

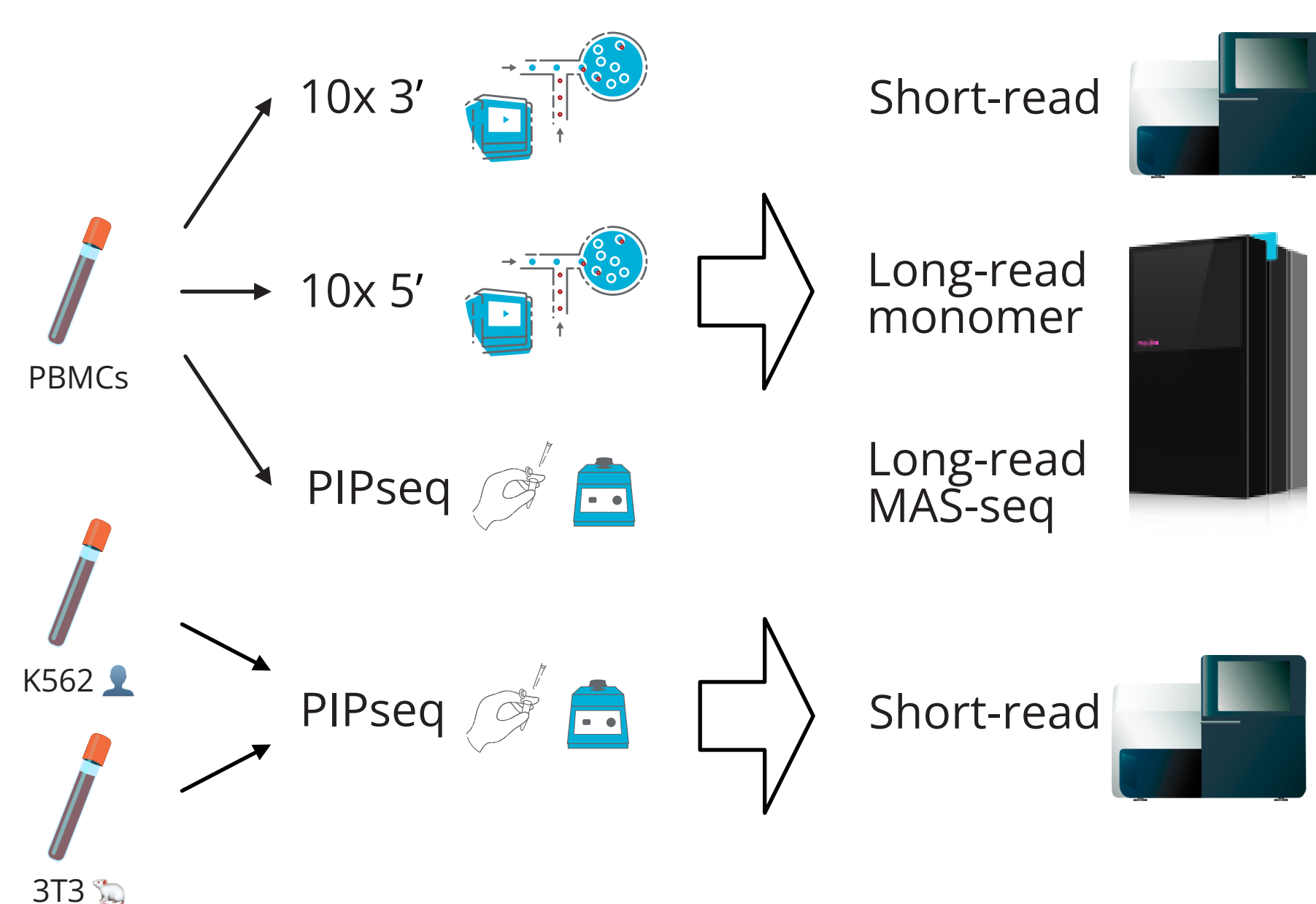
## Introduction

Long-read transcriptomics has moved into the mainstream, driven by transformational advances in long-read sequencing platforms, library construction methods, and computational tools. Increases in sequencing throughput have brought into focus the many types of artifacts introduced during single-cell cDNA synthesis. These artifacts complicate isoform identification and quantification, and can interfere with the detection of gene fusions. To determine the impact of single-cell platform on artifact formation we evaluated PIPseq V4 (Fluent Biosciences), 10x 3' V3.1, and 10x 5' V2.1 (10x Genomics) for their ability to make high-quality full-length cDNA libraries for isoform sequencing.

We prepared single-cell cDNA libraries from PBMCs with each method and found that 10x 3' and 5' short-read libraries yielded more UMIs per cell compared to PIPseq but with a higher cost per cell overall. We then prepared Iso-seq and MAS-seq libraries from the full-length cDNA and sequenced them on the PacBio Sequel IIe and Revio sequencers.

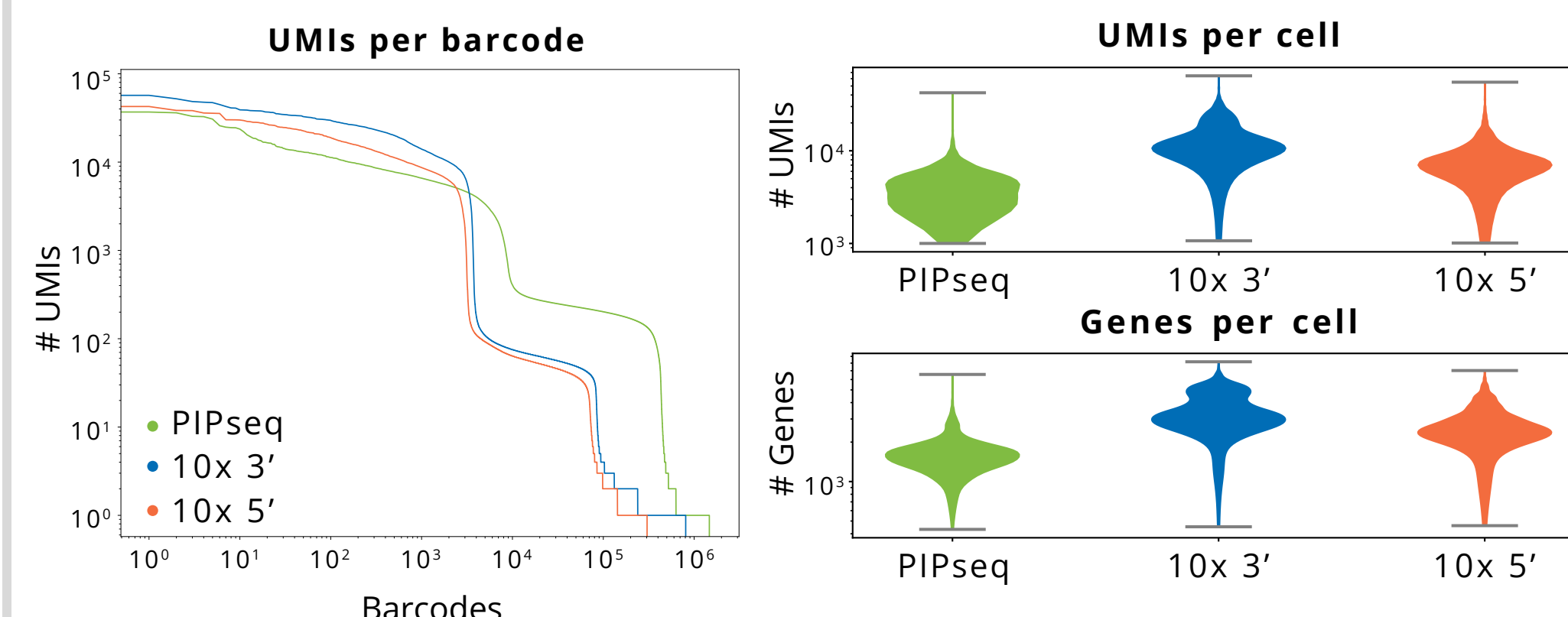
PIPseq produced fewer artifacts during cDNA creation. In addition, reads generated by PIPseq were more likely to end with a polyA motif and less likely to flank genomic polyA sites that can lead to internal priming artifacts. We also observed significant mRNA degradation in some cell types, suggesting that RNase activity is higher during the PIPseq protocol. These results indicate that PIPseq can be an effective platform for isoform sequencing from single cells, but further methods development is needed.

