High throughput phenotyping and plant modelling : two legs for combined physiological and genetic approaches ?

F. Tardieu, C. Welcker, C. Fournier, B. Muller, Th Simonneau, C. Granier











Phenotyping is the bottleneck in genomics.

- Sequences of Arabidopsis, Rice, Sorghum, poplar, tomato, maize etc
- Genotyping capacities for 1000s of genotypes
- Large collections of RILs, mutants, accessions

Is this the proper way to address the problem ? (defined by a need, not by a biological question) Are "genotyping" and "phenotyping" parallel activities ?

The 'problem' is that we can now measure traits in1000s of plants in a robotised way



CropDesign

Granier et al. 2006 New Phyt

Trends in pl sc 2008

We need to refine biological questions avoiding reinventing the wheel Modelling can help...

1 Phenotyping The problem : controversial statements (food for discussion)

Confusions of effects (examples with panels of lines)

Phenotyping : the most expensive way to discover genes of flowering time while looking for other targets ?



1 Phenotyping The problem : controversial statements (food for discussion)

Confusions of effects (examples with panels of lines)

Phenotyping : the most expensive way to discover genes of flowering time while looking for other targets ?



Also valid for leaf area and cell size in At (Cookson et al 2007 Ann Bot 99:703) effect of candidate genes on cell cycle (plenty of examples) etc. C. Granier



Confusions of effects (examples with panels of lines)

Plant water content, an expensive way of measuring plant size ? *NIR sensor in plants which are rehydrating in a black chamber*





IPK-Gatersleben

Lemnatec website

Confusions of effects (examples with panels of lines)

Plant water content, an expensive way of measuring plant size ? *NIR sensor in plants which are rehydrating in a black chamber*



Small plants recover water status more rapidly (this effect can be tested with a model)

1 Phenotyping The problem : controversial statements (food for discussion)

Phenotyping platforms in controlled conditions not appropriate for plant performance (small pots, low light...)

Field phenotyping platforms will never be numerous enough for a proper network of experiments (GxE, QTL x E)

... Phenotyping platforms are for something else

- detecting / analysing heritable traits involved in yield
- in a modeller's language : parameters of models



Lemnatec / CropDesign

Phenopsis

1. Phenotyping : The problem Partial conclusion

Roles for modelling ?

"Model assisted phenotyping" for :

- Optimising designs to the biological question (variables, precision...)
- Identifying heritable traits from massive (and messy) datasets
- Scale up from individual traits to whole-plant performance in the real world.



Biomass = $\int_{0}^{1} \frac{\text{Incident light}}{0} * \%$ intercepted * Radiation Use Efficiency (RUE)

Incident light Spatial variability of environmental conditions : acceptable or not ?

Modelling light at plant level

Modelling phenotypic consequences



Biomass = \int_{0}^{0} Incident light * <u>% intercepted</u> * Radiation Use Efficiency (RUE) Architecture : Which variables for a genetic and GxE analyses











Digitizing



Genetic / environmental analyses of parameters



QTL analysis

Karin Chenu and Christian Fournier, unpublished

Biomass = \int_{0}^{1} Incident light * % intercepted * Radiation Use Efficiency (RUE)

Stomatal conductance

(impossible to measure at high throughput with gas exchange equipment)



Partial conclusion

- Relevant variables for biological questions may need to be calculated from models (hidden variables, not raw phenotypes),
- Modelling : testing environmental scenarios / variability

- Experiments finished (different growing periods with different climates)
- Environmental conditions differ between periods, genotypes, stages
- 10⁴ to 10⁶ datapoints per experiment, not that clean

Handling data prior to genetic analyses



Lemnatec, CropDesign

Phenopsis

Biomass = \int_{0}^{t} Incident light * % intercepted * Radiation Use Efficiency (RUE)

How to manage time in fluctuating conditions ?

