Improving Nitrogen Use Efficiency in Wheat by Genome Wide and Candidate Genes Targeted Association Studies

Agathe Mini, Gaetan Touzy, Catherine Ravel, Titouan Bonnot, Etienne Paux, Jacques Le Gouis, Stéphane Lafarge

The BreedWheat Consortium
BreedWheat: French research project on wheat

2011-2020

15 public research laboratories (INRA, Universities…)

1 technical institute (Arvalis)

10 Breeding companies

1 competitiveness cluster (Céréales Vallée)
Nitrogen Use Efficiency: capacity for a plant to use nitrogen inputs

- Improve NUE in wheat to better valorize fertilizer inputs
  - For ecological issue (leaching, fertilizer production…)
  - For economical issue: nitrogen input account for 25% of operational cost
- For wheat, NUE is around 65%
  - A target for breeding to have a higher NUE for a better yield and grain protein content

Genetic of wheat adaptation to biotic and abiotic stresses

Analyse genetic diversity and use genomic data to better understand wheat adaptation to biotic and abiotic stresses
Genetics of wheat adaptation to nitrogen stresses

Candidate genes
Markers associated with trait(s) linked to yield stability under stress (GWAS results)
Knowledge about French/European germplasm
New phenotypic traits/methods

Candidate genes from -omics data  GWAS
GWAS

Genotypic Data
420k Axiom array

Genomic areas of interest
Tools for breeding

Phenotypic Data

Associated markers
Search for new variabilities in diversity panel
Field validation
GWAS: Materials

• Panel Composition
  – 220 winter wheat elite varieties

• Genotypic data
  ⇒ 197K SNPs available for GWAS (polymorphic…)
  ⇒ With physical position on Wheat RefSeqV1

• Analysis of structuration, LD
  – Mean = 2.4 cM
  – K & Q matrix for GWAS analysis
GWAS: Phenotypic Data

- 26 trials / 3 years, dedicated to:
  - Nitrogen: 12 => Opt. & Stress conditions
  - Drought: 6
  - FHB: 4
  - Septoria: 4

- 26 trials / $1.10^6$ data points
  - Plant phenology
  - Yield, yield components, GPC…
  - Biomass, nitrogen content, senescence…
Environmental characterization
Environmental classification

• For each environment, determination of
  – Timing of stress
  – Length of stress
  – Stress intensity

• Using
  – Environmental data
  – Plant development informations
  – Phenotypic data

10 Field trials validated and kept for the global analysis
Genotypes characterization

- Stress tolerance index => knowledge on genotypes behaviour regarding Nitrogen stresses
  - Calculate for each field trials
  - Classification made by type of stress: mean of index tolerance
    - High
    - Medium
    - Low
Analysis from the raw data to GWAS results

- Workflow analysis
- R script
- Common way to analyze field trials
- Graphical and statistical analysis

Capacity to

- Analyze phenotypic field data in order to:
  - Determine heterogeneity, errors, outliers...
  - Genotype x Treatment interaction
  - Heritabilities
  - ...
- Calculate adjusted means corrected by spatial effects
- Produce GWAS results
GWAS: Numbers

<table>
<thead>
<tr>
<th>Traits Classification</th>
<th>Traits Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrogen</td>
<td>17</td>
</tr>
<tr>
<td>Biomass</td>
<td>5</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
</tr>
<tr>
<td>Height</td>
<td>1</td>
</tr>
<tr>
<td>Disease</td>
<td>11</td>
</tr>
<tr>
<td>Phenology</td>
<td>4</td>
</tr>
<tr>
<td>GPC</td>
<td>5</td>
</tr>
<tr>
<td>Yield and Components</td>
<td>11</td>
</tr>
<tr>
<td>Senescence</td>
<td>22</td>
</tr>
<tr>
<td>Lodging</td>
<td>10</td>
</tr>
<tr>
<td>Tolerance Index</td>
<td>9</td>
</tr>
</tbody>
</table>

10 trials

1 075 traits x treatments x trials

196 464 SNP
Usable for GWAS

211 198 800 tested associations

1st selection: 73 LD blocks selected on traits, robustness...

SNP associated in at least 2 trials and for the same traits (624 LD blocks)

91151

Associated SNPs with traits of interest

16193
GWAS: examples

- A robust areas detected in 10 trials

- Selection of genetic areas
- Definition of boundaries using LD decay
- Candidate genes in the area
  - Need to densify => Sequence capture on genes in the confidence interval based on RefSeqV1 prediction
Allelic effect

- Effect on traits of the minor allele in the panel

![Graphs showing Allelic effect on GPC, Yield, and GPD traits](image-url)
• Densification
  – To have better associated markers (causal SNPs)
  – SequenceCapture targeted to genes in the area => SNP discovery & association studies

• Identification of candidate gene
  – A priori => by annotation analysis
  – Without a priori => transcriptomic data
Transcriptomic Data:

- **NUE:**
  - Response of grain protein composition to N and S supply in *Triticum monococcum*

**StageCond** analyzed x 3 rep (54 samples):
- 2 stages (100, 200° Cd)
- 4 stages (300, 400, 500, 600° Cd) x 4 nutritions

- Paired-End sequencing, HiSeq 2500 – Illumina
- **Expression:** 22 443 genes expressed
Analysis
- GO: stress response, amino acids transport...
- Differential expression regarding:
  - Time points
  - Nutrition
- GeneNetwork:
  - RulNet: a tool to make connections
    - has been developed to infer regulatory networks from high throughput –omics data, phenotypic data...
    - Identification of master regulators linked to traits of interest...
  - In combination with proteomic/metabolomic data
    - Genes, proteins, metabolites responding similarly to the nutrition?
    - Genes, proteins, metabolites strongly co-expressed /co-accumulated?
    - Central genes in the response to the nutrition?
Transcriptomic Data:

- Analysis
  - 152 Candidate genes
    - Understanding of specific stress mechanism
    - SNP discovery / Association studies
    - Co-localization between transcriptomic candidate genes and GWAS results

OverExpressed during grain filling
Hub-Highly connected
ABC transporter
GWAS: Conclusions

• GWAS
  – A very interesting set of phenotypic data and GWAS results
  – Need to valorize all these data-set:
    • Meta-analysis, GxE, combination of positive alleles etc…

• ToDo
  – Densification to have access to better associated markers
  – Analysis of natural variabilities in a diversity panel from Core collection (450 acc.)
  – Validation in the field of the most promising regions
BreedWheat - Conclusions

• Candidate genes:
  – Gene lists, knowledge on molecular/physiology mechanisms of nitrogen stress response…

• Tools for breeding, transgenesis/genome editing etc…
  – Associated-markers
  – Genotypes classification
  – Identification of interesting candidate genes

For NUE, but also drought, FHB and Septoria…
Thanks to all people involved in these work

**Gaetan Touzy**, Katia Beauchêne, Jean Pierre Cohan

**Agathe Mini**, Guylaine Besnier-Hebert, Sébastien Praud

Sylvie Dutriez, Alain Lacassagne

**Jacques le Gouis**, Catherine Ravel, Titouan Bonnot, Etienne Paux, Emmanuel Heumez

All teams involved in research projects

Franck Lacoudre, Jayne Stragliati, Celine Duque, Jérémy Derory

Stephen Sunderwith, Benoît Foucault, Patrice Senellard

Olivier Lucas, Laure Duchalais, Christophe Michelet

Pascal Giraudeau, Sébastien Ducroquet

Sebastien Caiveau, Odile Argilier
Thank You