Evaluating sequencing strategies for ATMP genomic risk assessment

Rafaella Buzatu

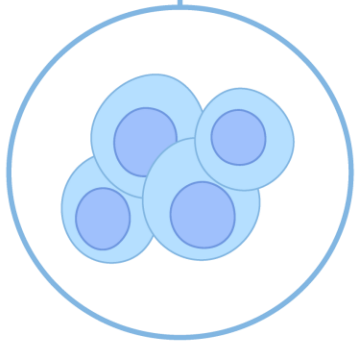
Prof. Micha Drukker
Asst. Prof. Christian Schröter
Tineke van den Hoorn
Marcel Hoefnagel

Universiteit Leiden

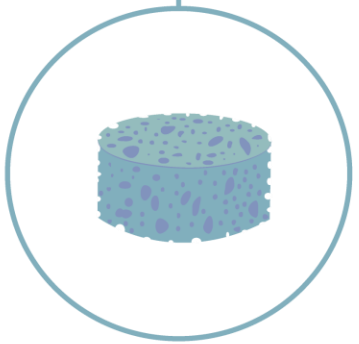
CBG MEB

Advanced Therapeutic Medicinal Products (ATMPs)

- Somatic cell therapies

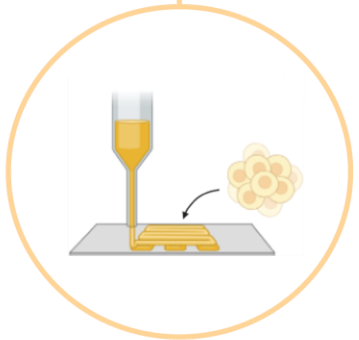


- Tissue engineered medicines

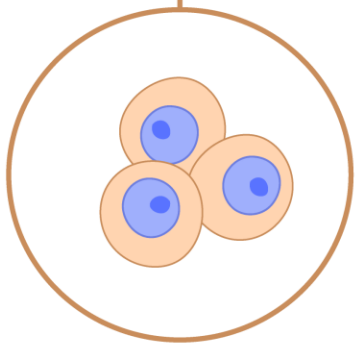


- Gene therapy medicines

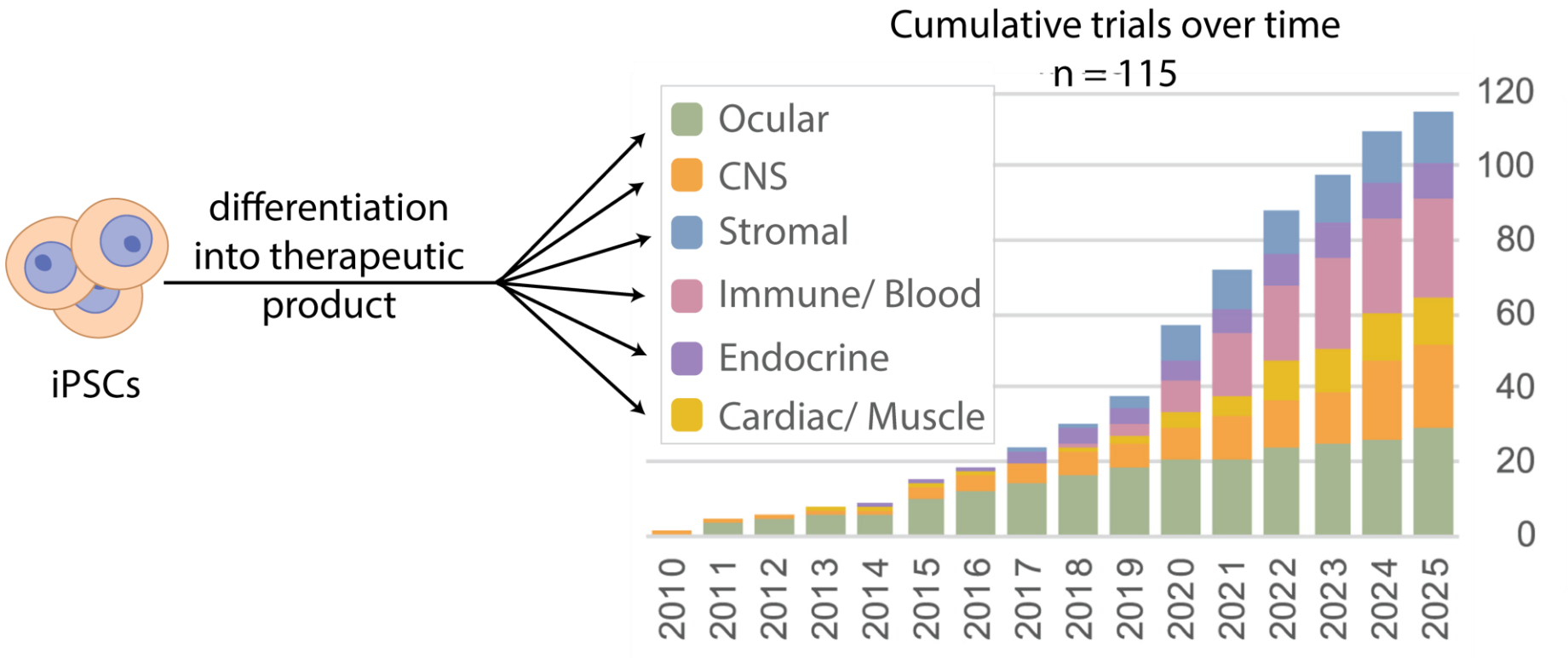
- Combined ATMPs

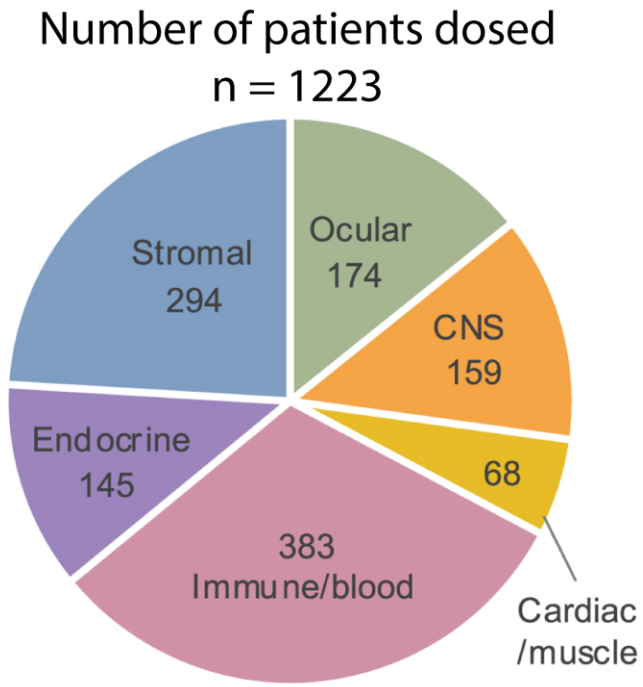


- iPSCs



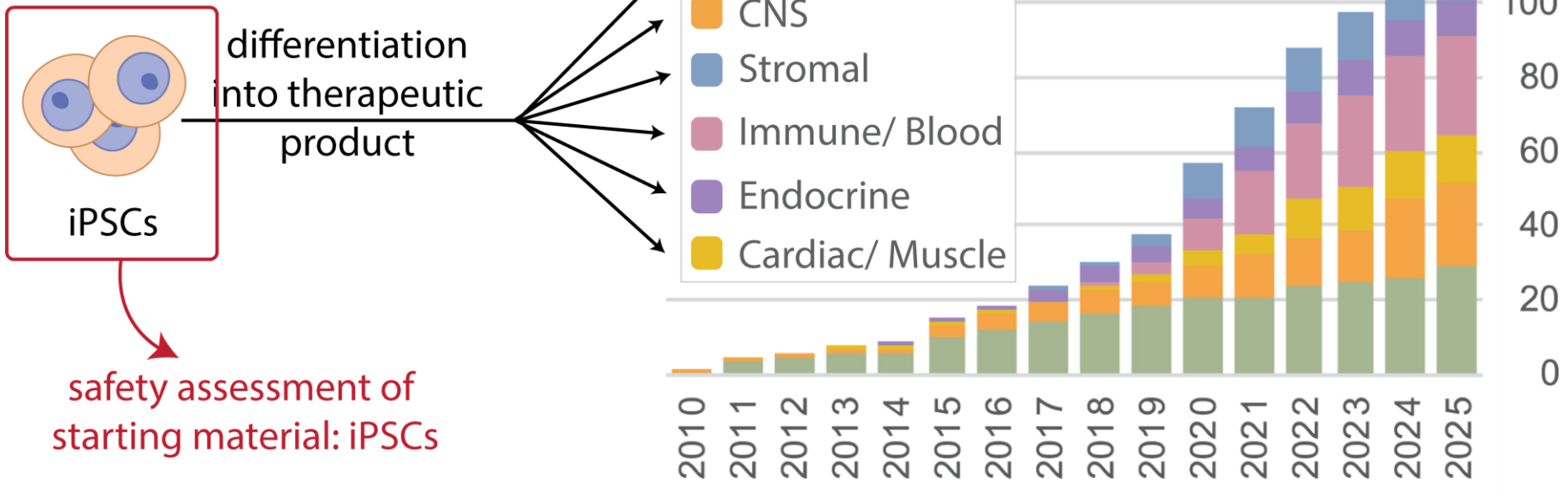
iPSC-derived ATMPs are actively being tested in clinical trials

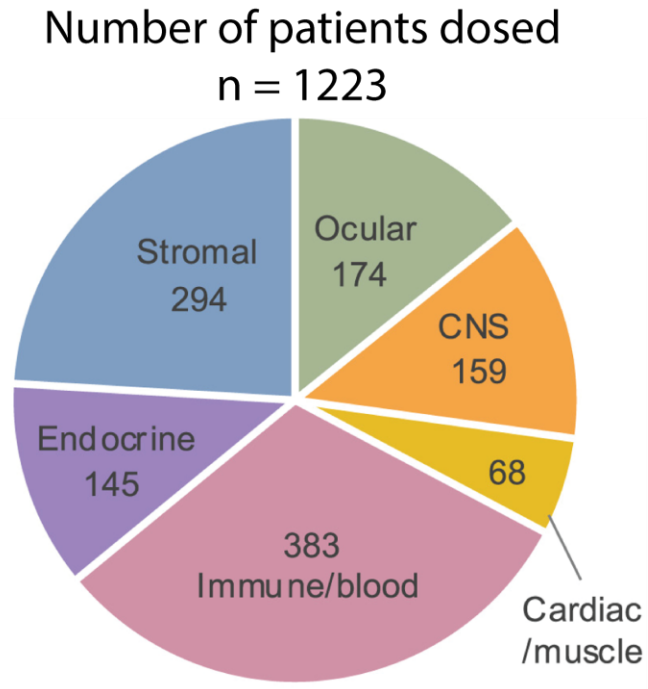




Adapted from Kirkeby A, Main H, Carpenter M. Pluripotent stem-cell-derived therapies in clinical trial: A 2025 update. *Cell Stem Cell*. 2025;32(1):10-37

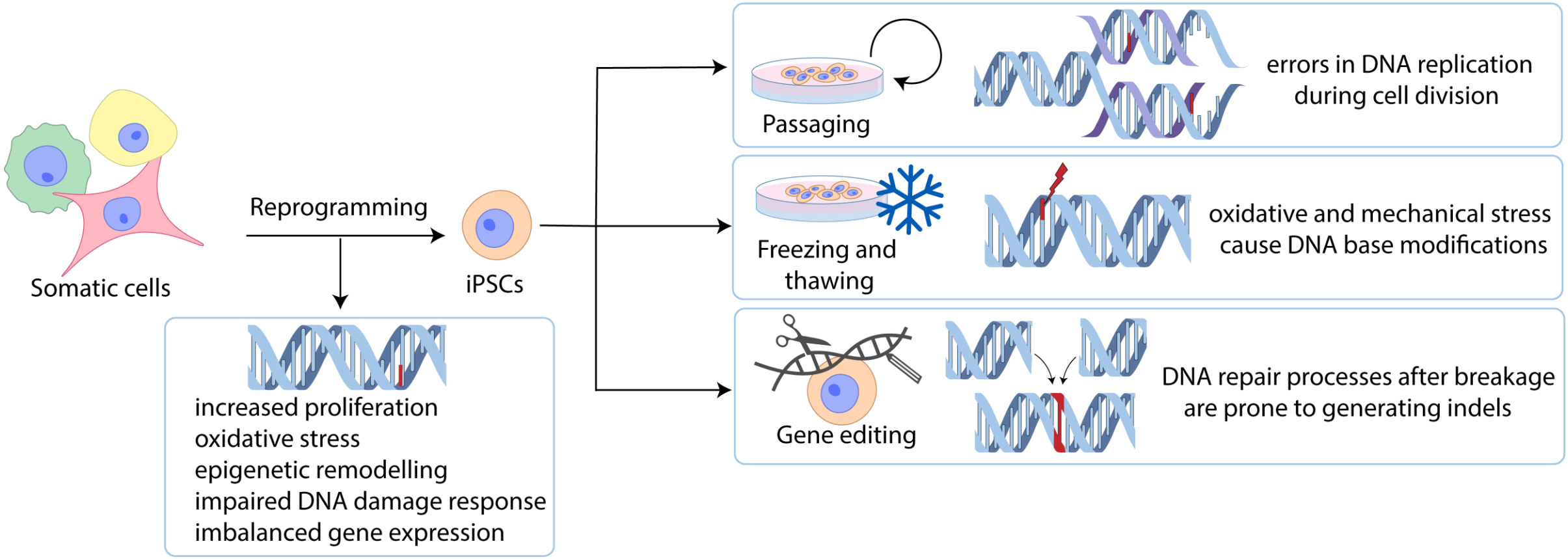
The iPSC starting material and final ATMP undergo rigorous safety assessment





Adapted from Kirkeby A, Main H, Carpenter M. Pluripotent stem-cell-derived therapies in clinical trial: A 2025 update. *Cell Stem Cell*. 2025;32(1):10-37

Genomic variants can occur in iPSCs during reprograming and culturing

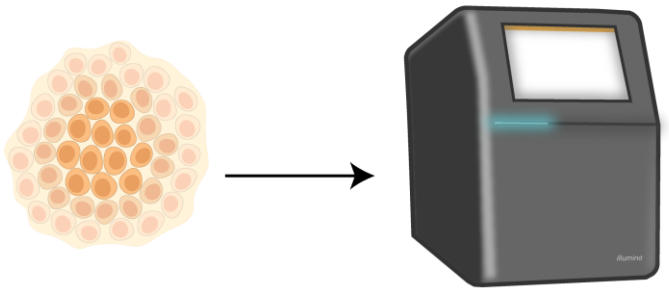


Standards for genomic safety assessment are actively being established

- (S.3.1) "**Genetic stability** should be evaluated for cell preparations that undergo extensive in vitro manipulation **using orthogonal methods**. When relevant, **cross reference to tumorigenicity studies in the non-clinical part** of the dossier can be made. " [1]
- (Recommendation 3.2.2.5) "Culture-acquired genetic abnormalities may be a significant risk and **should be part of in process and/or final product testing** for stem cell products that have undergone extensive expansion in vitro." [2]

