IOP Conference Series: Earth and Environmental Science

Mass genome sequencing of crops and wild relatives to accelerate crop breeding: the digital rice genebank

To cite this article: K L McNally *et al* 2020 *IOP Conf. Ser.: Earth Environ. Sci.* **482** 012005

IOP Conf. Series: Earth and Environmental Science **482** (2020) 012005 doi:10.1088/1755-1315/482/1/012005

Mass genome sequencing of crops and wild relatives to accelerate crop breeding: the digital rice genebank

K L McNally[1]*, R P Mauleon[1], D Chebotarov[1], S P Klassen[1], A Kohli[1], G Ye[1,2], H Leung[1], R S Hamilton[1] and R A Wing[1,3]

[1] International Rice Research Institute (IRRI), Pili Drive, Los Baños, Laguna 4031, Philippines
[2] IRRI-CAAS Joint Laboratory, Agricultural Genomics Institute of Shenzhen, Chinese Academy of Agricultural Sciences (CAAS), Shenzhen 518120, China
[3] Arizona Genomics Institute, University of Arizona, 1657 E Helen St, Tucson, AZ 85705, USA

*E-mail: k.mcnally@irri.org

Abstract. The advent of next generation sequencing, and more recently third generation sequencing, has enabled researchers to begin interrogating the genomic information of thousands of accessions of conserved genetic resources as well as cultivated varieties. Here, we describe recent results from the 3,000 Rice Genomes Project (3K RGP) and summarize a few of the projects for other crop species. The 3K RGP has served as the catalyst to create a digital rice genebank with the intent to sequence many more of the conserved accessions held in the International Rice Genebank Collection in the coming years. We are progressing with generating high quality reference builds for the 15 subpopulations defined by the 3K RG analyses. These along with reference builds of the wild relatives will allow access to the unique genomic regions specific to particular cultivated types and wild relatives. We have also begun efforts to sequence a further 10,000 accessions of rice in collaboration with partners in China. Yet, in-depth sequence data and initial comparative bioinformatic analyses are not enough to promote efficient use. Hence, high throughput phenomic screening in multiple environments, development of novel genetic populations, computational genetics and modeling will be necessary to understand the link between genotype to phenotype and its environmental control. The combination of these approaches underpinned by computational analyses will allow identification of novel genes and alleles that can be deployed into elite varieties for sustainable crop improvement.

Keywords: genetic resources, high throughput phenomics, next generation sequencing, rice.