

## Abstract

The PacBio Iso-Seq method produces high-quality, full-length transcripts and can characterize a whole transcriptome with a single SMRT Cell 8M.

We sequenced an Alzheimer whole brain sample on a single SMRT Cell 8M on the Sequel II System. Using the Iso-Seq bioinformatics pipeline followed by SQANTI2 analysis, we detected 162,290 transcripts for 17,670 genes up to 14 kb in length. More than 60% of the transcripts are novel isoforms, the vast majority of which have supporting cage peak data and polyadenylation signals, demonstrating the utility of long-read sequencing for human disease research.

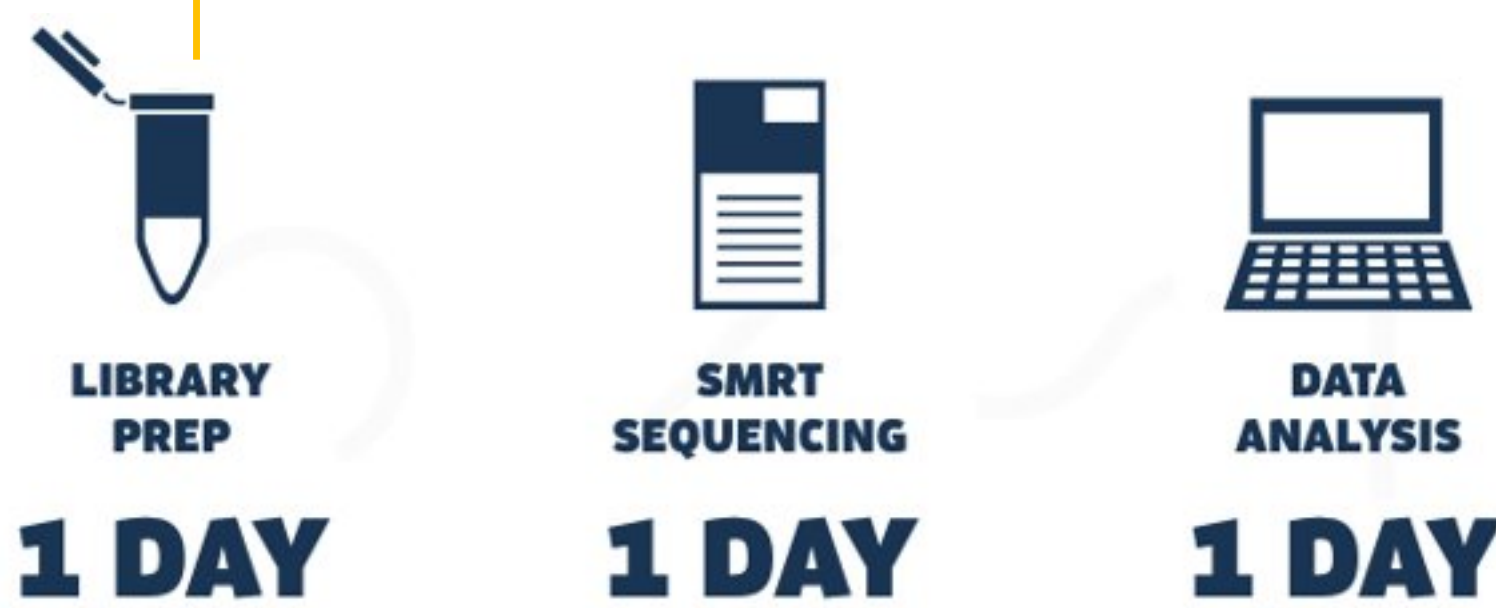
The data is publicly available at:

[https://downloads.paccloud.com/public/dataset/Alzheimer2019\\_IsoSeq/](https://downloads.paccloud.com/public/dataset/Alzheimer2019_IsoSeq/)

