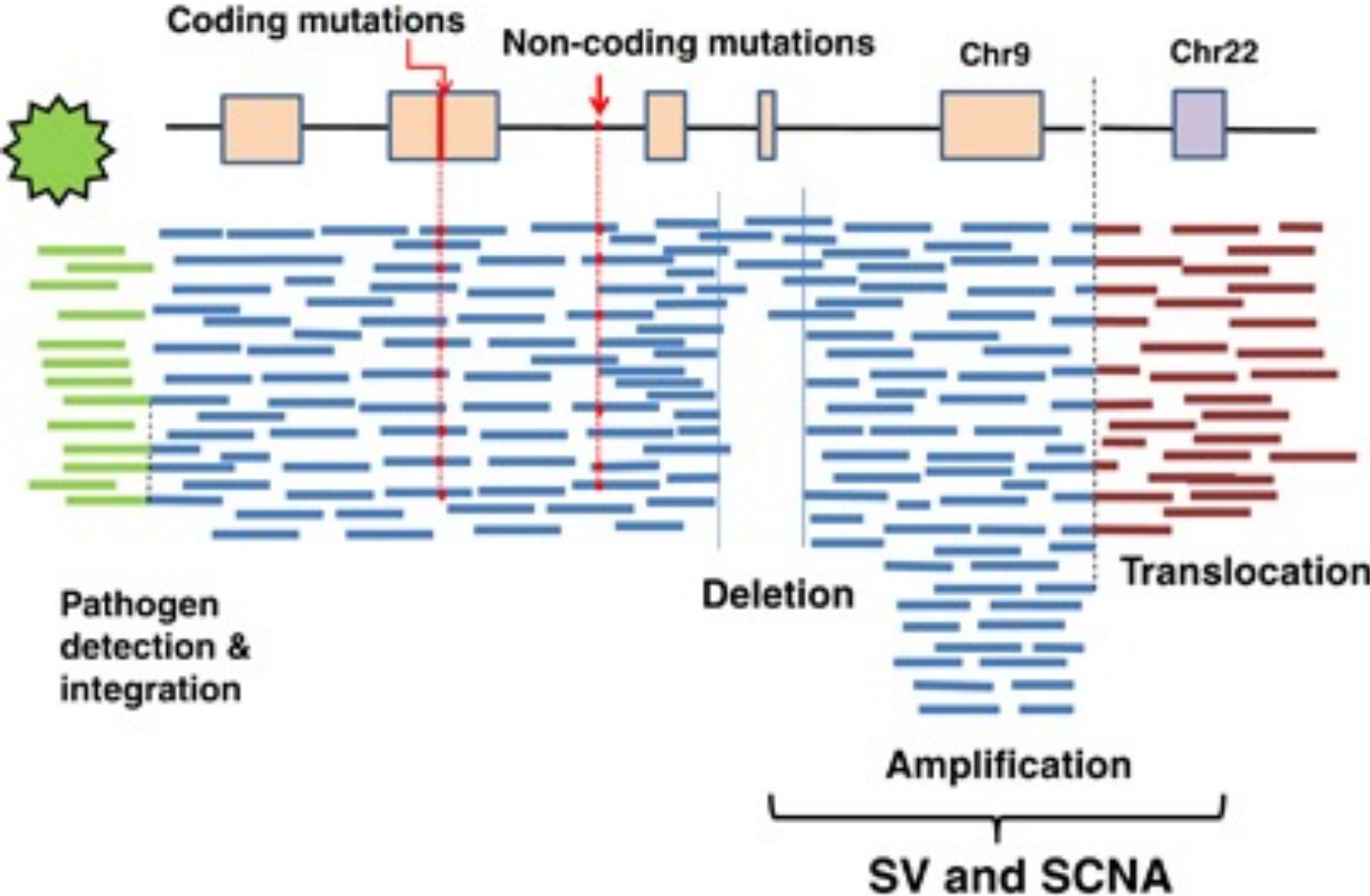
Lecture 4 : Structural variants

Vojta Bystry
vojtech.bystry@ceitec.muni.cz

NGS data analysis



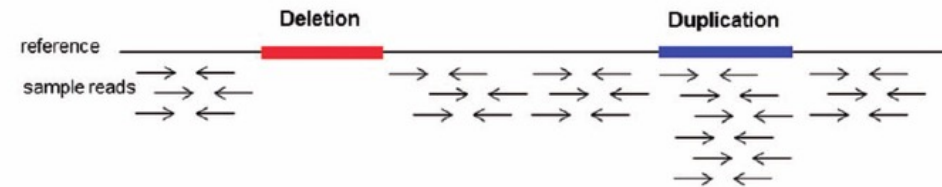
Structural variants calling



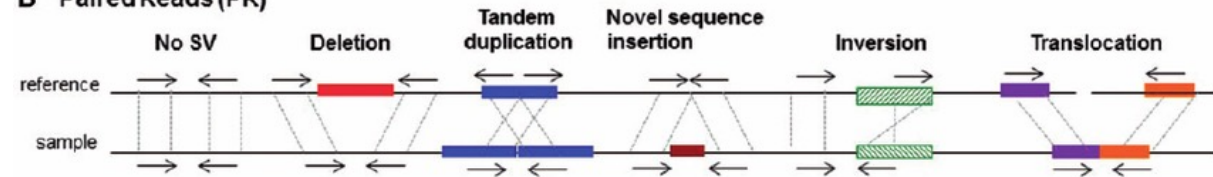
Structural variants calling

- Copy number variants (CNV)
 - Copy number analysis (CNA)
- Structural variants (SV)
 - Discordant reads analysis
 - Mainly from WGS
- Fusion genes analysis
 - From RNA-seq
 - SV in genes
 - Medical application

