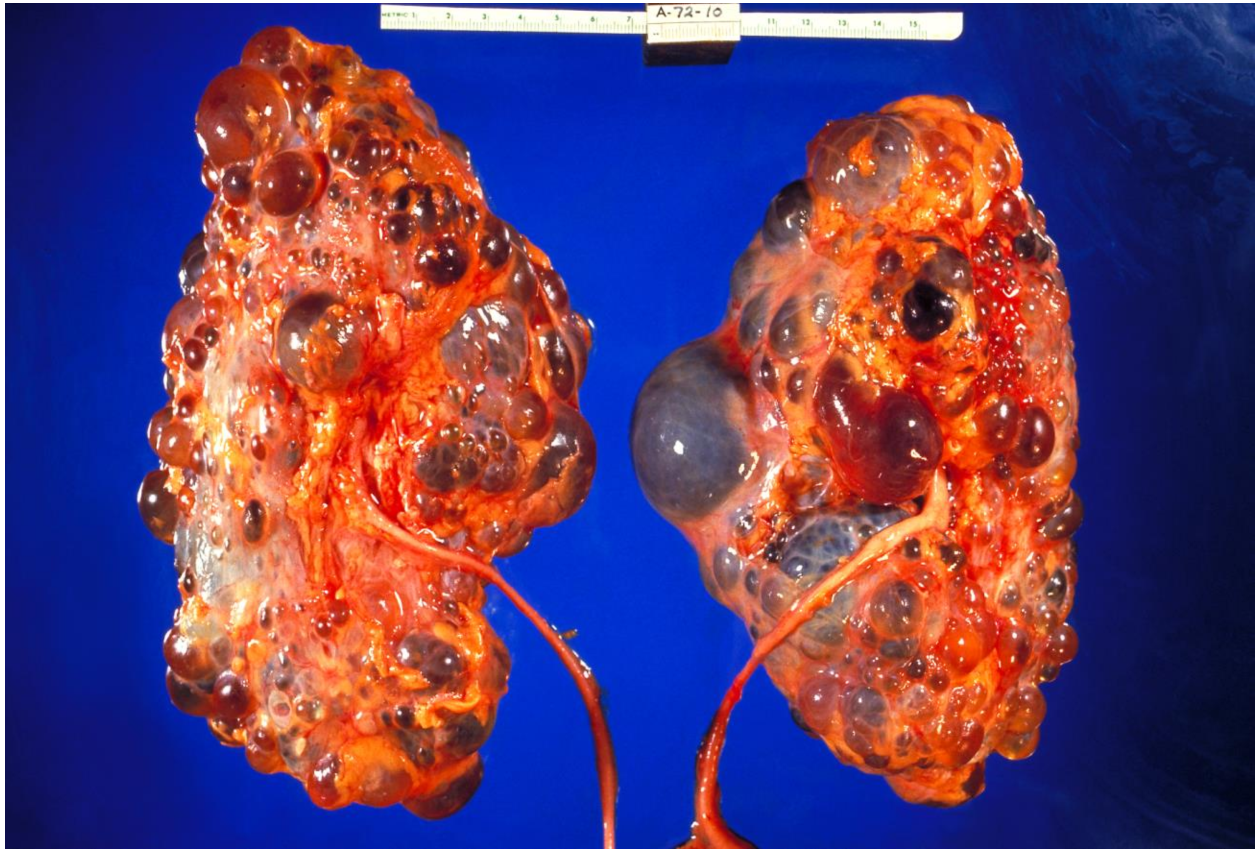
Detecting pathogenic variants in chronic kidney disease using long-read sequencing