## Full-Length RNA-Seq on Sequel II System

### Iso-Seq Express kit<sup>[1]</sup>

- 300 ng total RNA
- Full-length cDNA
- Multiplexing support



### Sequel II System

- 1 SMRT Cell 8M for whole transcriptome
- Up to 4 million full-length reads

### Bioinformatics

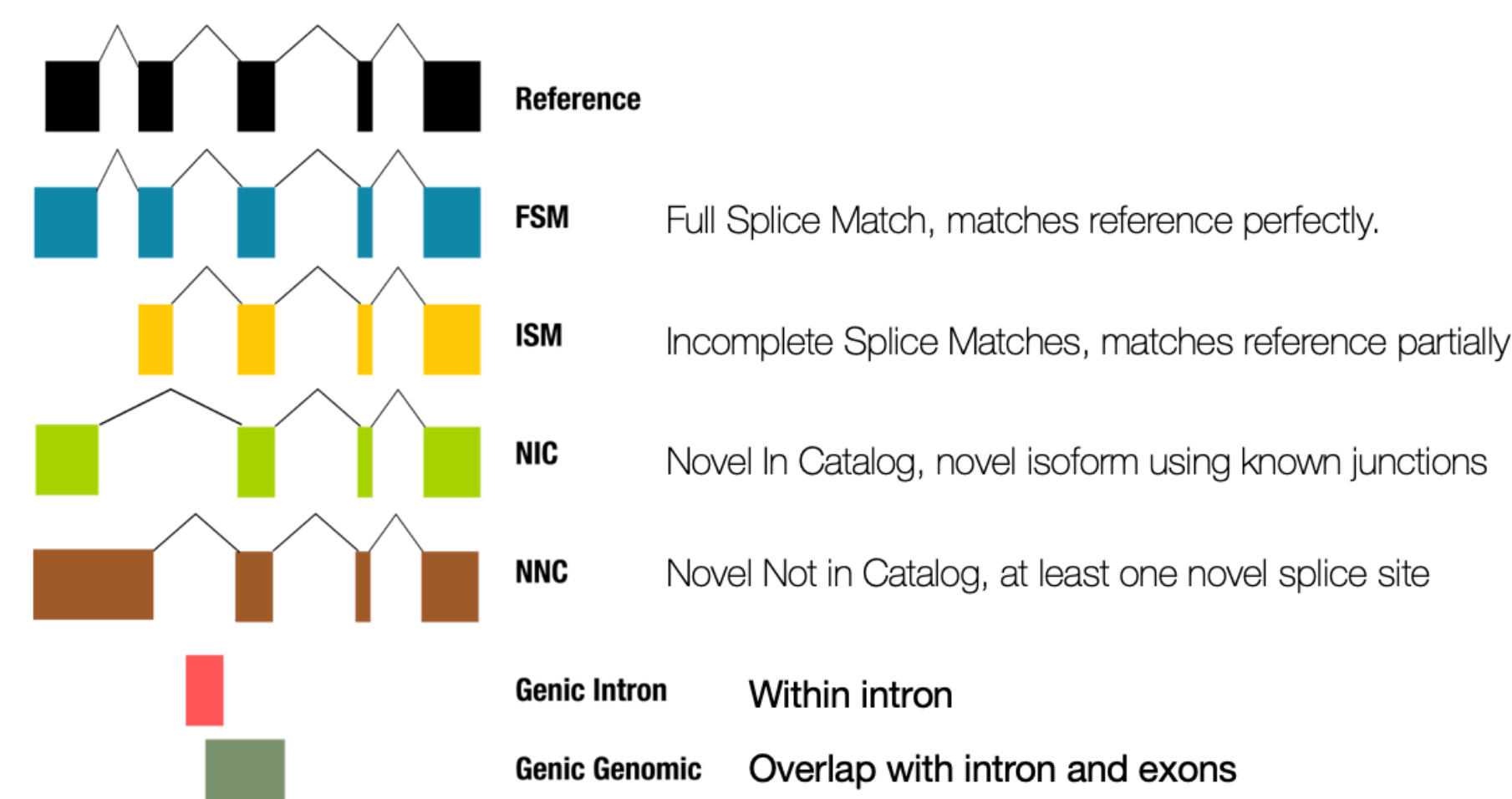
- Iso-Seq3 in SMRT Analysis
- Reads to ORF in 1 day
- Downstream community tools<sup>[2]</sup>

## Full-Length Isoforms using Iso-Seq

|                 | Known  | Novel   | Total          |
|-----------------|--------|---------|----------------|
| <b>Genes</b>    | 17,051 | 619     | <b>17,670</b>  |
| <b>Isoforms</b> | 51,660 | 110,630 | <b>162,290</b> |

**Table 1. Number of genes and isoforms detected in the Alzheimer brain sample after Iso-Seq analysis and SQANTI2 filtering.** All post-filtered isoforms have all junctions supported by public (Intropolis) RNA-seq data.