## Experimental Design

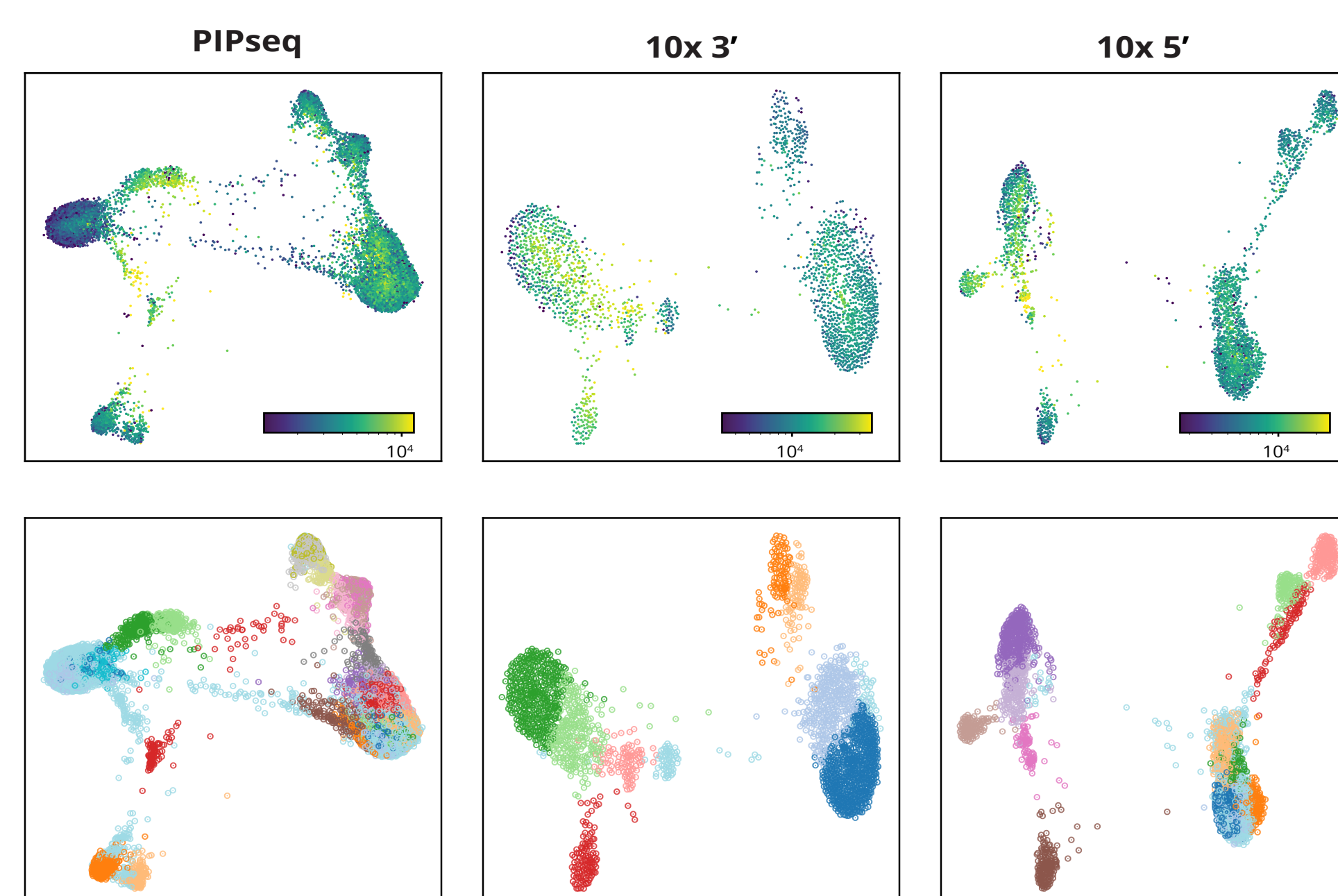


- For artifact and isoform analysis, PBMCs were prepared using the 10x 3', 10x 5', and PIPseq protocols
- PBMC cDNA was used as input for short-read sequencing on the Illumina NovaSeq X, as well as monomer (Isoseq) and MAS-seq sequencing on the PacBio Sequel IIe and Revio instruments
- For doublet analysis, a mix of K562 and 3T3 was prepared with PIPseq, sequenced with short-reads, and compared to publicly available 10x data

