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

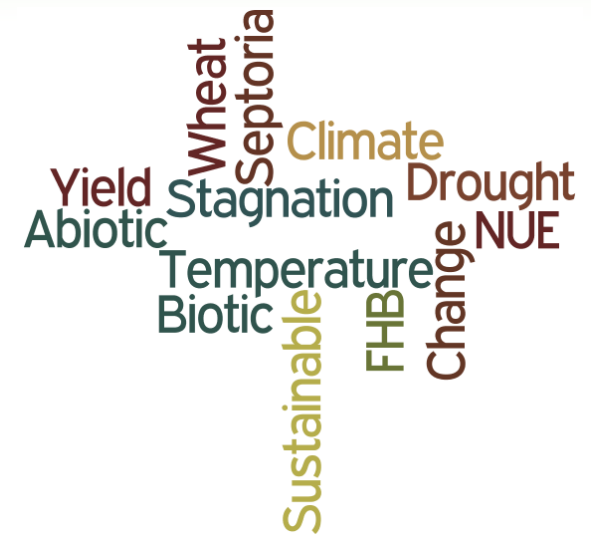**15 public research laboratories (INRA, Universities…)**

1 technical institute (Arvalis)

10 Breeding companies

1 competitiveness cluster (Céreales Vallée)

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

Yield Abiotic Wheat Septoria Stagnation Climate Drought Temperature NUE Biotic Sustainable FHB Change

- 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

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 germsplasm
New phenotypic traits/methods

Candidate genes from -omics data

GWAS

GWAS

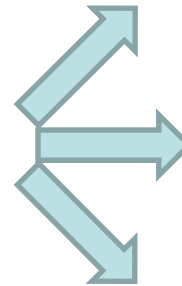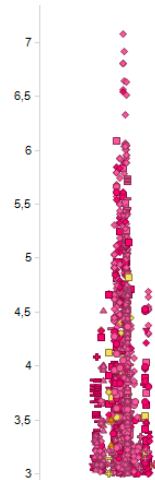**Genotypic Data**
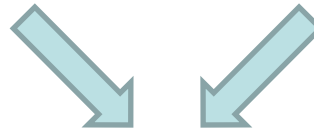420k Axiom array

Phenotypic Data

Genomic areas of interest

Tools for breeding

Associated markers

Search for new variabilities in diversity panel

Field validation

GWAS: Materials

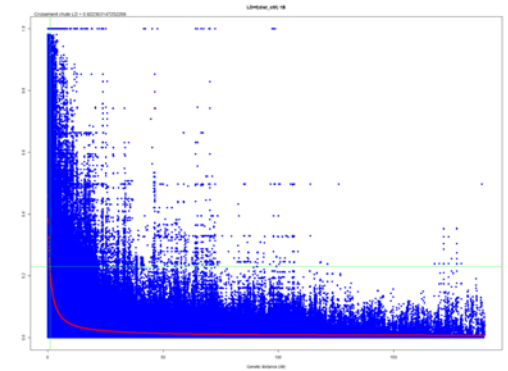- Panel Composition
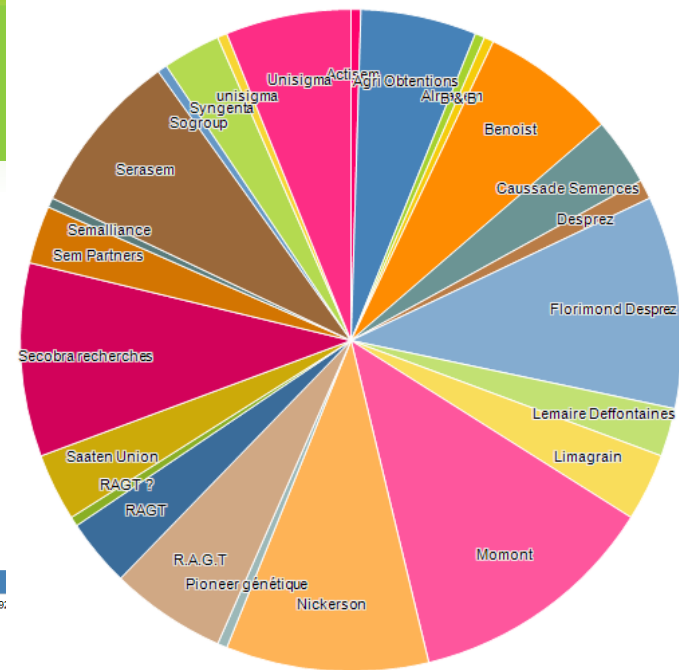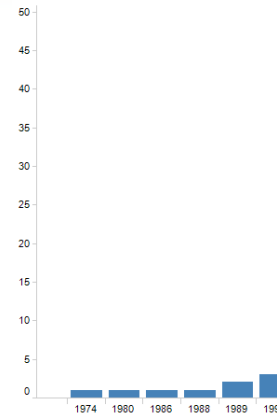 - 220 winter wheat elite varieties