- different growing periods with different temperatures
- estimation of rates and durations biased by temperature fluctuations

Temperature-compensated time and rates ('thermal time', 'degree days' 'GDD')



3 Organising the mess for genetic analyses : temperature compensation

Fluctuating temperature (field, greenhouse) :



- We do not know the mechanisms of response
- We do not know the mechanisms for coordination of processes
- But extremely robust "metamechanism"



- Leaf elongation
- Development Yin 1996
- Germination



 $Rate = \frac{k.T.exp(-Ea/RT)}{1 + exp(\Delta S/R - \Delta H/RT)}$



Very low genetic variability of the response :

- Rice, 7 lines
- maize 350 lines , (Sadok et al. 2007, PCE)

3 Organising the mess for genetic analyses : temperature compensation

Model

- Extremely robust
- No idea of the mechanisms
- A way to express time "as if" T was 20°C

```
Rate = \underline{k.T.exp(-Ea/RT)} \\ 1 + exp(\Delta S/R- \Delta H/RT)
```





Parent et al J. Exp Bot 2010.

3 Organising the mess for genetic analyses : temperature compensation

Temperature compensation in a large range of T : rapidly fluctuating developmental variables (leaf elongation rate)





Time after germination (compensated)

Serves to predict development of any leaf or to know the status of all developing leaves at a given time

3 Organising the mess for genetic analyses : temperature compensation



Temperature-compensated rates or durations allow joint analyses of several experiments and to identify patterns :

essential for high throughput phenotyping (controlled and field

Granier, Massonet, Turc, Muller and Tardieu, 2002, Annals of Bot 89:595-604

3 Organising the mess for genetic analyses : rapidly fluctuating variables

Some key variables with genetic variability fluctuate in minutes

How to deal with them ?



3 Organising the mess for genetic analyses : rapidly fluctuating variables

Example : leaf elongation rate



x 200 RILs, 3-4 experiments



detail : http://bioweb.supagro.inra.fr/phenodyn/

Response curve of each RIL of mapping populations each genotype, one set of parameters (field, chamber, greenhouse)



3 Organising the mess for genetic analyses : rapidly fluctuating variables



3 Organising the mess for genetic analyses : rapidly fluctuating variables



Reymond et al. 2003 Plant Phy



Strategy for data mining : we need specific methods aimed at answering questions

Temperature compensation, a key question (to which variables ?)

Rapidly fluctuating variables : "hidden variables" can

- synthesise genotypic behaviours
- be analysed genetically
- allow reconstruction of the original phenotype.

Data mining for response curves in litterature H Poorter's talk

4. From platform to fields





I have QTLs or associations for traits in phenotyping platforms (controlled or field)

Can they predict performance and G x E in a large range of scenarios ?

Needs field experiments BUT

- Never enough situations to test the interest of an allele
- Usually partly fails : where and why ?

→ Predicting the allele effect with model AND comparing with fields

4. From platform to fields





Chenu et al. 2008 Plant Cell Environment 31, 378– Col. Hammer, Chapman U. Queensland

4. From platform to fields



Observed (symbols) and simulated (lines) leaf lengths under contrasting climatic scenarios in the field

Chenu et al. 2008 Plant Cell Environment 31, 378– Col. Hammer, Chapman U. Queensland



Climatic data

virtual plant / genotype (with effect of QTLs)

calculated feedbacks of plants on environment (e.g. soil depletion)

effect of allelic composition on plant performance

Biomass = Incident light * % intercepted * Radiation Use Efficiency (RUE)

Effects of leaf QTLs on simulated yield



Alleles which maintain growth under stress

Tester Alleles which decrease growth under stress



Maintaining leaf area most often favourable except in terminal stress

Chenu et al. 2009 Genetics

Effects of leaf QTLs on simulated yield

Test requires

- Adequate environmental measurements in the field
- Methods for field phenotyping + model
- Multi-environmental tests

F. Baret's presentation

CONCLUSION

- A big risk if phenotyping is not considered via biological questions (phenotyping a new name for whole plant physiology ?)
- Phenotyping plaforms are not 'easier fields' : best use for identifying heritable parameters or 'hidden variables' used in models.
- Analysis of large messy datasets needs a framework of analysis
 - temperature-compensated rates (joint analysis of experiments
 - Hidden variables which synthesise time courses
- Platform to field : needs a theoretical framework
 - helps interpretation
 - allows 100s of virtual experiments in different scenarios