

Title:

***Indica* Rice Genome Assembly and Annotation using second and third generation sequencing technologies**

Rice is a second most important cereal crop in the world and half of global population depends on rice for daily food. India is the center of origin and diversity of rice. More than 80% of the area under rice cultivation in India belongs to *indica* type. However, much of genomic resources are available for *japonica* rice and not enough genomic information is available for *indica* rice cultivars. This study was aimed to develop genomic resources for *indica* rice by whole genome sequencing of *indica* cultivar HR-12 by next generation sequencing technologies. We are the first to sequence *indica* rice genome using both second (Illumina) and third (PacBio) generation sequencing technologies. The use of second and third generation technologies has significantly improved genome size coverage and gene annotation for *indica* rice.

Considering the global impact of rice blast disease, we have also sequenced three *indica* rice cultivars; Co-39 (susceptible to blast disease), Tetep and Tadukan (resistant to blast disease), using illumina-sequencing technology. Most of R genes were either fragmented or mutated in rice blast susceptible varieties (HR-12 and Co-39). The genome sequences generated from our study are valuable resource for screening of novel R genes and functional genomics studies in rice. A comparative analysis between *indica* and *japonica* subspecies genomes has revealed a large number of *indica* specific variants including SSRs, SNPs and InDels. This information will be highly useful for resistance breeding programme and diversity studies of *indica* rice germplasm.