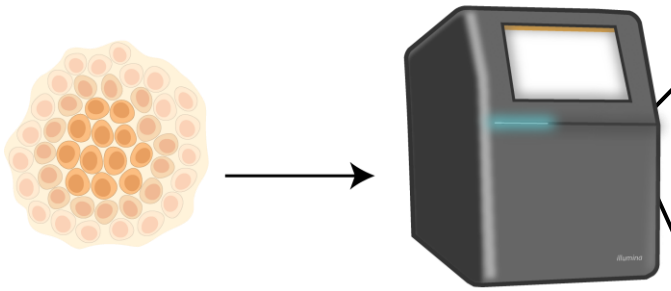
Goal: *Quantify information from NGS strategies (WGS, WES, RNAseq) relevant for safety assessment of ATMPs*

Next-generation sequencing allows large-scale sequencing of cell products

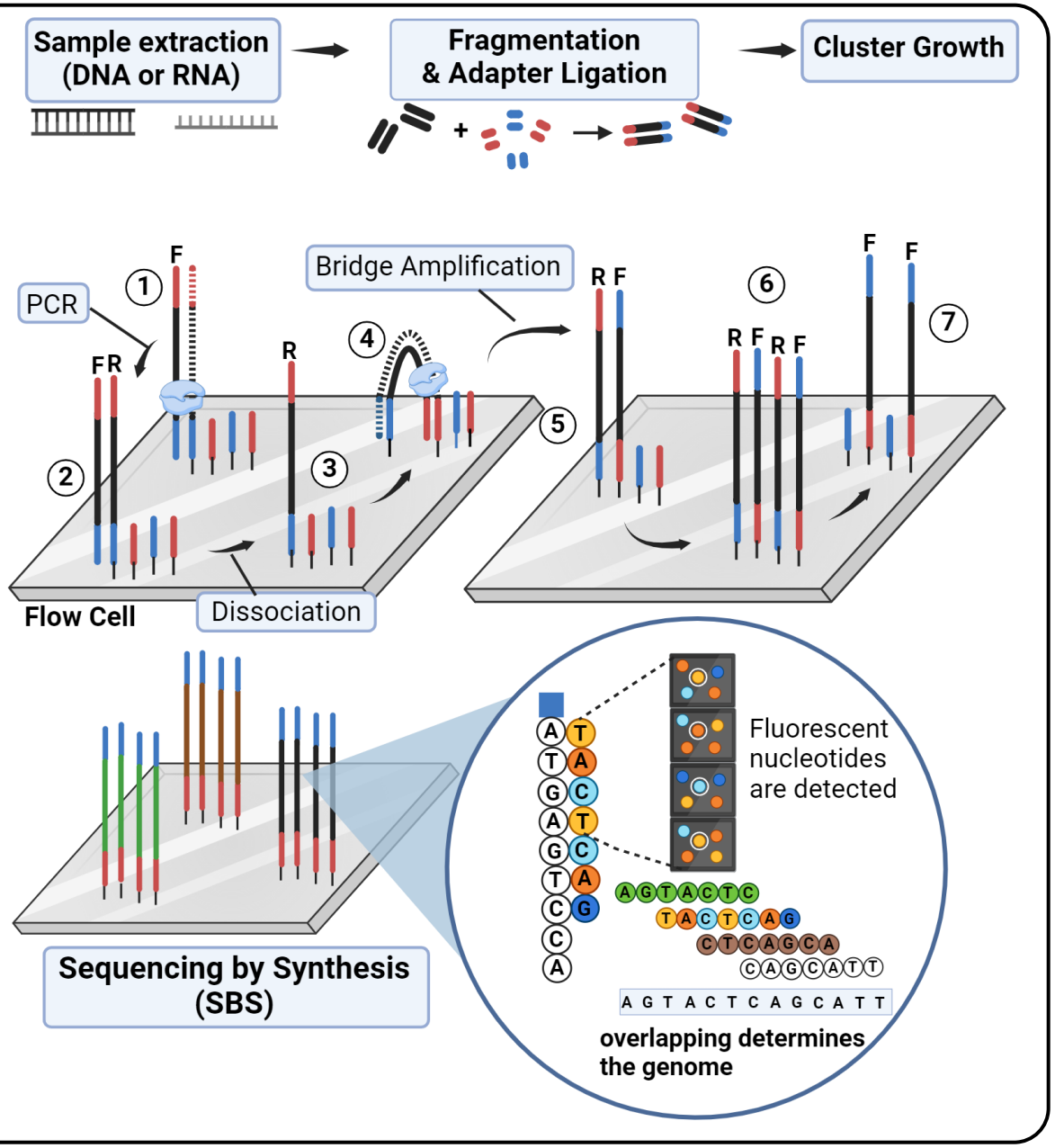


Sequencing of product

NGS Workflow

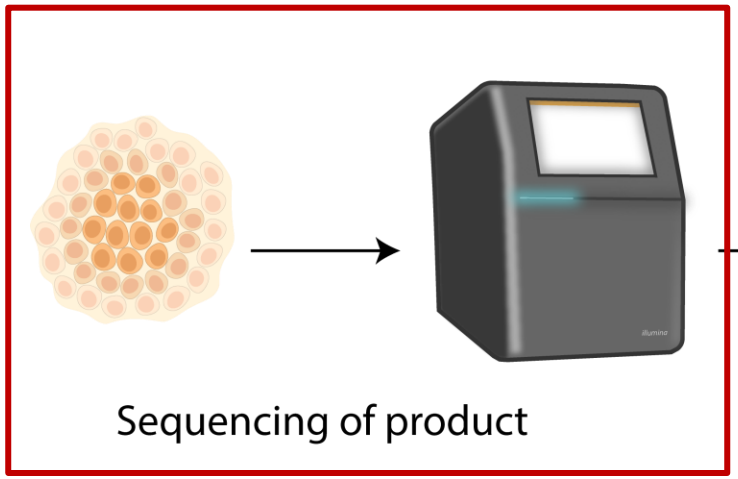


Sequencing of product

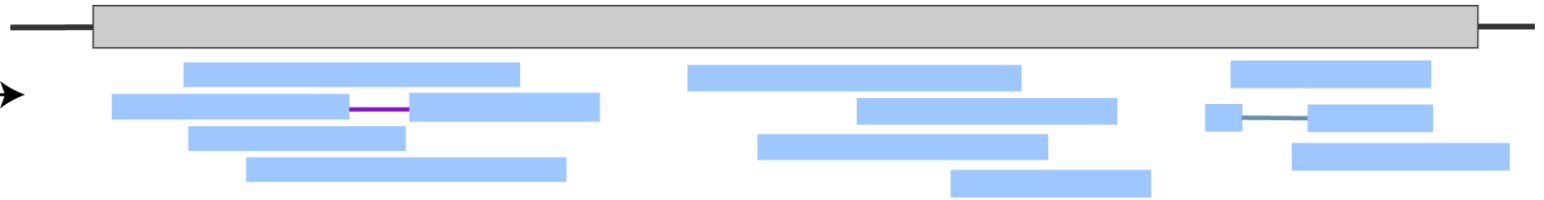


Xu, Z. (2025). Next-generation Sequencing (Workflow). https://app.biorender.com/biorender-templates/details/t-65f7c398cf888d0af6f7ec16-next-generation-sequencing-workflow

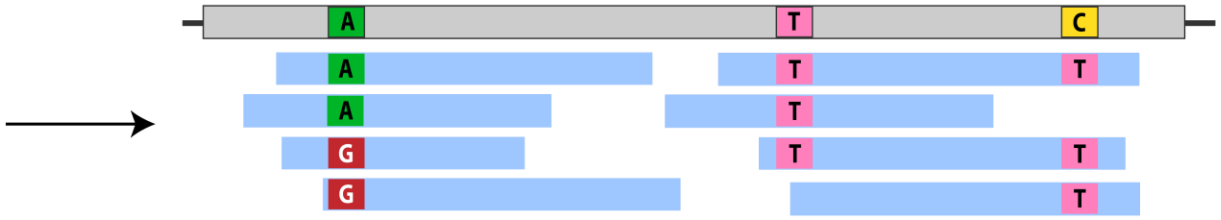
Genomic risk assessment includes variant calling and interpretation



