

Introduction

Climate change and the rapidly growing global population are increasingly straining world food production.

Long-read sequencing is being used in large-scale agricultural initiatives to help improve crop yields and combat pests and disease to meet increased agricultural demand.

The PacBio Revio system will enable researchers to meet the throughput demands for agrigenomics while maintaining best-in-class quality.

- **High throughput** → 1,000s of plant genomes per year
- **Low cost** → ~\$10/Gb
- **High quality** → >90% of bases ≥Q30

To complement the potential of Revio, a high-throughput end-to-end workflow was developed and validated.

Plant DNA extraction challenges

Extraction of high-quality DNA from plants is often time-consuming and difficult, due to:

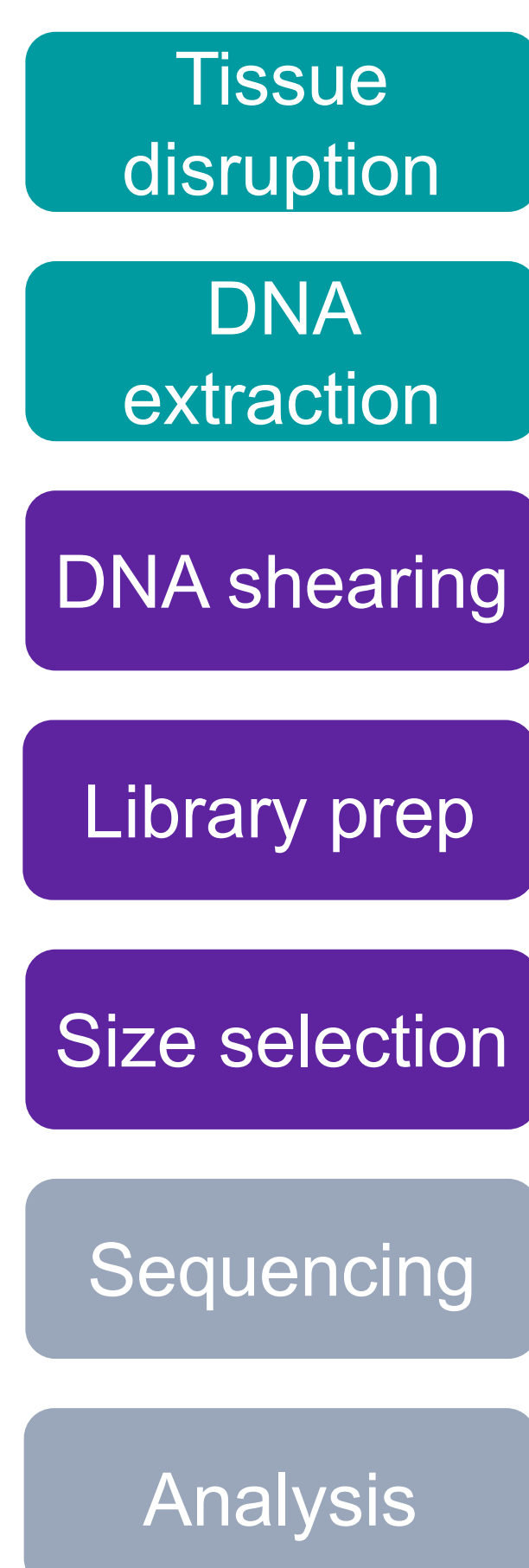
- Challenges with tissue and cell wall disruption
- Presence of nucleases, polysaccharides, and secondary metabolites
- Variable sample age and quality

Manual protocols typically utilize liquid nitrogen disruption, nuclei isolation, and CTAB precipitation which are not amenable to high-throughput applications.

High-throughput workflow

Corteva and PacBio collaborated to develop an end-to-end workflow for generating and processing thousands of HiFi libraries per year for plants.

Corteva/PacBio high-throughput workflow



High-throughput tissue disruption

A high-throughput tissue disruption method was developed to enable processing of samples in a 96-well plate format.