Victor Grentzinger, PhD student

BVN-SBN Annual Meeting 2023 @Living Tomorrow Vilvoorde

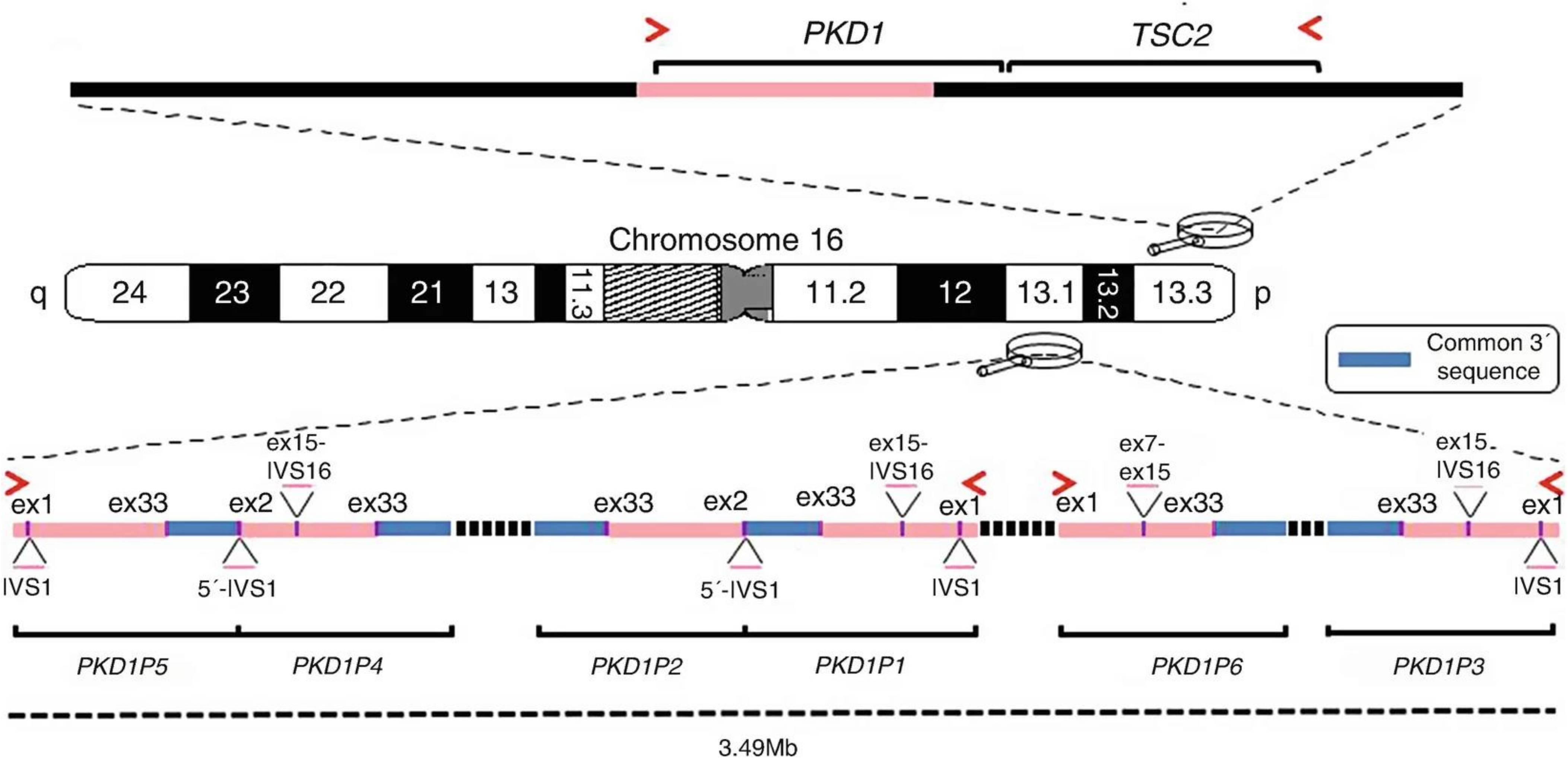
Identifying PKD1/PKD2 variants

- ADPKD: most commonly inherited renal disease
- Incidence: 1/1000
- Caused by mutations in PKD1 (70-75%) and PKD2 (25-30%)



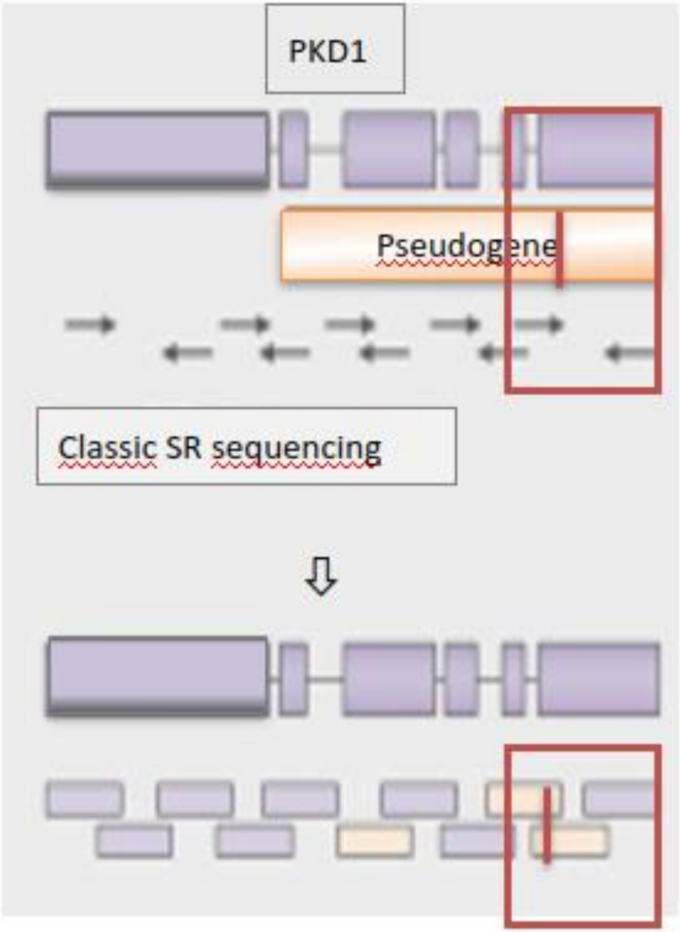
Lanktree, M.B., Haghighi, A., di Bari, I., et al. (2021). Insights into Autosomal Dominant Polycystic Kidney Disease from Genetic Studies. CJASN 16, 790–799.

PKD1 pseudogenes



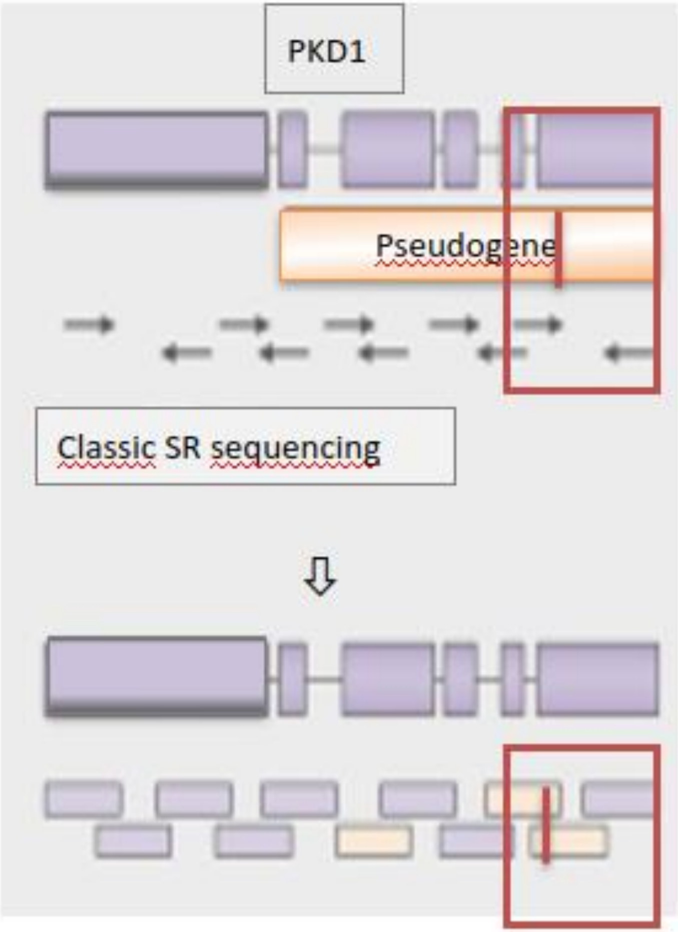
Cowley, B.D., and Bissler, J.J. eds. (2018). Polycystic Kidney Disease (Springer New York)

