APPLICATION NOTE

Modernizing molecular breeding with low-pass whole genome sequencing.

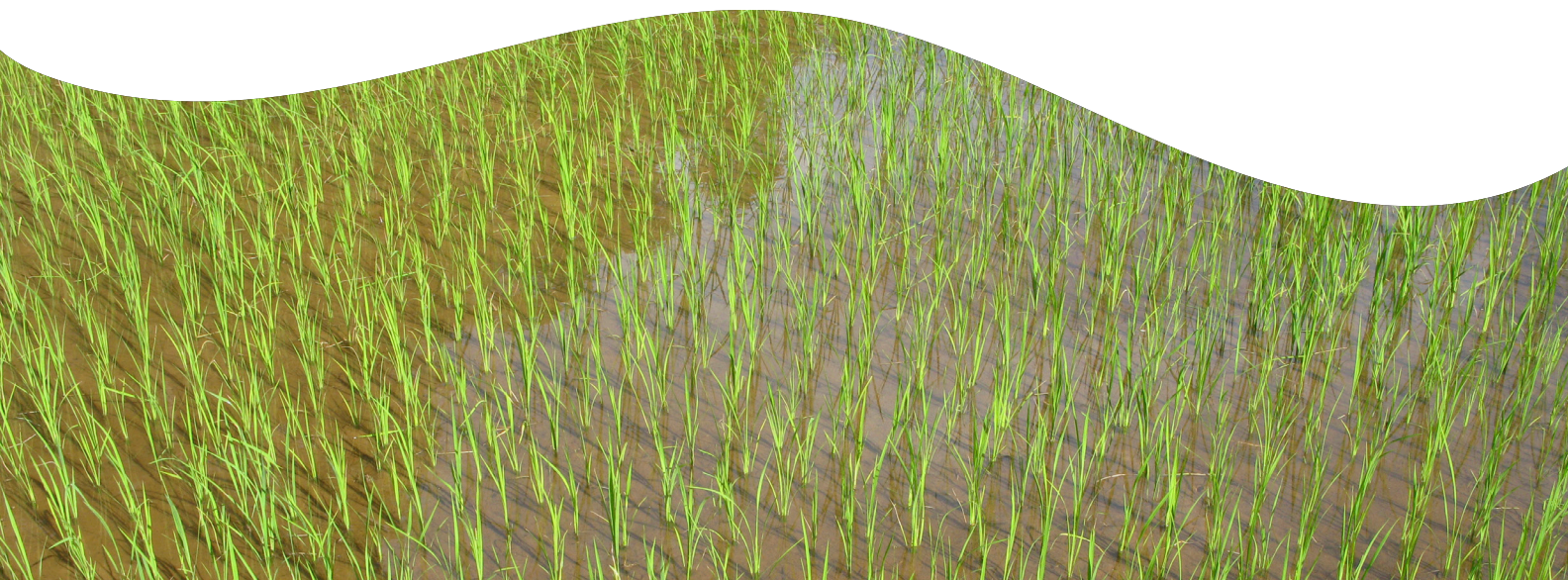
In collaboration with

CURIO

Low-pass whole genome sequencing (LP-WGS) offers a cost-effective approach to sequencing entire genomes at coverages ranging from 0.1x to 10x. This technique, followed by imputation, is becoming the preferred method for genotyping among plant and animal breeders. LP-WGS captures a broader spectrum of genetic variation compared to SNP arrays and is not limited by fixed content, speeding up the selection process and improving the reliability of predictions. However, the high computational demands associated with processing and imputing LP-WGS data make it less accessible for many research programs and breeding initiatives.

Methods

Sample preparation and sequencing were performed at Texas A&M AgriLife (TX, USA). Sixteen rice samples from clonal plants corresponding to the Presido cultivar (tropical japonica) were homogenized mechanically using the Omni Bead Ruptor™ 96 bead mill homogenizer (Revvity), followed by DNA isolation using the chemagic™ 360 instrument with the chemagic DNA Plant 100mg Kit (Revvity, CMG-795). Libraries were prepared using the NEXTFLEX™ HT Agrigenomics Low-Pass WGS Kit (Revvity) according to the manufacturer's instructions. Sequencing was performed on the Illumina® NovaSeq® X. FastQ files were uploaded to the CURIO™ platform for QC, alignment, and imputation analysis, using the Rice (*Oryza sativa japonica*) – IRGSP 1.0 assembly and Plant-Impute DB Project's reference panel.