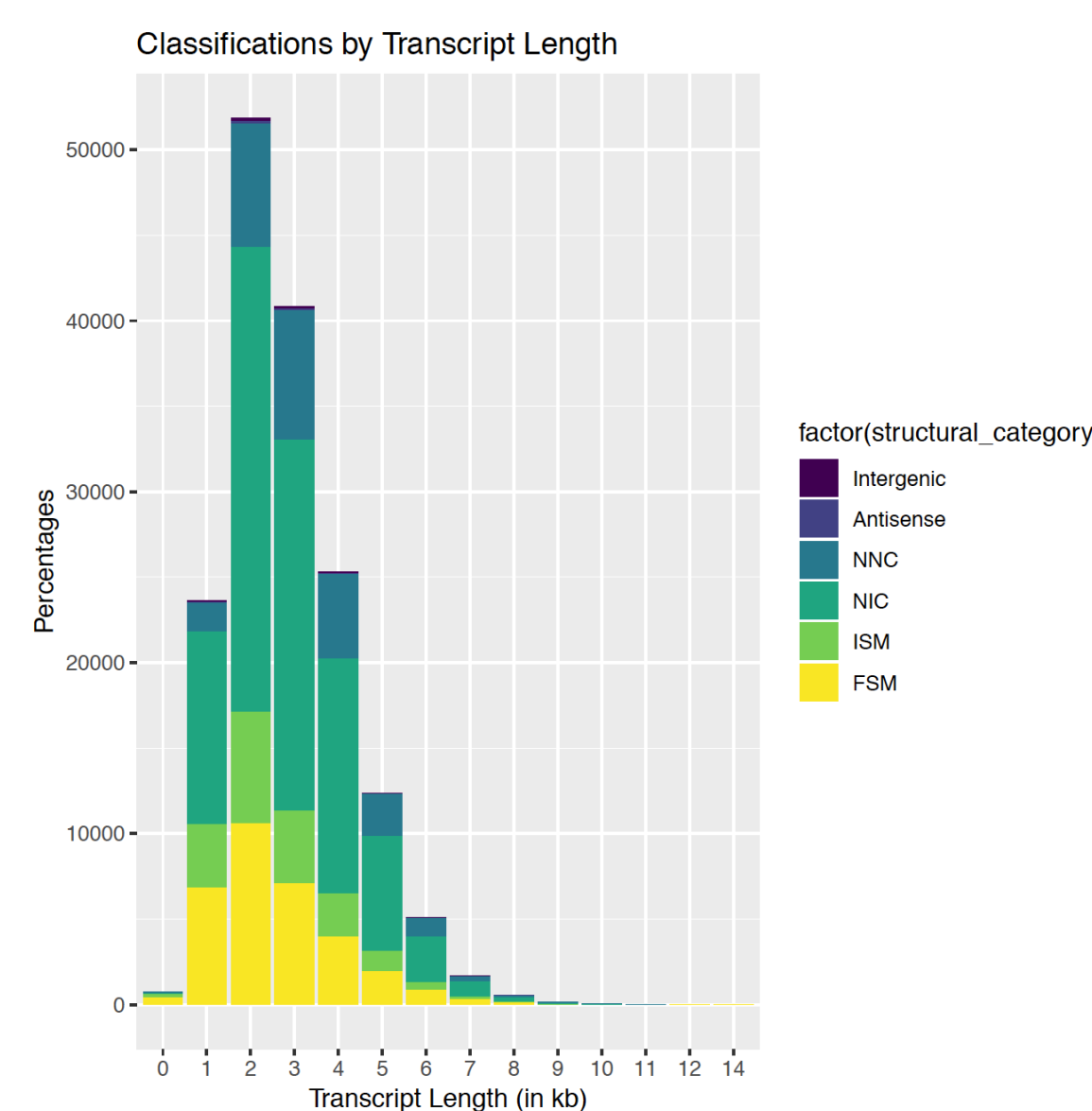
### How SQANTI Classifies Transcripts



### Comprehensive Isoform Detection

