Multiplexed IRRI's Rice Custom Amplicon Assay (RiCA) V5 Release Note: Updated with Higher Genomic Coverage and Additional Traits

AgriPlex Genomics collaboration

Introduction:

According to the Food and Agriculture Organization (FAO) of the United Nations, rice is one of the most important staple crops in the world providing a major source of calories and nutrients for billions of people. Global rice consumption is projected to reach 500 million metric tons by 2029 (FAO, 2021b). The increase in rice consumption is driven by population growth, urbanization, and changing dietary habits in many parts of the world.

Rice is cultivated globally; Asia, where the consumption of rice is a dietary tradition for many cultures, accounts for over 90% of the global production and consumption of the crop, with China and India alone producing more than half of the world's rice (FAO, 2021a). In Africa, rice production has been growing steadily, currently accounting for approximately 3% of global production. Latin America and the Caribbean account for 2%, and North America and Europe account for the remaining 1% of global rice production (FAO, 2021a, FAO, 2021b).

The increase in consumption requires a matching upturn in production. Molecular breeding can accelerate the process of developing new rice varieties with improved yield, quality, and disease resistance. Breeders can use molecular markers to identify genes or regions of the genome associated with desirable traits and increase the efficiency and the precision of selecting for these traits while decreasing the time and resources otherwise required by traditional breeding methods. Many new rice varieties with increased yield potential, grain quality, disease resistance, and pest resistance were developed using molecular markers. Some notable examples of successful molecular breeding programs include the development of Sub1 rice varieties that are tolerant to flooding (Xu et al., 2006), and improved disease resistance such as to bacterial blight (Jiang et al., 2012) and blast (Xiao et al., 2019).

The International Rice Research Institute (IRRI) has made significant contributions to the genetic research and molecular breeding of rice over the past several decades. IRRI's employment of molecular markers as a research tool has resulted in the development of new breeding technologies, high-yielding varieties, and the identification of useful traits. One of these contributions is The Rice SNP Consortium Array (RiCA): a high-throughput genotyping tool developed to improve rice molecular breeding efforts (Arbelaez et al., 2019). In 2020, IRRI in collaboration with AgriPlex Genomics constructed a multiplexed version of the RiCA panel on AgriPlex Genomics' PlexSeq platform. The multiplexed RiCA panel has been made available to the rice breeding community as RiCA version 2 and RiCA Version 4 (RiCAv4). In

2023, we have introduced the next version of this multiplexed rice SNP panel, RiCA Version 5 (V5).

The RiCA Version 5 SNP Panel

The performance of the markers in the existing RiCA V4 was evaluated and ranked based on success rate and robustness. The lowest performing 74 SNPs were purged from the panel.

126 New markers were added to the panel. Markers were chosen to close gaps in genome coverage and optimize the genomic distribution of the panel. These improvements are displayed in Table 1 that compares the genomic distribution of the new RiCA panel with the previous version. The average distance between adjacent markers of RiCA V5 decreased to 345 Kb and the average number of markers per chromosome increased to 90 SNPs. The 126 new markers and the 949 markers carried over from RiCA V4 panel were multiplexed and validated together to construct the new RiCA v5 panel, which consists of 1075 markers and maintains the basic structure of the panel, now made of:

- Genomic screen: 858 SNPs originating from the Cornell 6KArray Infinium Rice chip (Thomson et al., 2017) and the 3,000 Rice Genomes Project (Alexandrov et al., 2015; Mansueto et al., 2017; Wang et al., 2018). The markers are evenly distributed throughout the genome and were selected based on call rate higher than 95% and high minor allele frequencies (MAF ≥ 0.4). The markers added in the 5th version of the panel, were chosen to fill in gaps and provide more even representation of diversity of elite germplasm.
- Trait-related markers: 197 SNPs associated with different important trait genes/QTLs. Table 2 lists the traits and gene targets by trait categories.
- Purity Markers: 20 SNPs chosen for their discriminatory ability amongst elite indica material, used in determination of genetic purity, identity, and hybridity.

The complete markers list together with their chromosomal location can be found in Appendix A.

	RiCA V4		RiCA V5	
Chromosome	**N**	**Average distance between adjacent SNPs (Mbp)**	**N**	**Average distance between adjacent SNPs (Mbp)**
1	118	0.366495	138	0.311552
2	94	0.384033	102	0.353614
3	90	0.401370	93	0.393461
4	88	0.404819	88	0.404819
5	72	0.409232	82	0.356779
6	112	0.276178	108	0.286502
7	68	0.434530	82	0.359426
8	91	0.310137	89	0.317186
9	68	0.327277	66	0.334589
10	57	0.383034	62	0.336861
11	94	0.308710	88	0.330792
12	71	0.385049	76	0.354650
Mitochondrial	1	-	1	-
Total	1024	-	1075	
Average/chromosome	85	0.365905	90	0.345019

Table 1. SNP numbers comparison between the RiCA V4 and updated RiCA V5 panel.