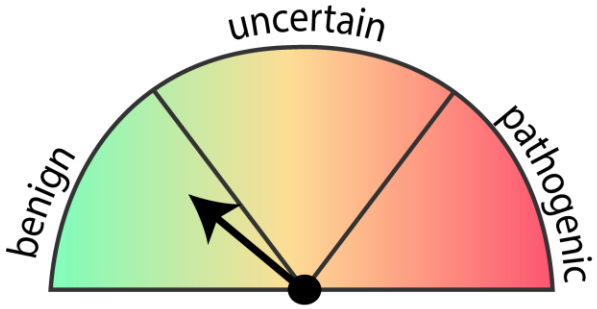
Sequencing of product



Alignment of sequencing reads to reference genome

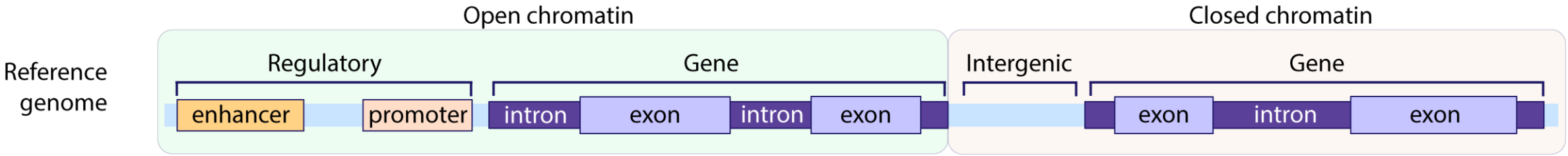


Variant calling

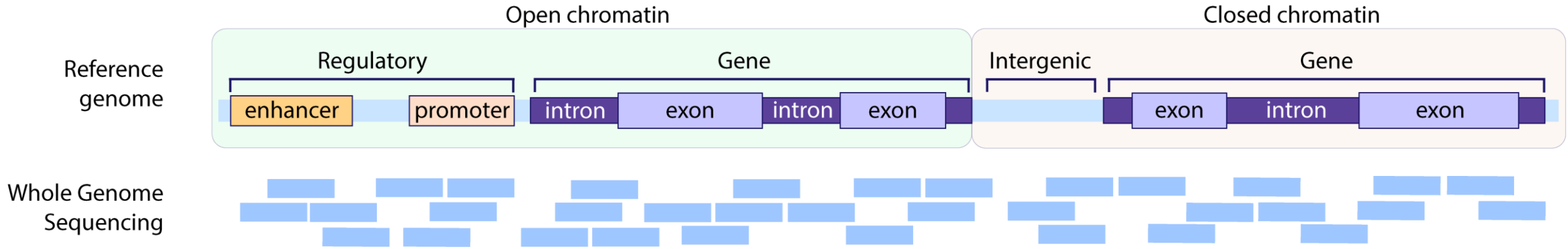


Variant interpretation

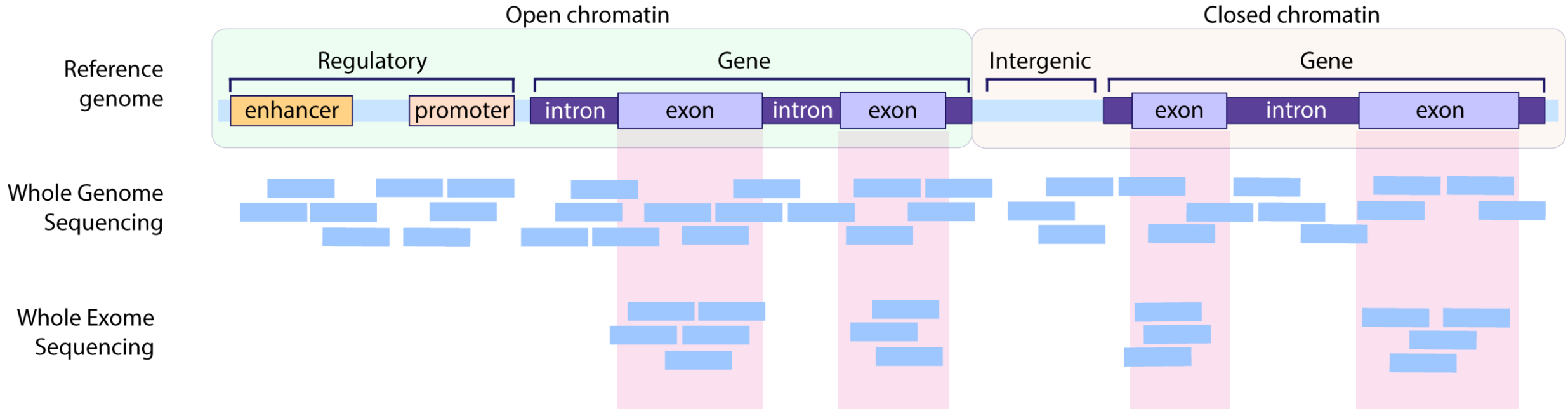
Sequencing methods differ in output



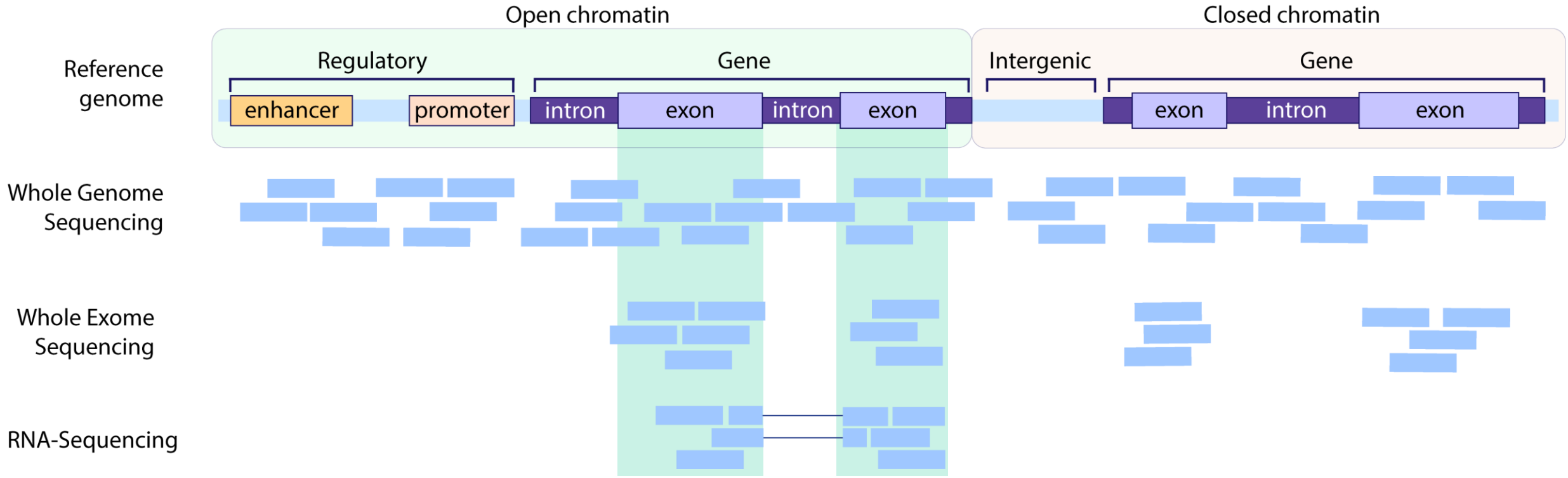
WGS uniformly covers the entire genome



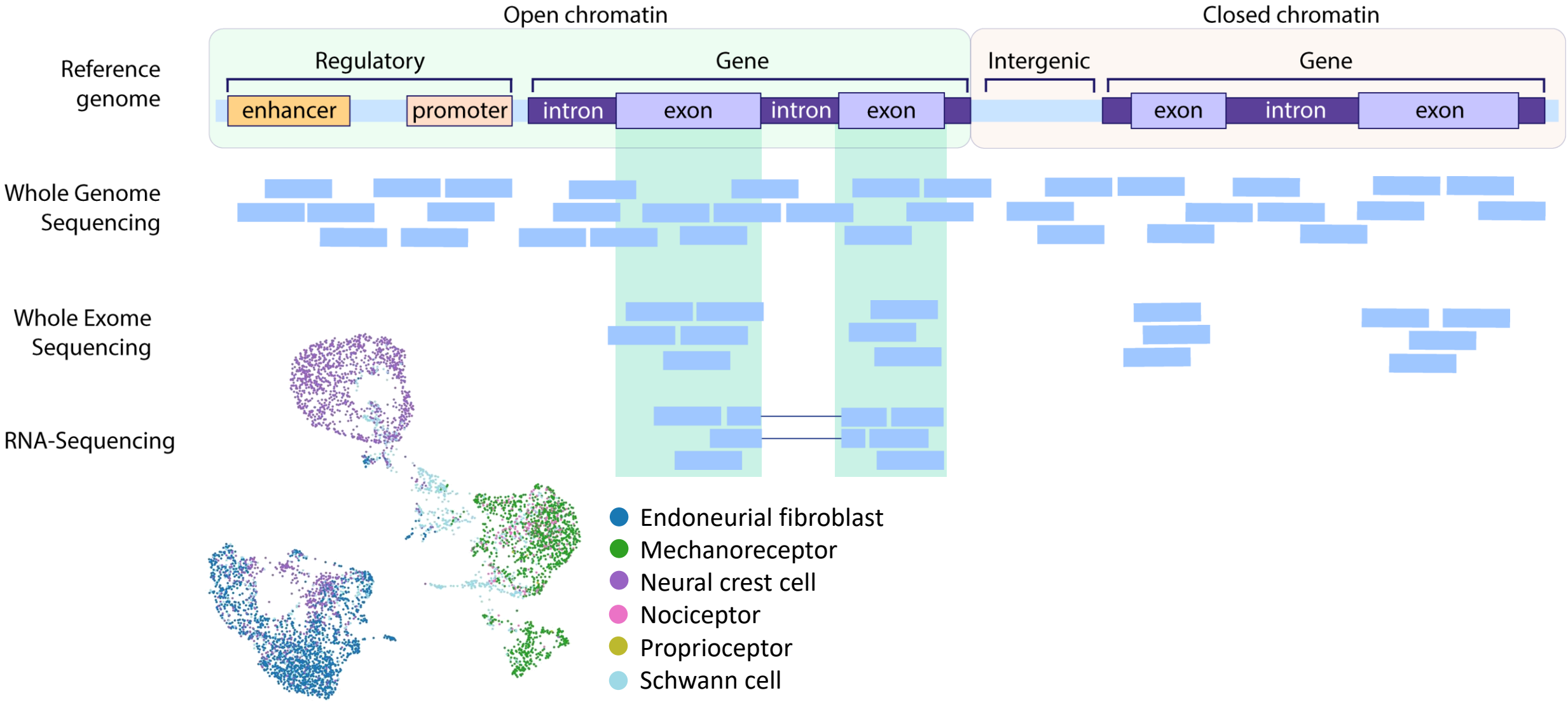
Exome sequencing enriches exonic regions through probe hybridization



RNAseq captures mature mRNAs through poly(A) tail capture



RNAseq also reveals gene expression patterns that define cell identity

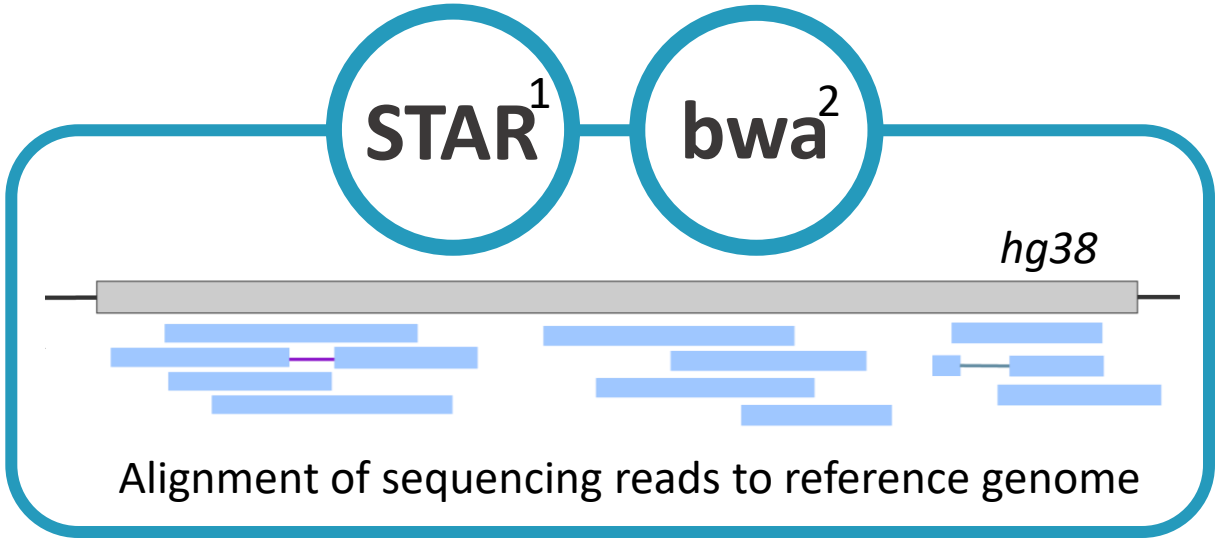


DNA and RNA-based sequencing reads are aligned with different methodologies

21 iPSC lines

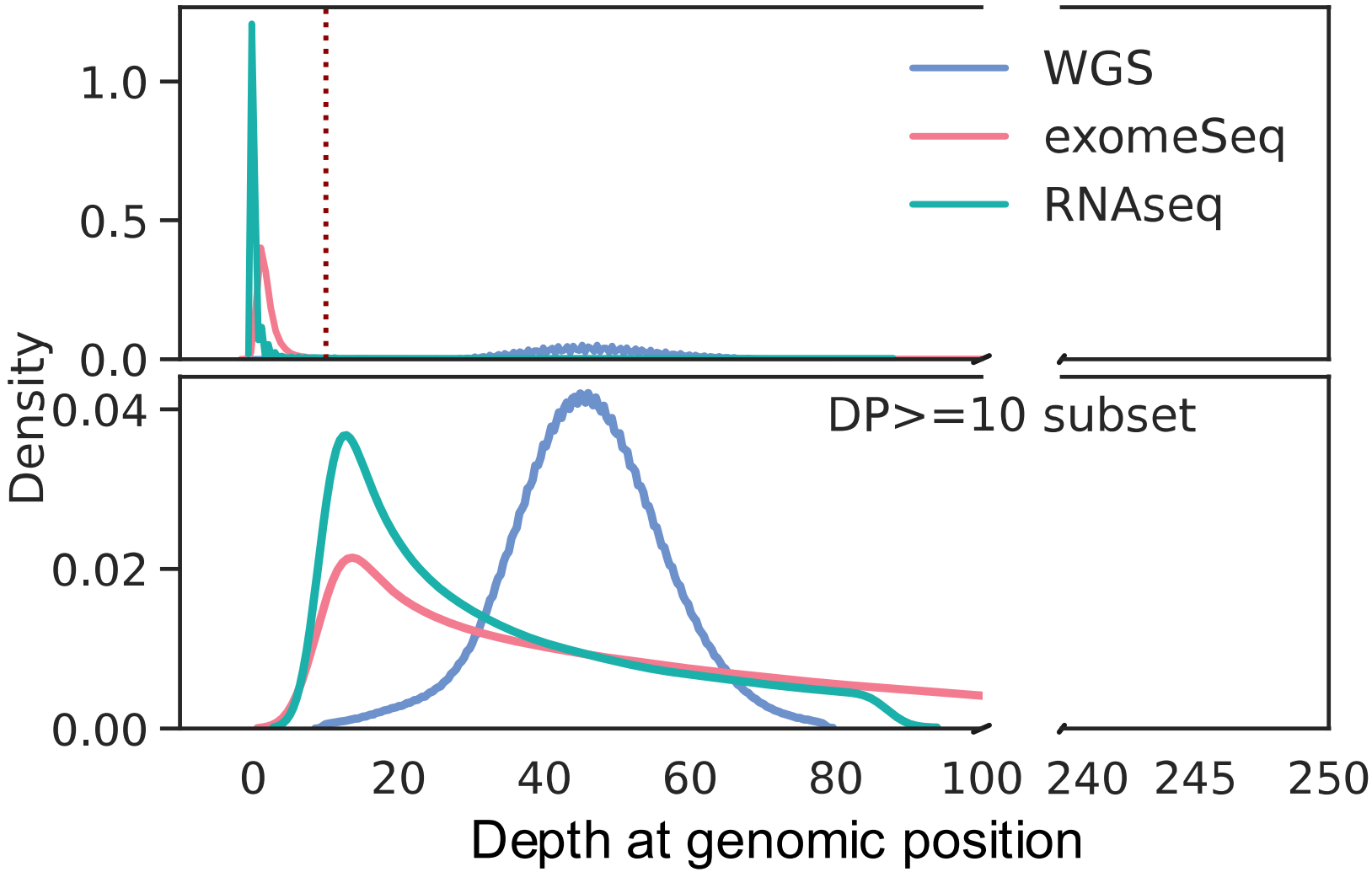
HipSci

Sequencing with
WGS, exomeSeq, RNAseq

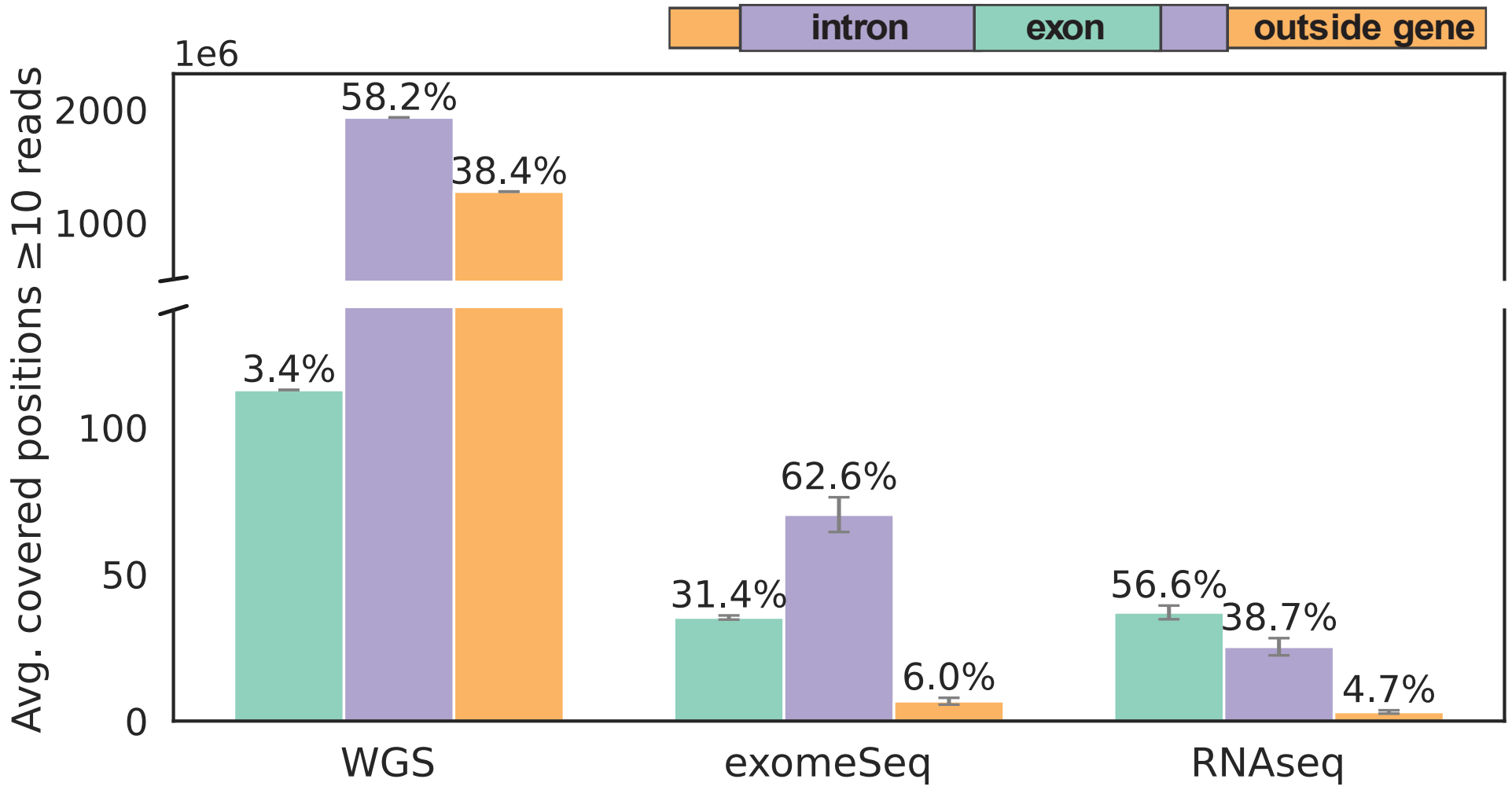


1. Dobin A. et al., *Bioinformatics* (2013), Li H. and Durbin R. , 2. *Bioinformatics* (2009), (3) Van der Auwera GA & O'Connor BD, *Genomics in the Cloud* (2020)

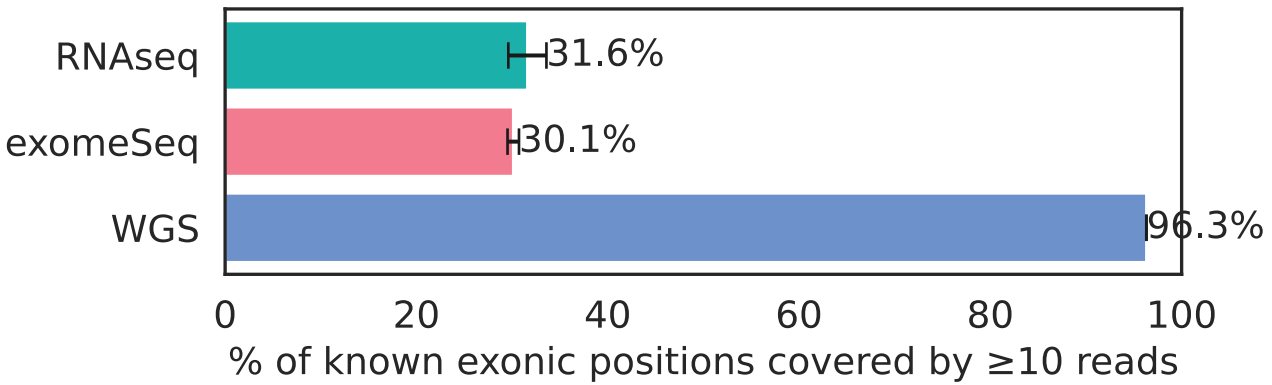
NGS displays a high number of low-coverage positions that are excluded