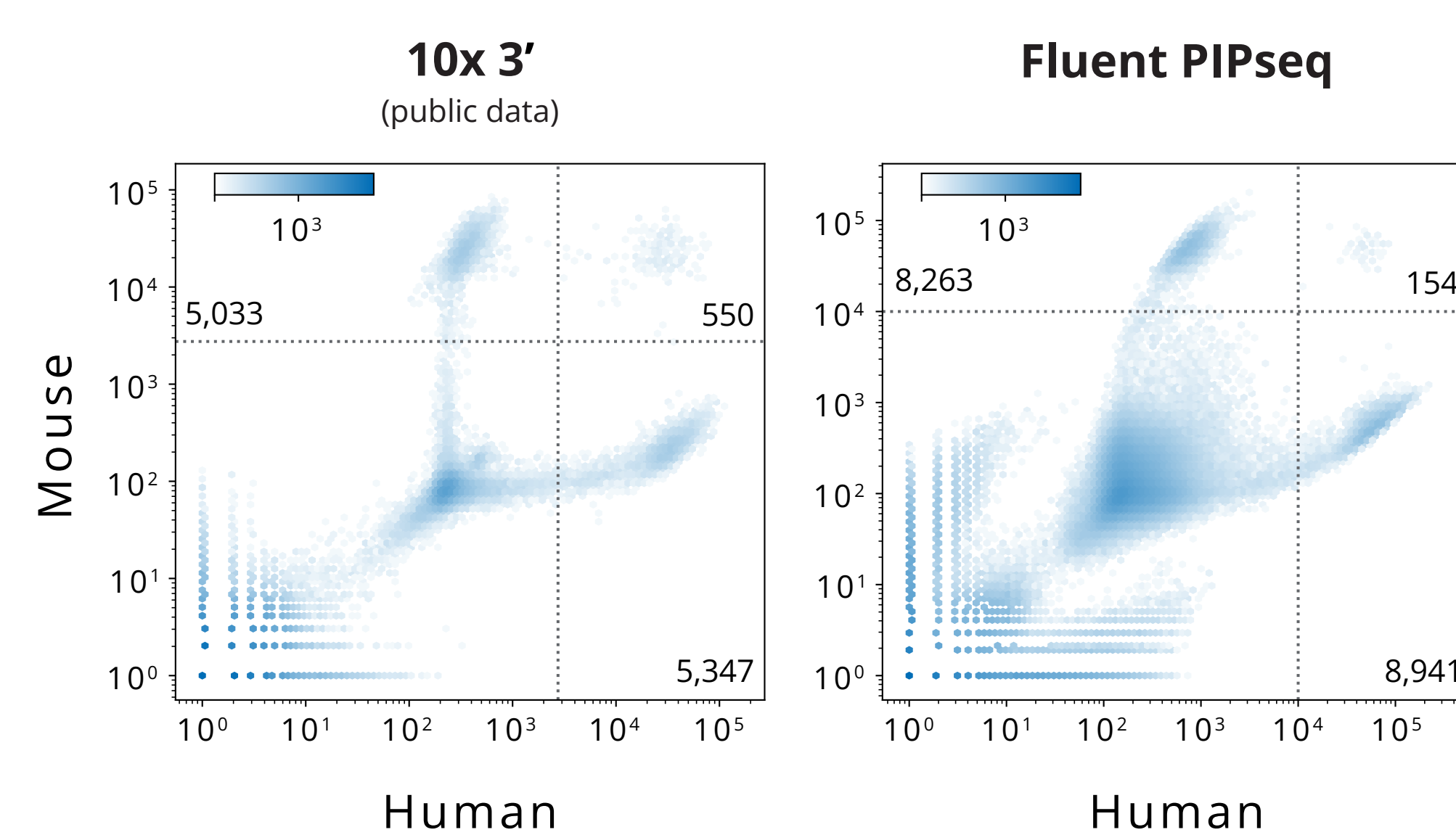
## Short-Read Sequencing



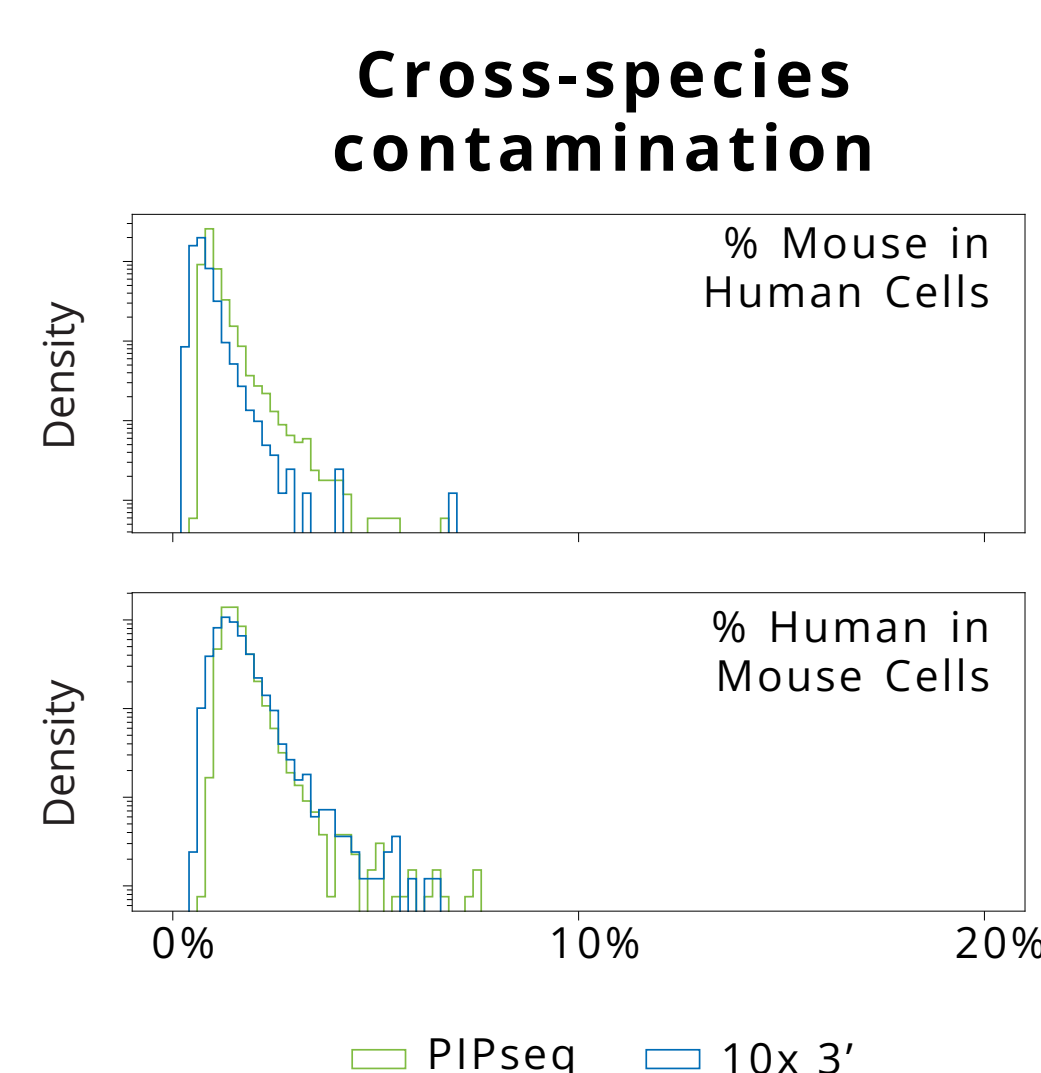
- In short-read analysis, PIPseq shows similar performance to 10x 3' and 5' libraries
- The method yielded more barcodes, but fewer UMIs and genes per barcode
- The three datasets led to similar clustering of known cell types



