



## SP3 Commissioned research

Optimizing marker-assisted breeding systems for drought tolerance in cereals through linkage of physiological and genetic models

Scott Chapman, David Bonnett, Greg Rebetzke  
(CSIRO Plant Industry)

Jiankang Wang (CIMMYT Beijing)

Graeme Hammer, Mark Dieters (University of Queensland)

Francois Tardieu, Claude Welcker (INRA)  
CGIAR plant breeders

# Why are we here?

## Optimism or arrogance

- Can we improve on direct selection for yield in 'random' and/or managed environments?
- How do we utilise/recombine known variation?
  - Candidate genes (inc. 'perfect markers')
  - Candidate traits
  - QTL
- How do we deal with sources of error in markers and QTL?
  - Choice of phenotype
  - Precision of phenotype (field design, analysis)
  - Complexity of genetic control (GxE, GxG)
  - Precision of marker/QTL





## Why use simulation?

- To control all sources of error
- To determine what level of ‘knowledge’ is required to enable advance that is *faster* than existing breeding methods?
  - How can we combine known genes into the same line (parent building)?
  - How useful is a marker or QTL for breeding?
  - How much genetic, environmental and physiological ‘noise’ can be overcome?



## Rationale and objective

- July 2005 to June 2007
- Exploit dynamic linkage of genetics and crop physiology to improve drought tolerance
- Develop some simulation case studies for
  - Parent building
  - QTL detection
  - QTL exploitation
  - Crossing and selection...



# Activities

- Wheat
  - Incorporate physiological model into existing sorghum framework
  - Case studies on marker-based selection
- Sorghum
  - identify genetic/trait information relevant to GCP/CGIAR breeding tasks
- Maize
  - case study on simple physiological model for leaf elongation
- Rice, legumes, clonal crops
  - evaluate needs and opportunities, initiate training where possible

# Topics

- Marker-based selection
  - Parent building
  - Selection on known markers or QTL
- Marker-assisted selection
  - QTL generated during selection process



# Marker-based selection

- Introduce the resistance gene (G1) in the donor parent P2 to the recurrent parent P1. The target genotype is shown below.

		Chromosome 1										
		M1	M2	G1	M3	M4	M5	M6	M7	M8	M9	M10
r (with the previous locus)		0.50	0.15	0.07	0.08	0.15	0.15	0.15	0.15	0.15	0.15	0.15
P1: recurrent parent		1	1	2	1	1	1	1	1	1	1	1
		1	1	2	1	1	1	1	1	1	1	1
P2: donor parent		2	2	1	2	2	2	2	2	2	2	2
		2	2	1	2	2	2	2	2	2	2	2
Target genotype		1	1	1	1	1	1	1	1	1	1	1
		1	1	1	1	1	1	1	1	1	1	1

		Chromosome 2										
		M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21
		0.50	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
		1	1	1	1	1	1	1	1	1	1	1
		1	1	1	1	1	1	1	1	1	1	1
		2	2	2	2	2	2	2	2	2	2	2
		2	2	2	2	2	2	2	2	2	2	2
		1	1	1	1	1	1	1	1	1	1	1
		1	1	1	1	1	1	1	1	1	1	1



# Marker-based selection – gene frequency

Selection method	Marker type	Distance between marker and gene		
		1cM	5cM	10cM
Homozygous selection in $F_2$ (co-dominant markers needed)	Single marker	0.991	0.954	0.910
	Flanking markers	1.000	0.998	0.990
Homozygous selection in $F_{10}$	Single marker	0.980	0.912	0.846
	Flanking markers	0.999	0.988	0.959
Enrichment selection in $F_2$ , homozygous sel in $F_{10}$	Single marker	0.982	0.847	0.847
	Flanking markers	0.999	0.963	0.963

- Flanking markers at 5 to 10cM better than single at 1cM
- Cheapest, fastest method is  $F_2$  gene enrichment selection, with flanking markers...