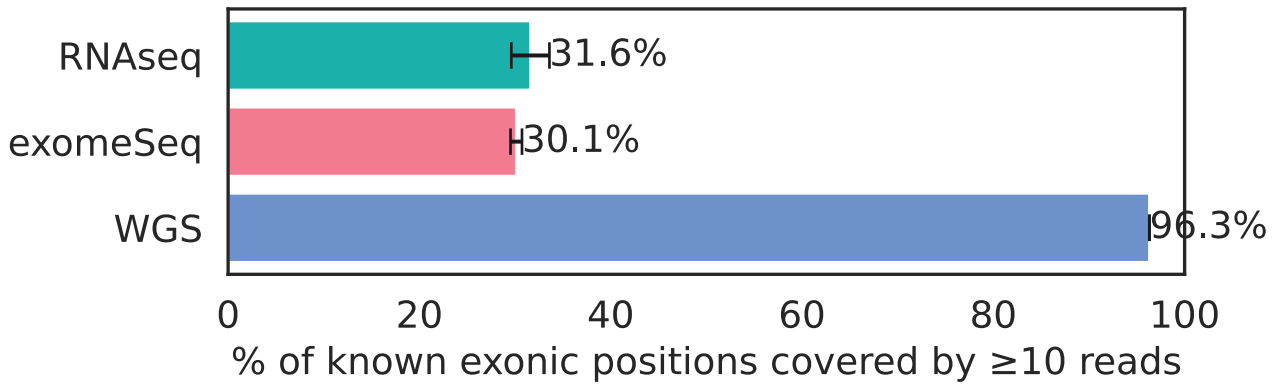
RNAseq data is enriched in exonic regions compared with ExomeSeq

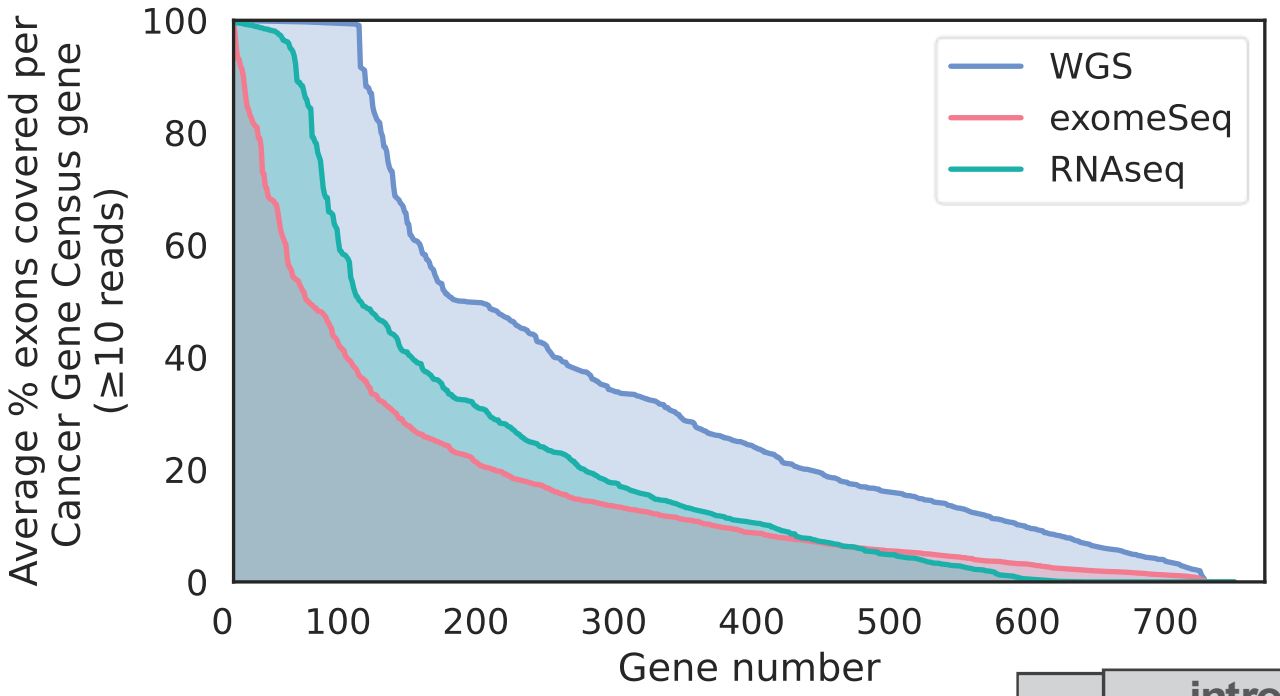


Exon coverage is similar between RNAseq and exomeSeq

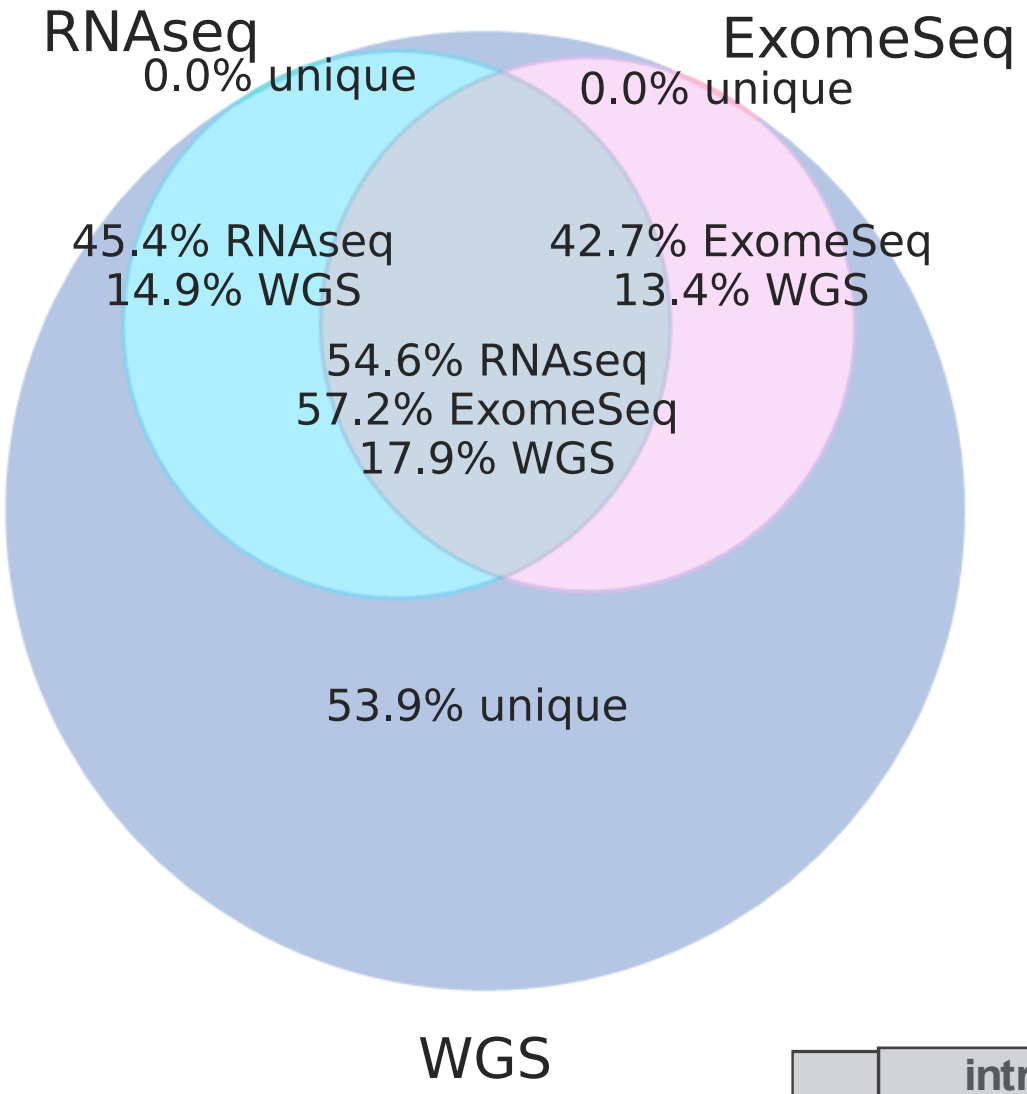


RNAseq reliably captures more cancer gene exons than exomeSeq





RNAseq and ExomeSeq cover different coding regions



Variant calling is performed using the same software

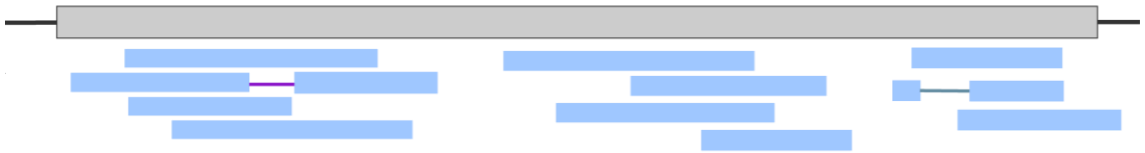
21 iPSC lines

HipSci

Sequencing with
WGS, exomeSeq, RNAseq

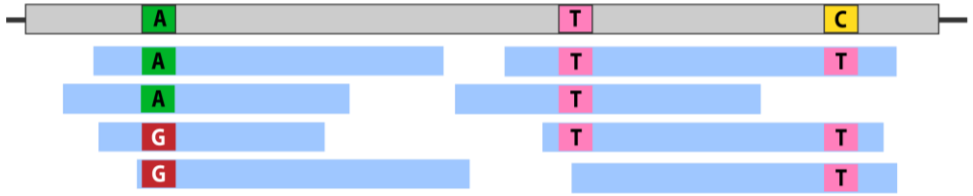
STAR[1] bwa[2]

hg38



Alignment of sequencing reads to reference genome

gatk[3]

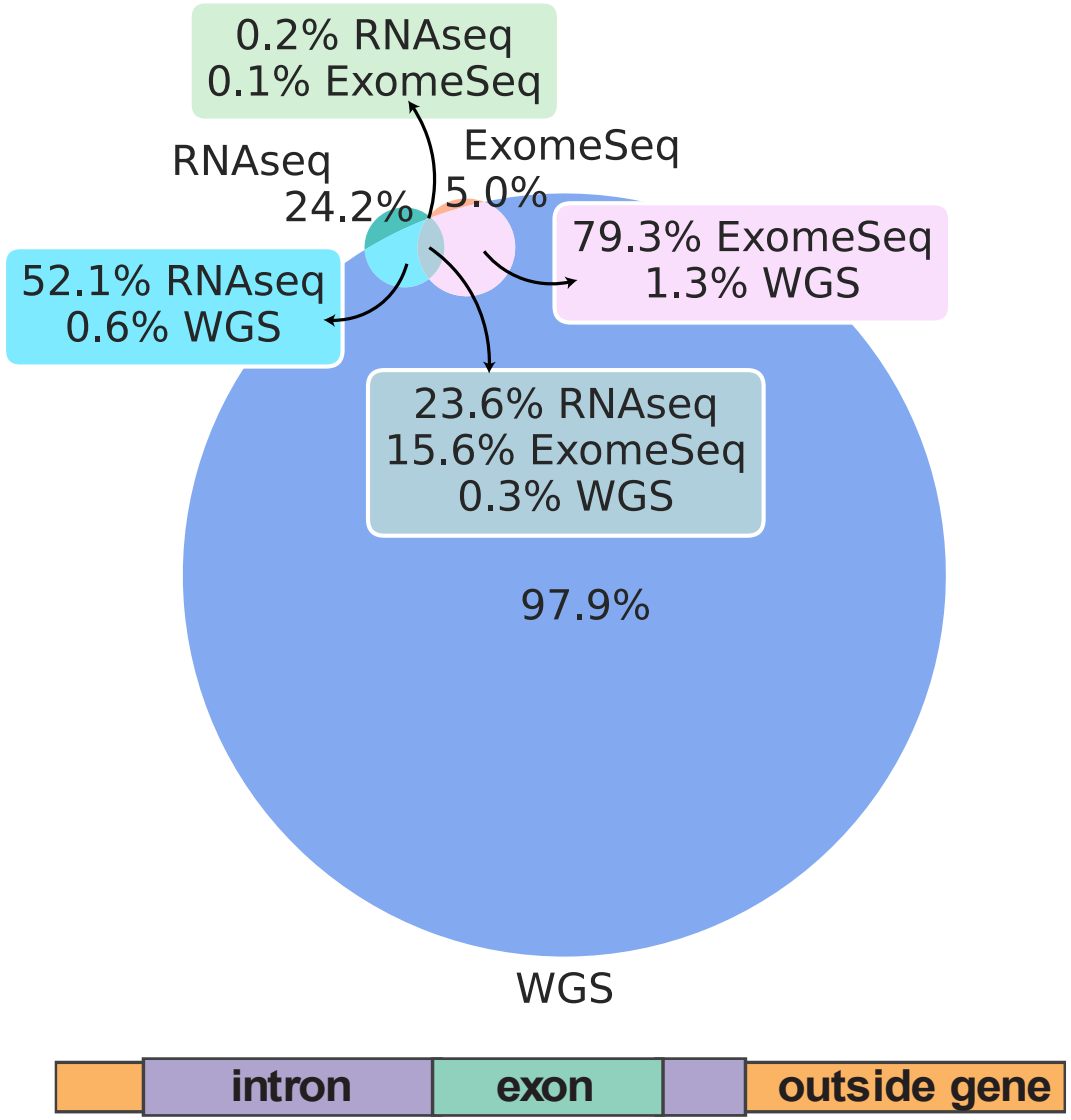


Variant calling

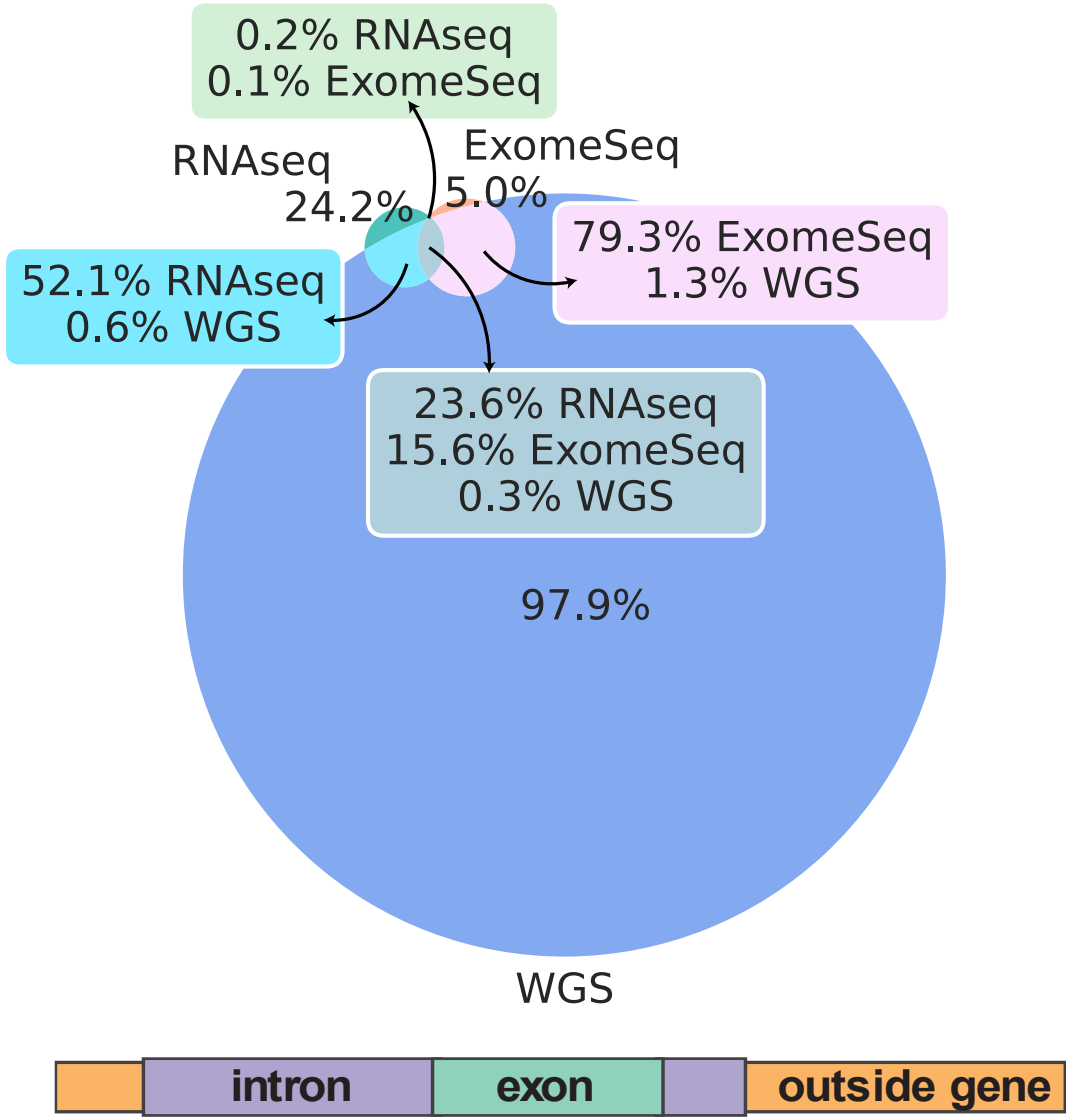
1. Marking duplicates
2. Correcting for sequencing errors
3. Calling the variants

