
Recovering the complexity in scientific knowledge: towards automatically reconstructing the maize floral transition gene regulatory network

Adrienne Ressayre^{*1}, Fabien Chardon², Nathalia Conde-Silva¹, Catherine Damerval¹,
Yves Deveaux¹, Arnaud Desbiez-Prat¹, Christine Dillmann¹, Elodie Marchadier¹,
Dominique De Vienne¹, Robert Bossy³, Estelle Chaix³, and Claire Nédellec³

¹Génétique Quantitative et Evolution - Le Moulon (Génétique Végétale) (GQE-Le Moulon) – Institut National de la Recherche Agronomique : UMR0320, Université Paris-Sud - Paris 11, AgroParisTech, Centre National de la Recherche Scientifique : UMR8120 – UMR de Génétique Quantitative et Evolution - Le Moulon, Ferme du Moulon F-91190 GIF SUR YVETTE, France

²Institut Jean-Pierre Bourgin (IJPB) – Institut National de la Recherche Agronomique : UMR1318, AgroParisTech – INRA Centre de Versailles-Grignon, Route de St-Cyr (RD10), 78026 Versailles Cedex, France

³Mathématiques et Informatique Appliquée du Génome à l'environnement (MaIAGE) – Institut national de la recherche agronomique (INRA) – INRA Unité MaIAGE. Bât. 233 et 210. Domaine de Vilvert, 78352 JOUY-EN-JOSAS CEDEX, France

Abstract

In annual plants, flowering time is a highly integrated life-history trait which summarizes growth and development. We used plants issued from two independent divergent selection experiments on flowering time in two maize inbred lines. One of the trait that was targeted by selection was the timing of floral transition, that occurs early during maize life-cycle. RNA-Seq data from shoot apical meristems of Early and Late genotypes sampled before, nearby and after floral transition revealed thousands of genes that changed their expression throughout time. The inventory of these genes was different between ancestral genetic backgrounds, but also between Early- and Late-flowering genotypes derived from the same ancestral inbred. The number of differentially expressed genes was much higher than expected from the level of genetic polymorphism between Early and Late populations, suggesting the existence of hubs in the maize floral transition Gene Regulatory Network (GRN).

The description of the GRN cannot be achieved through statistical analyses because of lack of power in transcriptome data, nor by functional analyses of gene ontologies, that do not account for regulations and interactions. Here we present a text-mining approach that aims not only at automatically collecting data from the scientific literature, but also at organizing them in terms of causal relationships. This approach relies on the design of a model of knowledge that enumerates possible entities like genes, metabolites, phenotypes and environmental factors that could belong to or alter the GRN, and describes their possible interactions. In collaboration with computer scientists, we are adapting an existing knowledge model developed for *Arabidopsis thaliana* seed development to address the question of *Arabidopsis* floral transition, the GRN of which is well known. After having proved the concept in *Arabidopsis*, the next step will be to adapt the model to maize.

*Speaker