- Protocols were developed for the SPEX Geno/Grinder, MP Biomedicals FastPrep-96, and SPEX MiniG
- All three instruments performed equivalently
- Able to disrupt 96 plant samples <10 min
- The same equipment can be utilized for DNA shearing



High-throughput DNA extraction

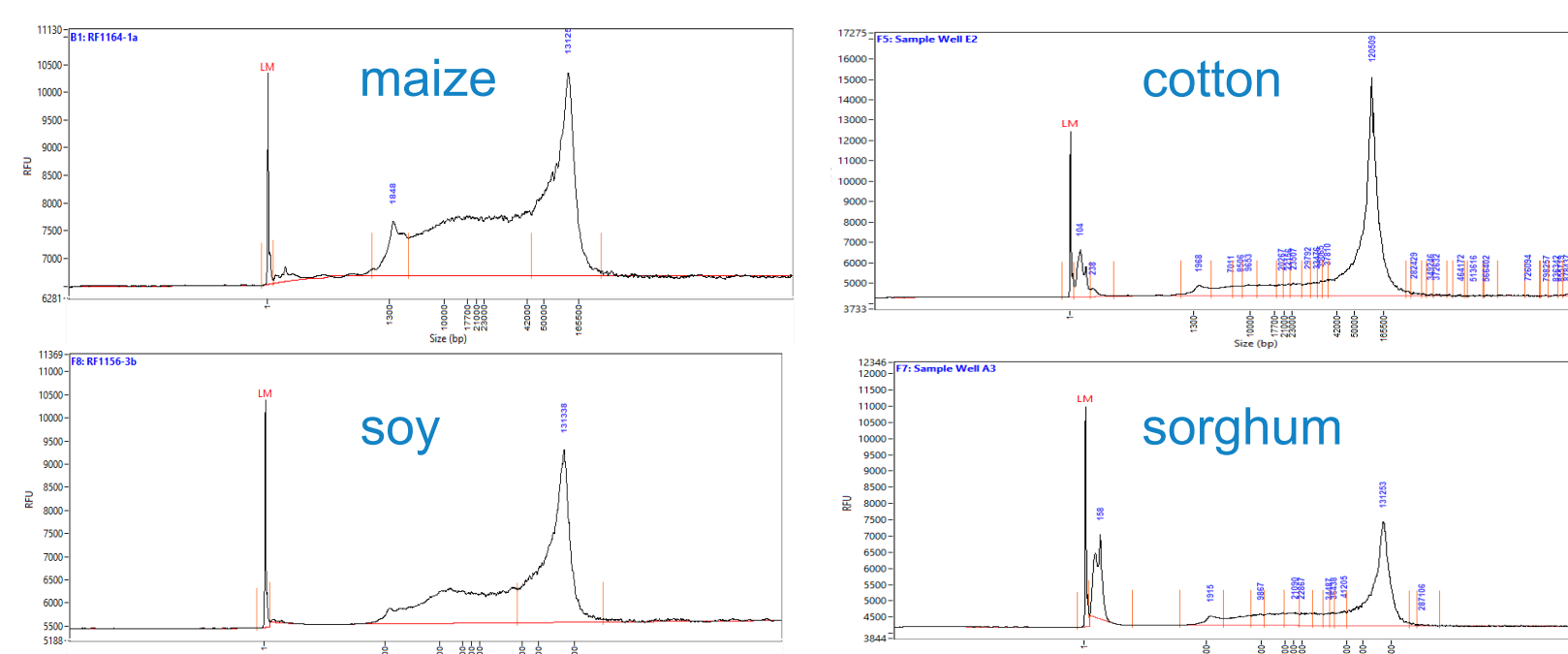
DNA extraction was automated on the Kingfisher DNA purification systems with PacBio-developed DNA extraction reagents and protocols.



| Sample type | Input mass (mg) | DNA yield (µg) | DNA mode size (bp) |
|-------------|-----------------|----------------|--------------------|
| Maize | 40 | 3.0–4.3 | 120,333 |
| Soy | 40 | 4.2–4.6 | 129,366 |
| Pepper | 50 | 3.7–4.4 | 151,137 |
| Wheat | 50 | 13.4–20.3 | 163,106 |
| Cotton | 40 | 0.6–3.6 | 120,509 |
| Sorghum | 40 | 0.9–2.0 | 131,253 |
| Sunflower | 40 | 0.9–2.1 | 127,224 |

Leaf tissue from several different plant types was disrupted using the SPEX Geno/Grinder and extracted using this protocol.

- DNA mode size as measured by Femto Pulse is >100 kb for all plants — example electropherograms are shown below



High-throughput DNA shearing

Two high-throughput shearing methods were developed.