## Real-world crosses with 9 genes....

Gene	<i>Rht-B1</i>	<i>Rht-D1</i>	<i>Rht8</i>	<i>Sr2</i>	<i>Cre1</i>	<i>VPM</i>	<i>Glu-B1</i>	<i>Glu-A3</i>	<i>tin</i>
Chromosome	4BS	4DS	2DL	3BS	2BL	7DL	1BL	1AS	1AS
Marker type	Codom	Codom	Codom	Codom	Dom	Dom	Codom	Codom	Codom
Distance between marker and gene (cM)	0	0	0.6	1.1	0	0	0	0	0
Sunstate <sup>a</sup>	<i>Rht-B1a</i>	<b><i>Rht-D1b</i></b>	<i>rht8</i>	<i>Sr2</i>	<i>cre1</i>	<b><i>VPM</i></b>	<b><i>Glu-B1i</i></b>	<b><i>Glu-A3b</i></b>	<i>Tin</i>
Silverstar+ <i>tin</i>	<b><i>Rht-B1b</i></b>	<i>Rht-D1a</i>	<i>rht8</i>	<i>sr2</i>	<b><i>Cre1</i></b>	<i>vpm</i>	<b><i>Glu-B1i</i></b>	<i>Glu-A3c</i>	<b><i>tin</i></b>
HM14BS	<i>Rht-B1a</i>	<i>Rht-D1a</i>	<b><i>Rht8</i></b>	<i>sr2</i>	<i>cre1</i>	<i>vpm</i>	<i>Glu-B1a</i>	<i>Glu-A3e</i>	<i>Tin</i>
Target genotype	<i>Rht-B1a</i>	<i>Rht-D1a</i>	<b><i>Rht8</i></b>	<i>Sr2</i>	<b><i>Cre1</i></b>	<b><i>VPM</i></b>	<b><i>Glu-B1i</i></b>	<b><i>Glu-A3b</i></b>	<b><i>Tin</i></b>

- All genes unlinked except *Glu-A3* and *tin* (3.8cM)
- Target genotype is
  - alternative dwarfing gene
  - disease resistance
  - good quality
- Design marker selection scheme to build target genotype

# Marker-based selection – parent building with 9 genes



- Gene enrichment selection is efficient
- Reduced genotypes and markers to about 15%
- 3500 marker assays vs 25 000!

Population stage	No gene enrichment selection in TCF <sub>2</sub>	2 genes in TCF1 and enrichment selection for all 7 genes in TCF <sub>2</sub>	Homozygous selected for <i>Rht8</i> , and enrichment selection for others in TCF <sub>2</sub>
TCF <sub>1</sub>	145 (x 5)	144	145
TCF <sub>2</sub>	-	37	114
DHs derived from TCF <sub>2</sub>	3440 (x 7)	408	286
<b>TOTAL</b>	<b>3585 (24805)</b>	<b>589 (3835)</b>	<b>545 (3525)</b>

Selection for 5 markers (2 homozy, 3 enriched) in TCF1,  
 Selection for 7 markers in TCF1, lines

# Marker-based selection – Some rules of thumb



- Within reasonable population size can select simultaneously for
  - 3 markers in F2
  - 7 markers in DH or RIL
  - 12 or 13 markers with F2 gene enrichment prior to DH or RIL
- Rules for
  - when a backcross is needed
  - in top cross, parent with largest number of favourable alleles should be last

