Allele Mining and its Significance in Crop Improvement

(*Munnesh Kumar and Kana Ram Kumawat)
Rajasthan State Seed and Organic Certification Agency, Jaipur, Rajasthan-302005
*muneshbhu94@gmail.com

Bioprospecting or Allele mining is a promising approach to dissect naturally occurring allelic variation at candidate genes controlling key agronomic traits which has potential applications in crop improvement programmes. It can be effectively used for discovery of superior alleles, through 'mining' the target trait from available germplasm. It can also provide the nucleotide variation associated with superior allele in accession with trait of interest. Allele mining may also pave way for introgression of novel alleles through Marker Assisted Selection. Therefore there is need to exploit the germplasm resources to further enhance the genetic potential of crop varieties for economically important traits through bioprospecting strategy.

Important Ex.
- *Sh4* for grain shattering in rice (Li and Sang, 2006)
- *Rc7* for grain pericarp color (Sweeney et al., 2007)
- *Wx* for granule-bound starch synthase (GBSS) (Wang et al., 1995)
- *GS3* grain size/shape (Fan et al., 2006)
- Bacterial leaf blight resistance gene *Xa21* from *Oryza longistaminata* (Khush et al., 1991)
- Blast resistance genes, *Pi9* from *Oryza minuta*, identified as superior allele (Sitch et al., 1989; Amante-Bordeos et al., 1992)
- *Pi40* from *Oryza australiensis* (Jeung et al., 2007)
- Grain filling allele *GIF1* (Wang et al., 2008a)
- Allele-mining of rice blast resistance genes at AC134922 locus (Wang et al., 2014)
- Allele mining for resistance to pepper veinal mottle virus. (Rubio et al., 2009).
- Improvement of marker-based predictability of Apparent Amylose Content in japonica rice through GBSSI allele mining (Biselli et al., 2014)

True allele mining
Earlier studies were focusing only on the single nucleotide polymorphism or insertion deletion occurs in exonic region since it is coding part. But recent studies shown that intron part will play a key role in regulatory mechanism. Hence any variation which in intronic region will be considered as true allele mining. Targeting a promoter region for harboring new allelic variant termed as promoter mining, which is becoming a future research area since till now intronic part we used to assume it is a junk DNA. But now we have been explored to its regulatory function.

Important example
1. Tubulin gene (Fiume et al., 2004)
2. Polyubiquitin gene, *rubi3* (Samadder et al., 2008) in rice
3. *VRN*-1 (affect vernalization response) in barley and wheat (Fu et al., 2005)

4. Mutation in 5′ splice site of the first intron of the *waxy* (*Wx*) gene had resulted in tenfold increase in the gene activity in rice (Isshiki et al., 1998)

Construction of core and mini-core collections

Development of a reduced representative set of a large base collection enable its enhanced use in crop improvement. Increased use of available genetic resources is a prerequisite for diversifying cultivar genetic base, enhancing current level of productivity and for continued genetic improvement to cater to the needs of changing consumer and end-user preferences and address biotic and a-biotic stresses.

Steps involved in allele mining procedure

1. Identification of accession with trait of interest
2. Selection of genes associated with trait of interest
3. Primer amplification of the target gene
4. Finding allelic variation of target gene
5. Primer design for target gene

Approaches for allele mining

(1) Modified TILLING (Targeting Induced Local Lesions in Genomes) procedure called Eco-Tilling:

TILLING uses mutagenized population whereas Eco-TILLING uses natural available germplasm to search the new allele, rest of the procedure remain same. TILLING is a reverse genetic tool from gene sequence to phenotype which have replaced the forward genetic tool from mutation through phenotype to the gene (Kurowska et al., 2011)

1. In TILLING seeds are mutagenized with chemical or physical mutagens mostly Ethyl methane sulfonate that causes G/C to A/T changes to produce M_1 plants.
2. M_1 plants are self to produce the M_2 from which DNA is extracted for analysis. The M_2 germplasm will allow recessive and lethal alleles to be recovered as heterozygotes, then DNA is extracted from the mutant population and pooled together.
3. Amplification products when heated and cooled will form mismatched heteroduplexes between the wild type and mutant DNA
4. Mismatches will be detected by a mismatch endonuclease enzyme that CEL I (extracted from celary) and cleaved into two different products, which will be detected in the 700 and 800 dye channel through LICOR DNA Analyzer.
5. The total size of the cleaved fragments should equal the total length of the entire product. Once the cleaved fragments and their respective polymorphic site are identified, these individuals are sequenced to verify the induced mutation.