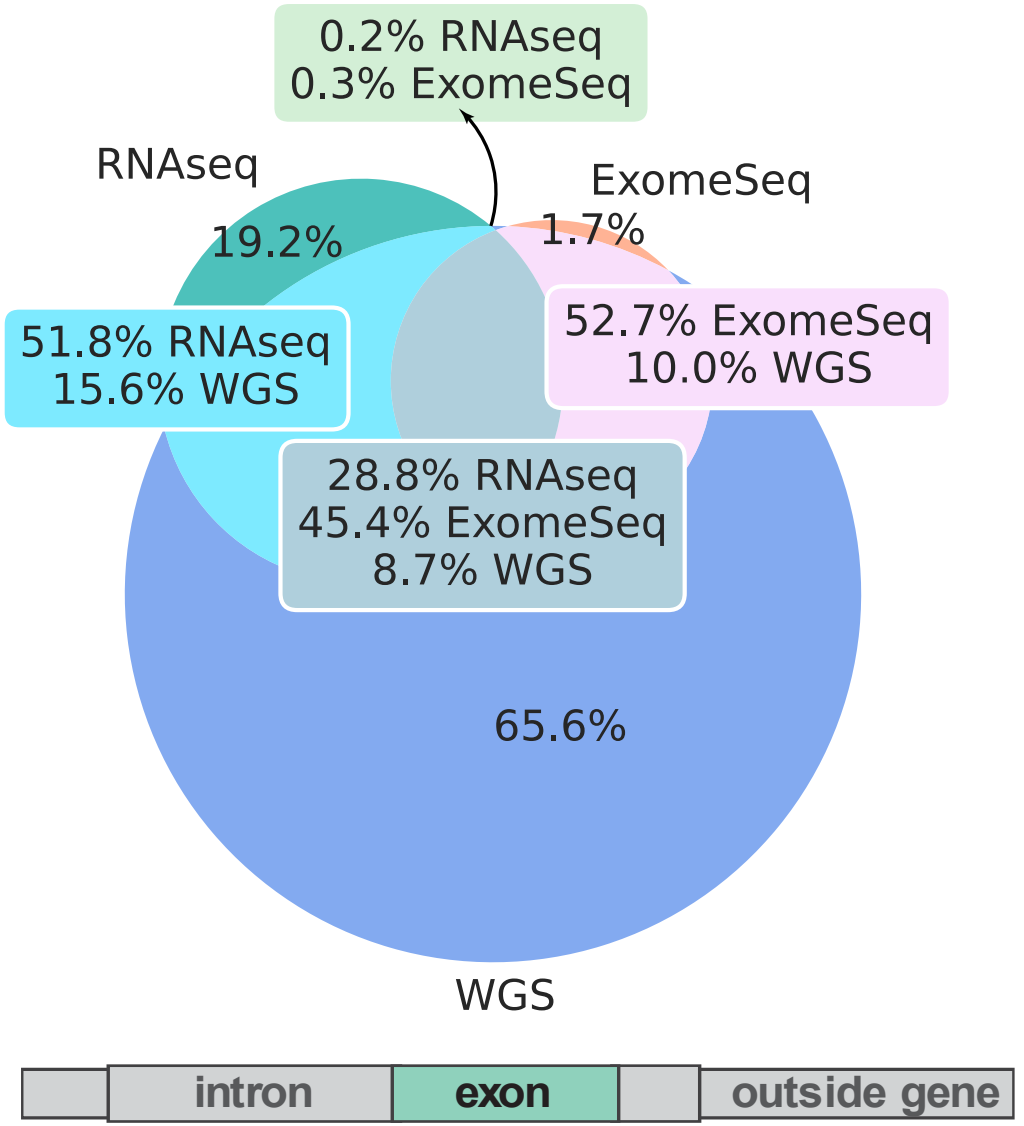
1. Dobin A. et al., *Bioinformatics* (2013), Li H. and Durbin R. , 2. *Bioinformatics* (2009), (3) Van der Auwera GA & O'Connor BD, *Genomics in the Cloud* (2020)