# Marker-assisted selection – An Integrated Approach Will Revolutionise Crop Improvement

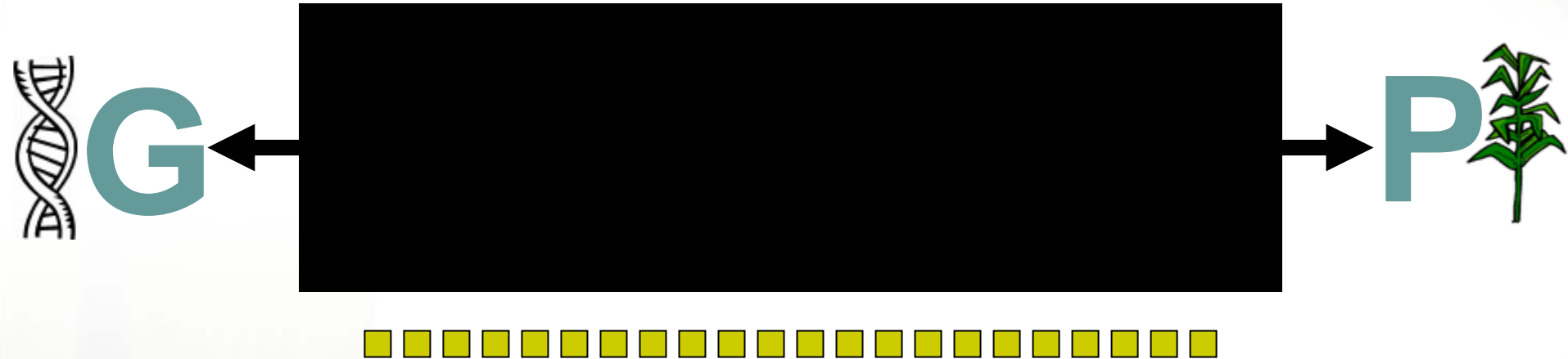


- Change the focus within breeding programs from the paradigm of ***identifying superior varieties*** to a focus on ***identifying superior combinations of genetic regions*** and packaging these regions into varieties
- Simulation modelling
  - links genotype to phenotype
  - identifies combinations of genotype characteristics and management systems that optimise resource capture in specific cropping environments.



How can we capture features of GP relationships for complex traits in models?

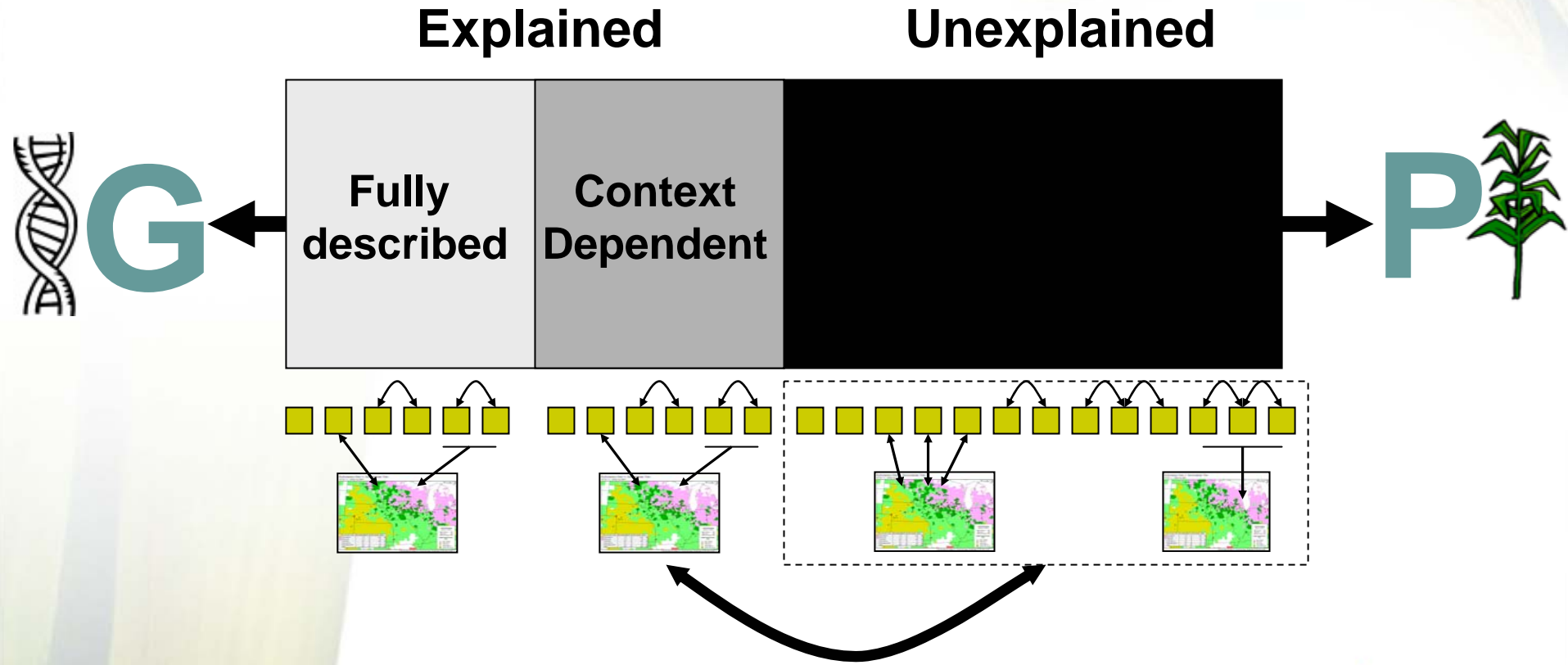
Unknown



- $P = G + G \times E + \varepsilon$  etc
- $P = G_{\text{Explained}} + G_{\text{Unexplained}} + \varepsilon$
- $P = \text{QTL}_{(\text{mod})} + E(NK) + \varepsilon$