- Genotypic data
  ⇒ **197K SNPs available for GWAS** (polymorphic…)
  ⇒ **With physical position on Wheat RefSeqV1**
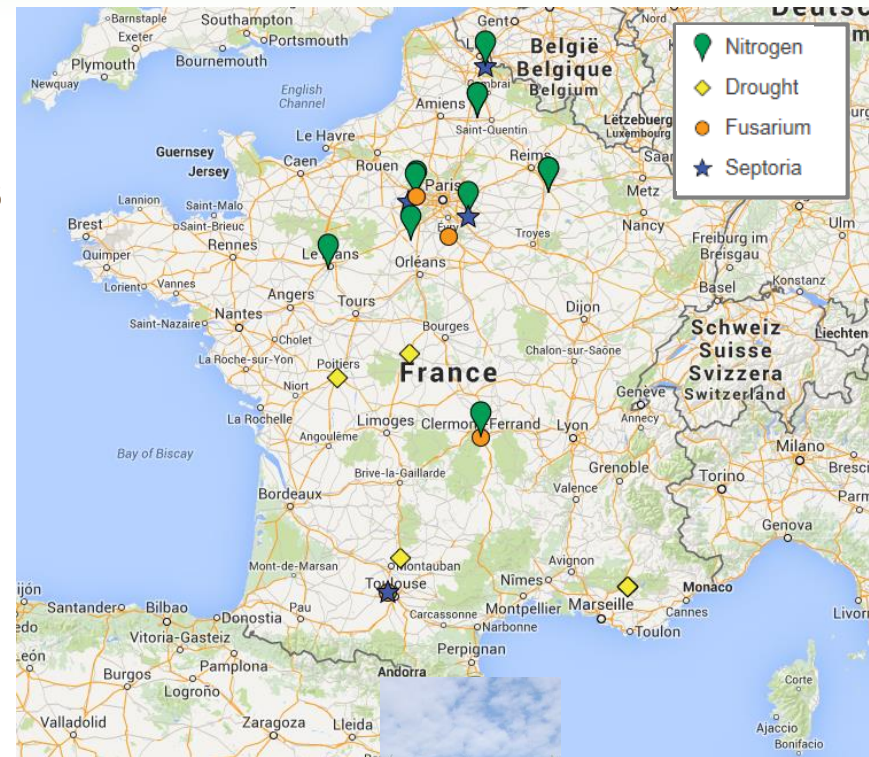
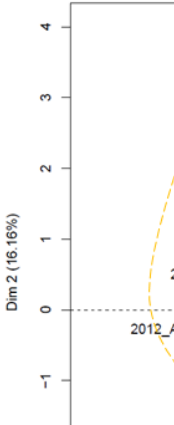- Analysis of structuration, LD
  - Mean = 2,4 cM
  - K & Q matrix for GWAS analysis

GWAS: Phenotypic Data

- 26 trials / 3 years, dedicated to:
  - <u>Nitrogen: 12</u> => 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…

- For each environment, determination of
  - Timing of stress
  - Lenght of stress
  - Stress intensity
- Using
  - Environmental data
  - Plant development informations
  - Phenotypic data