162,290 transcripts

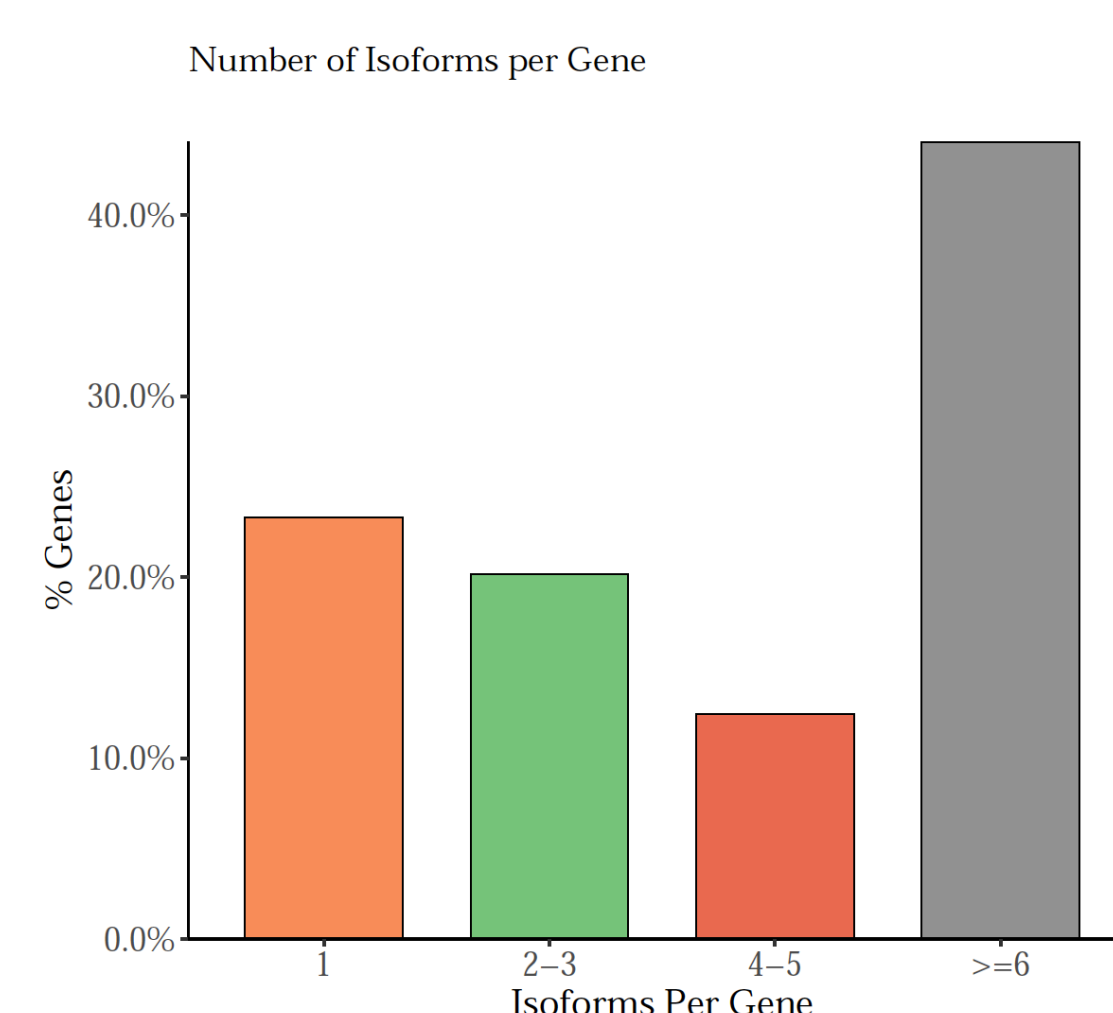
Min: 80 bp  
Max: 14,288 bp  
Mean: 3,347 bp