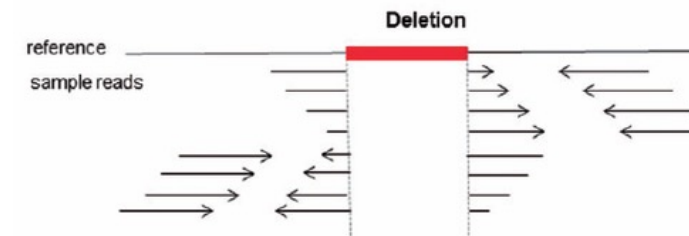
A Read Depth (RD)



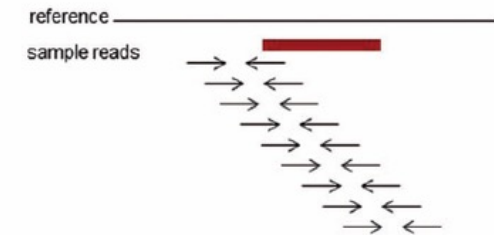
B Paired Reads (PR)



C Split Reads (SR)



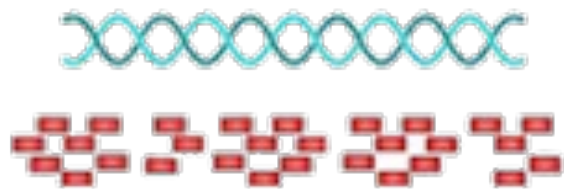
D. De Novo Assembly (AS)



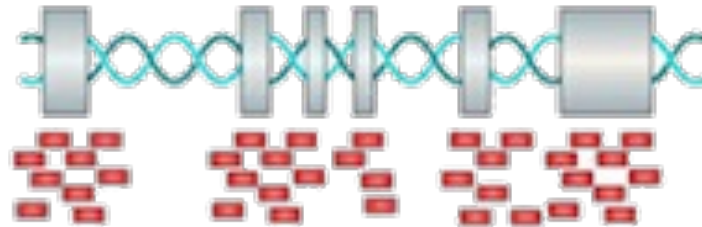
Copy number variants

- Not-PCR amplified (WGS)
- PCR amplified (WES, targeted)

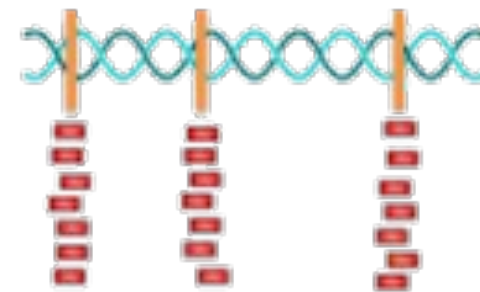
Whole genome sequencing



Whole exome sequencing

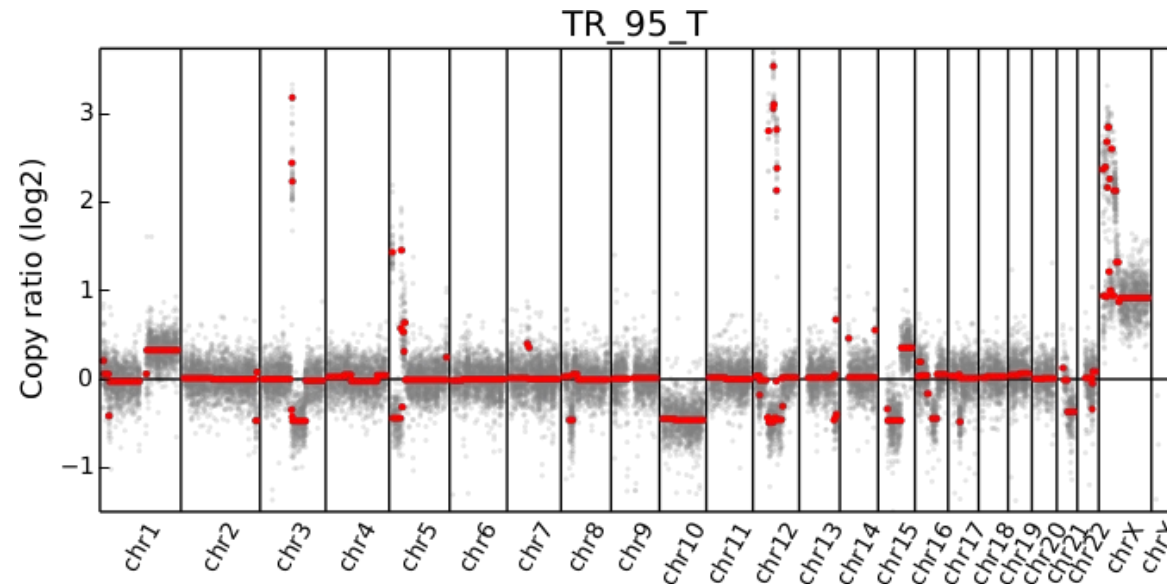
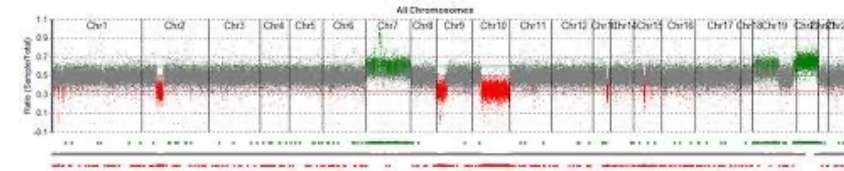


Targeted sequencing

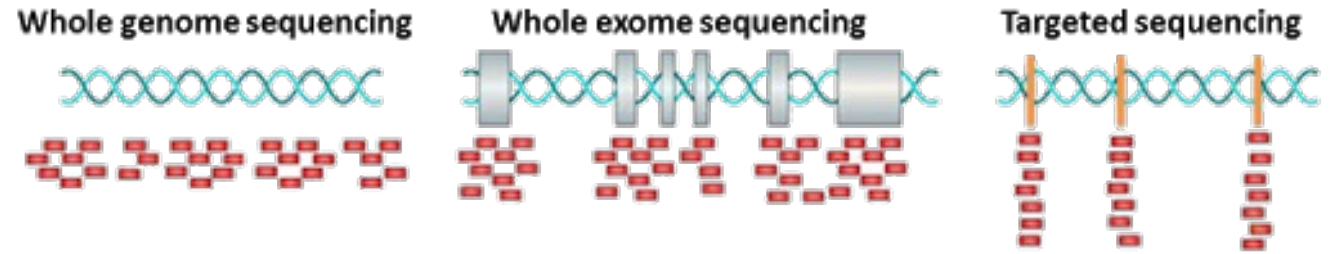


WGS copy number variants

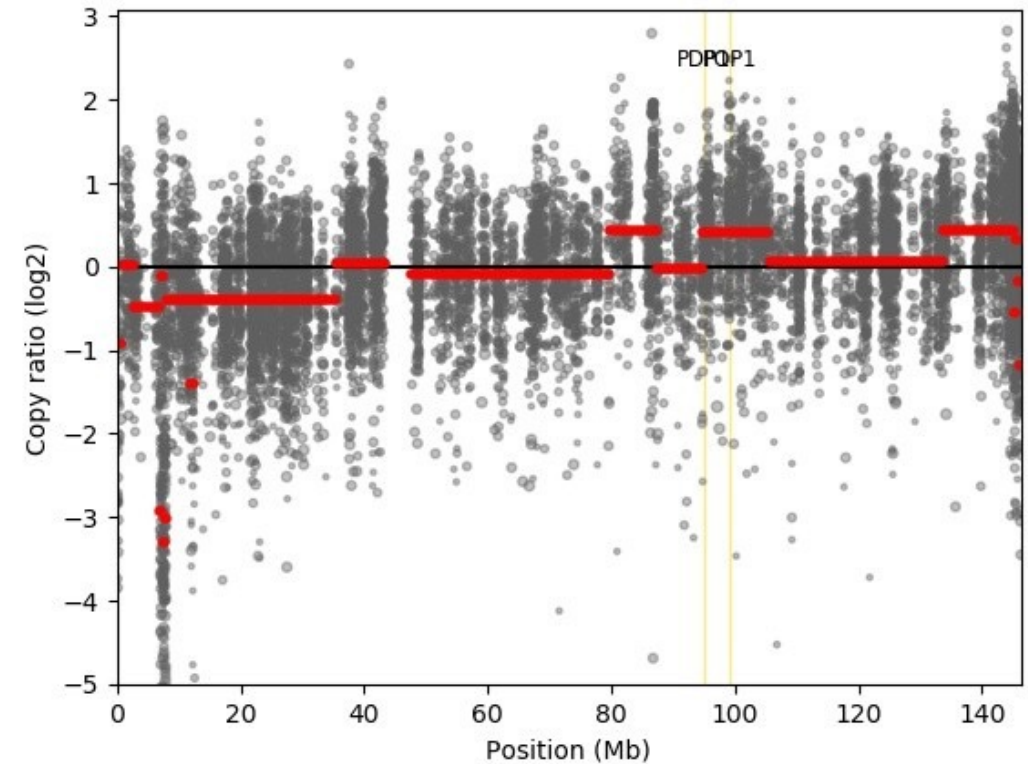
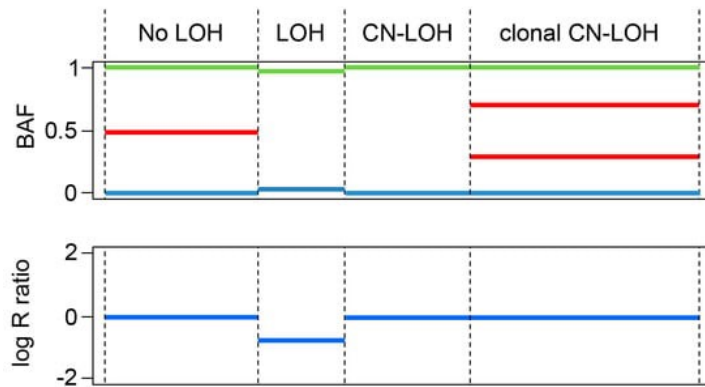
- Running window
- Normalization to the absolute coverage
- Low-coverage WGS CNV
 - Over large regions
 - Good for somatic low tumor purity
 - Minimal variant size vs. coverage trade-off