Individuals factor map (PCA)
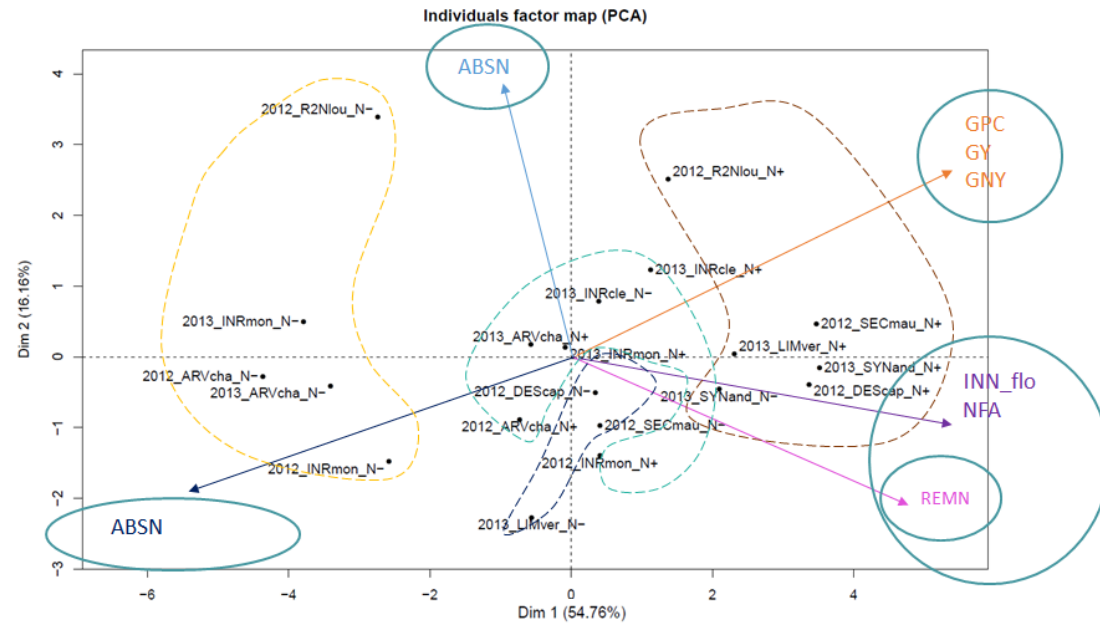
10 Field trials validated and kept for the global analysis

		Intensity		
		High	Medium	Low
Duration	High	4		
	Medium	2	1	
	Low			3

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

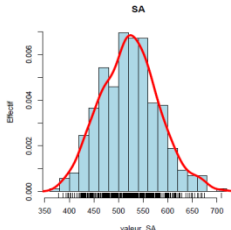| Génotype	GY.RN 2012_ARVcha	GY.RN 2012_DEScap	GY.RN 2012_INRmon	GY.RN 2012_R2Nbou	GY.RN 2012_STCmau	GY.RN 2013_ARVcha	GY.RN 2013_INRcle	GY.RN 2013_INRmon	GY.RN 2013_LIMVer	GY.RN 2013_SYNand	MOYENNE des résidus normalisés	Ecart type résidus normalisés	NOMBRE DE FOIS OU RN >0
PBW030	0.13	1.30	1.48	1.06	0.75	1.78	0.41	1.20	3.26	1.57	1.29	0.82	10
PBW184	0.78	1.29	0.34	-0.06	2.59	1.94	2.48	2.19	-0.11	1.40	1.29	0.96	8
PBW011	1.00	1.94	2.31	-0.94	0.91	1.82	0.89	0.10	1.14	1.85	1.10	0.92	9
PBW130	1.44	0.59	0.68	2.83	0.95	2.50	0.63	1.12	0.16	0.09	1.10	0.87	10
PBW012	2.50	1.81	1.43	1.61	0.09	1.43	-0.96	1.83	-0.54	0.85	1.01	1.07	8
PBW060	-1.30	1.57	1.23	1.32	0.50	1.30	-0.73	2.67	0.77	1.39	0.87	1.09	8
PBW142	1.18	1.82	0.31	1.22	1.27	1.09	-0.93	1.59	0.48	0.54	0.86	0.75	9
PBW015	1.86	1.28	0.95	-1.06	0.73	1.35	0.62	1.10	0.26	1.36	0.84	0.76	9
PBW129	1.52	1.18	2.85	1.74	0.90	-0.37	-0.22	0.14	1.72	-1.07	0.84	1.14	7
PBW125	2.04	2.31	0.44	-0.12	0.10	0.74	-0.06	0.75	0.68	1.23	0.81	0.79	8
PBW195	-0.10	1.88	0.42	-0.27	2.45	0.66	1.72	0.80	-0.32	0.26	0.75	0.91	7
PBW159	0.17	0.61	0.43	-0.09	0.80	0.03	0.47	2.13	2.02	0.58	0.71	0.73	9
PBW156	0.68	0.82	0.65	-1.20	0.50	2.64	0.88	-0.05	1.57	0.62	0.71	0.94	8
PBW114	0.85	-0.24	0.02	1.18	1.17	1.75	-1.29	1.17	1.47	1.01	0.71	0.88	8
PBW008	0.78	2.07	1.48	0.40	1.26	-0.80	0.64	0.73	-0.89	1.14	0.68	0.89	8
PBW033	0.38	0.40	2.02	0.50	1.25	-1.15	0.11	1.34	0.13	1.50	0.65	0.86	9
PBW160	0.07	0.52	1.01	-0.11	1.50	0.04	1.58	0.87	1.15	-0.17	0.65	0.63	8
PBW090	-0.65	0.45	0.78	1.78	0.86	0.34	-1.01	0.86	1.53	1.36	0.63	0.85	8
PBW128	0.84	-0.37	0.37	0.09	1.53	1.68	0.37	1.46	-0.46	0.54	0.61	0.73	8
PBW021	0.10	1.45	0.78	-0.05	0.15	1.29	0.34	-0.51	1.61	0.67	0.58	0.67	8

INRA Mons - 2012

$r = 0,75$
$R^2 = 57\%$

PBW065

NR+ → genotype perform better than the « overall mean »

PBW154

NR- → genotype perform worse than the « overall mean »

GY (kg/ha) N-

GY (kg/ha) N+
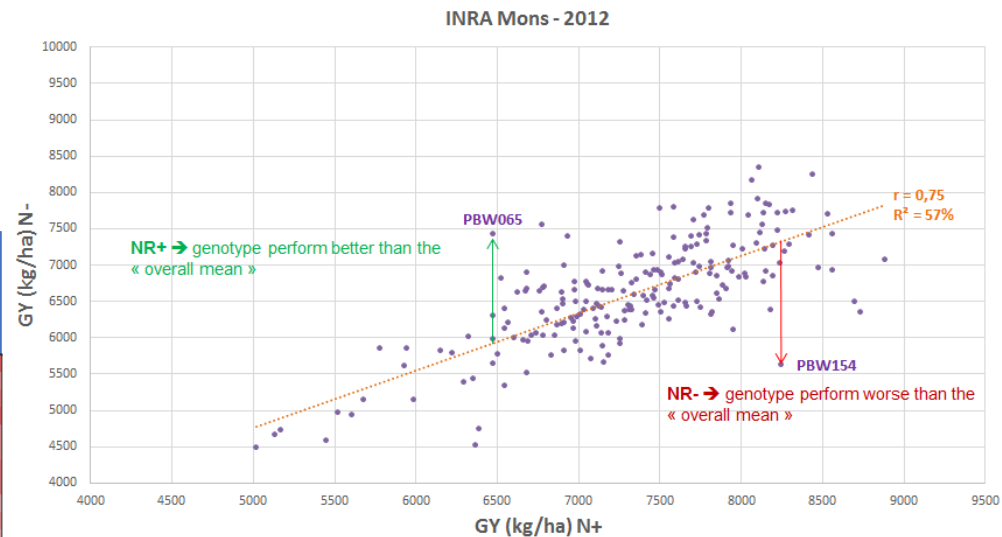
GWAS: a pipe line of analysis

- ## 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
    - Heritabilites
    - …
  - Calculate adjusted means corrected by spatial effects
  - Produce GWAS results
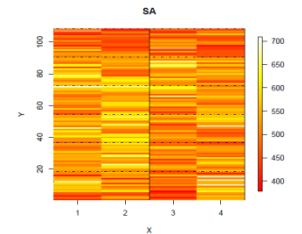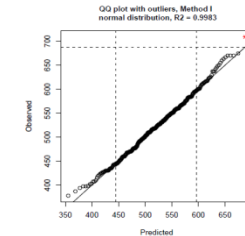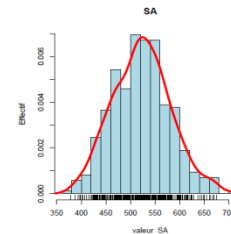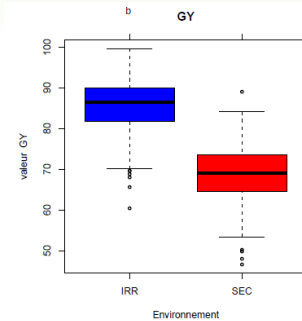
GWAS: Numbers

10 trials

Traits Classification	Traits Number
Nitrogen	17
Biomass	5
Others	3
Height	1
Disease	11
Phenology	4
GPC	5
Yield and Components	11
Senescence	22
Lodging	10
Tolerance Index	9

In HN & LN

181 traits

1 075 traits x treatments x trials

196 464 SNP Usable for GWAS

211 198 800 tested associations
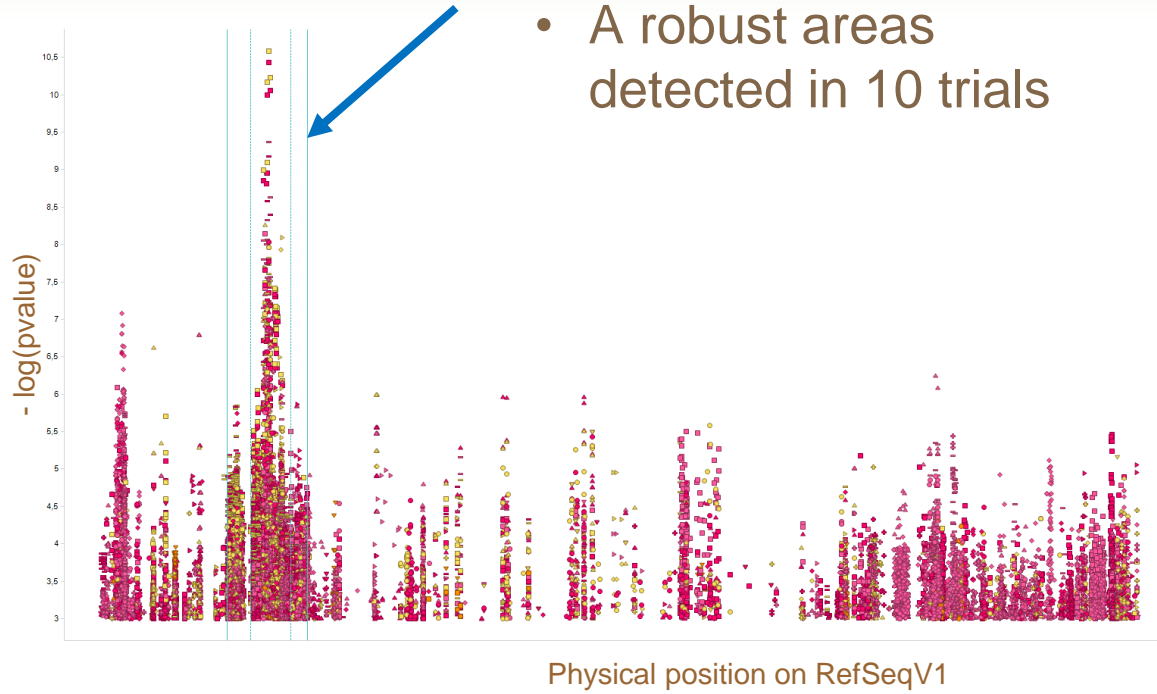
1st selection: 73 LD blocks selected on traits, robustness…

16193

91151

Associated SNPs with traits of interest

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

GWAS: examples

• A robust areas detected in 10 trials

- log(pvalue)

Physical position on RefSeqV1

Gogat
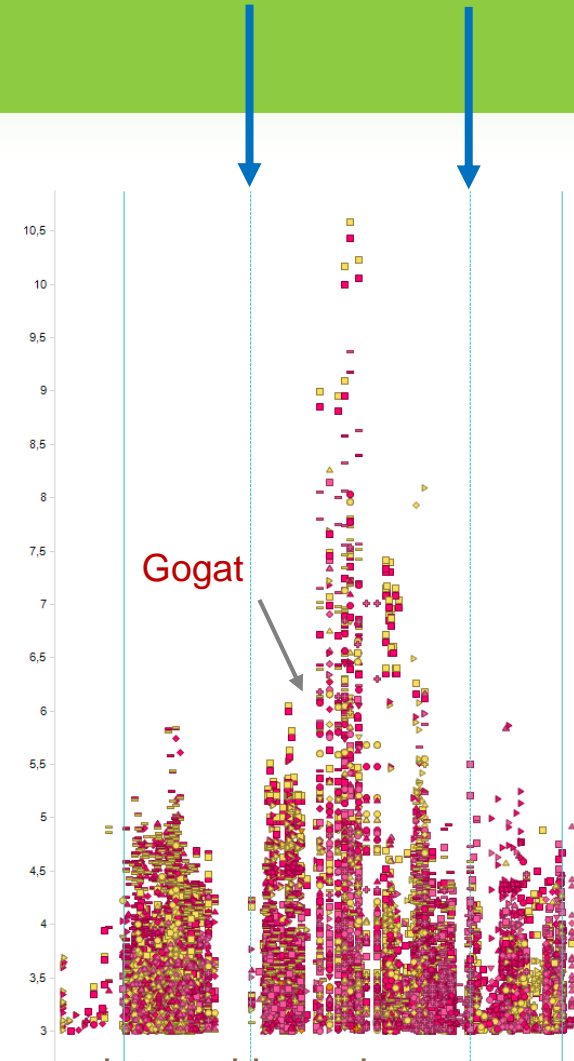
• Selection of genetic areas

• 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

Breed Wheat

Allelic effect

- Effect on traits of the minor allele in the panel

GPC   Yield   GPD

- log(pvalue)

Effect
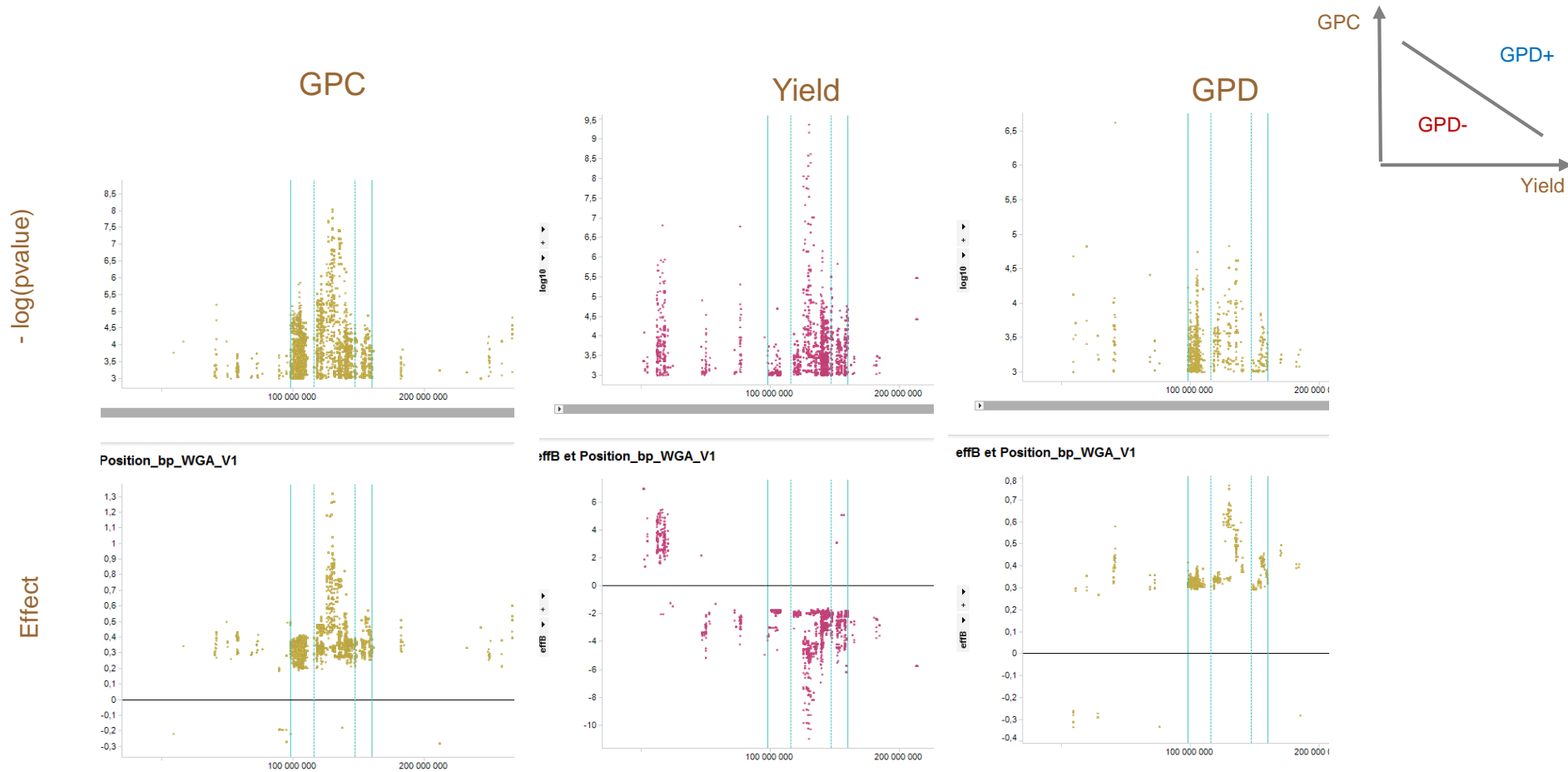
Next ?
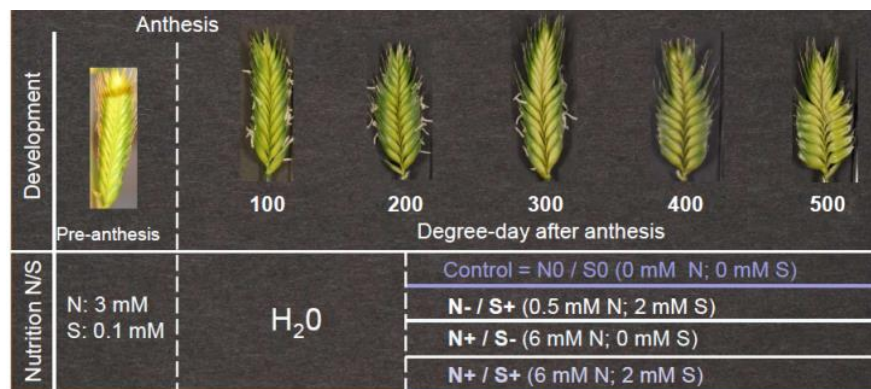
- Densification
  - To have better associated markers (causal SNPs)
  - SequenceCapture targeted to genes in the area => SNP discovery & association studies

- Identification of candidate gene
  - *A priori* => by annotation analysis
  - Without *a priori* => trancriptomic data

Transcriptic Data:

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

(Bonnot et al. in prep)

- **18 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

Transcriptomic Data:

- ## 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
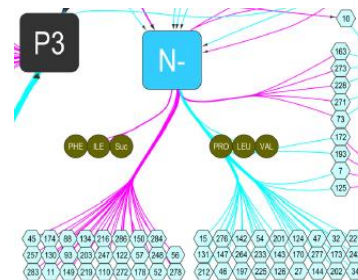   - Genes, proteins, metabolites responding similarly to the nutrition?
      - Genes, proteins, metabolites strongly co-expressed /co-accumulated ?
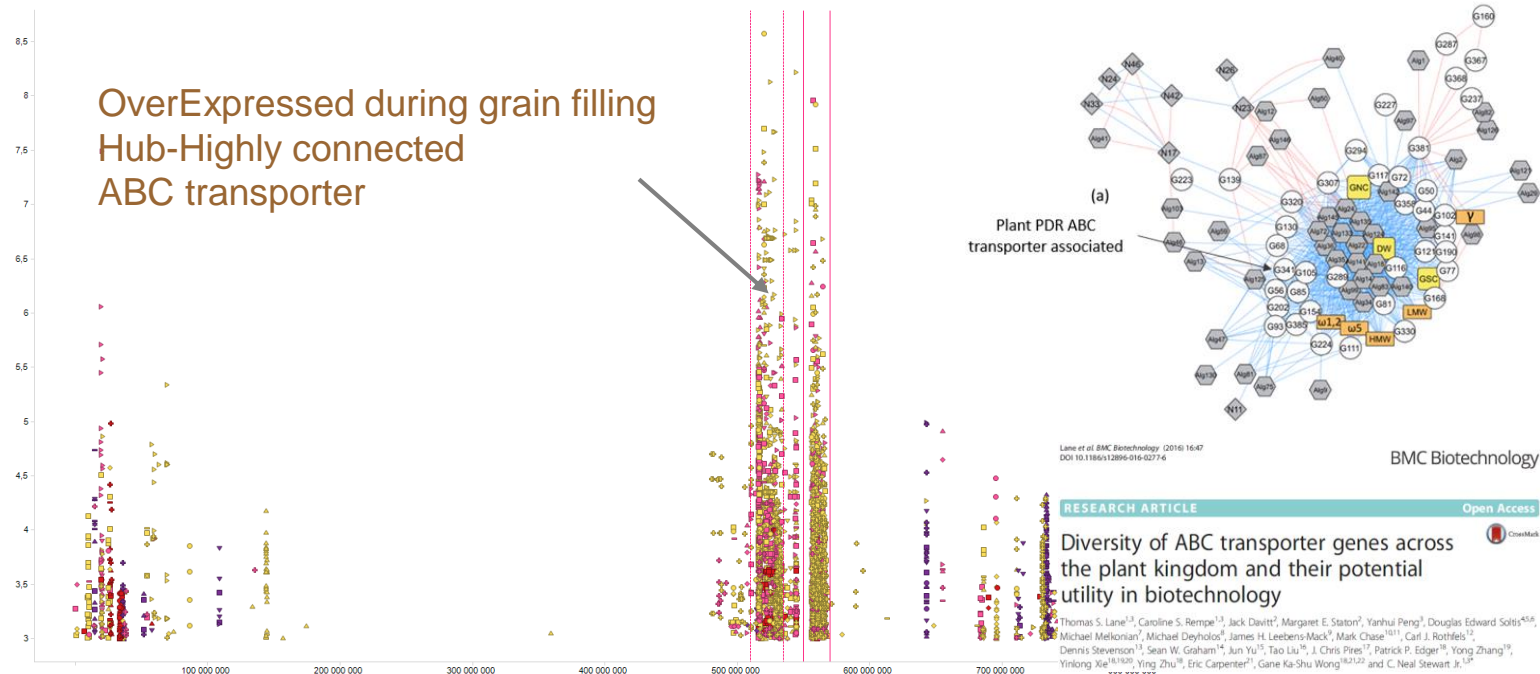   - Central genes in the response to the nutrition ?

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

Plant PDR ABC
transporter associated

(a)

Lane et al BMC Biotechnology (2016) 16:47
DOI 10.1186/s12896-016-0277-6

BMC Biotechnology

RESEARCH ARTICLE   Open Access

Diversity of ABC transporter genes across
the plant kingdom and their potential
utility in biotechnology

Thomas S. Lane[1,3], Caroline S. Rempe[1-3], Jack Davitt[2], Margaret E. Staton[2], Yanhui Peng[3], Douglas Edward Soltis[4,5,6], Michael Melkonian[7], Michael Deyholos[8], James H. Leebens-Mack[9], Mark Chase[10,11], Carl J. Rothfels[12], Dennis Stevenson[13], Sean W. Graham[14], Jun Yu[15], Tao Liu[16], J. Chris Pires[17], Patrick P. Edger[18], Yong Zhang[19], Yinlong Xie[18,19,20], Ying Zhu[19], Eric Carpenter[21], Gane Ka-Shu Wong[18,21,22] and C. Neal Stewart Jr.[1,3*]

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
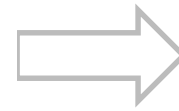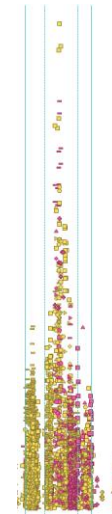
BreedWheat - Conclusions

- 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…*

P 276
Gaetan Touzy

Thanks to all people involved in these work

Gaetan Touzy, Katia Beauchêne, Jean Pierre Cohan

Celine Zimmerli, Guenter Weltz, Michael Schmolke

Agathe Mini, Guylaine Besnier-Hebert, Sébastien Praud

Sylvie Dutriez, Alain Lacassagne

Denis Beghin, Ellen Goudemand, Olivier Robert

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, Benoit Foucault, Patrice Senellard

Olivier Lucas, Laure Duchalais, Christophe Michelet

Pascal Giraudeau, Sébastien Ducroquet

Sebastien Caiveau, Odile Argilier,

Thank You