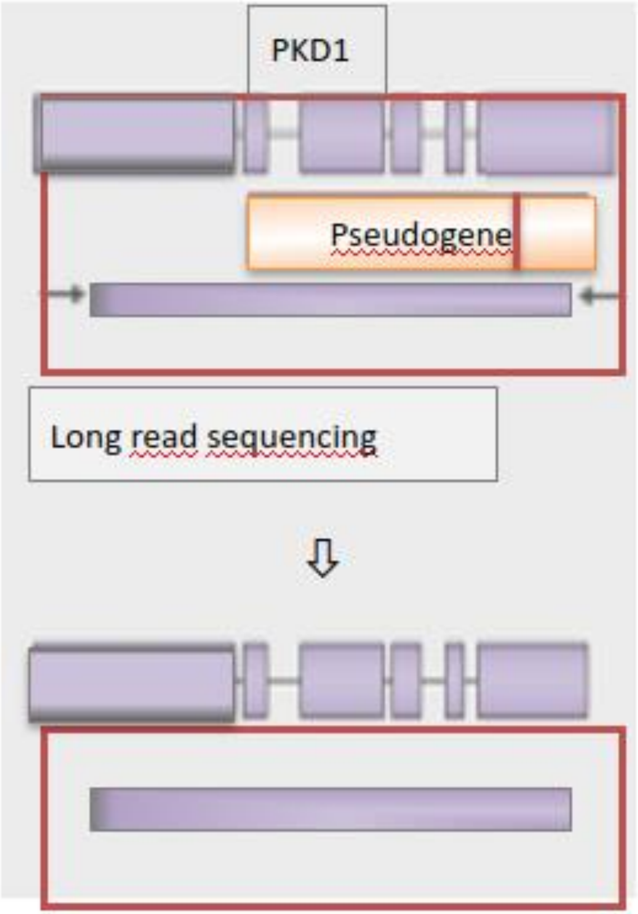
Short-read sequencing limitation



Wang, Y., Zhao, Y., Bollas, A., et al. (2021). Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol 39, 1348–1365.

Long read sequencing

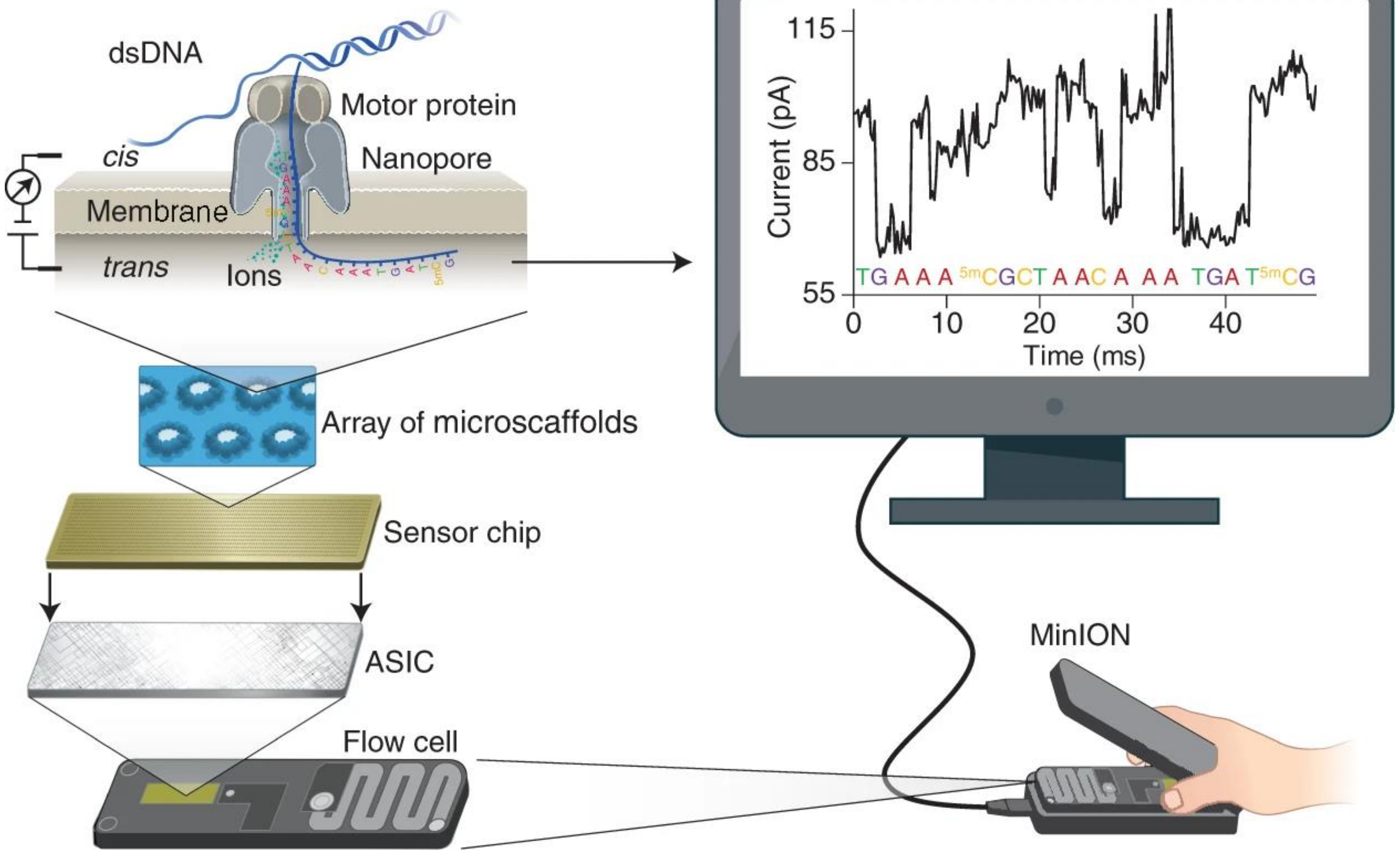




Wang, Y., Zhao, Y., Bollas, A., et al. (2021). Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol 39, 1348–1365.