### Most Isoforms Near CAGE peak + polyA Signal

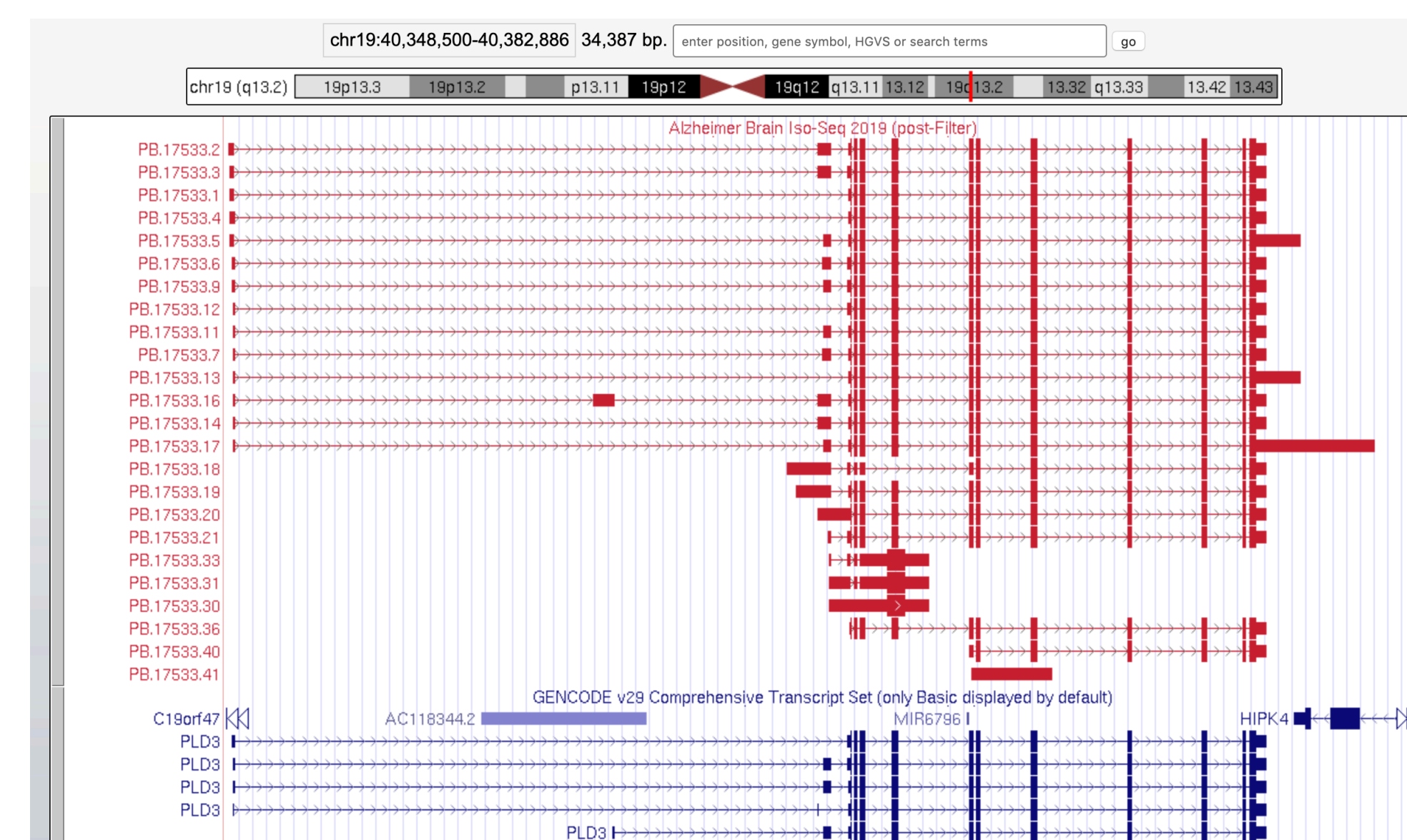
| Category          | Count  | CAGE peak within 50 bp | polyA Motif Detected |
|-------------------|--------|------------------------|----------------------|
| <b>FSM</b>        | 32,649 | 70%                    | 72%                  |
| <b>ISM</b>        | 19,011 | 37%                    | 62%                  |
| <b>NIC</b>        | 84,610 | 36%                    | 55%                  |
| <b>NNC</b>        | 25,323 | 57%                    | 72%                  |
| <b>Antisense</b>  | 321    | 24%                    | 43%                  |
| <b>Intergenic</b> | 376    | 24%                    | 38%                  |