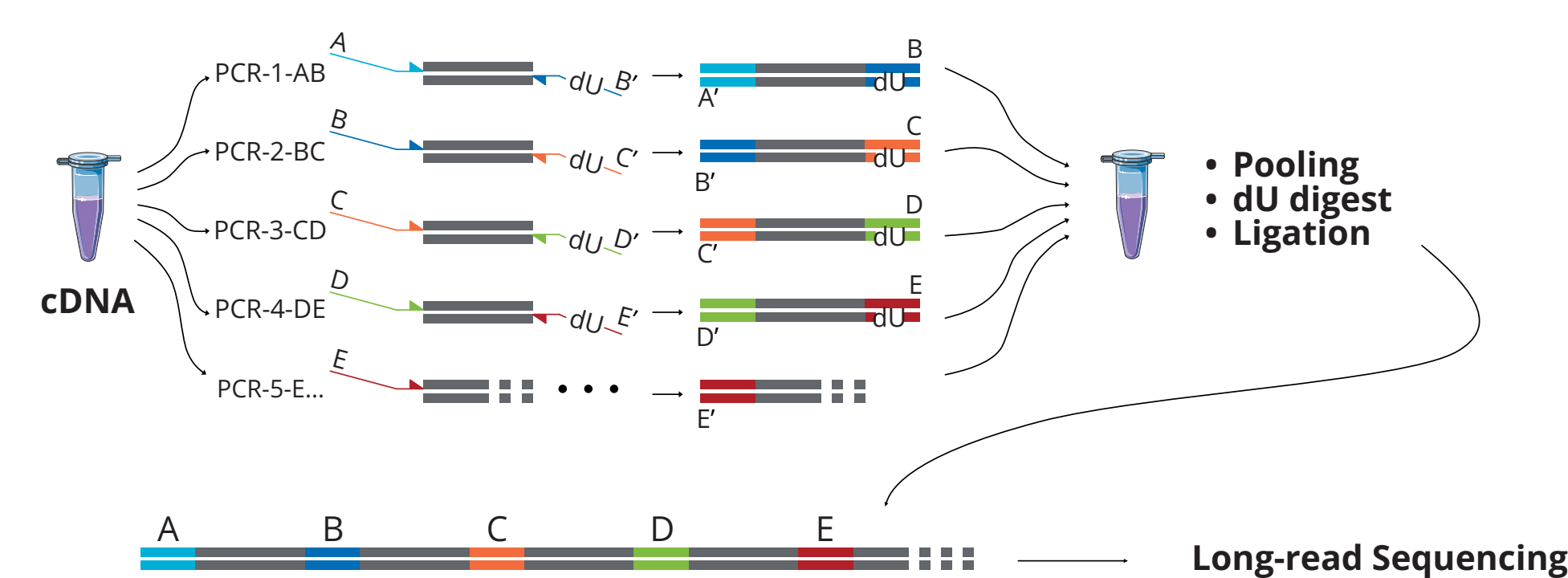
## Barnyard Experiments



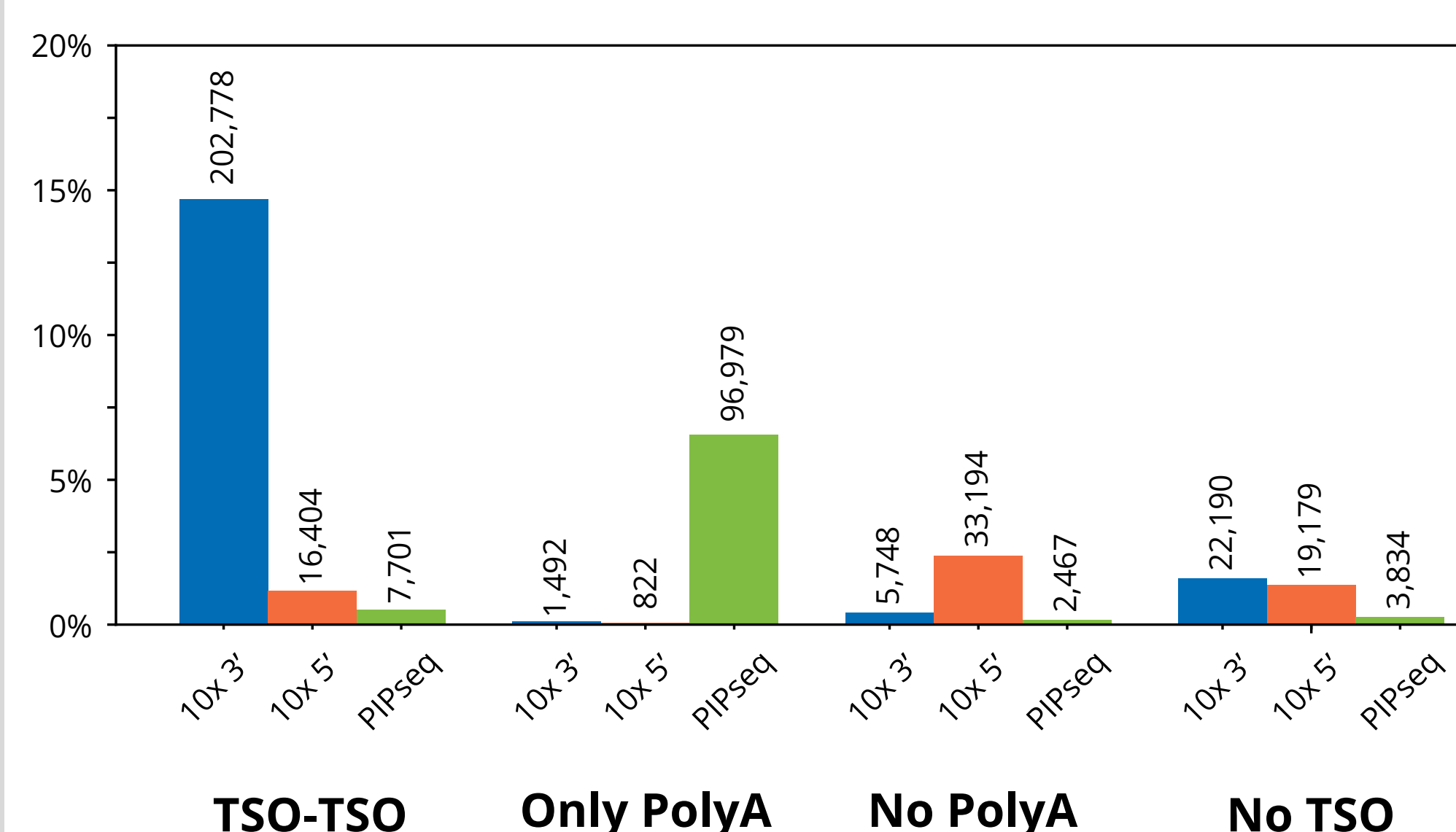
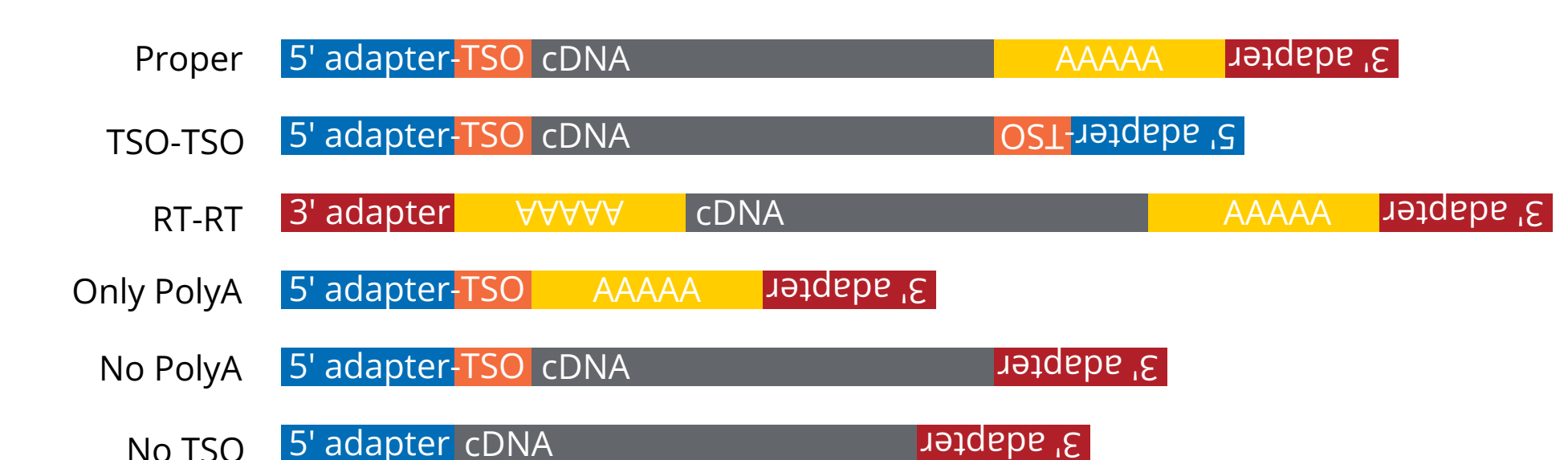
- PIPseq shows a higher rate of ambient mRNA in empty droplets, but a much lower rate of doublets (<1%).
- Using cutoffs of >10,000 UMIs from one species and <1000 from the other, we see a similar rate of cross-species contamination in both experiments.