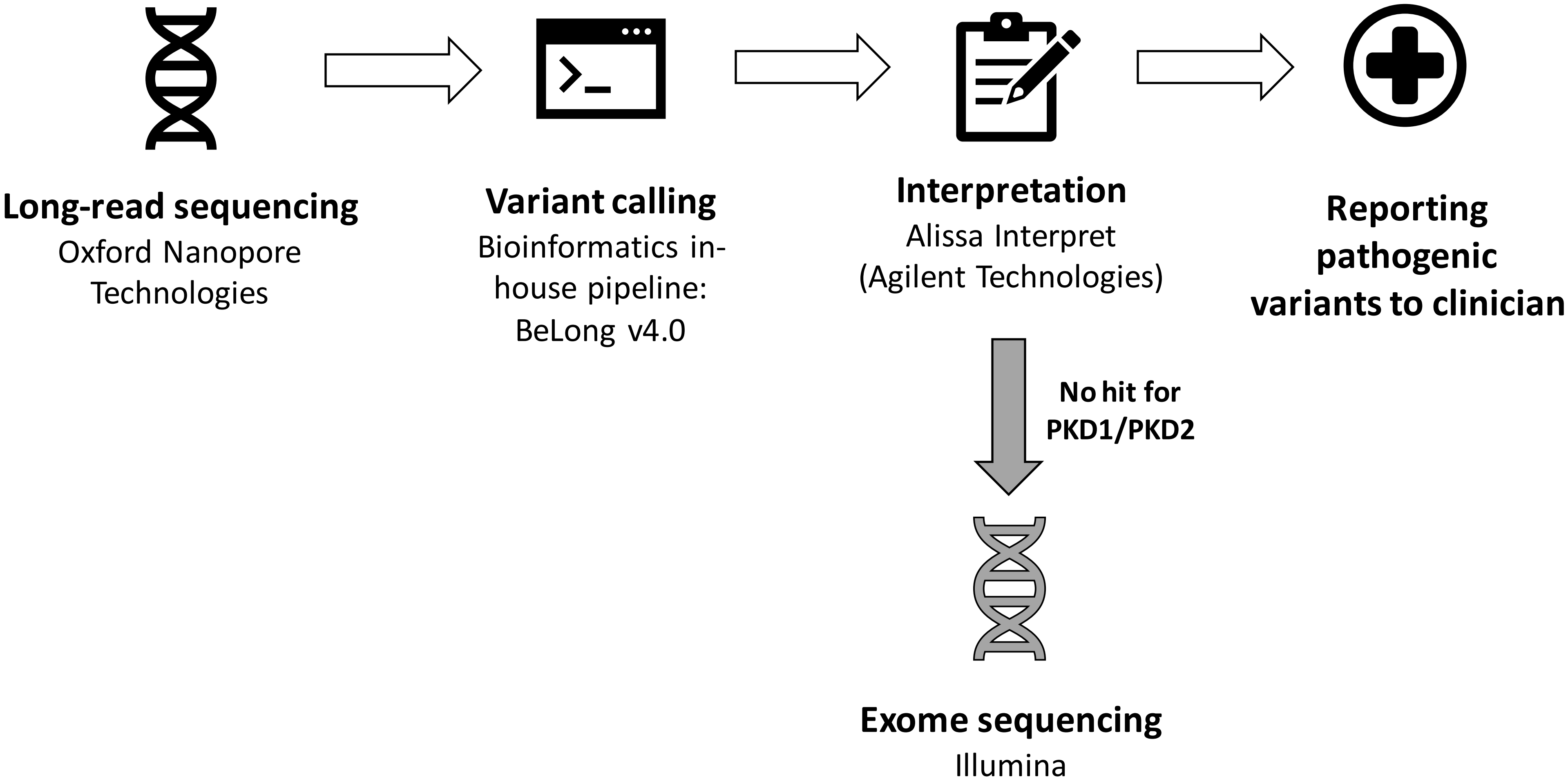
Nanopore sequencing

- DNA strand passes through a nanopore
- Record electrical current changes
- Read length up to 4Mb
- Real-time analysis and monitoring



Wang, Y., Zhao, Y., Bollas, A., et al. (2021). Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol 39, 1348–1365.

ADPKD analysis workflow (in diagnostics)



Long-read sequencing
Oxford Nanopore Technologies

Variant calling
Bioinformatics in-house pipeline: BeLong v4.0

Interpretation
Alissa Interpret (Agilent Technologies)

Reporting pathogenic variants to clinician

No hit for PKD1/PKD2

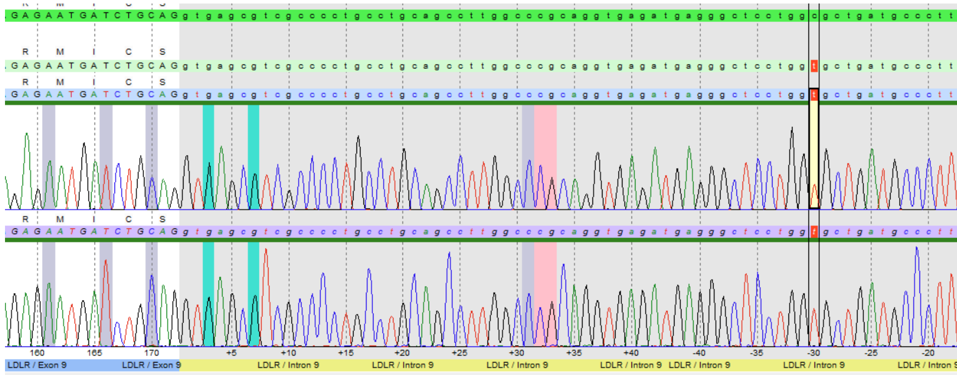
Exome sequencing
Illumina

Long read validation strategy

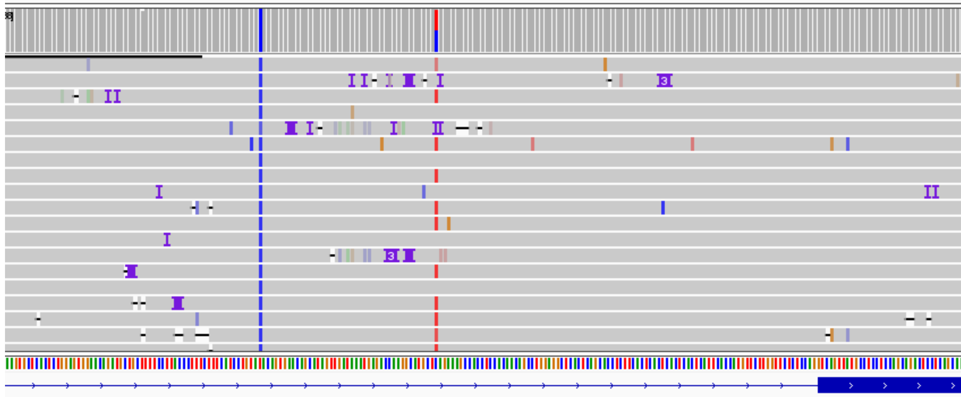
6 SHOX amplicons ; 3 patients
28 CFTR amplicons ; 6 patients
26 ABCB4 amplicons ; 5 patients
1 SLC6A8 LR-amplicon ; 4 patients
13 CASR amplicons ; 3 patients
10 CPA1 amplicons ; 2 patients
8 CTRC amplicons ; 2 patients
4 PRSS1 amplicons ; 1 patient
4 SPINK1 amplicons ; 1 patient
1 LDLR amplicon ; 1 patient
4 different ABCB11 amplicons ; 4 patients
4 different ABCB4 amplicons ; 4 patients