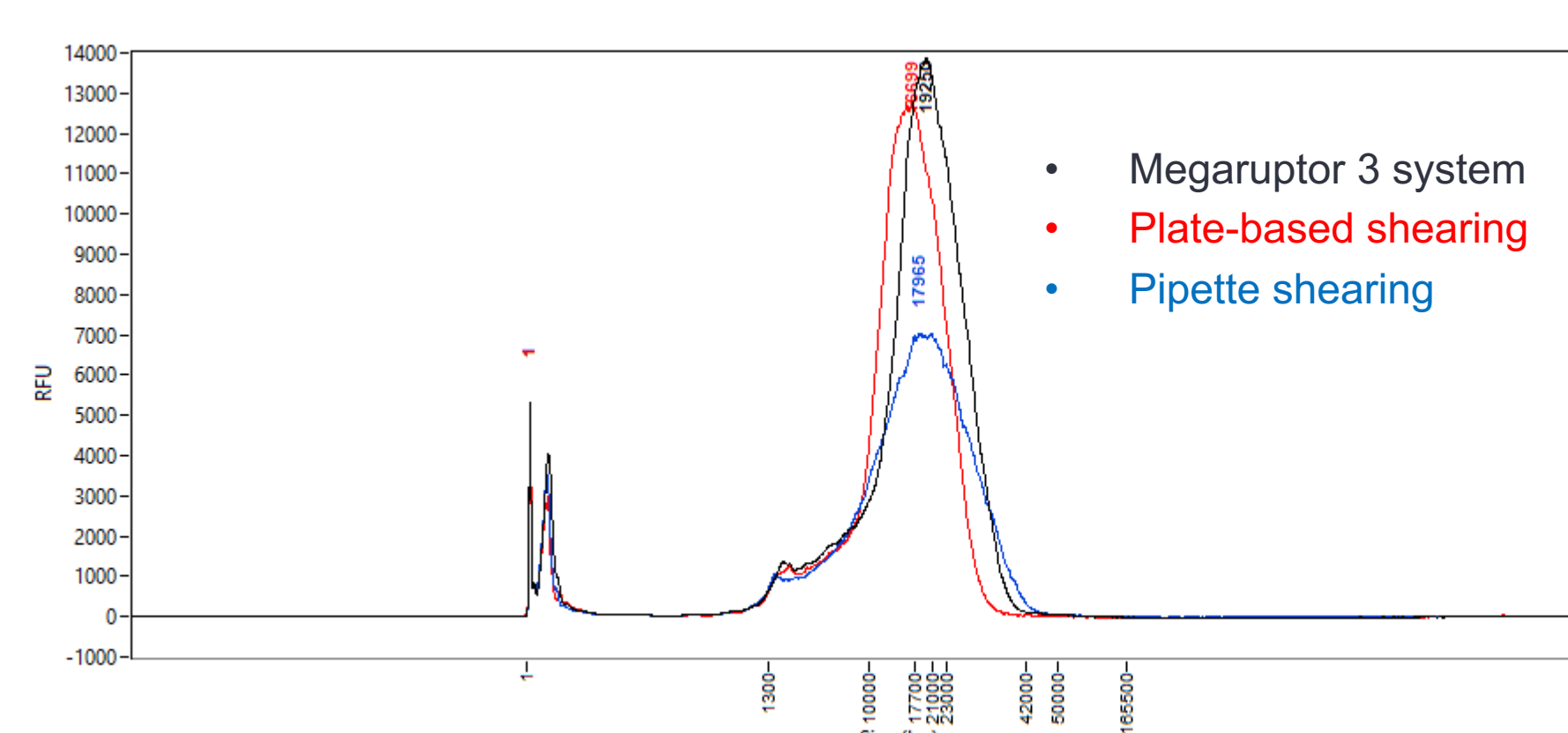
Plate-based shearing

- Protocols developed for FastPrep-96, Geno/Grinder, and MiniG
- Process 96 samples in <10 min
- Same equipment as used for tissue disruption



Pipette shearing

- Protocols developed for Hamilton NGS Star and Hamilton NIMBUS but are easily adaptable to other liquid handlers
- Process 96 samples in 15 min
- Requires only standard 300 µL pipette tips
- Can be incorporated into automated library prep