PCR amplified CNV

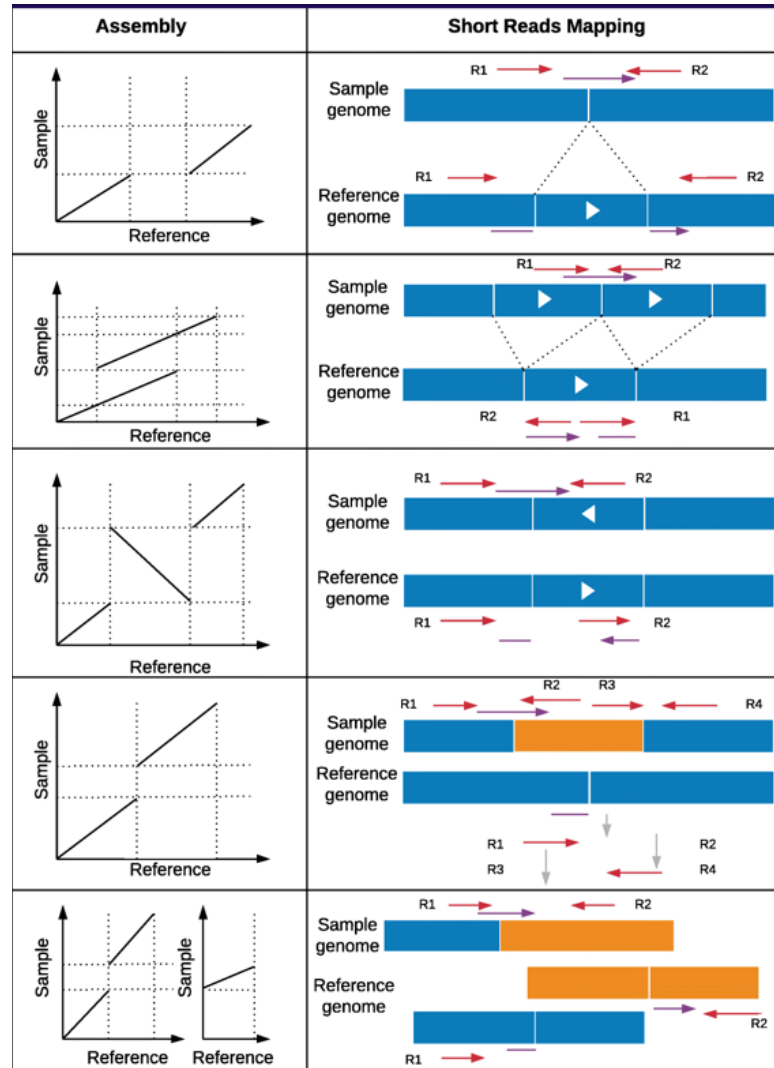


- Individual analysis problematic
- Panel of (normal) samples
- BAF frequency
 - Help CNA
 - Can call “loss of heterozygosity” LOH