Trait category	Trait	Gene targets
Disease	Blast	*Pi9 (3 alleles), Pik (11 alleles), Pii, Pita, Ptr, pi21, Pi35, Pi54, qPi33, Bsr-D1*
	Bacterial blight	*qXa4, xa5, qXa7, xa13, Xa21, Xa23, qXa26, Sweet14*
	Brown planthopper	*Bph9, Bph17, Bph32*
	Other insects	*qGm4*
	Virus	*TSV1, TBV1, rymv1-2, rymv1-4, rymv1-5, rymv2, RYMV3, STV11*
Grain quality	Amylose	*Waxy* (6 different alleles distinguished)
	Chalkiness	*PGC8.2, GW5*
	Aroma	*fgr-1*
	Gelatinization temp.	*Alk (3b* and *4)*
	Other grain quality	*NAS3 (grain zinc), LOX3 (spoilage), SBE-I, SBE-IIb, SBE-III, GPT1*
	Grain size/shape	*GS3, TGW6, SLG7, GW8, GS9*
Abiotic stress	Drought	*qDTY1.1, qDTY2.2, qDTY3.1, qDTY3.2, qDTY 4.1, qDTY12.1, Dro1*
	Cold	*Cold1, qSCT1, qCST10, qPSST6*
	Heat	*TT1*
	Salinity	*qSIS1.2, Saltol, qSOR1*
	Anaerobic germination	*qAG1, qAG3*
	Submergence	*Sub1, Snorkel1-Snorkel2*
	Other abiotic stresses	*BET1 (boron toxicity)*
Yield components	Heading date	*Ghd7, DTH8, Hd2/PRR37, Hd1, Hd6a, RFT1, Hd3a, ehd1*
	Miscellaneous	*Sd1, DEP1, NAL1, WFP1, HIS1, GNP1, GFR1, FZP1, NGR5, SCM2, LP, Tb2*
Hybrid		*Rf3, Rf4, ms-IR36, WA-CMS, tms5*

Table 2. List of traits and genes represented on the AgriPlex Genomics RiCA V5 Panel

PlexSeq™: The mid-density, multiplexed, SNP genotyping platform

The revised SNP collection was again multiplexed and validated as a PlexSeq panel. The most important attribute of PlexSeq is its simplicity. PlexSeq™ workflow consists of:

- Primer design and multiplex prediction: A proprietary algorithm that uses artificial intelligence to identify the optimal sets of compatible primers that can be mixed in one PCR amplification.
- Crude DNA isolation.
- Thermocycling: highly multiplexed Primary PCR followed by secondary barcoding PCR.
- Pooling and Sequencing: barcoded amplicons are combined, purified, quantitated, and uploaded on an NGS sequencers.
- Once the sequencing is complete, a proprietary genotype calling analysis software, provides an automated sequencer to data report workflow.

This workflow is amenable to automation; all steps are carried out on liquid handlers and high-capacity thermocyclers. The efficiency and usefulness of the panel is further supported by AgriPlex Genomics vast collection of barcode combinations. Thus, effectively, the number of individuals genotyped simultaneously is only limited by the sequencer's capacity.

In addition, the fact that PlexSeq panels are a collection of PCR primers not tethered to a surface (as for example in chips), provides the flexibility to dynamically customize and alter the composition of the markers in the panel so that it may best fit the germplasm or the application.

The 1k RiCA V5 is available as a service from AgriPlex Genomics. The panel and corresponding software are also available as a kit to be used by in-house genotyping laboratories.

Applications

Genomic selection:

The panel enables genomic selection in indica breeding programs. On average, any two elite indica parents will display between 300 – 500 polymorphic markers distributed across the genome (Figure 2), a density sufficient to enable imputation back to the full genome level. The combination of rapid turnaround time and low cost enables cost-effective genotyping of a prediction population during the last generation of line fixation, saving money on the cost of field space for seed increase and allowing rapid recycling of progeny as parents.

Marker-assisted backcrossing:

The RiCA V5 panel provides an excellent low-cost option for background recovery estimates in marker-assisted backcrossing programs. The combination of highly informative background markers together with a wide selection of peak trait markers allows accurate estimation of background recovery, ensures valuable genes from the recipient line are recovered, and may provide additional confirmation that a target gene is carried by the selected progeny. Background selection can reduce the number of backcross generations required by 2 or more to achieve >95% recipient recovery.

QTL profiling:

The SNP panel at AgriPlex contains 197 trait markers targeting high-value trait targets (table 2) covering a range of traits related to disease resistance, grain quality, abiotic stress tolerance, heading date, hybrid rice production and others. These markers are designed to give accurate profiling of the targeted genes and QTLs across all *Oryza sativa* genomic diversity, including both indica and japonica. The same, or part of these QTLs and traits can be interrogated by specific, single plex assays, providing a mechanism for initial selection based on major loci followed by validation and full genome profiling on the RiCA array.