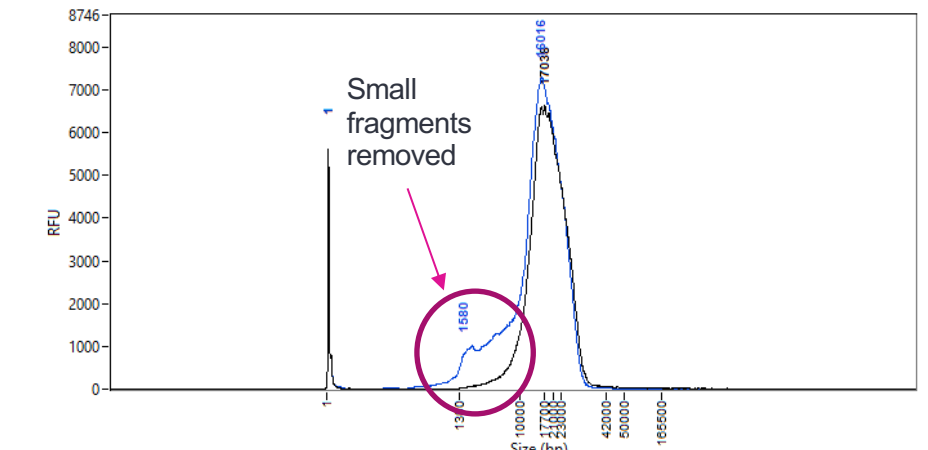


Maize DNA sheared using the two high-throughput shearing methods produced similar size distributions as DNA sheared with the Megaruptor 3 system.

Automated library preparation and size-selection

Library preparation, utilizing the SMRTbell prep kit 3.0, was automated onto a Hamilton NGS star.

- A new SMRTbell size-selection wash was developed to improve removal of small fragments



Sequencing on PacBio Revio system

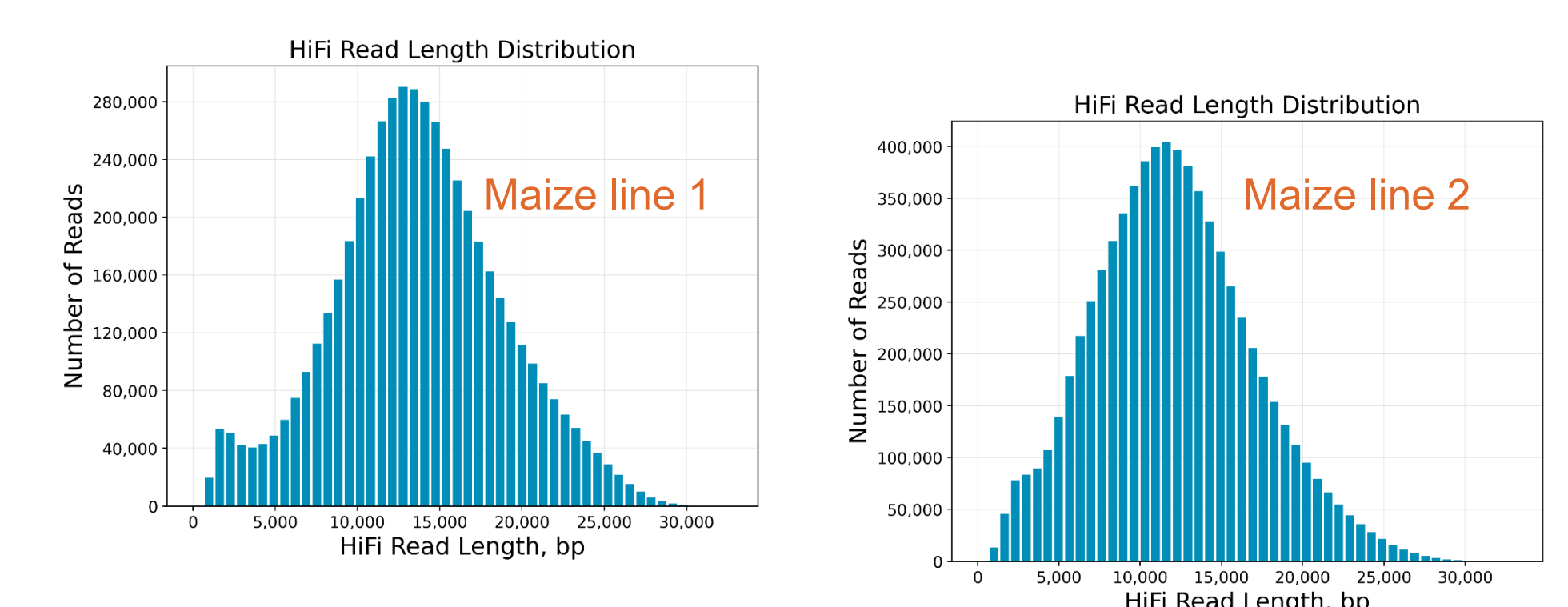
Two maize lines were processed using the Corteva/PacBio high-throughput sample prep workflow and sequenced on the PacBio Revio system.



Methods

- Disruption on SPEX Geno/Grinder
- Extraction on KingFisher Apex
- DNA shearing on Hamilton NIMBUS
- Library preparation with SMRTbell prep kit 3.0
- SMRTbell size-selection wash
- 24-hour sequencing movie

| Sample | Mean HiFi read length (bp) | HiFi yield (Gb) |
|--------------|----------------------------|-----------------|
| Maize line 1 | 12,064 | 86.8 |
| Maize line 2 | 13,662 | 70.9 |



Comparable HiFi read length distributions were obtained for the two maize lines.