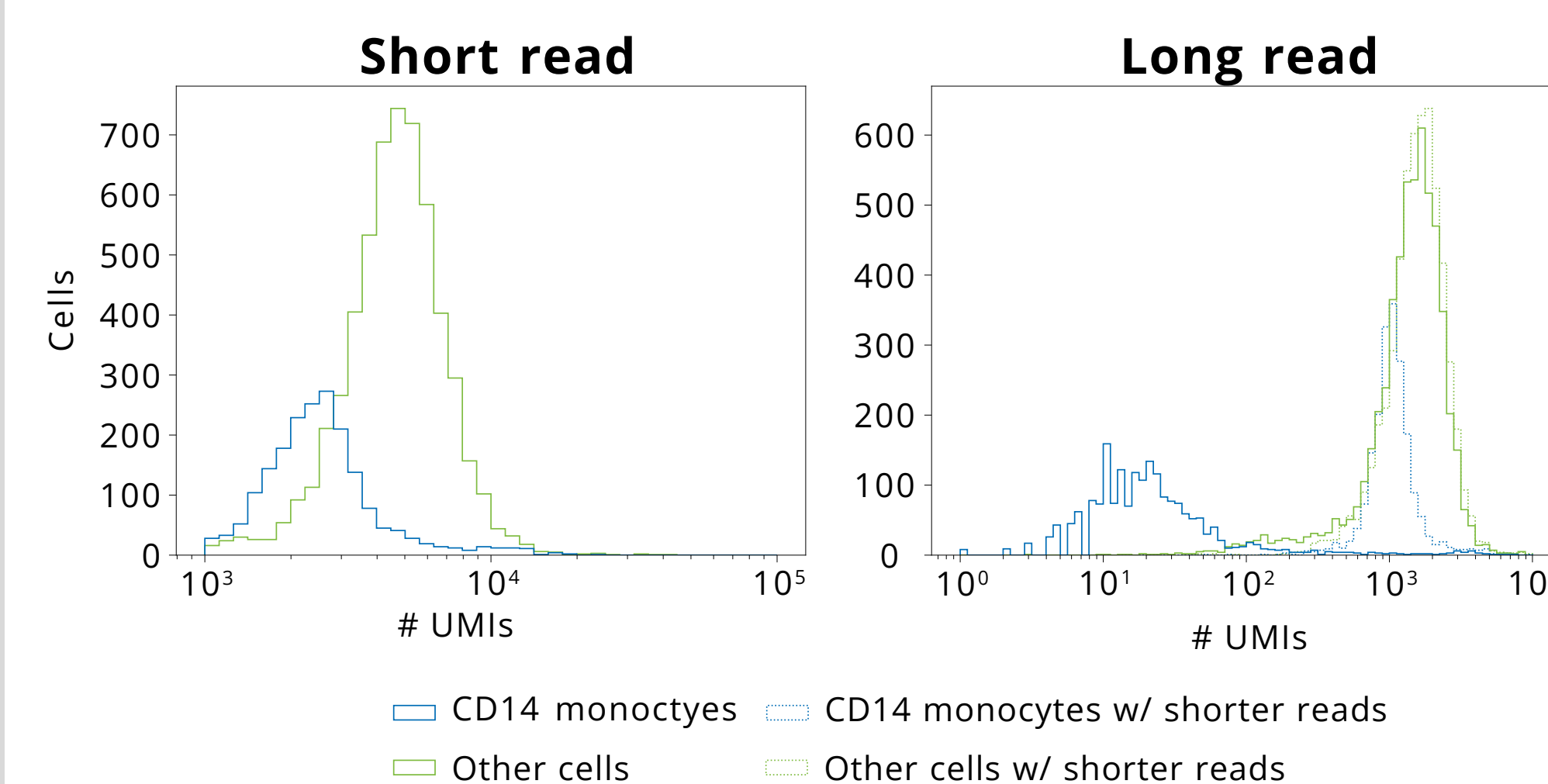
## Long-Read Sequencing



### Common Artifact Types

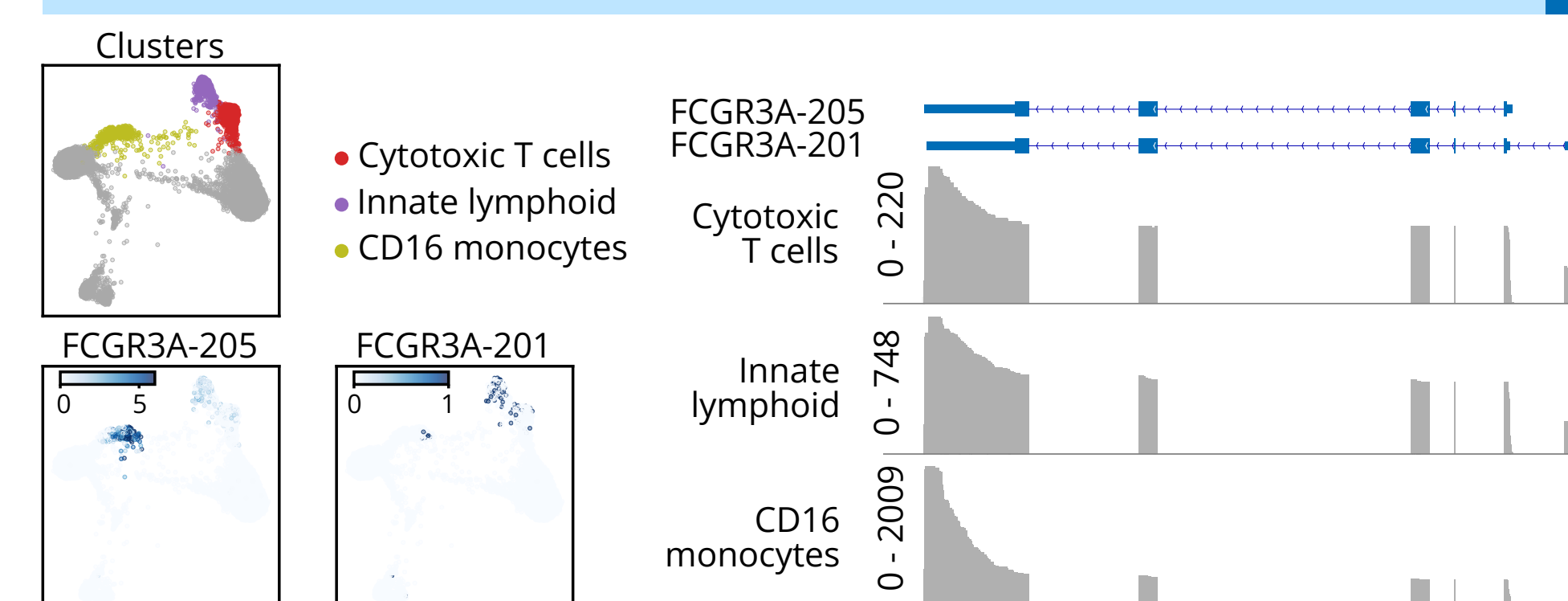


## mRNA Degradation Artifacts

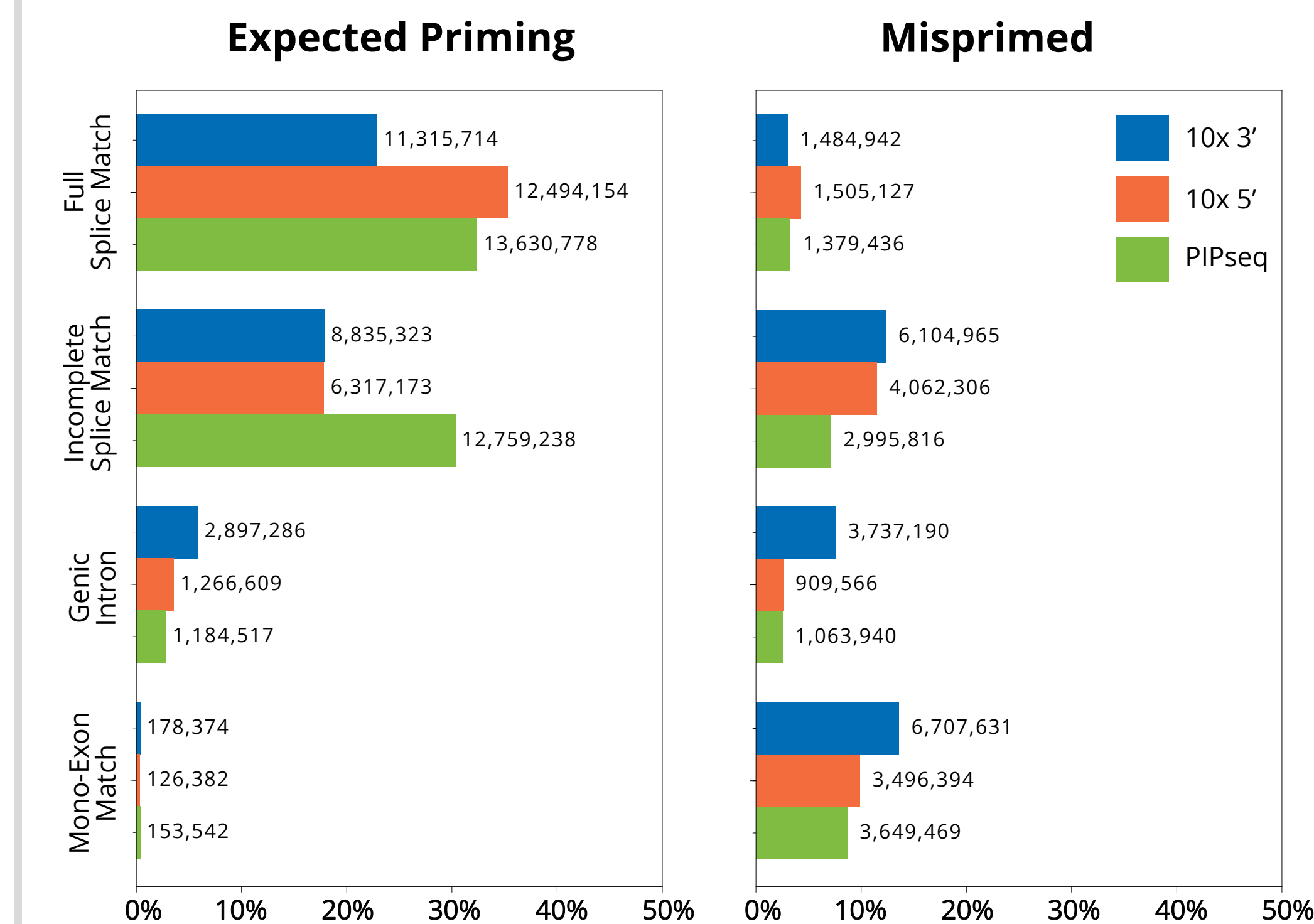


In PIPseq data, high RNase activity in CD14 monocytes leads to mRNA degradation and lower UMI counts. In the context of long reads, this can significantly impair isoform assignment and quantification.

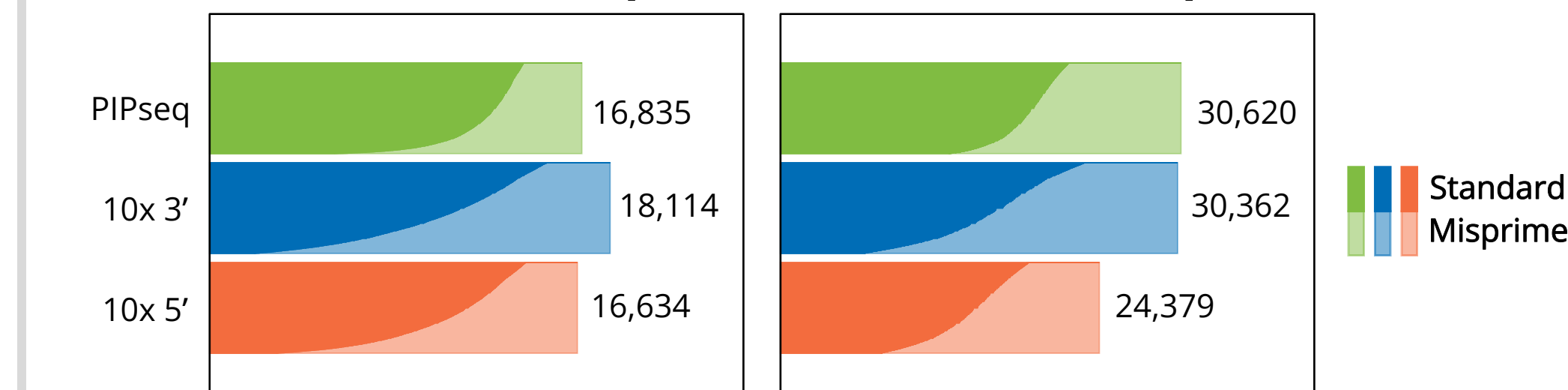
## Differential Isoform Usage



## Mis-priming artifacts



### Known Transcripts



## Conclusion

- PIPseq shows promise as a new platform for high-throughput single-cell RNAseq at low cost
- Artifacts produced during library preparation can impact MAS-seq array formation, but will also affect short-read counts
- Internal priming is observed in all three platforms, and improperly primed reads can lead to "novel" isoform calls

## Acknowledgements

- PIPseq library preparation and analysis was done with assistance from Fluent Biosciences
- PacBio provided MAS-seq kits and Revio sequencing reagents as part of a collaboration agreement
- Data used in this poster was generated at the Broad Institute, for more information visit: <https://genomics.broadinstitute.org>

## References

- Clark, I. C. et al. Microfluidics-free single-cell genomics with templated emulsification. Nat Biotechnol 1–10 (2023) doi:10.1038/s41587-023-01685-z.
- Al'Khafaji, A. M. et al. High-throughput RNA isoform sequencing using programmed cDNA concatenation. Nat Biotechnol (2023) doi:10.1038/s41587-023-01815-7.
- Fontanez, K. et al. High-throughput single cell analysis with Particle-templated Instant Partitions (PIPseq™).
- Li, H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100 (2018).
- Prjibelski, A. D. et al. Accurate isoform discovery with IsoQuant using long reads. Nat Biotechnol 1–4 (2023) doi:10.1038/s41587-022-01565-y.
- Barnyard data for 10x 3' was downloaded from [www.10xGenomics.com](http://www.10xGenomics.com)