(2) Sequencing based allele mining:

Sequencing is a method that allows information of the arrangement using a precise sequence of nucleotides in a sample of DNA (Ashkani et al., 2015). In TILLING and Eco-TILLING, sophisticated procedure is used to detect novel allele compare to sequenced based allele mining. In sequenced based allele mining we already have information regarding the base sequences of target allele or else we can sequence the target trait. Afterword's we can screen the population and match with the reference sequences will lead to identification of single nucleotide polymorphism (SNP). Set of SNP refers to haplotype. This sequence based allele mining will help in crop improvement by identifying SNP and haplotype structure responsible for abiotic and biotic stress. This is a cost effective method to find out novel allelic variant of known candidate on agarose gel and its utility in gene mapping (Raghavan et al. 2007).

Applications of allele mining

1. Discovery of superior alleles, through 'mining' the target trait from diverse genetic resources.
2. Provide insight into molecular basis of novel trait variations and identify the nucleotide sequence changes associated with superior alleles.
3. Evolution of alleles
4. Allelic similarity/dissimilarity at a candidate gene and allelic synteny with other members of the family can also be studied.
5. Molecular discrimination among related species.
6. Development of allele-specific molecular markers.
7. Introgression of novel alleles through marker assisted selection or deployment through genetic engineering.

Challenges

There are number of accessions which are held in various gene banks. The aim of breeder is to find out novel allele from a wealth of genetic resources. So the challenge is to efficiently identify and exploit the useful variation for crop improvement. Here, some of the challenges described in order to understand them and finding out way to overcome from it.

❖ Identification of accession with target trait
❖ Development of core/mini core collections
❖ Precise phenotyping
❖ Handling genomic resources
❖ Higher sequencing costs
❖ Feasibility in using bioinformatics tools

Conclusion

Bio prospecting new allele which can be essential to determine and utilize novel alleles, hidden in genetic diversity. Which ultimately can be exploited in crop improvement by incorporating those novel allele into an elite variety. In future sequence based allele mining will become a popular strategy since now a days NGS platform becoming low cost. Thereby, searching a new allele for manifestation of crop genotype to improve yield will be done through exploitation of high throughput technique like allele mining.

References

1. Amante-Bordeos, A., Sitch, L.A., Nelson, R., Damacio, R.D., Oliva, N.P. and Aswidinnoor, H. (1992) Transfer of bacterial blight and blast resistance from the tetraploid wild rice *Oryza minuta* to cultivated rice *Oryza sativa*. Theor Appl Genet.84, 345–54.
2. Ashkani, S., Yusop, R.M., Shabanimofrad, M., Azadi, A., Ghasemzadeh, A., Azizi, P. and Latif, M.A. (2015) Allele Mining Strategies: Principles and Utilisation for Blast Resistance Genes in Rice (*Oryza sativa* L.). Curr. Issues Mol. Biol. 17, 57-74. DOI: http://dx.doi.org/10.21775/cimb.017.057
3. Biselli, C., Cavalluzzo, D., Perrini, R., Gianinetti, A., Bagnaresi, U.S., Orasen, G., Desiderio, F., Lupotto, E., Cattivelli, L. and Valè, G. (2014) Improvement of marker-based predictability of Apparent Amylose Content in japonica rice through GBSSI allele mining, Rice, 7:1.
4. Brown, A. H. D. (1989a) A case for core collections. In: Brown, A. H. D., Frankel O. H., Marshall, D.R. and Williams, J. T. (eds), The Use of Plant Genetic Resources. Cambridge: Cambridge University Press, pp. 136–156.
5. DeVicente, M.C. and Tanksley S.D. (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics, 134:585–96.

6. Fan, C., Xing, Y., Mao, H., Lu, T., Han, B. and Xu, C. (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet, 112:1164–71.
7. Fiume, E., Christou, P., Gianı, S. and Breviario, D. (2004) Introns are key regulatory elements of rice tubulin expression. Planta, 218:693–703.
8. Frankel, O. H. (1984) Genetic prospective of germplasm conservation. In: Arber, W., Limensee, K., Peacock, W.J., Starlinger, P., (eds.), Genetic Manipulation: Impact on man and society. Cambridge university press, Cambridge, pp. 161-170.
9. Fu, D., Szucs, P., Yan, L., Helguera, M., Skinner, J.S. and von Zitzewitz, J. (2005) Large deletions within the first intron in VRN-1are associated with spring growth habit in barley and wheat. Mol Gen Genomics, 273:54–65.