

Detection of genomic deletions in rice using oligonucleotide microarrays

Myron Bruce¹, Ann Hess¹, Jianfa Bai², Ramil Mauleon³, M. Genaleen Diaz^{1,4}, Nobuko Sugiyama³, Alicia Bordeos³, Guo Liang Wang⁵, Hei Leung³, Jan E. Leach¹

¹Colorado State University, Fort Collins, CO 80523-1177 USA, ²Kansas State University, Manhattan, KS 66502 USA, ³International Rice Research Institute, Manila, Philippines, ⁴Institute of Biological Sciences, University of the Philippines Los Baños, Philippines, ⁵The Ohio State University, Columbus, OH 43210 USA

Introduction

The induction of genomic deletions by physical or chemical agents is an easy and inexpensive method to generate a genome-saturating collection of mutations. Different mutagens allow for a collection of mutants with varying deletion sizes. This allows for identification of deletions in single gene models or deletions spanning regions that may collectively govern a trait (e.g., quantitative trait loci, QTL). However, deletion mutants have not been widely used in functional genomics because the mutated genes are not tagged and therefore, are difficult to identify. We present a microarray-based approach to identify deleted genomic regions in rice mutants selected from a large collection of gamma ray (G) and fast neutron (N) generated deletion lines developed at IRRI.

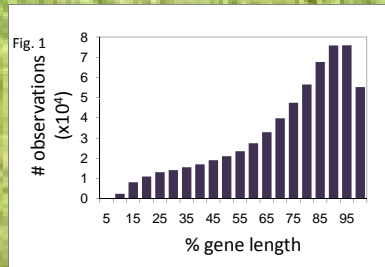
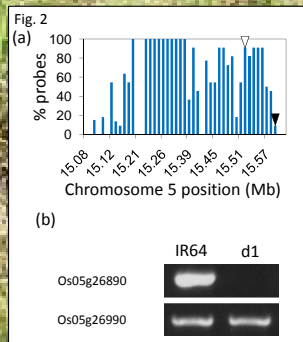


Fig. 1. Distribution of probes on the Affymetrix Rice GeneChip® is biased to the 3' end of gene models. Because absolute lengths of genes vary, the genes are represented as percentage of length. The Affymetrix probes were binned into 5% intervals along the gene length. The y axis represents the number of probes on the array within a bin. The arrays exhibit a 3' bias in gene representation.

Fig. 2. Mutant line d1 contains a ~500 kb deletion on chr 5 encompassing the RGA1 gene. a) Gene models in the region show a high percentage of probes with log₂ (mutant probe intensity/wild type probe intensity) = -0.8, indicating a large deletion (see analysis below). b) PCR confirmation of the deletion of RGA1 (Os05g26890) relative to wild type (indicated by an open arrowhead in part a) and PCR confirmation of the right border of the deletion (Os05g26990) relative to wild type (indicated by a closed arrowhead in part a). The left border was not resolved.



Analysis used for deletion discovery allows flexibility, depending on the end-user's tolerance for false positives or negatives. The specific parameters (log ratio and proportion) were selected using Table 1 as a guide.

- The log ratio $[\log_2(\text{mutant probe signal intensity}/\text{wild type probe signal intensity})]$ for each probe is first determined and log ratios for flagging probes are selected at less than -0.6 or -0.8.
- Probe sets that have more than a defined proportion of probes (0.4-0.5), i.e., those with a log ratio < -0.6 or -0.8 , are called as potential gene model deletions.
- Using different combinations of log ratio ($\log_2(\text{PM}_{\text{mutant}}/\text{PM}_{\text{wild type}})$) gives varying degrees of sensitivity and accuracy.
- These values are adjustable according to the researcher's desire for eliminating false positives or detecting more potential deletions.

Table 1. True and false positive rates (TPR and FPR, respectively) for different log ratio ($\log_2(\text{mutant PM probe intensity}/\text{wild type PM probe intensity})$) and proportion (probes flagged/total probes in probe set) combinations.^a

Log ratio	Proportion	TPR ^b	FPR1 ^c	FPR2 ^d
-0.6	0.4	0.833	0.051	0.0015
-0.6	0.5	0.800	0.020	<0.0001
-0.6	0.6	0.767	0.020	0
-0.6	0.7	0.600	0.020	0
-0.8	0.3	0.833	0.020	0.001
-0.8	0.4	0.833	0.020	0.0002
-0.8	0.5	0.767	0.010	<0.0001
-0.8	0.6	0.633	0.010	0
-1	0.3	0.833	0.010	0.0002
-1	0.4	0.800	0.010	<0.0001
-1	0.5	0.667	0.010	0
-1	0.6	0.600	0.000	0

^aAnalysis based on PCR confirmation 30 deletions and 10 non-deletions.
^bTPR was calculated as the proportion of PCR-confirmed deletions that were correctly called by the analysis.
^cFPR1 is the proportion of PCR-confirmed non-deletions, that are correctly called deleted by the analysis.
^dFPR2 is the proportion of probe sets meeting defined log ratio and proportion combinations for $\log_2(\text{WT}/\text{WT})$ and $\log_2(\text{WT}/\text{WT})$.

Labeled genomic DNA hybridized directly to the Affymetrix Rice GeneChip® allows rapid localization of deleted regions in rice mutants.

- The arrays exhibit a 3' bias in gene representation (Fig. 1). This bias means that deletions in the 5' regions will be underrepresented in the dataset.
- A known deletion containing the RGA1 gene on chr 5 in rice mutant d1 was detected by hybridization of genomic DNA onto the arrays (Fig. 2).



Fig. 3. Alleles of *sp1* show lesion mimic phenotype.