Structural variants calling

- RP = Read pair
 - Pair-end sequencing
- SR = Split reads



Structural variants calling

- RP = Read pair
 - Pair-end sequencing
- SR = Split reads

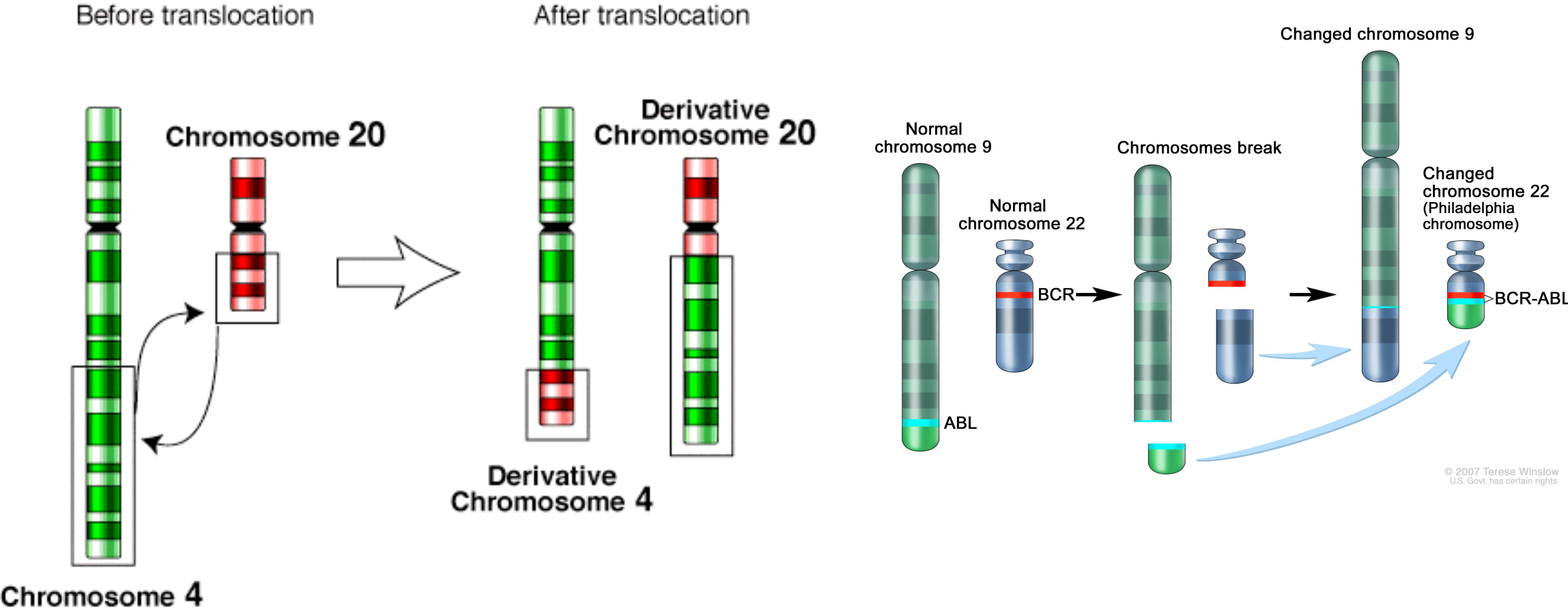
	Assembly	Short Reads Mapping
Deletion		
Duplication		
Inversion		
Insertion		
Translocation		

Structural variants calling

- RP = Read pair
 - Pair-end sequencing
- SR = Split reads
- Tools:
 - Delly, Manta, GRIDSS