109 Amplicons
12 genes
36 patients

Individual amplicons Sanger sequenced (« Gold Standard »)



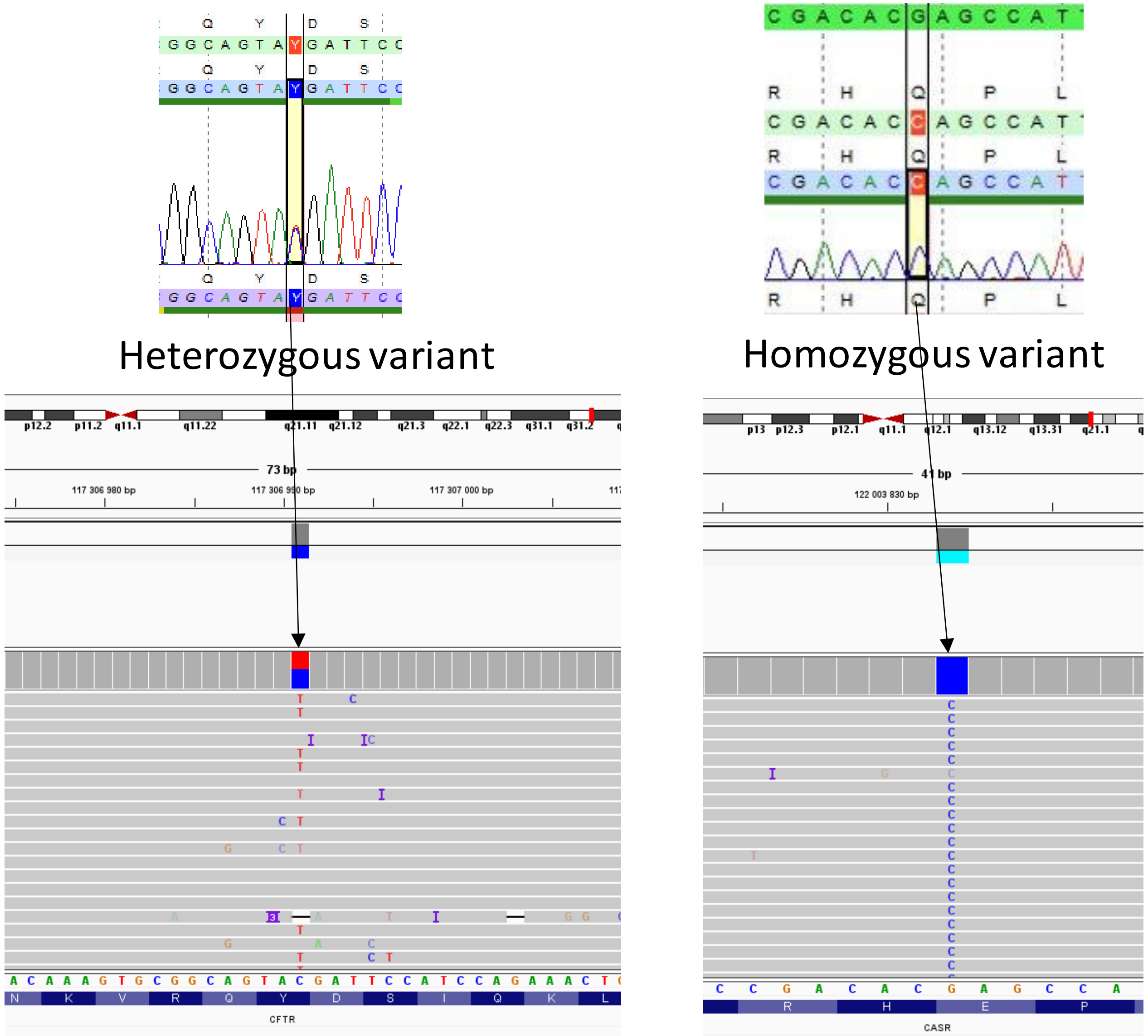
Long read sequencing (Oxford Nanopore Technologies) per patient pooled amplicons



Analytical performance

Sanger Sequencing
« Gold Standard »

Long read Data



Heterozygous variant

Homozygous variant

Sensitivity: 100% (n=83)

Precision: 96,57% (n=166)