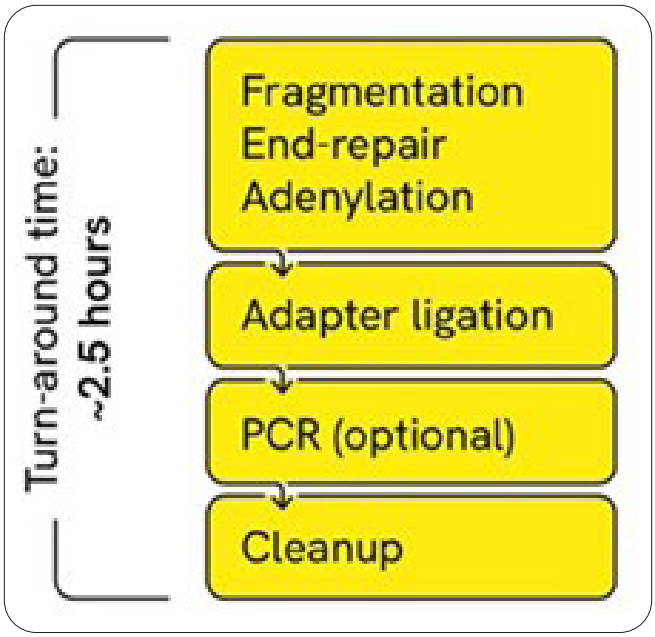


Figure 1: NEXTFLEX HT Agrigenomics Low-Pass WGS Kit workflow

Results

The CURIO™ platform completed the analysis of the FASTQ data (27 Gbp per sample) within 30 minutes. The average depth was 41.7x, corresponding to 97% of covered positions in the rice genome (Figure 2). Equal coverage of chromosomes is crucial for variant detection and reliable imputation. The average depth and percentage of coverage positions were comparable in chromosomes 1 to 12, but the chromosomes of mitochondria and chloroplast showed lower coverage. There are reports stating mechanical homogenization might favor the extraction of nuclear DNA, which is more abundant and robust compared to the smaller and more fragile mitochondria and chloroplasts.

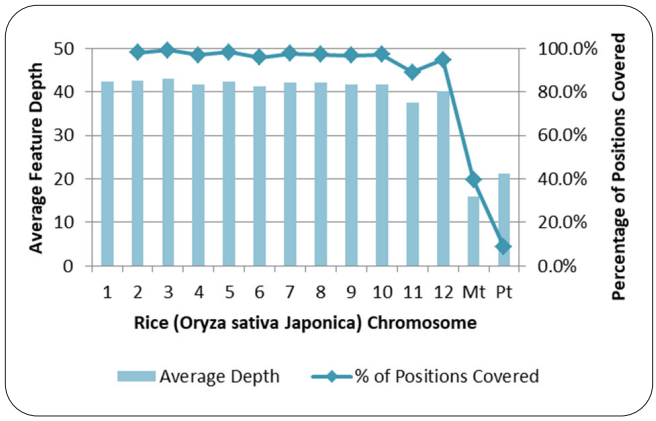


Figure 2: Average depth per chromosome without any down-sampling. Comparable percentage of covered annotated gene positions is observed across chromosome 1-12, with evident drop for mitochondrial and chloroplast chromosomes

Variant analysis

Single nucleotide variants (SNVs) from homozygous or heterozygous positions with depth ≥ 2 were reported. The imputation algorithm used was Beagle 5.4. The imputation analysis compared the LP-WGS data to a reference panel to predict missing genotypes using statistical algorithms that exploit linkage disequilibrium patterns. The result is a larger set of genotypes that enhances the utility of genomic data for downstream analysis.

Figure 3: Variant density (~40x average depth) across the genome (A) and in chromosome 1 (B). A region of 1 Mb (chr1:8,445,656-9,445,655) is randomly selected to show the identified SNVs, which in this case are 993 (C).

Imputation analysis

Imputation analysis compares the LP-WGS data to a reference panel to predict missing genotypes by using statistical algorithms that exploit linkage disequilibrium patterns. The result is a larger set of genotypes that enhances the utility of genomic data for downstream analysis. The imputation algorithm used here was Beagle 5.4.

Concordance calculates the proportion of genotypes that are correctly imputed compared to the true genotypes. High concordance means that the imputed data closely matches the true genetic data. A related metric is the imputation quality score (IQS), that compares the imputed genotypes to a reference set of known genotypes. While both concordance and IQS assess the reliability of the imputation, IQS is more comprehensive as it considers various factors such as allele frequency and genotype uncertainty. Both metrics are automatically calculated by the CURIO™ platform (Table 1).

Table 1: Number of SNV detected before and after imputation analysis. Please note how similar are the observed concordances at 4x and 41x.

Average depth	SNV number (variant analysis)	SNV number (imputation analysis)	Concordance	IQS
41.7	1,636,318	4,897,277	90.30%	0.739
4.0	1,014,042	4,897,277	90.18%	0.691
2.0	603,490	4,897,277	86.69%	0.535
1.5	455,145	4,897,277	85.55%	0.488

The CURIO™ Genomics Platform delivers a visual representation of the imputation analysis with a single-click (Figure 4).