Small variant overlap between genomics and transcriptomics methods

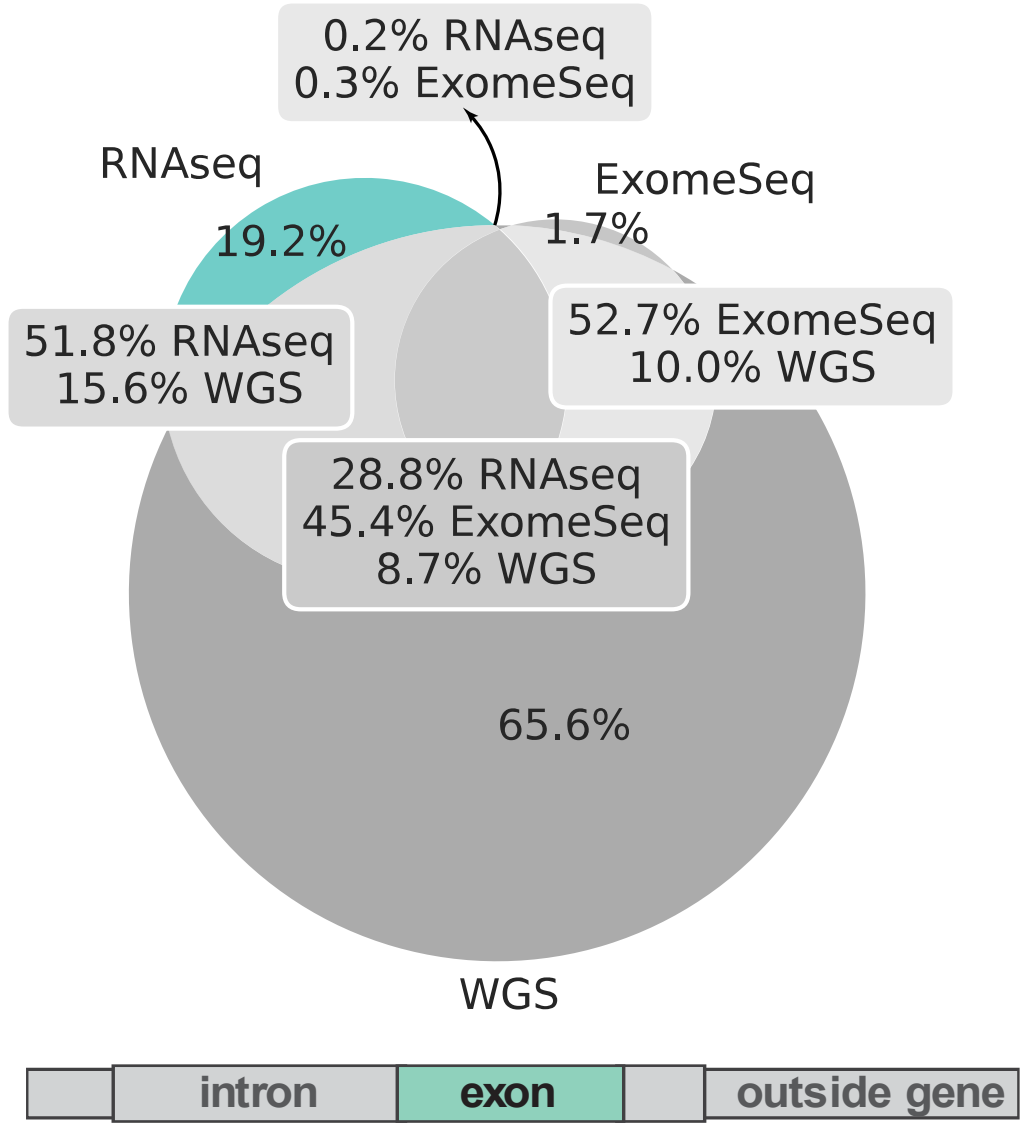


RNAseq captures 23% of WGS variants, while exomeSeq only 18%

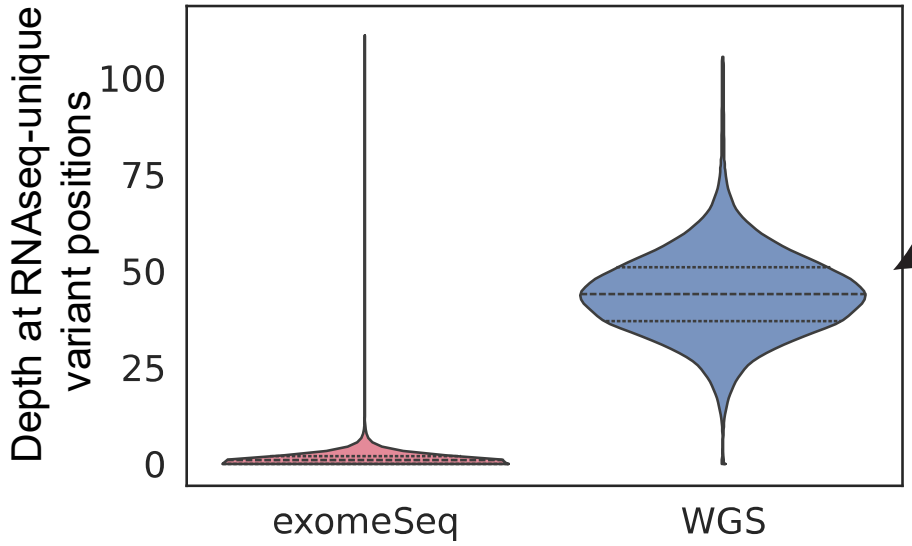


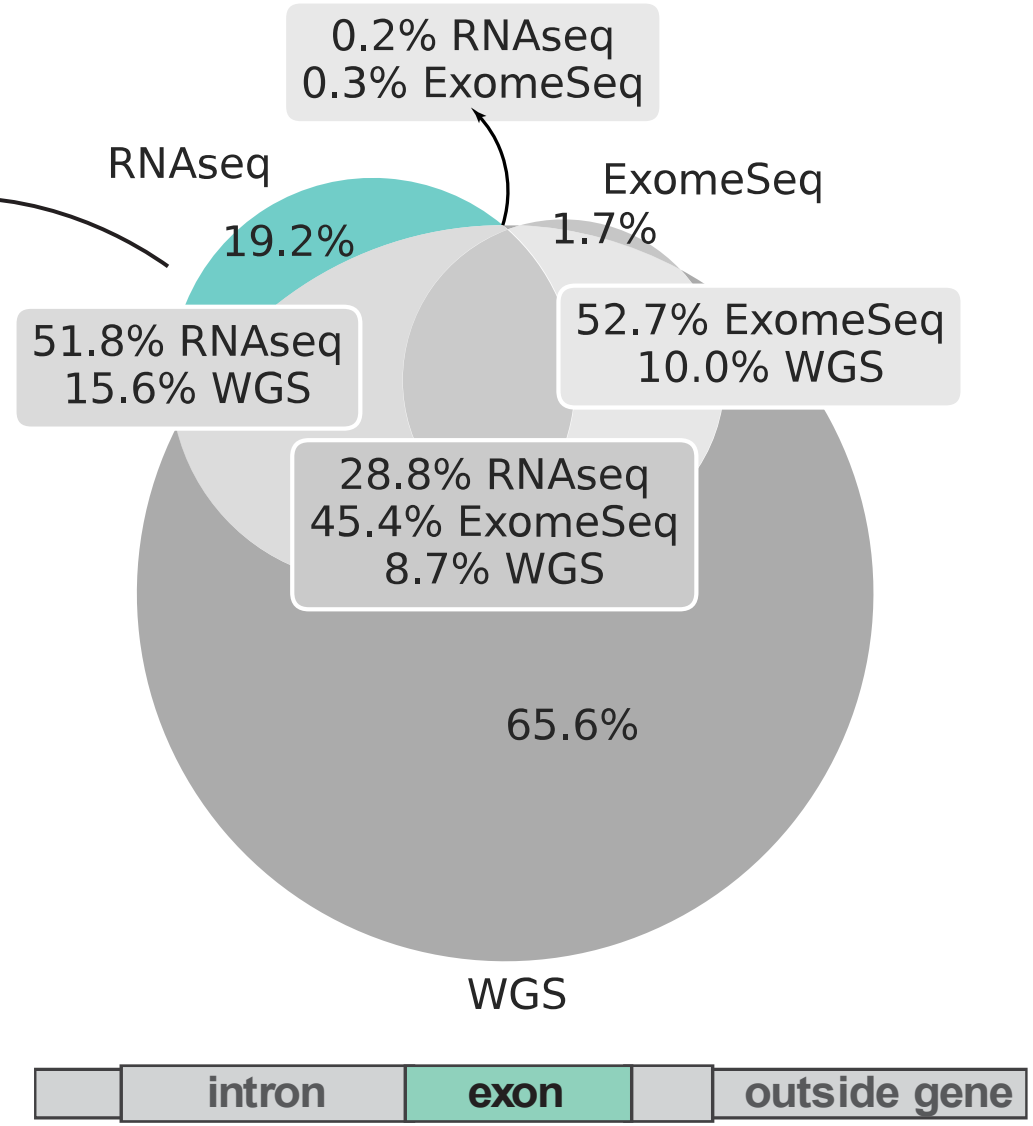


19% of variants captured using RNAseq are specific to this method

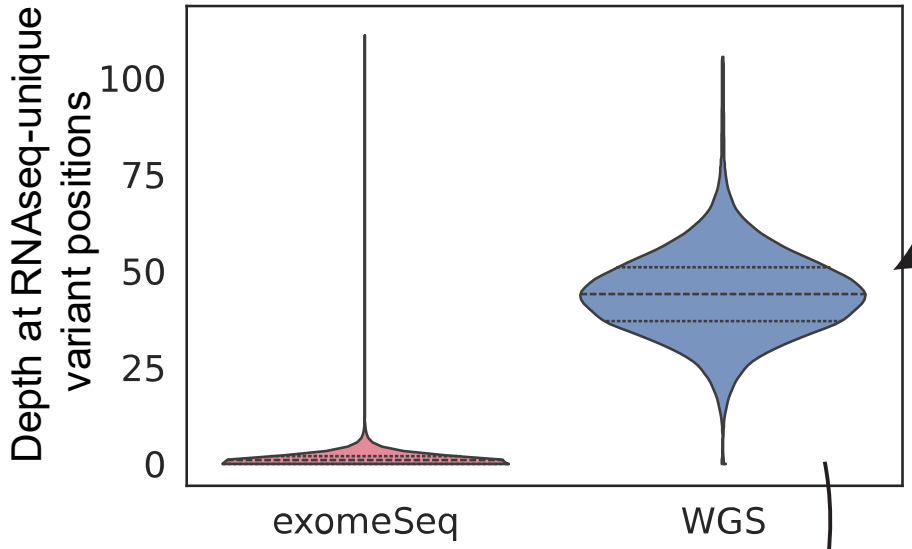


RNAseq-unique variant positions are covered in WGS but not in exomeSeq

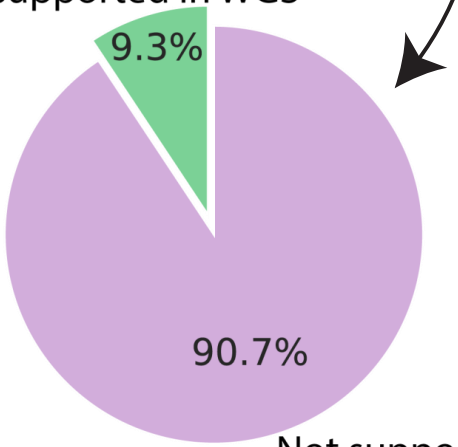




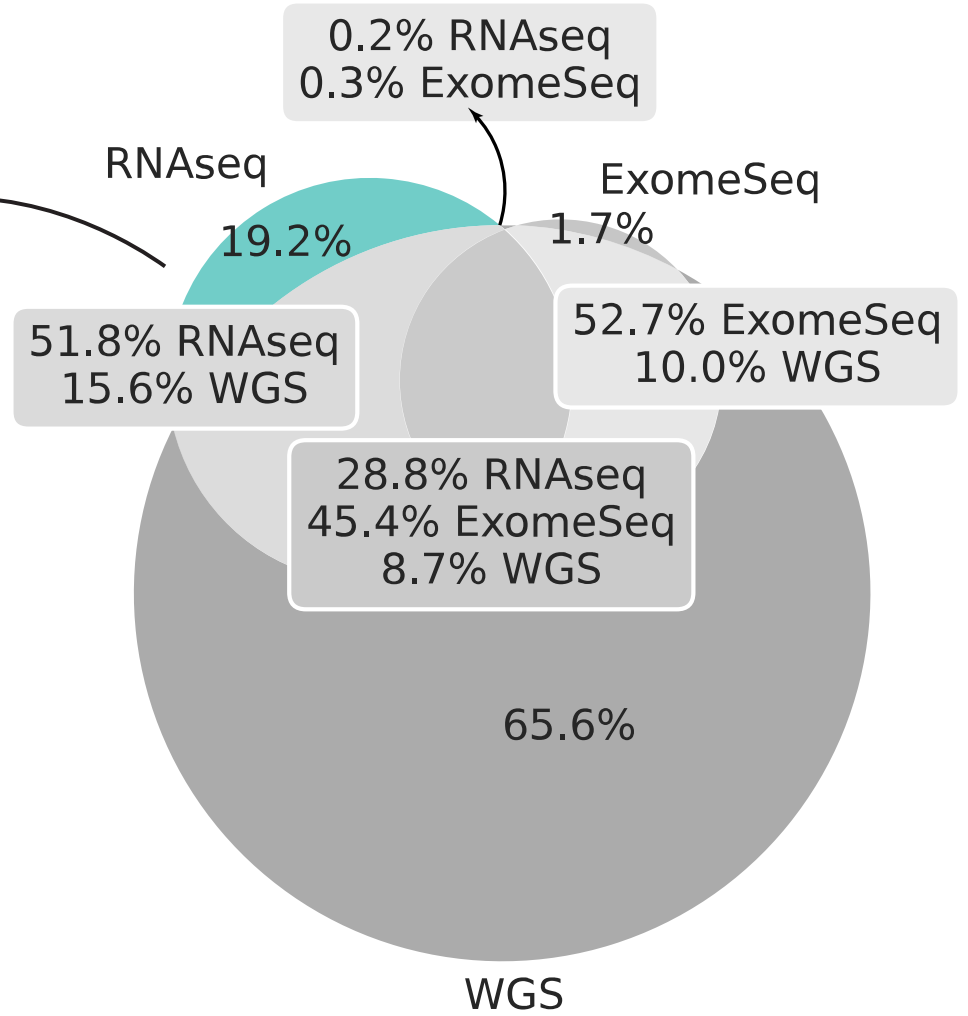
Over 90% of RNAseq-specific variants are not confirmed by WGS



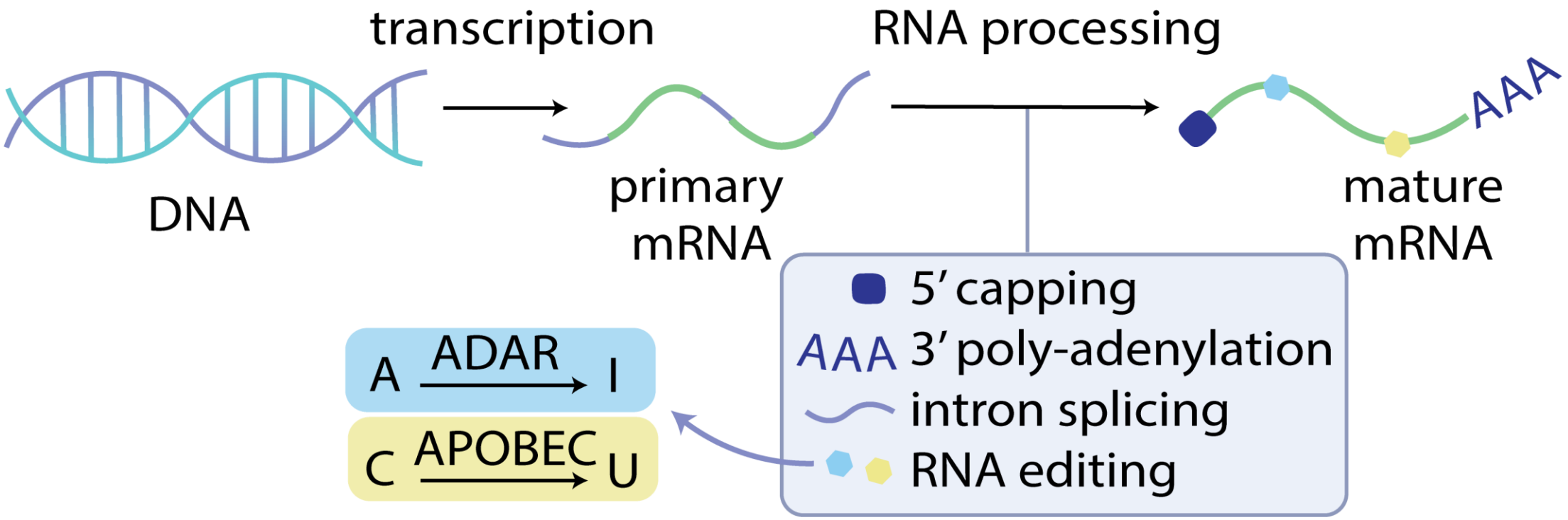
RNAseq-unique variants supported in WGS



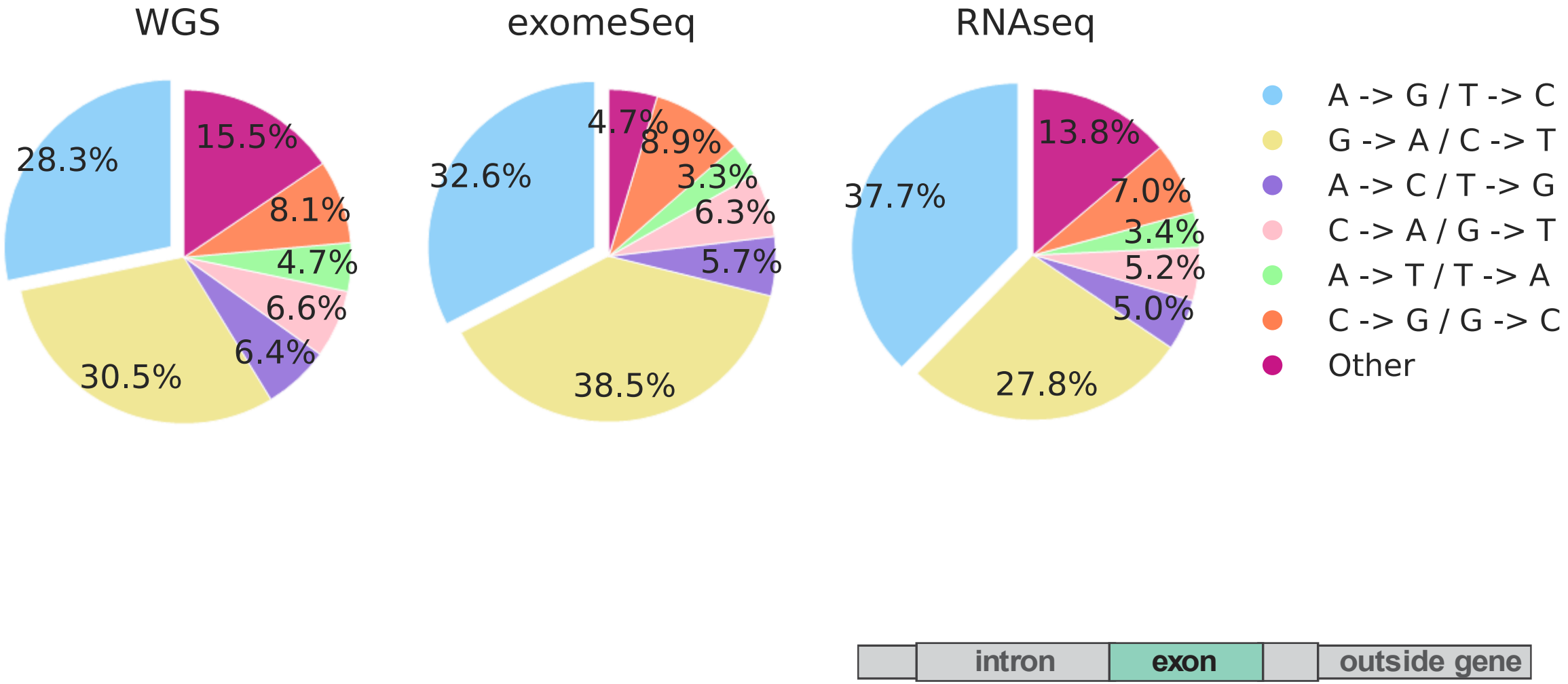
Not supported (ALT DP = 0)



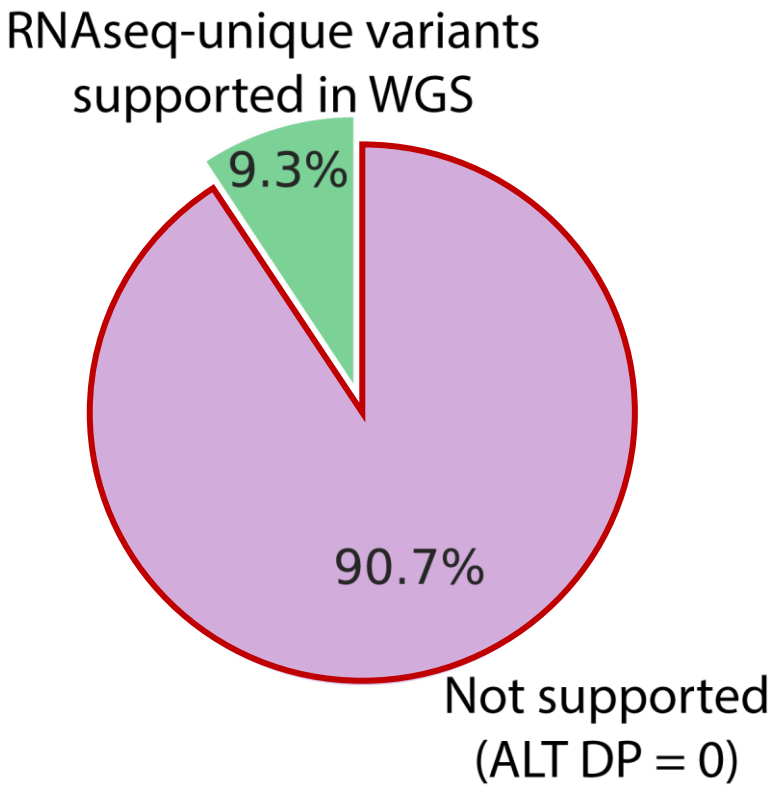
RNA editing can be detected by RNAseq, not by DNA-based sequencing methods



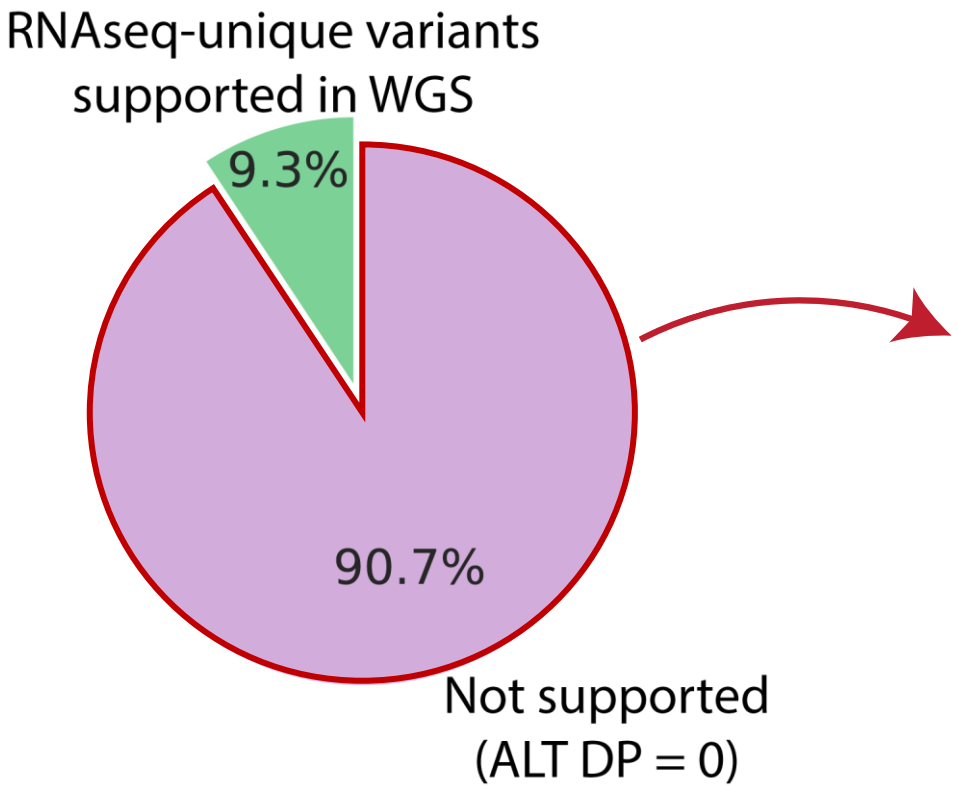
The rate of variation for different base changes is similar among methods

