Annual Project Report

Project title: Towards improved crop photosynthesis efficiency: elucidation and validation of genes underlying Arabidopsis photosynthesis QTLs

Post-doc: Dr Thu-Phuong Nguyen

Anticipated results 2019

Numerous data have been collected for photosystem II efficiency in Arabidopsis under various suboptimal growth conditions, such as low nitrate nutrition, low phosphate nutrition, cold temperature and low-high irradiance. We anticipated that in-depth analyses of the existing data would provide a good overview of the genetic landscape (QTLs) underlying PSII efficiency variation. The outcome would indicate whether there is a common genetic mechanism and (or) specific ones that contribute to observed variation in different stress conditions. These analyses allows us to select a small number of interesting QTLs, those are strongly significant and (or) with common effect, to further characterized with mutants.

Results anticipated to be obtained by the end of the year, would be the selection of QTLs to further characterize, and a T-DNA mutant analysis for candidate genes underneath those selected QTLs.

Achieved results 2019

+ Photosystem II efficiency was analyzed systematically for 4 data sets of comparable experiments
+ GWAS using univariate model for single timepoints: 239 in total
(62 for nitrate nutrition, 102 for phosphate nutrition, 57 for cold temperature and 18 for low-high irradiance)
+ GWAS using multivariate model for time series within experiments: 46 in total
(13 for nitrate nutrition, 18 for phosphate nutrition, 18 for cold temperature and 7 for low-high irradiance)
+ GWAS using multivariate model across 3 experiments (low nitrate, low phosphate nutrition and cold temperature) for data points within a plant age of 15 to 20 day): 18
+ These GWAS reveal almost a hundred loci involved in photosynthesis above the significant LOD of 5.6, calculated based on the predicted linkage disequilibrium (LD) blocks in Arabidopsis genome. Most interesting QTLs that are worth following within this project timeline: two closely linked QTLs are responsive to multiple stresses, and one highly significant QTL for Nitrate nutrition deficiency
+ Further investigation into 2 QTLs: with T-DNA mutants ordered for genes underneath the QTLs, but are not yet analyzed
+ Project was on hold for 4 months due to maternity leave

Anticipated results 2020

+ To further characterize the two or three most interesting QTLs with the following steps:
1. LD and Haplotype block analysis of the QTL interval: this helps in narrowing down the selection of candidate genes underlying the QTL
2. In silico analysis of candidate genes to prioritize candidates
- check sequence of gene in other natural accessions (with resequencing data): to check if the gene structure in Col reference is correct and see if there interesting sequence/motif in the candidate genes.
- check the expression in database: which tissues, which conditions, co-expression
- Gene ontology: eg. gene family, gene response to, related to photosynthesis or not
- Predicting mode of action of the gene: level of expression or protein structure
3. T-DNA mutant analysis of the candidate genes:
- Genotyping: primer design, DNA extraction, PCR, gel electrophoresis
- Phenotyping of homozygous T-DNA lines
- Expression analysis: RNA isolation, qPCR
4. Allelic complementation:

- Crossing 2 extreme accessions with contrasting phenotype associated with the SNP of the involved gene with both a T-DNA knock-out mutant (Col background) of the same gene, and with Col. Comparing the phenotype of F1s of the crosses with WT and mutant, in case of the causal gene being affected, this should lead to a phenotypic difference

5. Check expression of candidate genes in extreme accessions which carry polymorphic SNPs under stress condition. It can be extended to an RNAseq experiment? Growing a few extreme accessions under most distinguishing suboptimal condition, e.g Nitrate deficiency and expression profiling of these accessions by RNAseq. This experiment will be useful to generate expression library, that will further support the selection of candidate genes, not only for this current work, but also for future work with other loci.
6. If there are no T-DNA available for promising candidate gene: CRISP/Cas mutant generation: will be a lot of cloning involved
7. Cloning cDNA of targeted gene to generate overexpression vector, or gDNA for complementation
8. Check if there is an available biparental mapping populations made from the accessions of interest to complementarily identify the causal gene.
9. There is also a Swedish GWAS population, for which photosystem II efficiency has been measured, and the genotype info is available, which can be used for GWAS analysis. This set of data introduces additional phenotypic and genetic variation that can reveal more loci and possibility to refine the QTL of interest.

+ Write a manuscript based on GWAS analysis of data: "Genetic landscape of photosystem II efficiency variation in Arabidopsis"

Products 2019 *(papers, reports, non-scientific papers, posters/presentations in conferences/workshops etc., media-presentations, other)*

+ Annual Meeting Experimental Plant Sciences, Lunteren 2019: Poster presentation
+ Keystone Symposia Climate Change - Linked Stress Tolerance in Plants, Hannover 2019: Poster presentation