Seed purity and hybridity:

The RiCA V5 panel includes 20 markers specifically chosen for their ability to distinguish elite material. Combined with the genome-wide and trait markers, this makes the panel an excellent choice for applications related to genetic purity, some examples are, purity testing (uniformity), F1 hybridity tests, and varietal identity fingerprinting.

Conclusion

The RiCA V5 panel offered by AgriPlex Genomics provides an excellent, cost-effective genotyping option for applications requiring a moderate to high sample throughput at modest SNP density. The size of the panel enables genomic selection, and the substantially expanded suite of trait markers can be useful in other Marker Assisted Selection applications. The panel fits in neatly with rapid line fixation protocols, due to its low cost-per-sample and fast turnaround time. These benefits enable major-locus selection and genomic selection to occur before field amplification of seed, saving both time and money.

The addition of purity markers expands its usefulness in a range of applications, going from safeguarding line identity along breeding programs and continuing into seed production. The flexibility of the PlexSeq platform also enables continual revision and upgrading of the

marker system, ensuring the technology keeps pace with current breeding and production needs.

References

FAO. (2021a). Rice Market Monitor. Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/economic/est/est-commodities/rice/rice-market-monitor/en/

FAO. (2021b). Rice Market Outlook. Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/worldfoodsituation/csdb/en/

Jiang, H., Feng, Y., Bao, L., Li, X., Gao, G., Zhang, Q., & Xiao, J. (2012). Improving blast resistance of Jin23B and its hybrid rice by marker-assisted gene pyramiding. Molecular breeding, 30(3), 1679-1688.

Xu, K., Mackill, D. J., & Ismail, A. M. (2006). Genetic variation for abiotic stress tolerance in rice. In Advances in molecular breeding toward drought and salt tolerant crops (pp. 301-312). Springer, Dordrecht.

Xiao W, Yang Q, Huang M, Guo T, Liu Y, Wang J, Yang G, Zhou J, Yang J, Zhu X, Chen Z, Wang H. Improvement of rice blast resistance by developing monogenic lines, two-gene pyramids and three-gene pyramid through MAS. Rice (N Y). 2019 Nov 4;12(1):78. doi: 10.1186/s12284-019-0336-4. PMID: 31686256; PMCID: PMC6828908.

Juan David Arbelaez, Maria Stefanie Dwiyanti, Erwin Tandayu, Krizzel Llantada, Annalhea Jarana, John Carlos Ignacio, John Damien Platten, Joshua Cobb, Jessica Elaine Rutkoski, Michael J. Thomson and Tobias Kretzschmar (2019) 1k-RiCA (1K-Rice Custom Amplicon) a novel genotyping amplicon-based SNP assay for genetics and breeding applications in rice. Rice 12-55. https://doi.org/10.1186/s12284-019-0311-0

Thomson MJ, Singh N, Dwiyanti MS, Wang DR, Wright MH, Perez FA, DeClerck G, Chin JH, Malitic-Layaoen GA, Juanillas VM, Dilla-Ermita CJ, Mauleon R, Kretzschmar T, McCouch SR (2017) Large-scale deployment of a rice 6 K SNP array for genetics and breeding applications. Rice 10-40. https://doi.org/10.11 86/s12284-017-0181-2

Alexandrov N, Tai S, Wang W, Mansueto L, Palis K, Fuentes RR, Ulat VJ, Chebotarov D, Zhang G, Li Z, Mauleon R, Hamilton RS, McNally KL (2015) SNP-seek database of SNPs derived from 3000 rice genomes. Nucleic Acids Res 43: D1023–D1027. https://doi.org/10.1093/nar/gku1039

Mansueto L, Fuentes RR, Borja FN, Detras J, Abriol-Santos JM, Chebotarov D, Sanciangco M, Palis K, Copetti D, Poliakov A, Dubchak I, Solovyev V, Wing RA, Hamilton RS, Mauleon R, McNally KL, Alexandrov N (2017) Rice SNP-seek database update: new SNPs, indels, and queries. Nucleic Acids Res 45: D1075– D1081. https://doi.org/10.1093/nar/gkw1135

Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F, Mansueto L, Copetti D, Sanciangco M, Palis KC, Xu J, Sun C, Fu B, Zhang H, Gao Y, Zhao X, Shen F, Cui X, Yu H, Li Z, Chen M, Detras J, Zhou Y, Zhang X, Zhao Y, Kudrna D, Wang C, Li R, Jia B, Lu J, He X, Dong Z, Xu J, Li Y, Wang M, Shi J, Li J, Zhang D, Lee S, Hu W, Poliakov A, Dubchak I, Ulat VJ, Borja FN, Mendoza JR, Ali J, Gao Q, Niu Y, Yue Z, Naredo MEB, Talag J, Wang X, Li J, Fang X, Yin Y, Glaszmann JC, Zhang J, Li J, Hamilton RS, Wing RA, Ruan J, Zhang G, Wei C, Alexandrov N, McNally KL, Li Z, Leung H (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature 557:43–49. https://doi.org/10.1038/s41586-018- 0063-

Appendices

Appendix A