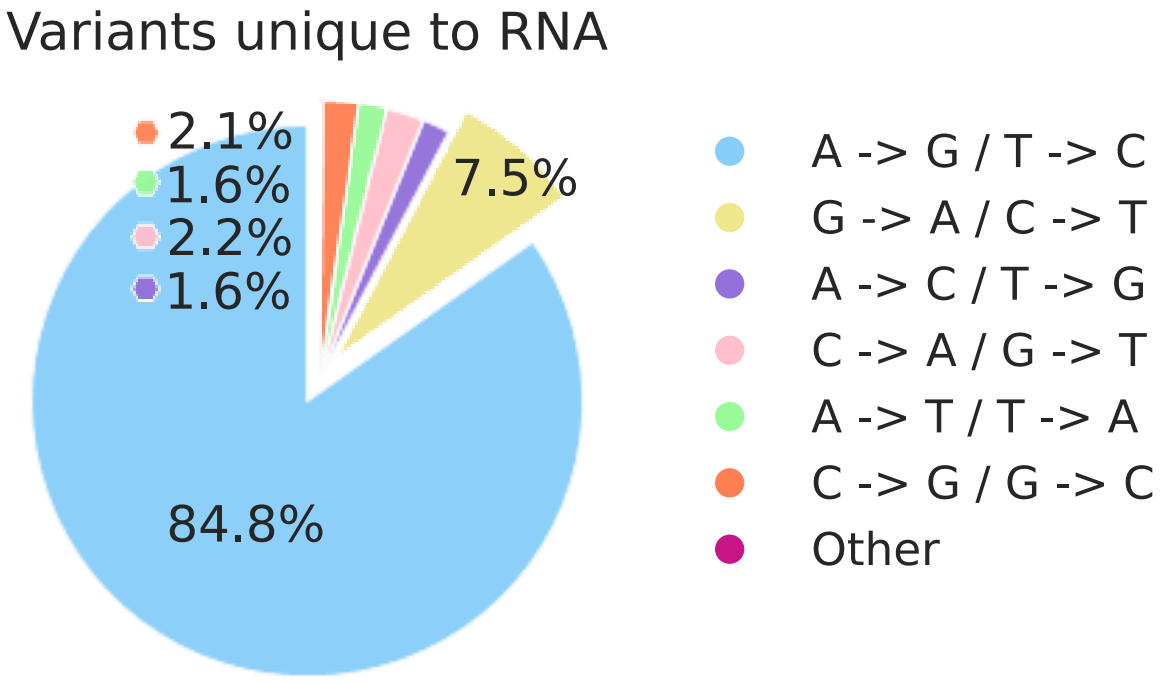


92% of variants unique to RNA follow RNA editing patterns



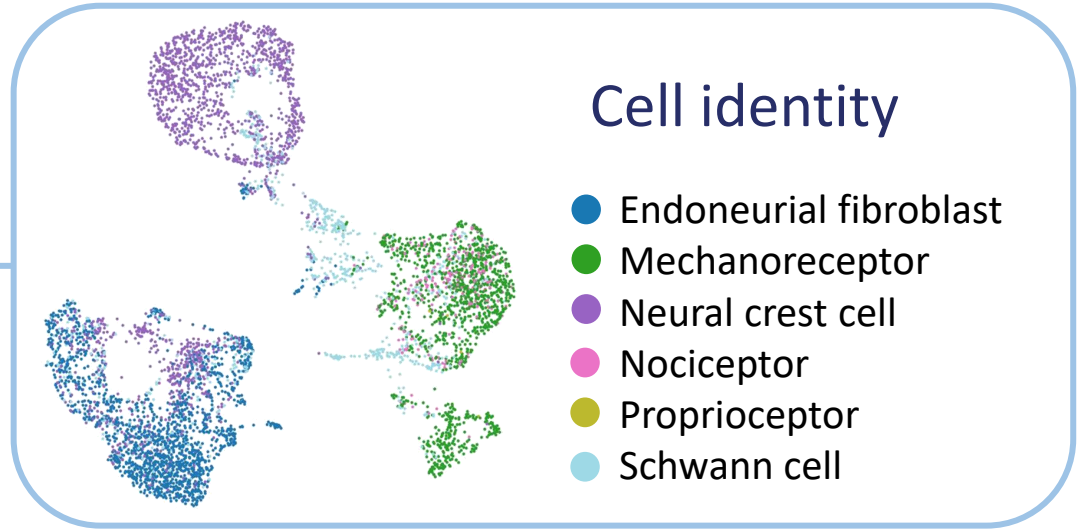
92% of variants unique to RNA follow RNA editing patterns





Take home messages

RNAseq



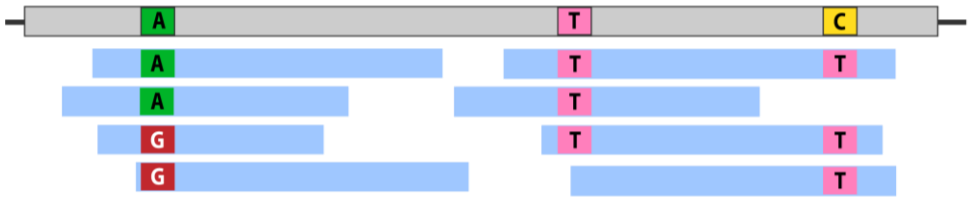
Take home messages

RNAseq

Cell identity

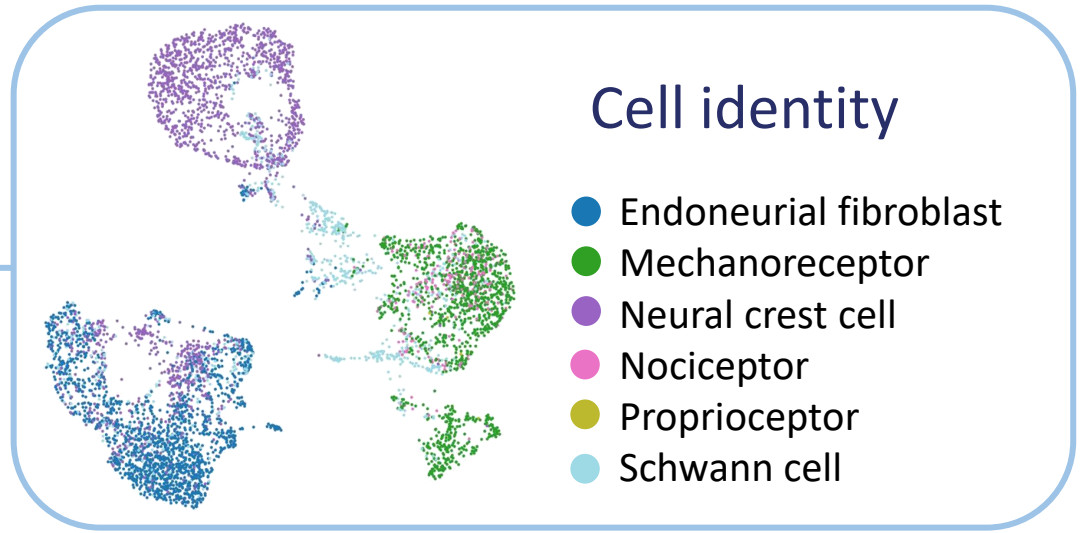
- Endoneurial fibroblast
- Mechanoreceptor
- Neural crest cell
- Nociceptor
- Proprioceptor
- Schwann cell

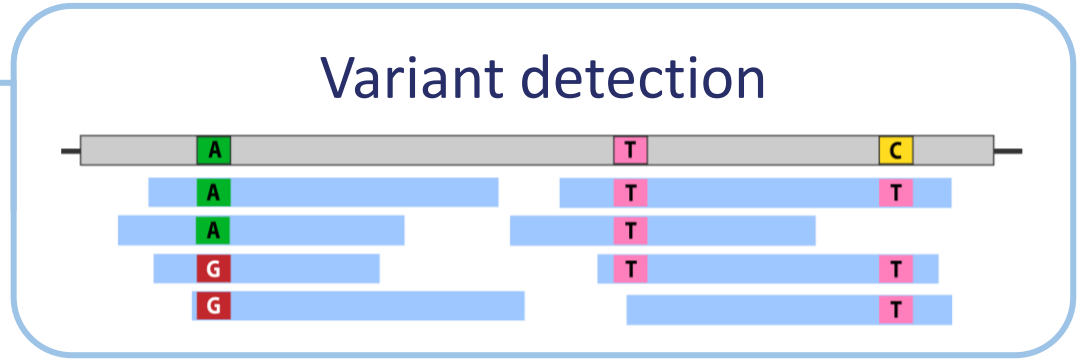
Variant detection



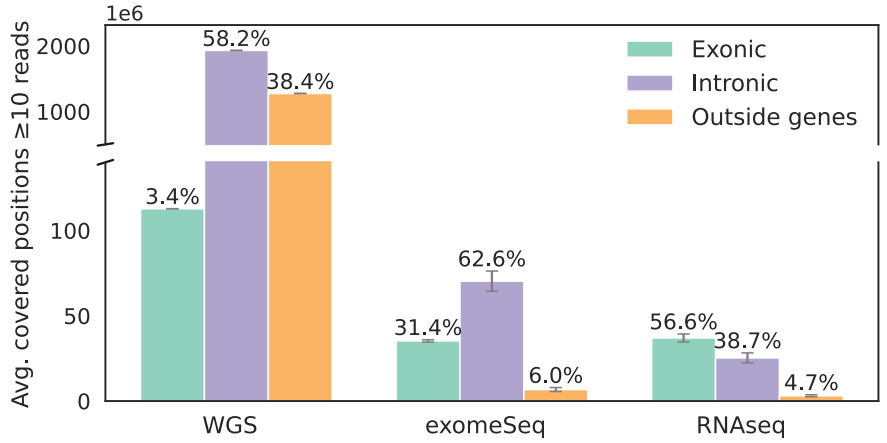
Take home messages

RNAseq



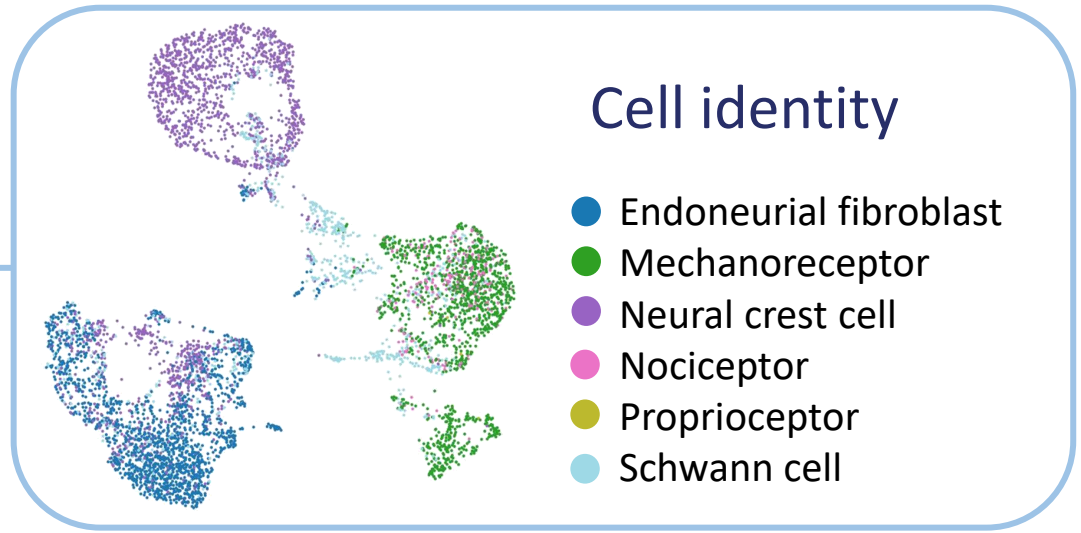


- RNAseq has a **higher relative focus on coding regions** that exomeSeq

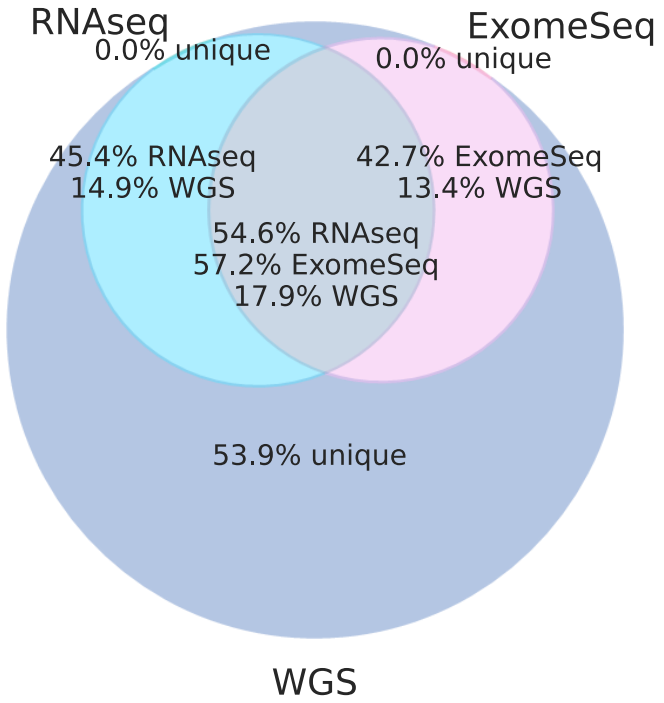


Take home messages

RNAseq

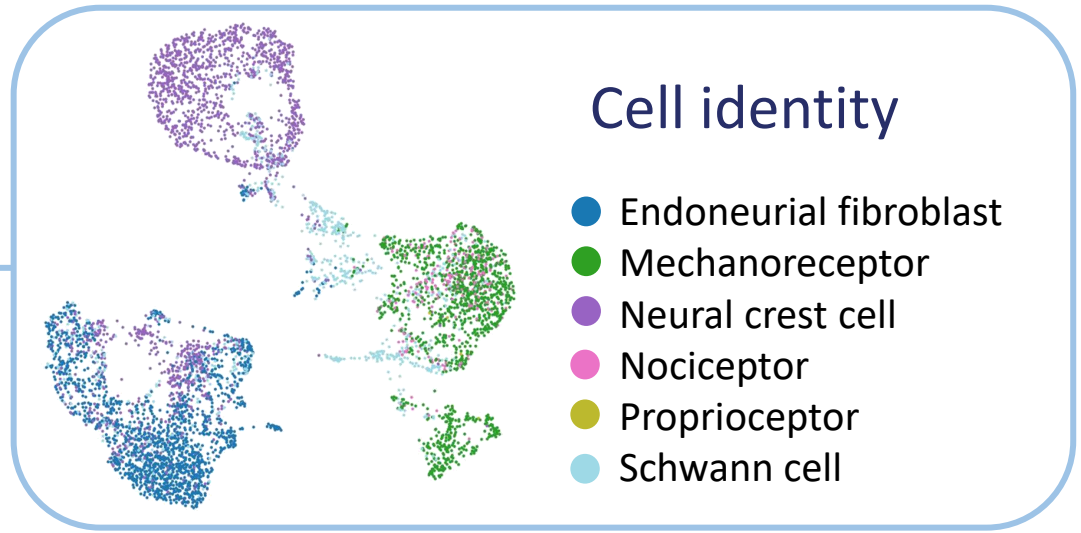


- RNAseq has a **higher relative focus on coding regions** that exomeSeq
- RNAseq and exomeSeq **share coverage** over **only ~50%** of their respective regions

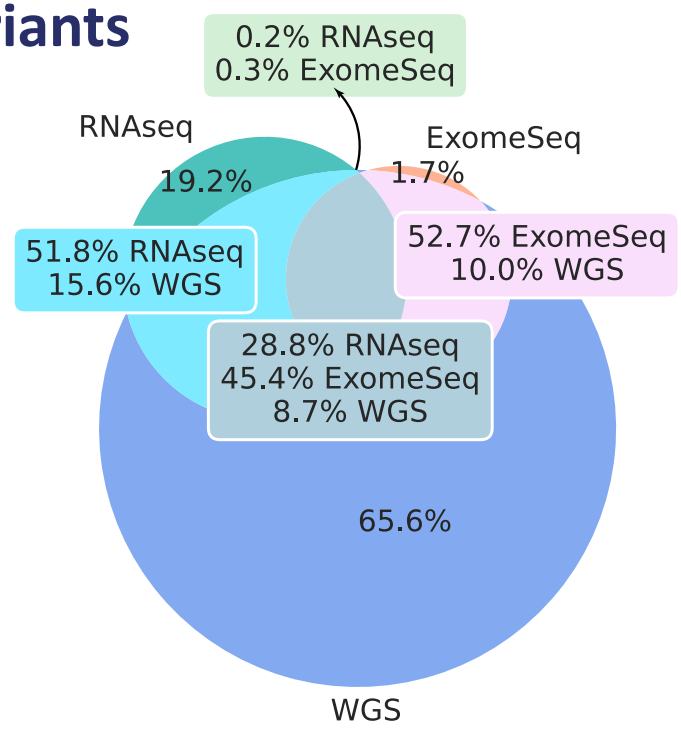


Take home messages

RNAseq

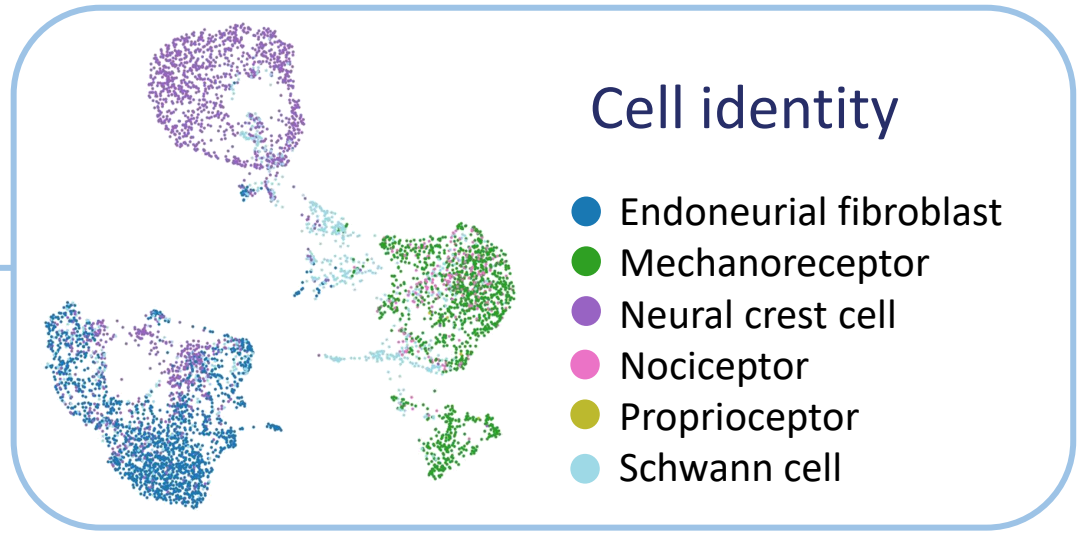


- RNAseq has a **higher relative focus on coding regions** that exomeSeq
- RNAseq and exomeSeq **share coverage** over **only ~50%** of their respective regions
- RNAseq captures **23 % of WGS coding variants**

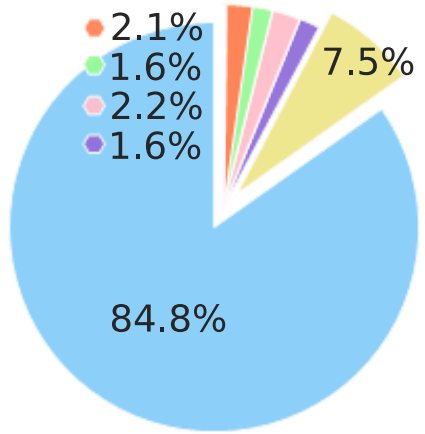


Take home messages

RNAseq



- RNAseq has a **higher relative focus on coding regions** that exomeSeq
- RNAseq and exomeSeq **share coverage** over **only ~50%** of their respective regions
- RNAseq captures **23 % of WGS coding variants**
- The genomic variants captured by RNAseq that cannot be recapitulated in WGS are likely **RNA editing events**



Acknowledgements

Universiteit Leiden

Prof. Micha Drukker, PhD
Christian Schröter, PhD

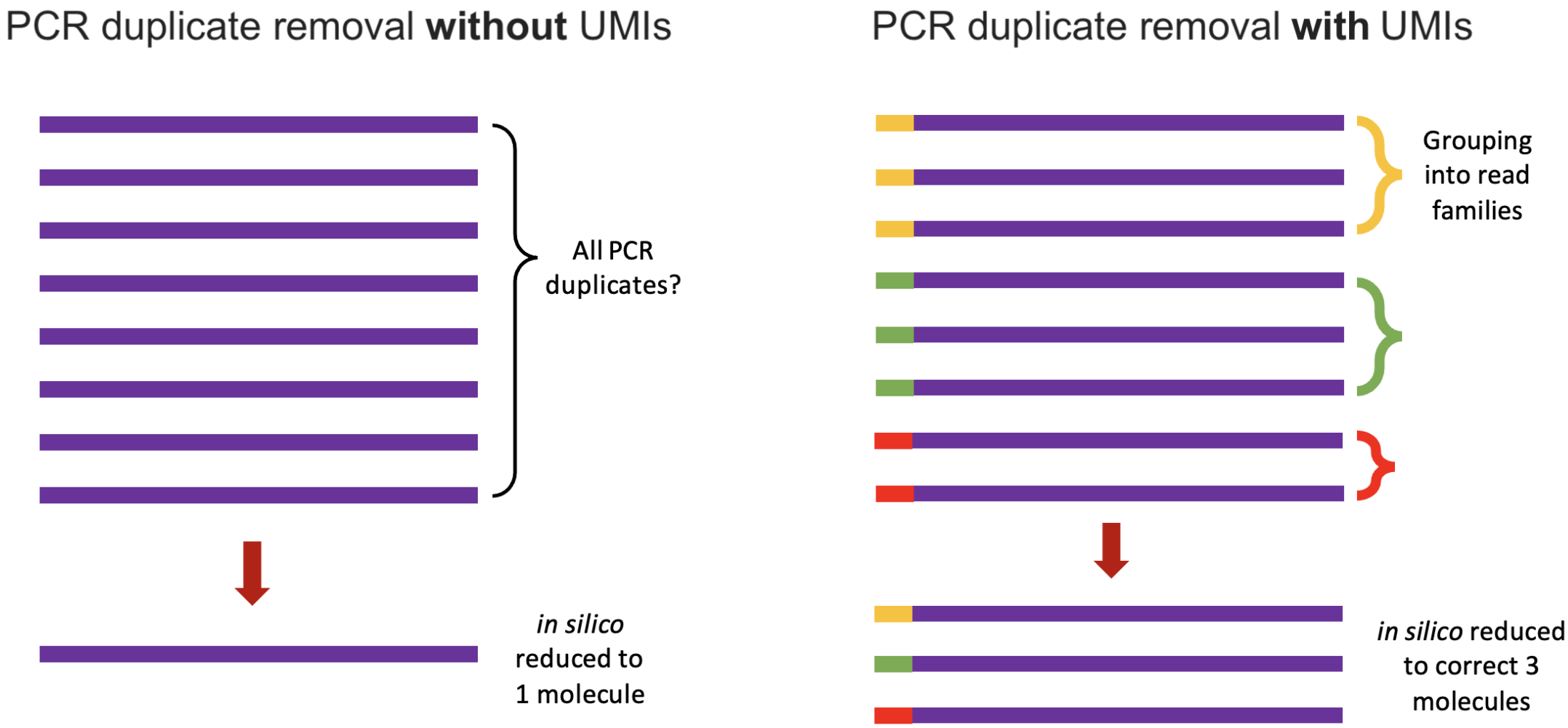
Andreea Iosif, MSc
Daan Vlemmings, Ir.
Cecile Herbermann, BSc
Nina Schulten, MSc
Jingchao Wu, PhD
Max Fernkorn, PhD
Benjamin Tak, MSc
Yuting Wang, MSc
Ruben de Vries, MSc
Eliane Züger, MSc
Joey Zuijdervelt, MSc
Yubin Guo, MSc
Ferdinand Teichert, Dr. Med. Univ.

CBG MEB

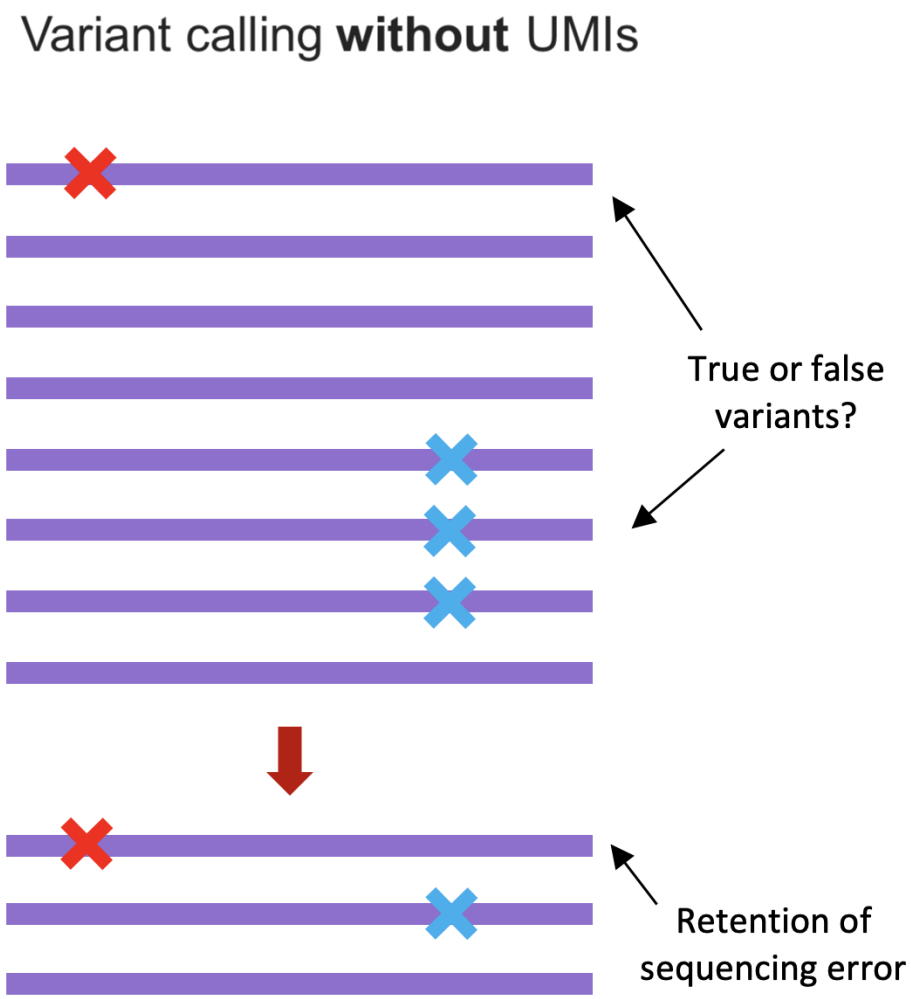
Tineke van den Hoorn, PhD
Marcel Hoefnagel, PhD

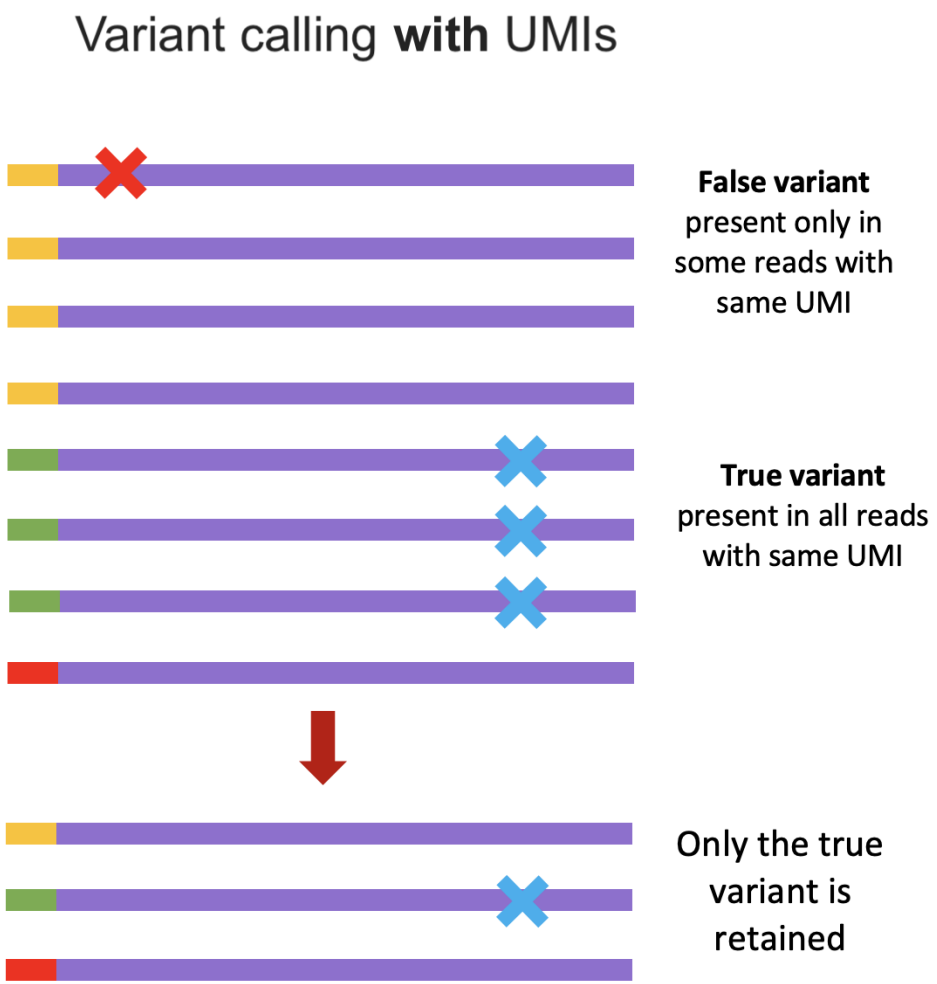
drukkerlab.nl

Extra: Unique Molecular Identifiers (UMIs) tag each sequence in a sample library



Extra slide: Variant calling using UMIs allows detection of sequencing errors





Extra slide: Variant overlap between the three methods