# Modeling G-P relationships: Parameterization with "real" data

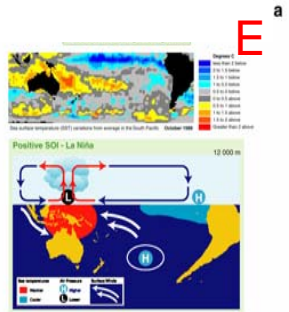


- $$P = G_{\text{Explained}} + G_{\text{Unexplained}} + \varepsilon$$

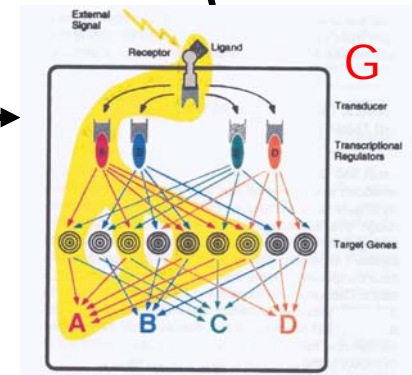
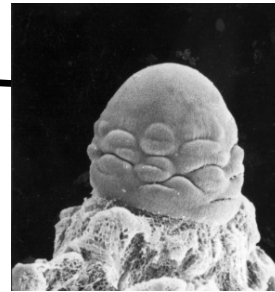
(Cooper et al ICSC 2004)



