PacBio

Whitepaper

Transforming plant and animal research with full-length and single-cell isoform sequencing

The ongoing effects of population growth and climate change present formidable challenges to global biodiversity and agriculture. Understanding the biology of plants and animals in response to environmental factors is critical to mounting solutions to these challenges. Transcriptome sequencing is a valuable tool for revealing the dynamic responses of different organisms to environmental stress such as drought and diseases.

Alternative splicing (AS) in eukaryotic species generates functional diversity by expressing different combinations of exons in the same gene. RNA sequencing of expressed transcripts has historically been performed using short reads, but this method suffers from inherent limitations. Read lengths are too short to capture entire transcripts from end to end, making it difficult to achieve isoform-level resolution. Because isoforms, not genes, are often the information required to assess functional impacts, short reads fail to capture the complexity of plant and animal transcriptomes.

The PacBio® Iso-Seq® method offers full-length isoform information using highly accurate long reads (Figure 1), with no assembly required. This isoform-level resolution enables characterization of complex alternative splicing, prediction of open reading frames (ORFs), allele-specific isoform expression, and cell type differentiation.

In this whitepaper, we discuss the features of full-length and single-cell isoform sequencing with the Iso-Seq method and their applications in plant and animal research. We show how combining the Iso-Seq method with Kinnex™ kits further increases throughput, making PacBio long reads a cost-effective RNA sequencing solution.

102-326-605 REV01 JAN2024

PacBio

How does Kinnex scale throughput?

Based on the Multiplexed Arrays Sequencing (MAS-Seq) (Al'Khafaji et al., 2023) method, Kinnex joins together full-length cDNAs into larger fragments, allowing for multiple transcripts to be sequenced on a single HiFi read. The result is higher throughput for cost-effective isoform sequencing. Compared to non-Kinnex methods, the *Kinnex single-cell RNA kit* increases throughput by 16-fold for single-cell RNA-Seq, while the *Kinnex full-length RNA kit* increases throughput by 8-fold for full-length RNA sequencing.

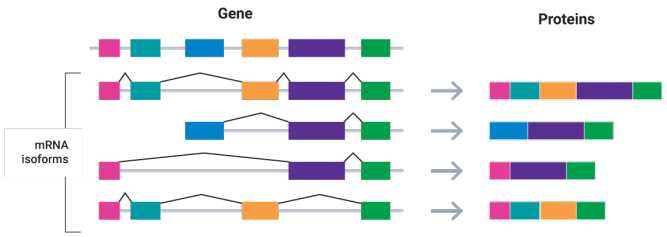

Figure 1. HiFi reads span the full length of individual transcripts without the need for assembly, which enables functional prediction of different isoforms.

What can you do with Kinnex?

- Generate comprehensive whole transcriptomes at gene and isoform-level resolution
- Build annotations for assessment of functional elements of the genome
- Discover key transcription factors and transcriptional signatures, alternative splice sites, and transcriptional fusion or readthroughs with high accuracy
- Predict ORFs for evaluation of protein or functional RNA coding regions
- Identify allele-specific isoform expression
- Reveal cell type-specific isoform expression and new cell clusters in single-cell transcriptomes

Kinnex end-to-end workflow

The analysis of HiFi reads to generate accurate, full-length transcript isoforms is streamlined and fully supported with Iso-Seq and Single-cell Iso-Seq analyses in SMRT® Link that are compatible with popular tertiary analysis tools.

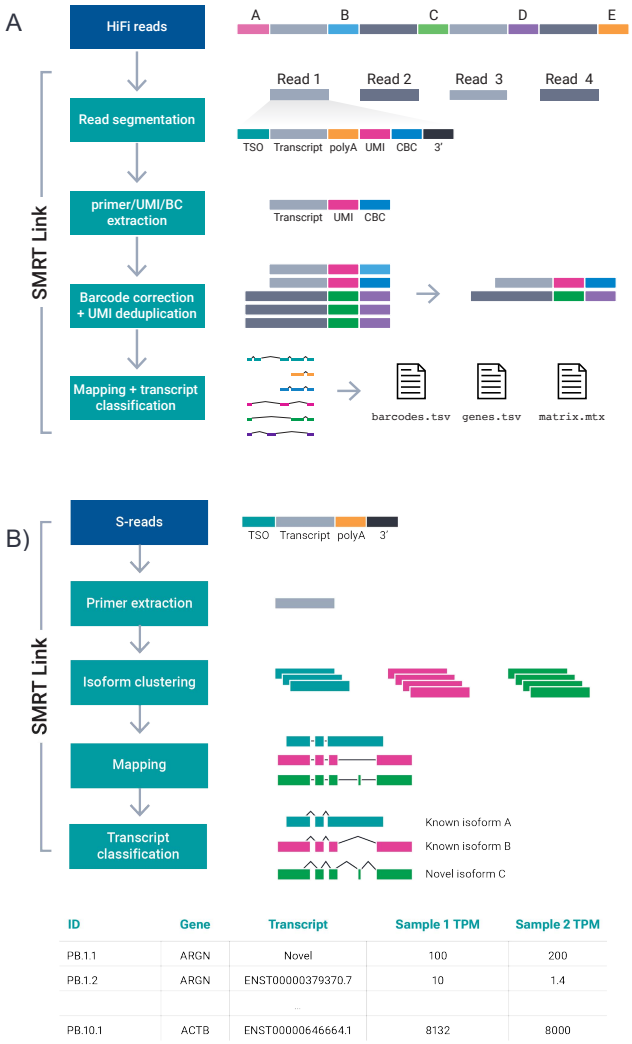

Figure 2. Iso-Seq analysis workflow through SMRT Link for a) single-cell and b) bulk isoform sequencing.

102-326-605 REV01 JAN2024

Transforming plant + animal research

More complete genome annotation for actionable insight

Genome annotation is the process of using sequence data to delineate the structural elements of a genome and identify the functional impacts of these regions. This process is extremely important because it translates raw sequence data to an organism's physiological functions and is what converts the potential of genomic information into actionable insight for researchers. While an annotation serves to complete the package of a genome, it is also valuable in itself, providing information about isoforms produced under varying conditions. Genome annotations are incredibly useful to plant and animal researchers who look to map responses to abiotic stressors like drought tolerance and reactions to variations in temperature and salinity. The high accuracy and long read lengths of HiFi sequencing demonstrate a distinct advantage over other technologies for generating high quality and more comprehensive genome annotations.

Iso-Seq in action

Xiao et al. (2023) performed transcriptome sequencing across floral development stages on the plant *Trachycarpus fortunei*. In addition to the plant's economic value, understanding the molecular regulation of this species' flower development is particularly important due to its unique reproductive system. In contrast to most other plants, *T. fortunei* is a dioecious species, which means each individual plant has only male or female flowers that develop from initially bisexual flowers. With the high-quality PacBio annotation, the authors identify 159 differentially expressed genes between male and female flowers corresponding to variable biochemical pathways and 28 genes as potential sex-associated SNPs. These findings point to specific mechanisms responsible for this unique reproductive system in plants and stress the **power of highly accurate long-read annotation methods for plant and animal research.**

PacBio provides solutions for full-length and single-cell RNA sequencing

	Full-length RNA sequencing	Single-cell RNA sequencing
PacBio advantages	• Achieve full-length isoform resolution (1–10 kb) without the need for assembly • Identify rare or novel transcripts and alternative splicing events, and gene fusions • Predict ORFs • Discover allele-specific isoform expression and usage	• Provide full-length isoform information (1–5 kb*) at the individual cell level • Reveal cellular heterogeneity and signals • Identify cell type-specific, allele-specific isoform expression
Example research topics	• Identify novel splicing events to investigate mechanisms of responses to biotic and abiotic stressors • Access transcript diversity for high value traits for crop breeding programs	• Explore species' cell atlases to characterize different cell types and understand the diversity of life at cellular resolution • Understand complex tissue development in endangered species to inform conservation plans
Supporting products & platforms	• Iso-Seq express 2.0 kit and *Kinnex full-length RNA* kit • Sequel® II/IIe and Revio™ systems	• *Kinnex single-cell RNA kit* • Sequel II/IIe and Revio systems

*Transcript length distribution limited by single-cell platforms used to generate single-cell cDNA

102-326-605 REV01 JAN2024

Single-cell sequencing across species

While long-read RNA sequencing is integral for developing more complete genome annotations, it is also becoming equally important in single-cell mapping studies, particularly for example, efforts to map cell types across diverse species. The transcriptomes made available by the Kinnex single-cell RNA kit make these studies more approachable with a 16-fold increase in throughput. Using highly accurate single-cell RNA sequencing to understand how tissues function across diverse species will help plant and animal researchers understand how these species have evolved and can help inform conservation decisions.