- The characteristic irregular orange lesions are shown in leaves of 4 *sp1* alleles used in this study.
 - This phenotype was first described in 1970 by S. Kiyosawa.
- Fig. 4. Array-based deletion discovery identifies allelic relationships among *sp1* mutants.**
- Hybridization of genomic DNA from two confirmed allelic *sp1* mutants (G650 and F1856) and two mutants showing the *sp1* lesion mimic phenotype (G9799 and F2045) identified overlapping deletions in all four lines on chr. 12.
 - A log ratio cutoff of < -0.8 for 50% or more of probes in a probe set was used.
 - Open arrowheads indicate deletions in gene models confirmed by PCR.
 - Closed arrowheads indicate gene models confirmed to be present by PCR.
 - Overlapping mutations allow resolution of region responsible for *sp1* phenotype down to about 70 kb, or ~20 candidate gene models.

Fig. 3. Alleles of *sp1* show lesion mimic phenotype.

- The characteristic irregular orange lesions are shown in leaves of 4 *sp1* alleles used in this study.
- This phenotype was first described in 1970 by S. Kiyosawa.

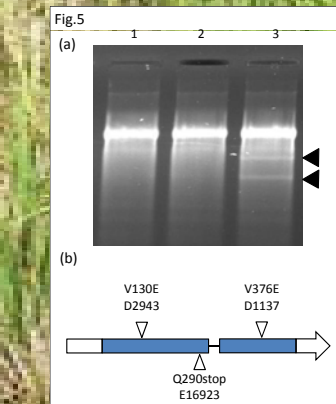
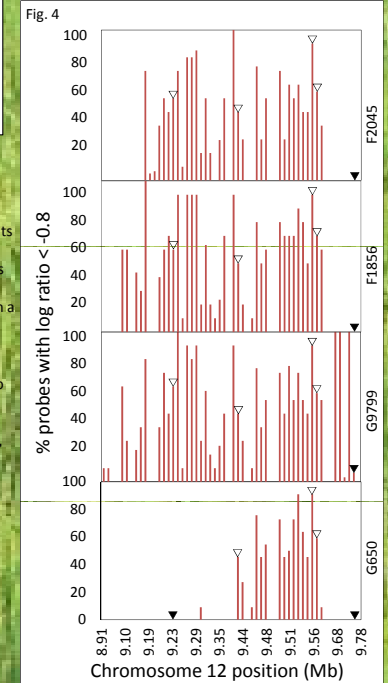


Fig. 5. A cytochrome P450 family member is a candidate for *Sp1*. Candidate genes located in the *Sp1* region by array hybridization (Fig. 4) were screened for SNPs in an EMS-generated mutant showing the *sp1* phenotype by TILLING.

Fig. 5. A cytochrome P450 family member is a candidate for *Sp1*. Candidate genes located in the *Sp1* region by array hybridization (Fig. 4) were screened for SNPs in an EMS-generated mutant showing the *sp1* phenotype by TILLING.

- (a) Detection of heteroduplex by TILLING between DNA for the rice mutant E16923 and wild type parent IR64. PCR products specific for LOC_Os12g16720 (a cytochrome P450 family member). Lanes 1 and 2 are CEL1 treatments of IR64 and E16923 amplicons, respectively. Lane 3 shows the activity of CEL1 enzyme on a heteroduplex generated between IR64 and E16923 amplicons.
- (b) The cytochrome P450 family member in *sp1* mutants contain SNPs. E16923 contains a SNP at position 290 that resulted in a stop codon. Two diploxybutane-generated *sp1* mutants, D1137 and D294B, show SNPs in LOC_Os12g16720 that caused amino acid changes.

Fig. 7. User-friendly rice genome browser developed to allow researchers to search for deletions in a gene or region of interest. In addition to deletions in individual gene models, an aggregation analysis was applied to the data, indicating larger deletions spanning many gene models.

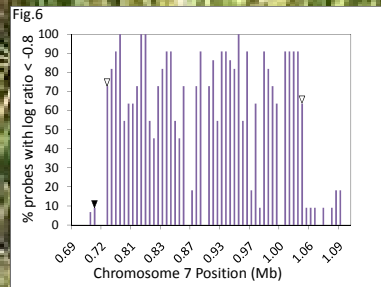


Fig. 6. Confirmation of a ~300 kb deletion on chromosome 7 in mutant line G282 as predicted by array hybridization using log ratio cutoff of < -0.8 for 50% or more of probes in a probe set. Open arrowheads indicate deletions in gene models confirmed by PCR. The closed arrowhead indicates a gene model confirmed to be present by PCR.

Fig. 6. Confirmation of a ~300 kb deletion on chromosome 7 in mutant line G282 as predicted by array hybridization using log ratio cutoff of < -0.8 for 50% or more of probes in a probe set. Open arrowheads indicate deletions in gene models confirmed by PCR. The closed arrowhead indicates a gene model confirmed to be present by PCR.

Conclusions

Hybridization of labeled genomic DNA directly onto the Affymetrix Rice GeneChip® allows rapid localization of deleted regions in rice mutants.

- Deletions are reliably predicted without prior knowledge (an unknown large deletion was detected in mutant G282 on chr 7 (Fig. 6)).
- Deletions ranged in size from one gene model to ~500 kb and were predicted on all 12 rice chromosomes.

The utility of the technique as a tool in forward genetics was demonstrated in combination with an allelic series of mutants to rapidly narrow the genomic region, and eventually identify a candidate gene responsible for a lesion mimic phenotype.

- A user-friendly rice genome browser with the location of deletions in 14 mutants aligned to the rice pseudomolecules allows researchers to identify deletions in regions of interest.

As the rice research community continues to perform hybridizations in the future, this resource will grow more valuable, moving towards describing a saturated deletion collection.



Travel funding provided by:



Research funding provided by:

