

Title: Study on the genetic diversity of soybean roots under field conditions

(圃場におけるダイズの根の遺伝的多様性に関する研究)

BUI The Khuynh

Thesis Summary

Root plays a central role in water and nutrient uptake of plants, however, to date, our understanding of the genetic control of the root system is still limited which hindered the success of breeding programs for root improvement in crops, including soybean, the most important legume crop. Our study was conducted to investigate the genetic diversity of root traits in a 200-accession panel of soybean, mostly consisting of the Japan and World mi-core collections under field conditions across three years (2017, 2019 and 2020), employing two irrigation treatments: irrigated and non-irrigated treatments. The condition in the field of Arid Land Research Center, Tottori University, which is characterized by the high sand content in the soil, allowed us to successfully collect the large-scale root phenome data with minimum root loss. The results from our study (in Chapter 2) revealed a high diversity in root traits across the whole panel under both irrigated and non-irrigated conditions. And of note, higher genetic diversity in most root traits was seen among genotypes from the world mini-core collection, compared with accessions from Japan mini-core collections. With the high diversity in the whole panel observed, I

have successfully identified promising soybean genotypes which showed large root trait improvement under irrigation compared with non-irrigated conditions. In addition, genotypes with high root performance and high stability across years were also found.

In chapter 3, I conducted a genome-wide association study (GWAS) to dissect the genetic control of root traits and then, identified genes associated with root-related traits under both irrigated and non-irrigated conditions. The large genetic diversity and the high heritability of root traits in our study have enabled us to detect 7 QTLs associated with root traits. These results were very promising as I further detected the candidate genes located in these QTLs. Interestingly, among candidates that I have identified, *Glyma.07G084300*, associated with total root length (TRL) under irrigated conditions, known as the *Arabidopsis* *CDK-activating kinase 1AT*, was reported for the roles in cell proliferation, cell expansion, and the regulation of root cell. Haplotype analysis also revealed the significant difference in TRL, as genotypes carrying the alternative allele compared with the reference had a significantly higher TRL.

In an attempt to predict root-related traits in soybean under field conditions (Chapter 4), I have used two approaches including (i) genomic prediction, and (ii) the prediction of root traits from shoot environment data using machine learning (ML). By fitting BLUP (the best linear unbiased prediction method) into 100

hundred thousand of SNPs selected across the whole genome of soybean, I have successfully predicted root traits of soybean across environments and the predictive ability in our study was higher compared with of previously reported. As I noticed the variation in prediction ability among different traits and different environments, therefore, in order to optimize the prediction ability of root traits across different environments, I aimed to incorporate the information of genotype and environment interaction into prediction models. As a result, the use of MxE and Reaction Norm models helped to improve the prediction ability of root traits across different environments compared with the single environment. The prediction of root traits from the available data (plant height, shoot fresh weight, and concentrations of 19 mineral elements in leaf) observed in the shoot environment using random forest (RF) outperformed other ML models in our study. Interestingly, the addition of leaf ionome data into RF model helped to improve the prediction ability of root traits in our study. Especially, the use of selected mineral elements with the highest variable importance reduced the error of prediction models when the validation data was found in different environments compared with the data used for training the model, suggesting that these elements can carry information about environmental changes and hence, can be seen as potential variables for prediction. The promising results obtained in root traits prediction in our study proved that it's feasible to estimate the underground performance of open field crops, especially in field-grown soybean

from the genomic data (SNP dataset) and the easily assessable data observed in the field.