Iso-Seq in action

Ayana, Zandecki et al. (2021) performed single-cell Iso-Seq on the brain of the *Nothobranchius furzeri* fish to investigate the mechanisms of brain regeneration in response to injury, aging, and disease. While the authors initially performed single-cell RNA sequencing with short reads, they also use the **full-length PacBio Iso-Seq method to characterize full-length isoforms across the entire transcriptome and improve read mapping by 53%.** This high-quality sequencing data allows the authors to identify 25 cell types in the brain with spatial and relational information, providing a fuller picture of brain cell type diversity in a promising new model organism.

Isoforms drive species diversity

In addition to genome-wide structural variation, alternative splicing is another important molecular mechanism for driving plant and animal evolution. The most comprehensive tool to examine alternative splicing is full-length RNA sequencing (the Iso-Seq method) with HiFi long-read sequencing. Given that molecular diversity translates to species diversity, HiFi long-read sequencing presents the path to understanding biodiversity.

Iso-Seq in action

Reese et al., 2023 used the Iso-Seq method to characterize the transcriptomic origin of diversity across species by generating 264 Iso-Seq libraries from human and mice samples. Of the 200,000 full-length transcripts generated, the authors identified 40% as containing novel exon junction chains. The authors show that 57.8% of orthologous gene pairs use a different isoform diversity mechanism between human and mice samples of matching tissues and that this occurs on a gene-by-gene basis. These comparative results demonstrate that transcript structure diversity drives the diversity between species**. This study lays the foundation of using long-read RNA sequencing to evaluate the evolution of differential mechanisms of isoform diversity across species** and impacts the interpretation of orthologous genes between humans and animal models.

Conclusion

These Iso-Seq research examples demonstrate the advantages of full-length isoform sequencing with Kinnex to transform plant and animal research. The capabilities of PacBio HiFi sequencing to provide full-length isoform resolution opens the door to the cutting-edge biological discoveries needed to address the intensifying global challenges to biodiversity and agriculture.

102-326-605 REV01 JAN2024

Resources and references

Resources

1. Procedure & checklist – Preparing Kinnex libraries using the Kinnex full-length RNA kit
2. Procedure & checklist – Preparing Kinnex libraries using Kinnex single-cell RNA kit
3. Application note – *Kinnex full-length RNA kit* for isoform sequencing
4. Application note – *Kinnex single-cell RNA kit* for single-cell isoform sequencing
5. Application note – Bioinformatics tools for full-length isoform sequencing
6. Iso-Seq Bioinformatics resources

References

Al'Khafaji et al. (2023). High-throughput RNA isoform sequencing using programmed cDNA concatenation. *Nature Biotechnology*, 1-5. https://doi.org/10.1038/s41587-023-01815-7

Ayana, R., Zandecki, C., Mariën, V., Seuntjens, E., & Arckens, L. (2021). Single-cell sequencing unravels the cellular diversity that shapes neuro-and gliogenesis in the fast aging killifish (*N. furzeri*) brain. *bioRxiv*, 2021-07. https://doi.org/10.1101/2021.07.04.450918

Reese, F., et al. (2023). The ENCODE4 long-read RNA-seq collection reveals distinct classes of transcript structure diversity. *bioRxiv*, 2023-05. https://doi.org/10.1101/2023.05.15.540865

Xiao, F., Zhao, Y., Wang, X., Mao, Y., & Jian, X. (2023). Comparative transcriptome analysis of dioecious floral development in *Trachycarpus fortunei* using Illumina and PacBio SMRT sequencing. *BMC Plant Biology*, *23*(1), 536. https://doi.org/10.1186/s12870-023-04551-x

Explore plant + animal genomics solutions:
https://www.pacb.com/plant-animal-sciences/

102-326-605 REV01 JAN2024