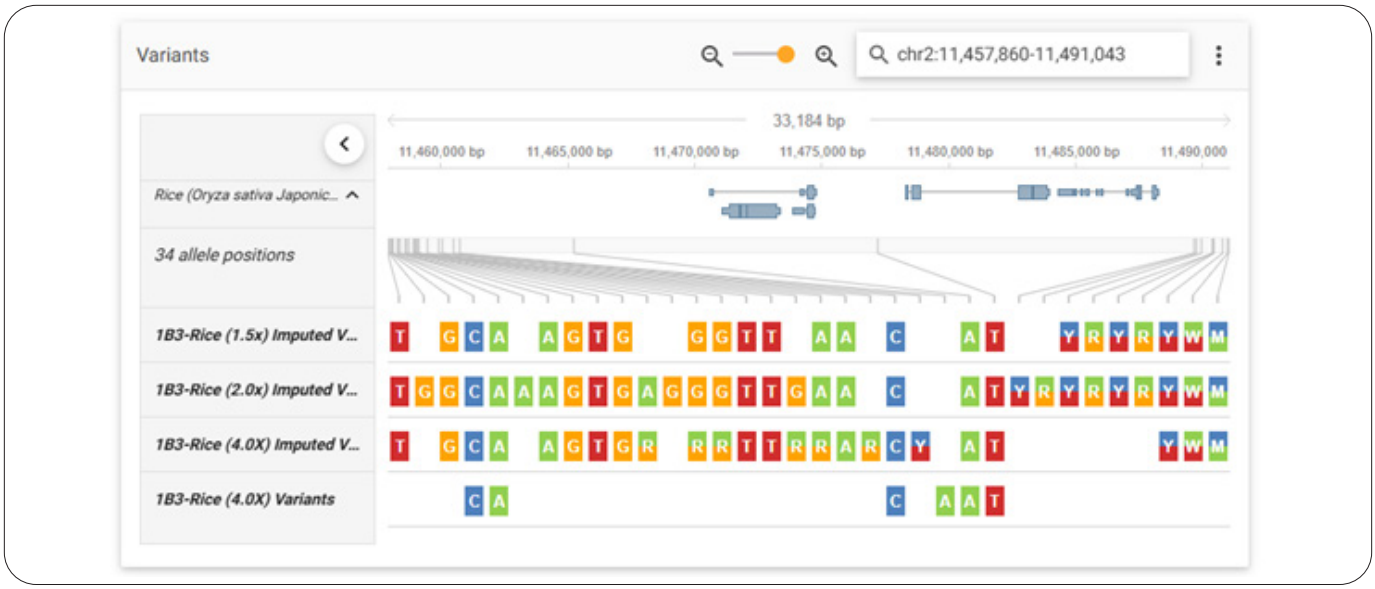


Figure 4: 33kb region in chromosome 2 showing the called variants at 4.0x versus resulting imputed variants at different depths.

Conclusion

The combination of the NEXTFLEX HT Agrigenomics Low-Pass WGS Kit with the CURIO™ platform offers a fast and robust solution for genotyping any plant or animal species with just a few clicks, eliminating the need for bioinformatics expertise or computational infrastructure.This streamlined approach simplifies and democratizes high-throughput plant and animal studies from start to finish.

References

Fragoso, C. A., et al. (2020). "Low-pass sequencing in crop genomics and breeding: Emerging applications." Plant Genomics, 14(1), 234-243.

Yang, J., et al. (2021). "Cost-efficient whole-genome sequencing-based genotyping for crop research and breeding." Genomics, 113(2), 869-875.

Hickey, J. M., et al. (2022). "Genomic selection in crops: Prospects and challenges with low-pass whole genome sequencing." Plant Biotechnology Journal, 20(2), 345-355.

Wang, M., et al. (2023). "Imputation accuracy in low-pass whole genome sequencing for agricultural crops." Frontiers in Plant Science, 11, 1057.

Edwards, D., et al. (2023). "Advancing crop genomics with low-pass sequencing: A review of current trends and future directions." Current Opinion in Plant Biology, 68, 102188.

Rice reference panel: gong_lab.hzau.edu.cn/Plant_imputeDB

Browning, B.L., et al (2021). “Fast two-stage phasing of large-scale sequencing data”. Am. J. Hum. Genet. 108(10):1880-1890.

Browning, B.L., et al (2018). “A one-penny imputed genome from next-generation Reference Panels.” Am. J. Human Genet. 103(3):338-348.

Stahl, K. et al (2021). “Assessment of Imputation Quality: Comparison of Phasing and Imputation Algorithms in real data.” Front. Genet., 12:724037.

Copyright ©2025, Revvity, Inc. All rights reserved.
For research use only. Not for use in diagnostic procedures.