PKD1/PKD2 validation

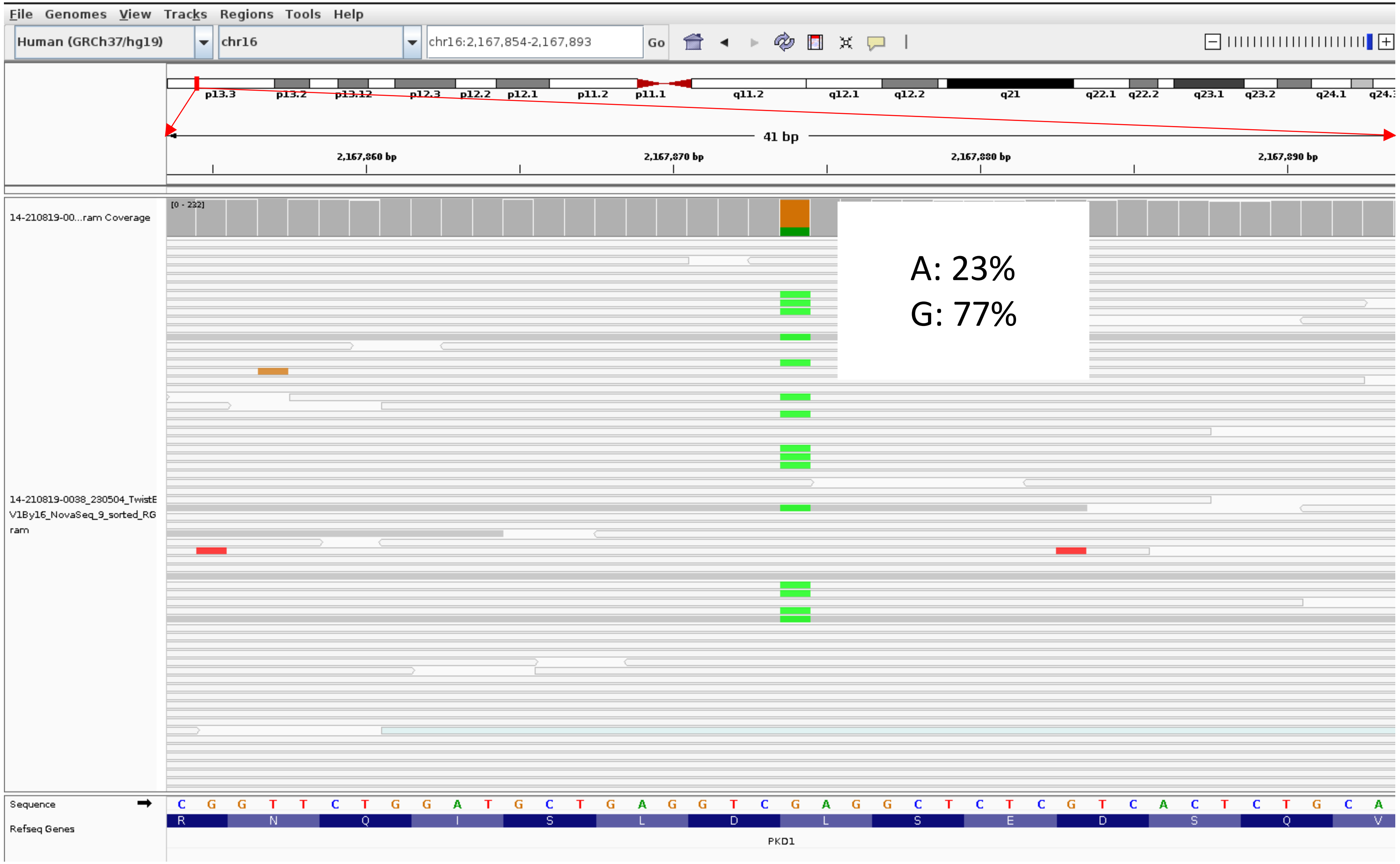
- 34 patients with 56 pathogenic variants confirmed by Sanger

- Analyzed both with:
  - long read: amplicon sequencing with Oxford Nanopore Technologies
  - short read: whole exome sequencing (WES) with Illumina