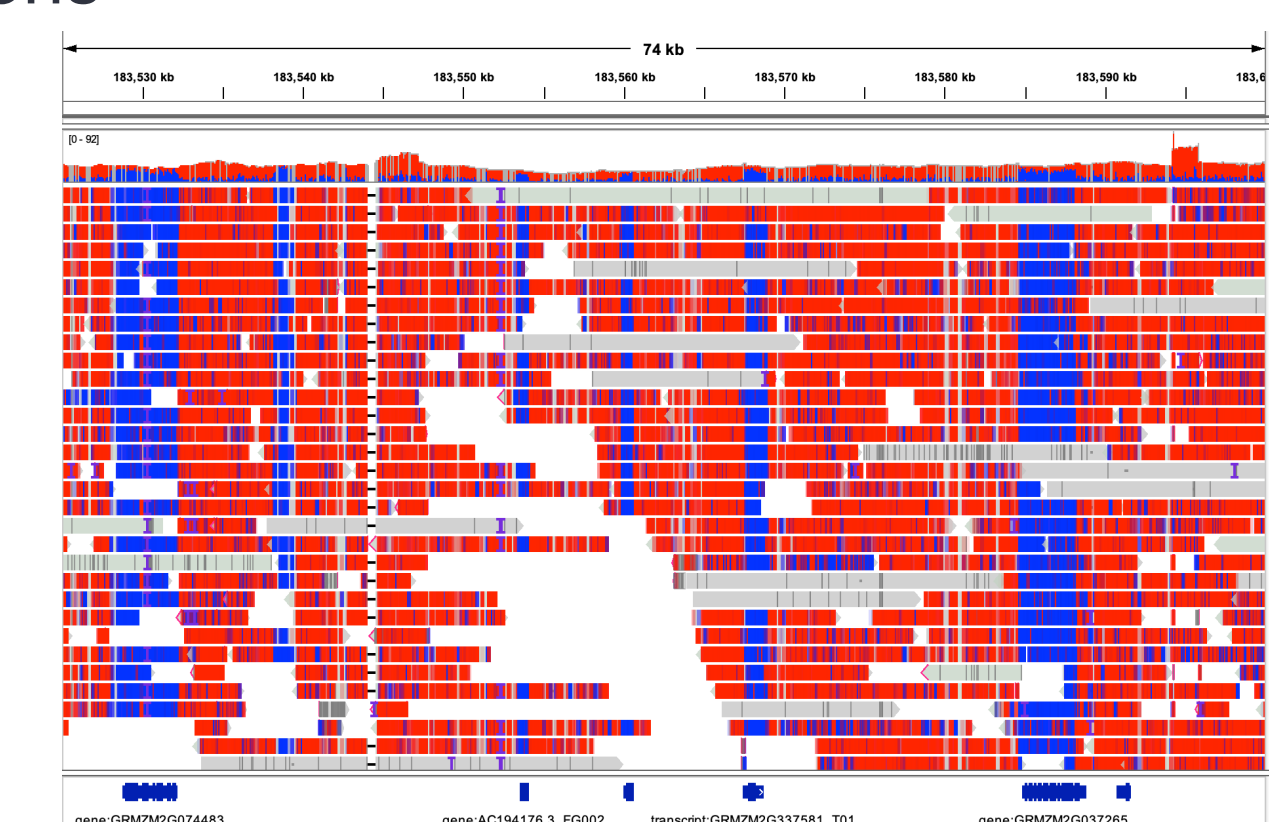
| Sample | Assembly size (Gb) | Contig N50 (Mb) | Contig N90 (Mb) |
|--------------|--------------------|-----------------|-----------------|
| Maize line 2 | 2.23 | 162.6 | 16.6 |

Assembly was performed on maize line 2 using hifiasm.

Methylation analysis

Revio performs 5-base sequencing (A, C, G, T + 5mC) by default with on-instrument basecalling and Google DeepConsensus. A representative region of the genome is shown.

- Red indicates hypermethylation
- Blue indicates hypomethylation
- Expressed genes indicated by blue, unmethylated regions



High-throughput capabilities

The Corteva/PacBio high-throughput workflow described has been validated by simultaneous processing of 96 plant samples with multiplexed sequencing on Revio. (Data not show)

Acknowledgements

The authors would like to thank the following for their data contributions: John Tenhundfeld, Zach Ireland, Paige Roos, and Justin Sibert