### Alzheimer Brain Shows High Splicing Complexity

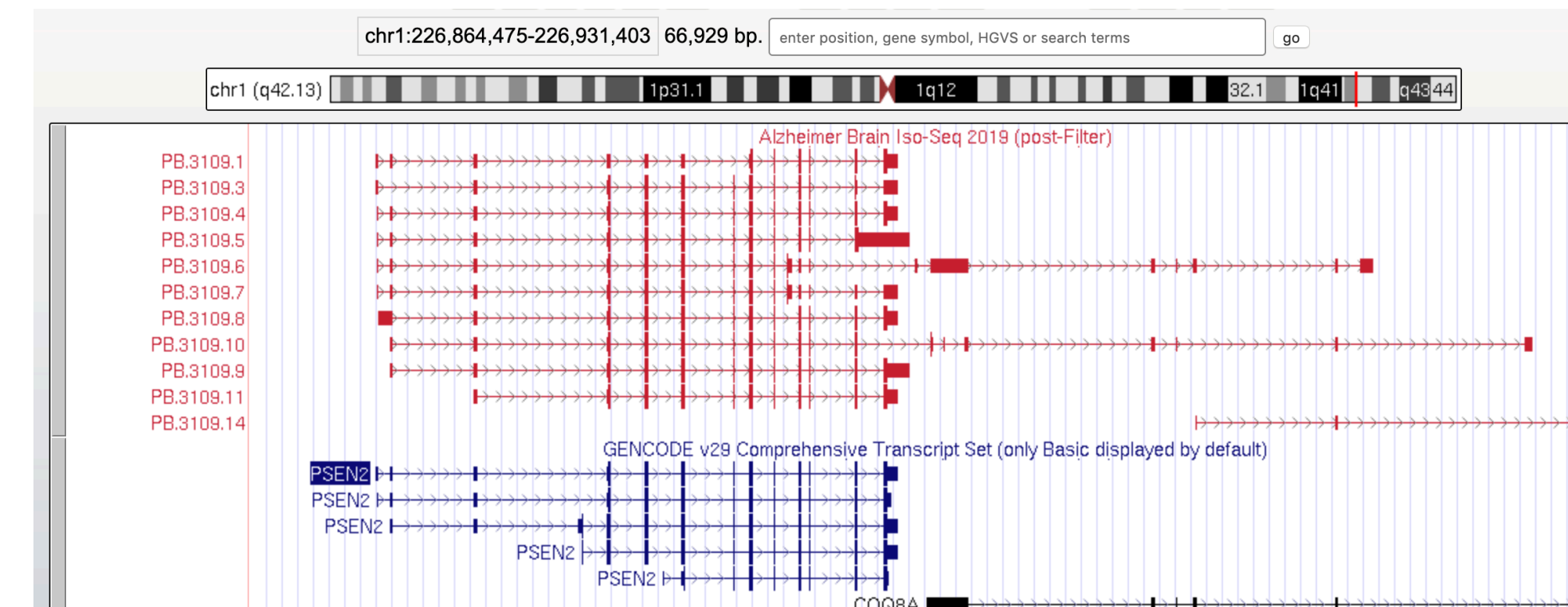


## Complex Splicing in Alzheimer Brain

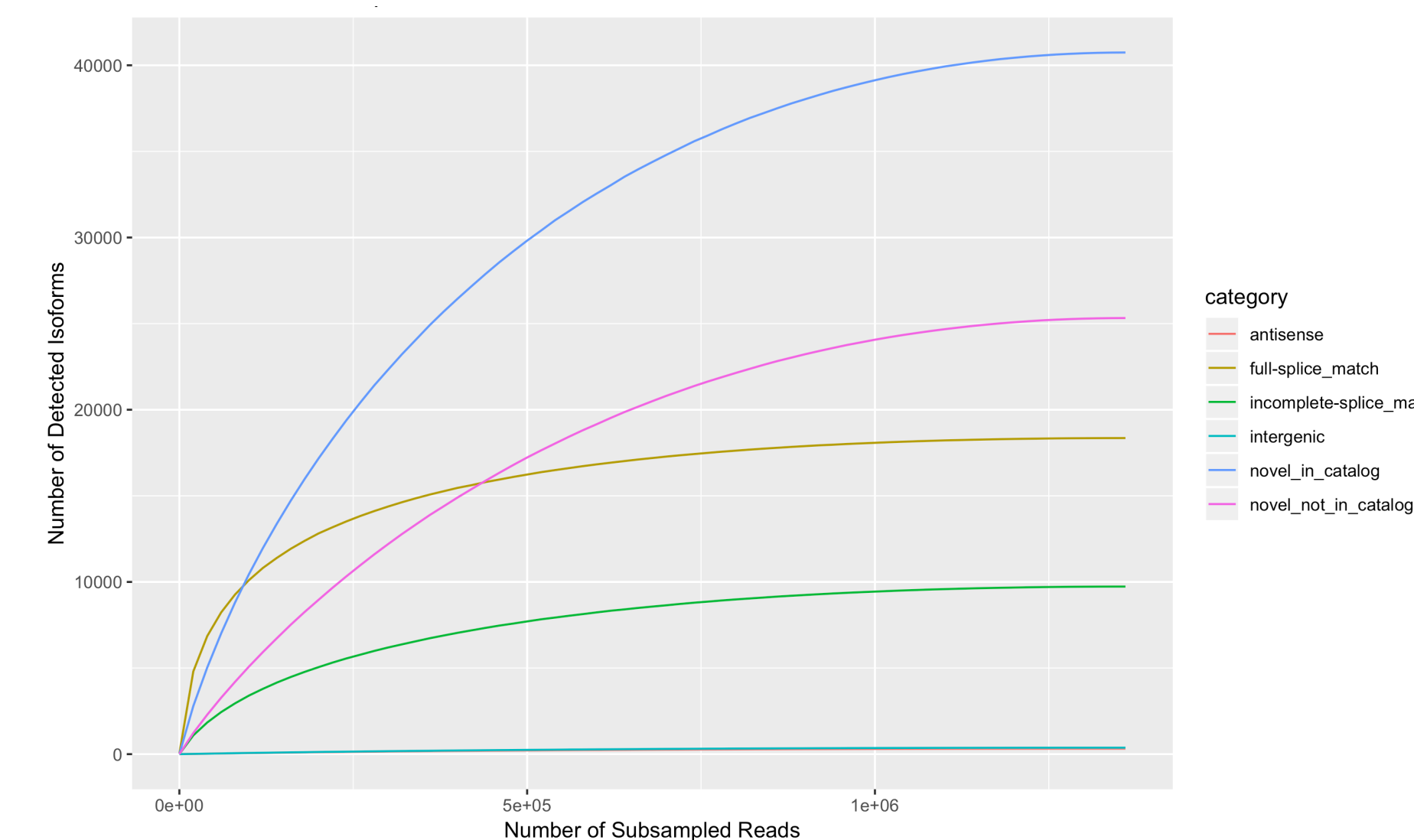
### PLD3 Isoforms with Alternative Starts/Ends



### PSEN2 Isoforms with Fusion to COQ8A



### Saturation of Known Transcripts



## Conclusions

- The Iso-Seq method produces high-quality, full-length transcripts up to 15 kb
- 1-day prep, 1-day sequencing, 1-day bioinformatics analysis
- Iso-Seq method reveals complex alternative splicing in Alzheimer brain sample

## References

- [1] <https://www.pacb.com/iseq>
- [2] Community Tools for Iso-Seq  
<https://github.com/magdoll/SQANTI2>  
[https://github.com/magdoll/cDNA\\_Cupcake](https://github.com/magdoll/cDNA_Cupcake)  
<https://github.com/genomerik/tama>