- All pathogenic variants are found in both long read and short read data

Short read sequencing in multimapped regions

WES
Short read

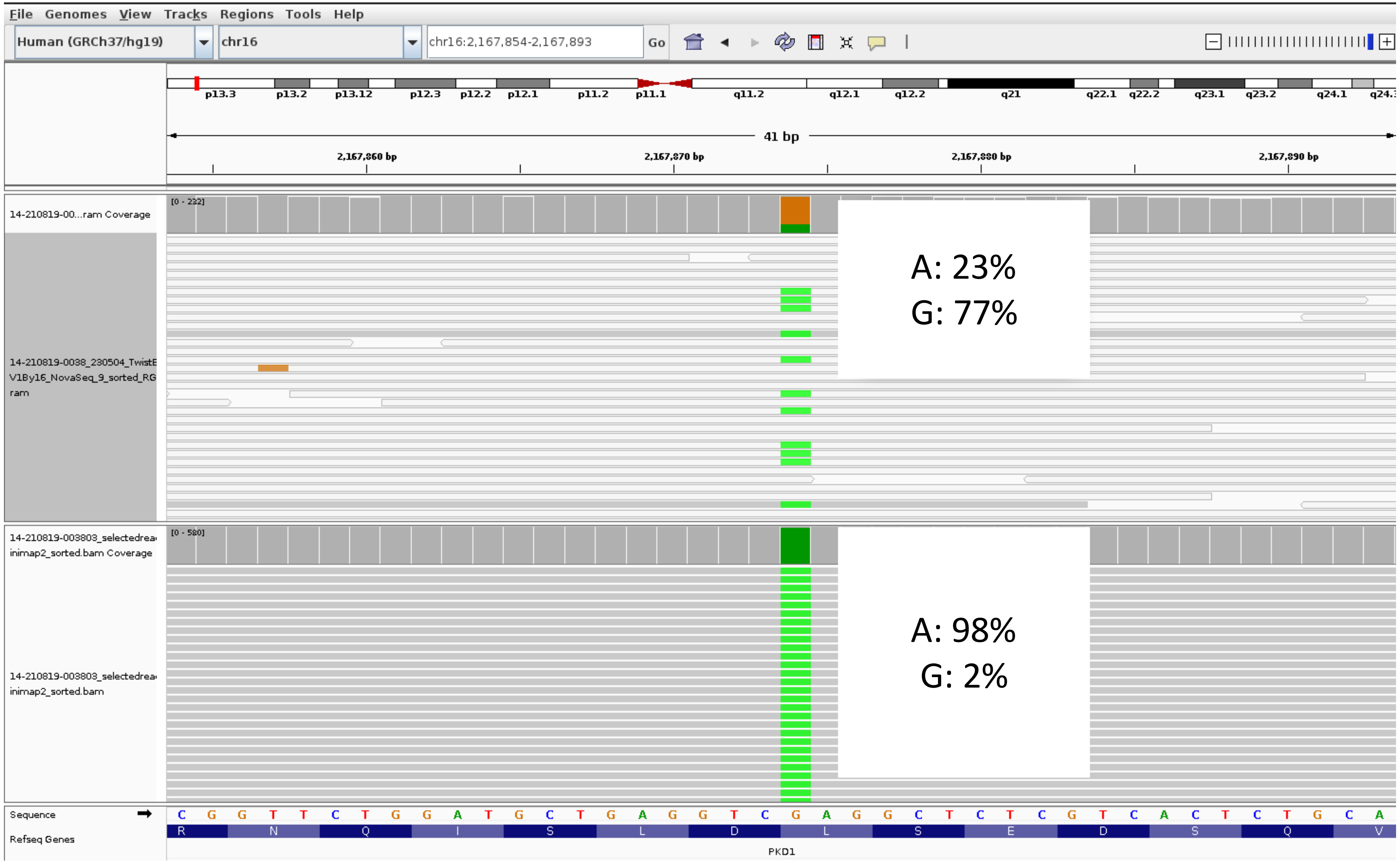


A: 23%
G: 77%

Long read solution to the multimapping limitation

WES
Short read

Amplicon
Long read

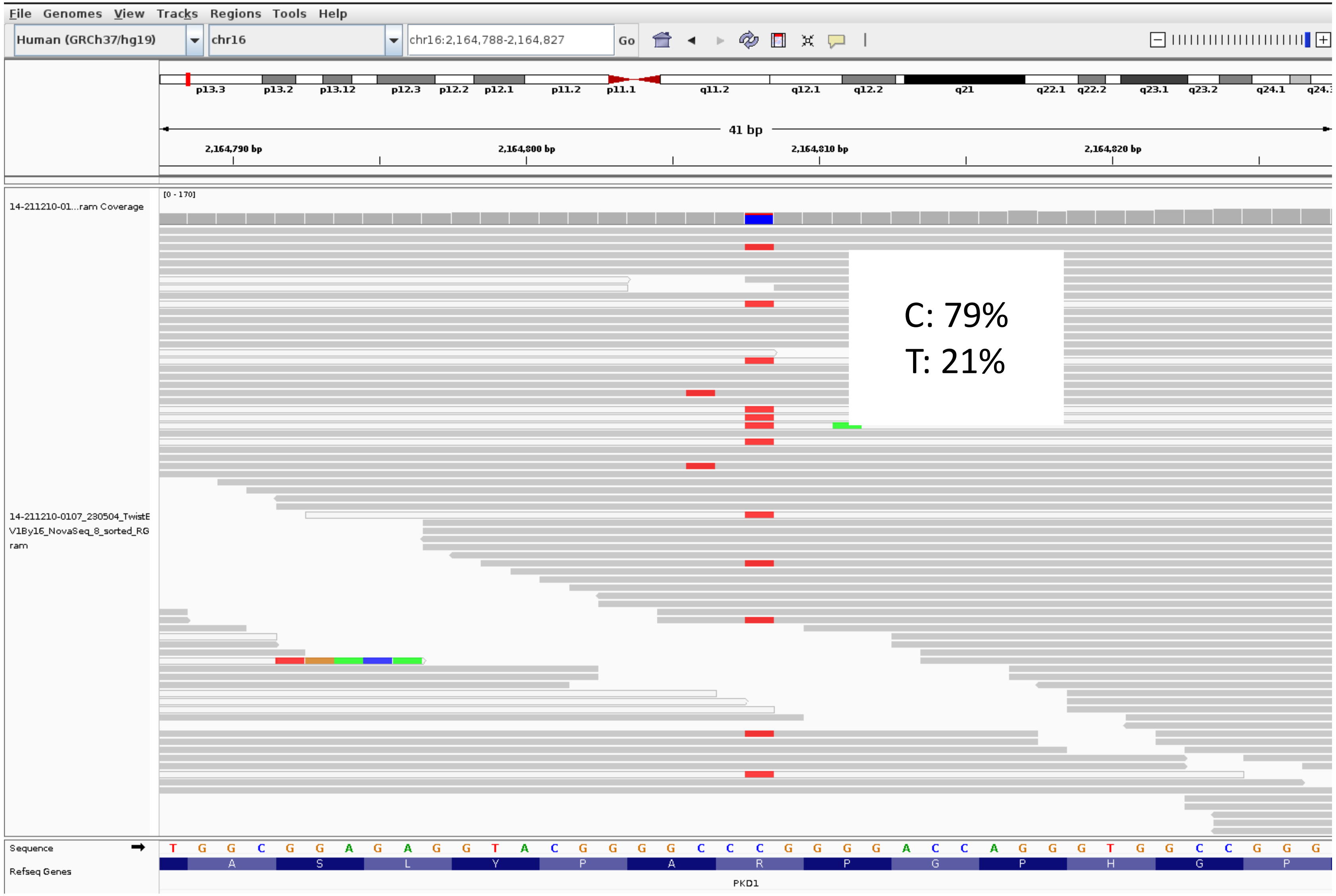


A: 23%
G: 77%

A: 98%
G: 2%

Variants with unbalanced allelic frequency

WES
Short read

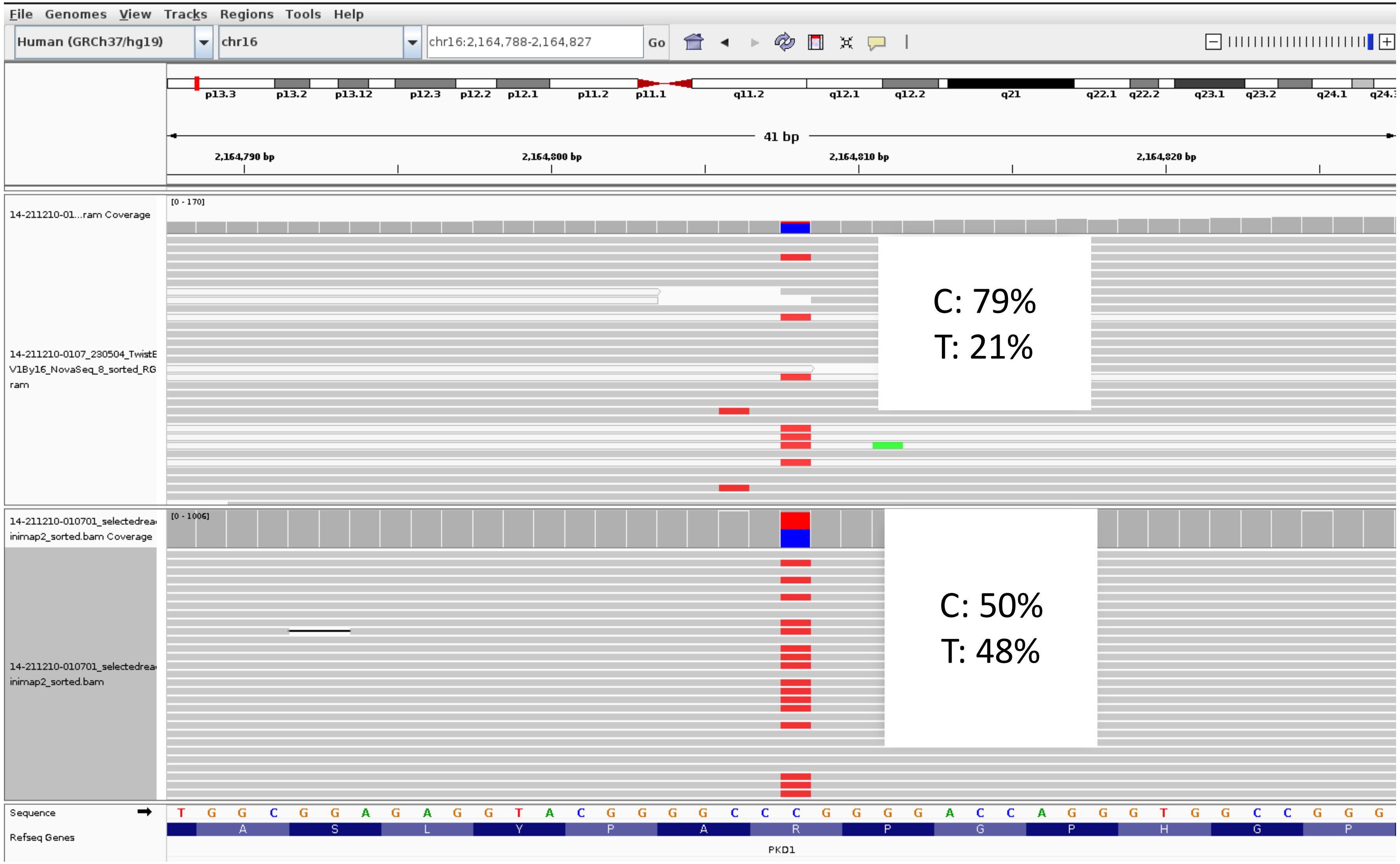


C: 79%
T: 21%

Better allelic frequencies using long read

WES
Short read

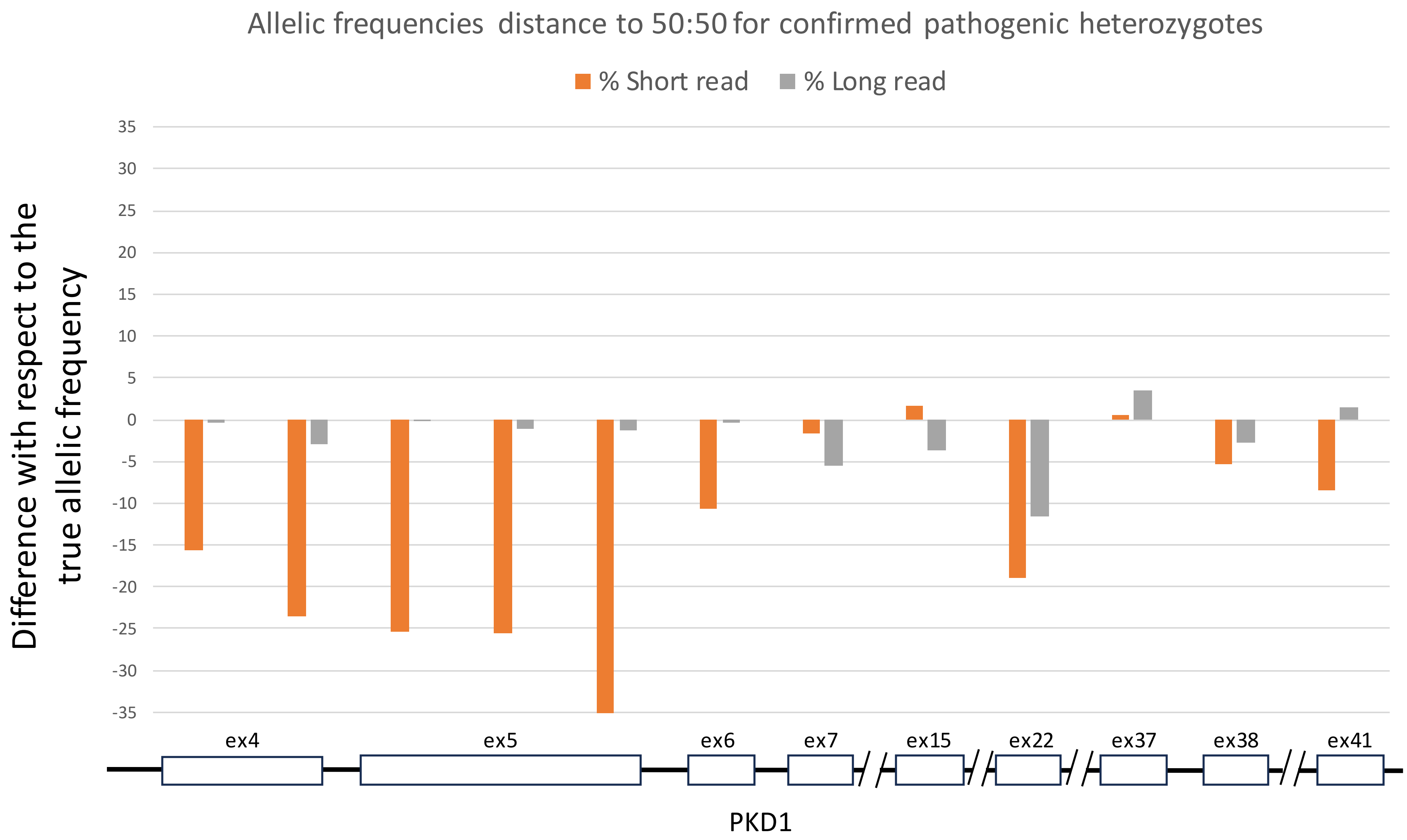
Amplicon
Long read



C: 79%
T: 21%

C: 50%
T: 48%

Allelic frequencies of pathogenic heterozygotes



Conclusion

- Overcoming short read sequencing limitation with long read sequencing
- Diagnosis pipeline with long read (Oxford Nanopore Technologies) shows reliable results
- Fast and cost-efficient alternative to Exome sequencing

Acknowledgement

CHU de Liège

LIÈGE université

Bernard Grisart
Maxime Meynsbrughen
Maria Artesi
Leonor Palmeira
Vinciane Dideberg

Benoit Charloteaux
Keith Durkin
Anne Brysse
Vincent Bours

Homopolymer stretches

- Region of repeated nucleic acid
- Can be mislabeled as deletion
- Low allelic frequency -> filtering on allelic frequency percentage