

RiceDBreeder: Integration and utilization of gwas and Korean rice phenotype, pedigree and resequencing data

Dong U Woo¹, and Yang Jae Kang^{1,2,*}

¹*Division of Bio & Medical Bigdata Department (BK4 Program), Gyeongsang National University,
Jinju 52828, Republic of Korea*

²*Division of Life Science Department, Gyeongsang National University, Jinju 52828, Republic of
Korea*

*Corresponding author: kangyangjae@gnu.ac.kr

Modern breeding has evolved from traditional phenotype-based methods to molecular breeding using biotechnology, and now to digital breeding, which leverages artificial intelligence and big data. However, in the case of Korean rice resources, phenotype and genomic data are managed separately, making integrated analysis challenging and limiting data accessibility and usability. Phenotype data is recorded individually, complicating integrated analysis, while next-generation sequencing (NGS) data is fragmented, making it difficult to find resequencing information for specific varieties. Moreover, pedigree information is often fragmented or provided in document format, making it inefficient to search for crossbreeding combinations.

To address these issues and promote the practical application of digital breeding, we developed RiceDBreeder. This web application integrates resequencing data, phenotype data, and pedigree information for 8,831 Korean rice varieties, along with genome-wide association study (GWAS)-based trait-linked marker data, allowing researchers to easily explore and analyze this information. Through RiceDBreeder, researchers can visually examine the status of phenotype data, the position of varieties in the pedigree, and compare markers and traits efficiently.

RiceDBreeder enables users to quickly identify phenotypic traits linked to the selected varieties and the corresponding single nucleotide polymorphism (SNP) alleles. Additionally, it allows genome-wide comparison of related varieties through pedigree analysis. These features empower breeders to efficiently and quickly select the most suitable varieties for their breeding programs.