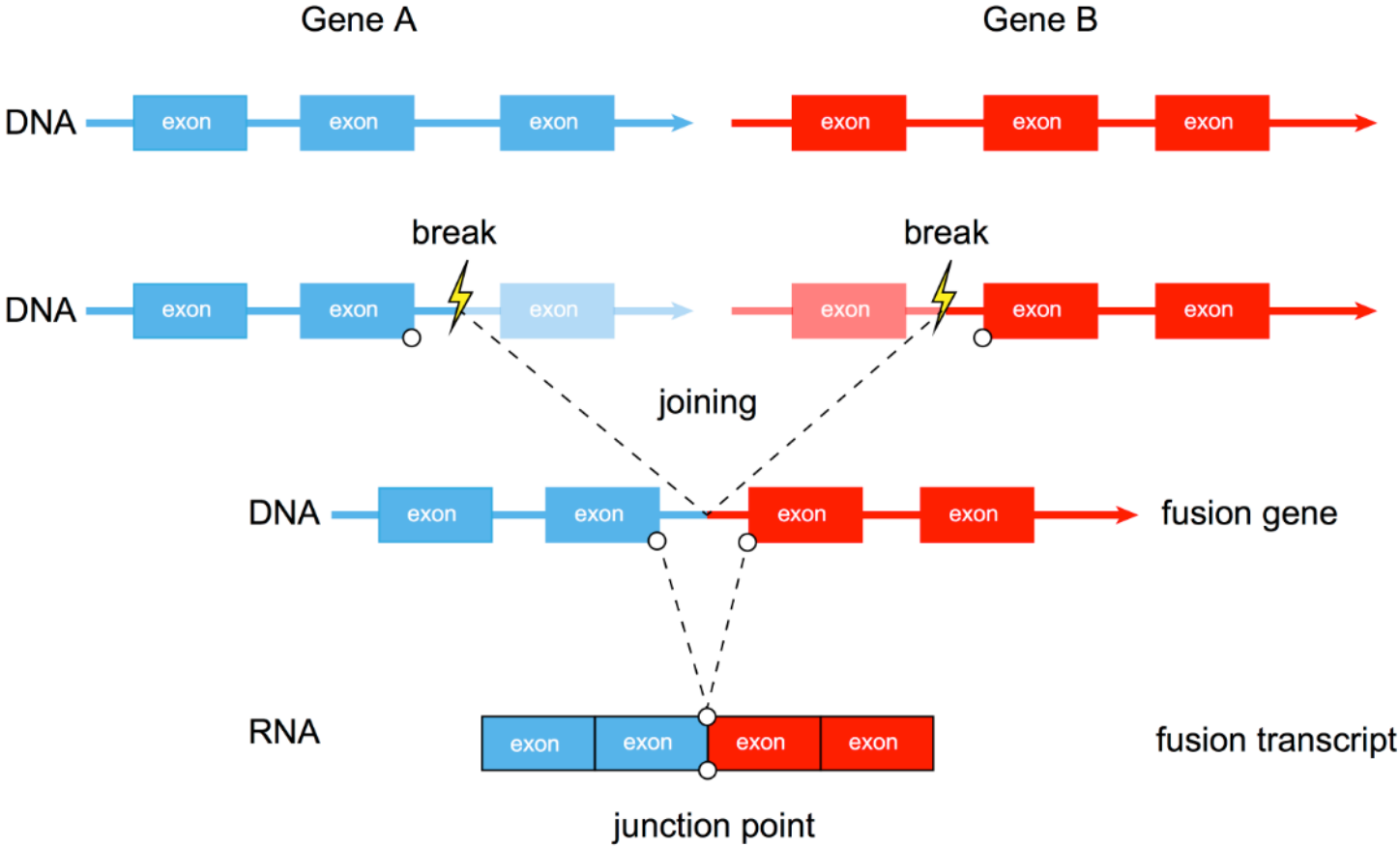
	Assembly	Short Reads Mapping
Deletion		
Duplication		
Inversion		
Insertion		
Translocation		

Fusion genes



© 2007 Terese Winslow
U.S. Govt. has certain rights

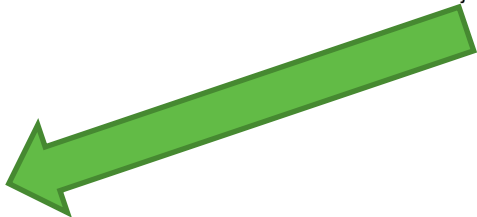
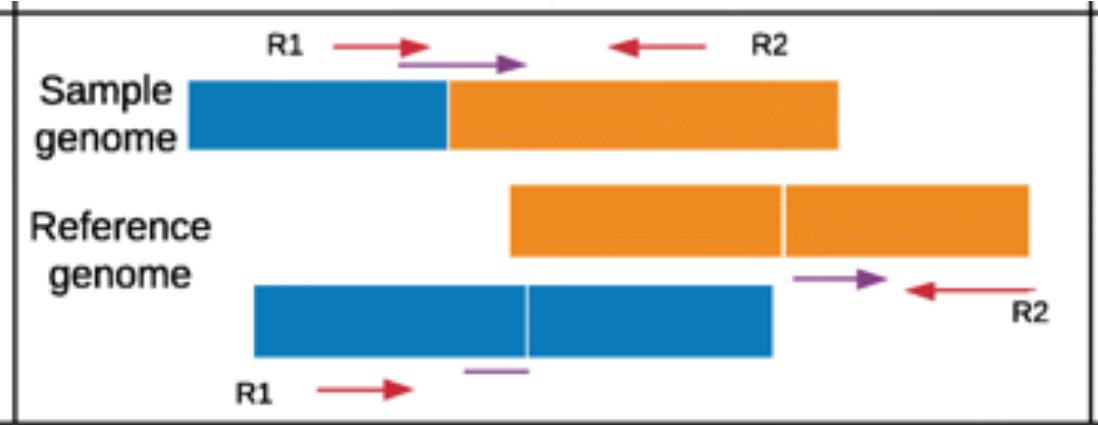
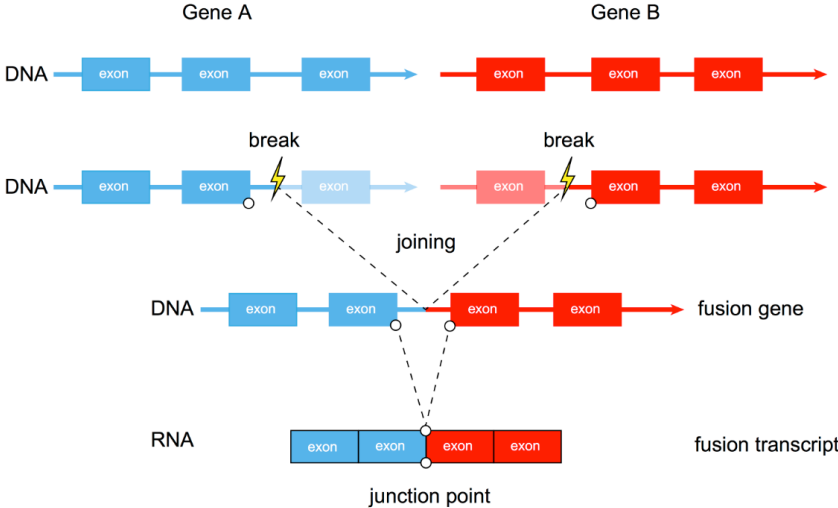
Fusion genes



Fusion genes

	Assembly	Short Reads Mapping
Deletion		
Duplication		
Inversion		
Insertion		
Translocation		

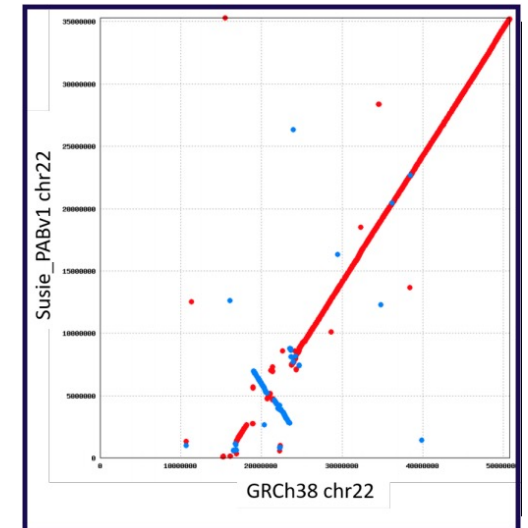
Fusion genes



Long reads

- Short reads
 - 1kb
- Long reads
 - 100 – 1000kb
- Hot topic

	Assembly	Short Reads Mapping
Deletion	<p>A graph with 'Sample' on the y-axis and 'Reference' on the x-axis. A diagonal line represents the identity. A vertical dashed line indicates a missing segment in the sample genome.</p>	<p>Sample genome: A blue bar with a gap. Reads R1 and R2 are shown above it, with arrows pointing to the reference genome below. Reference genome: A blue bar with a gap. Reads R1 and R2 are shown below it, with arrows pointing to the reference genome. Dotted lines connect the ends of the reads to the reference genome, showing they skip the deleted region.</p>
Duplication	<p>A graph with 'Sample' on the y-axis and 'Reference' on the x-axis. A diagonal line represents the identity. A horizontal dashed line indicates an extra segment in the sample genome.</p>	<p>Sample genome: A blue bar with an extra segment. Reads R1 and R2 are shown above it, with arrows pointing to the reference genome below. Reference genome: A blue bar with a gap. Reads R1 and R2 are shown below it, with arrows pointing to the reference genome. Dotted lines connect the ends of the reads to the reference genome, showing they map to both the original and duplicated regions.</p>
Inversion	<p>A graph with 'Sample' on the y-axis and 'Reference' on the x-axis. A diagonal line represents the identity. A segment of the diagonal line is inverted, showing a reversed order of segments.</p>	<p>Sample genome: A blue bar with a segment that is inverted. Reads R1 and R2 are shown above it, with arrows pointing to the reference genome below. Reference genome: A blue bar with a gap. Reads R1 and R2 are shown below it, with arrows pointing to the reference genome. Dotted lines connect the ends of the reads to the reference genome, showing they map to the inverted region.</p>
Insertion	<p>A graph with 'Sample' on the y-axis and 'Reference' on the x-axis. A diagonal line represents the identity. A horizontal dashed line indicates an extra segment in the sample genome that is not in the reference.</p>	<p>Sample genome: A blue bar with an orange inserted segment. Reads R1, R2, R3, and R4 are shown above it, with arrows pointing to the reference genome below. Reference genome: A blue bar with a gap. Reads R1, R2, R3, and R4 are shown below it, with arrows pointing to the reference genome. Dotted lines connect the ends of the reads to the reference genome, showing R3 and R4 map to the inserted region.</p>
Translocation	<p>A graph with 'Sample' on the y-axis and 'Reference' on the x-axis. A diagonal line represents the identity. A segment of the diagonal line is broken and rearranged, showing a translocation.</p>	<p>Sample genome: A blue bar with an orange segment that has moved from another location. Reads R1 and R2 are shown above it, with arrows pointing to the reference genome below. Reference genome: A blue bar with a gap. Reads R1 and R2 are shown below it, with arrows pointing to the reference genome. Dotted lines connect the ends of the reads to the reference genome, showing they map to different locations in the sample genome.</p>



Long reads

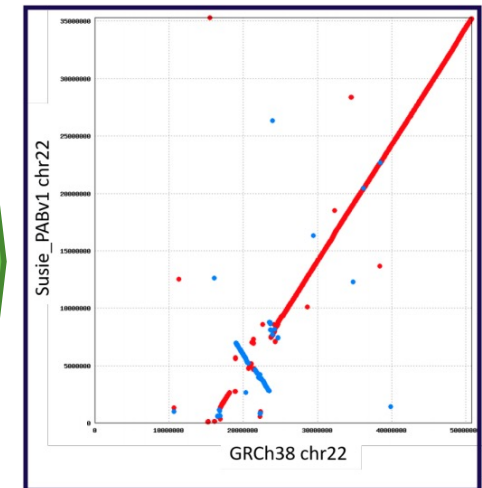
- Short reads
 - 1kb
- Long reads
 - 100 – 1000kb
- Hot topic

	Assembly	Short Reads Mapping	Long Reads Mapping
Deletion			
Duplication			
Inversion			
Insertion			
Translocation			

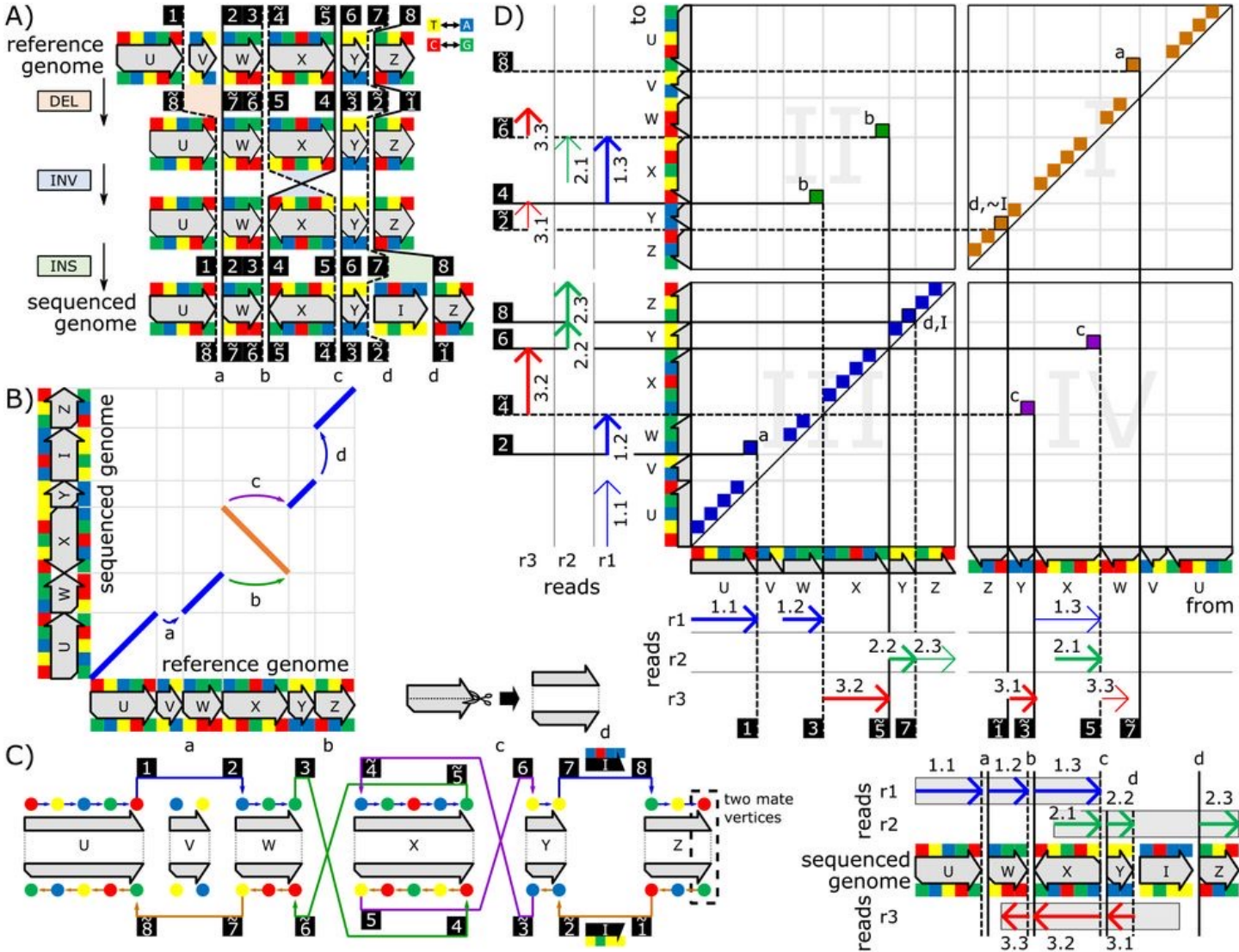
Long reads

- Short reads
 - 1kb
- Long reads
 - 100 – 1000kb
- Hot topic

	Assembly	Short Reads Mapping	Long Reads Mapping
Deletion			
Duplication			
Inversion			
Insertion			
Translocation			



Complex variants





CEITEC



@CEITEC_Brno

Thank you for your attention!