# Reducing complexity by capturing GxExM interactions

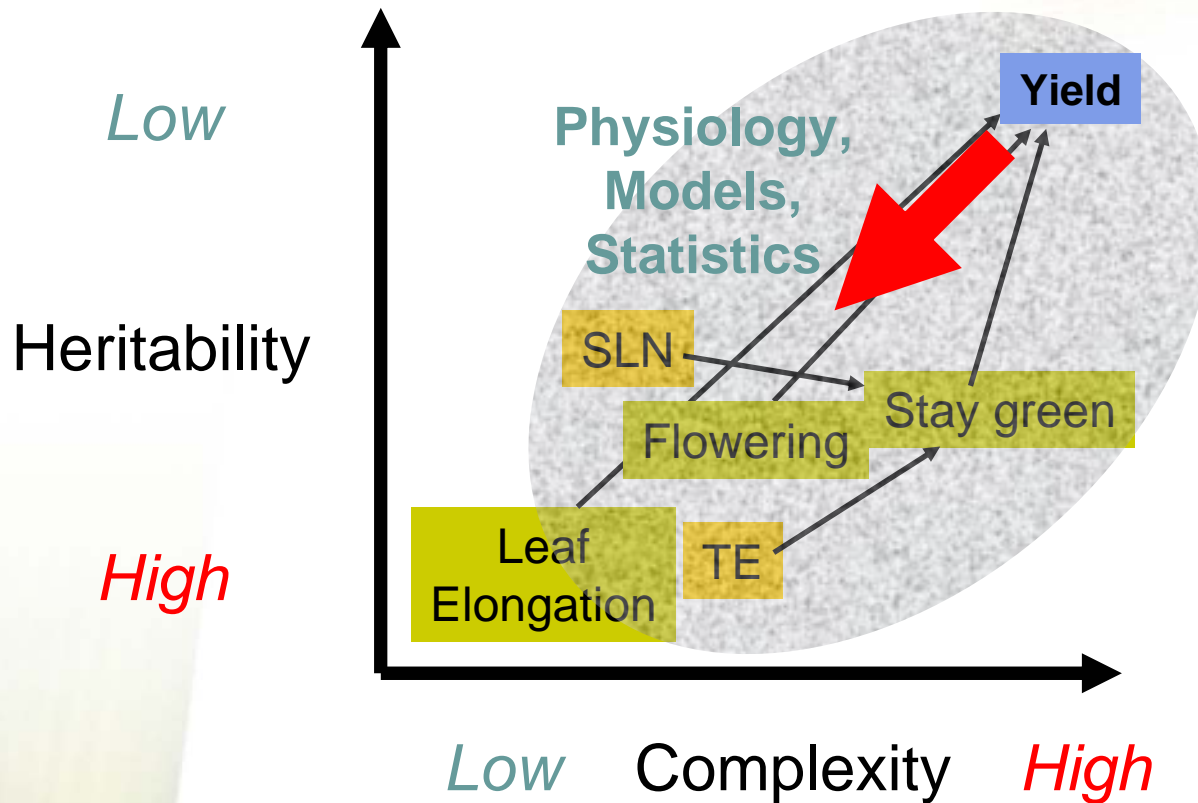


- Information flow and control
- Sensing, signalling, responding





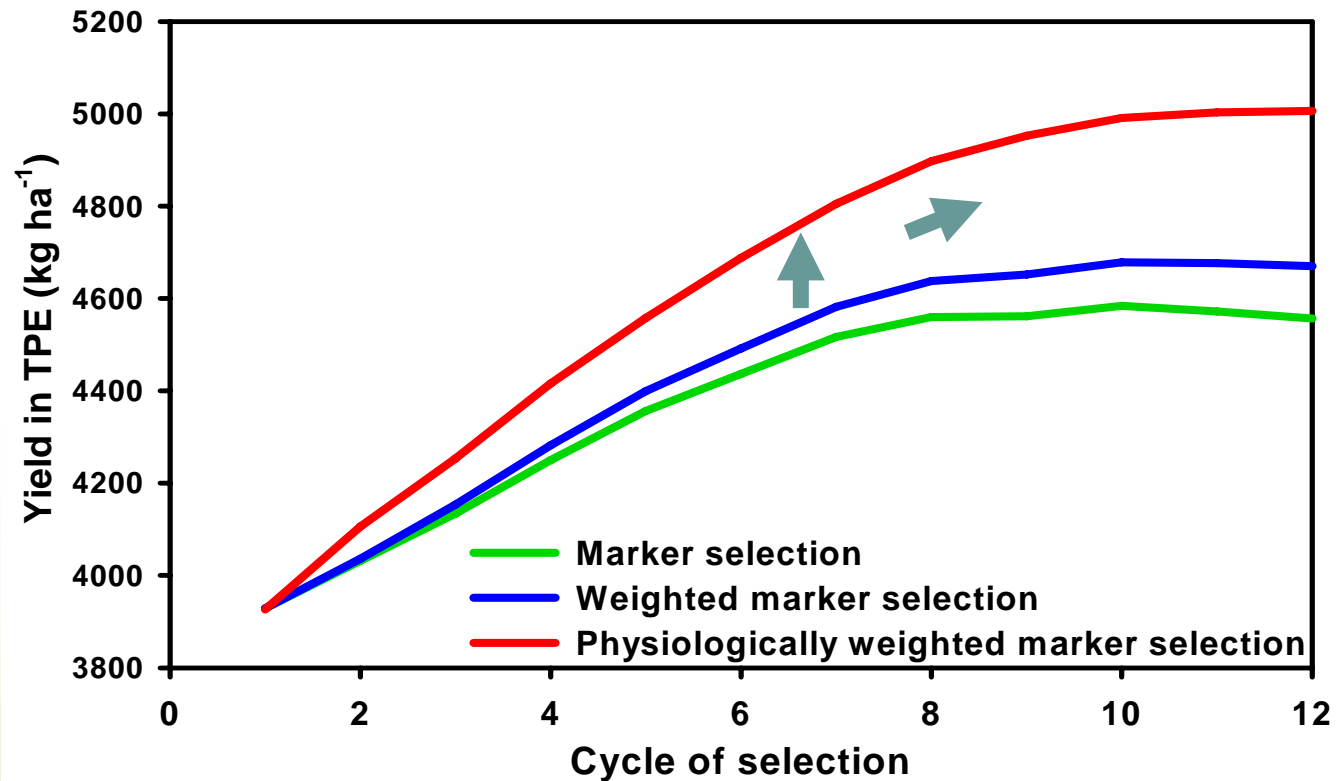
# Aim 1: To reduce trait complexity...



.... by accounting for 'unexplained variance'



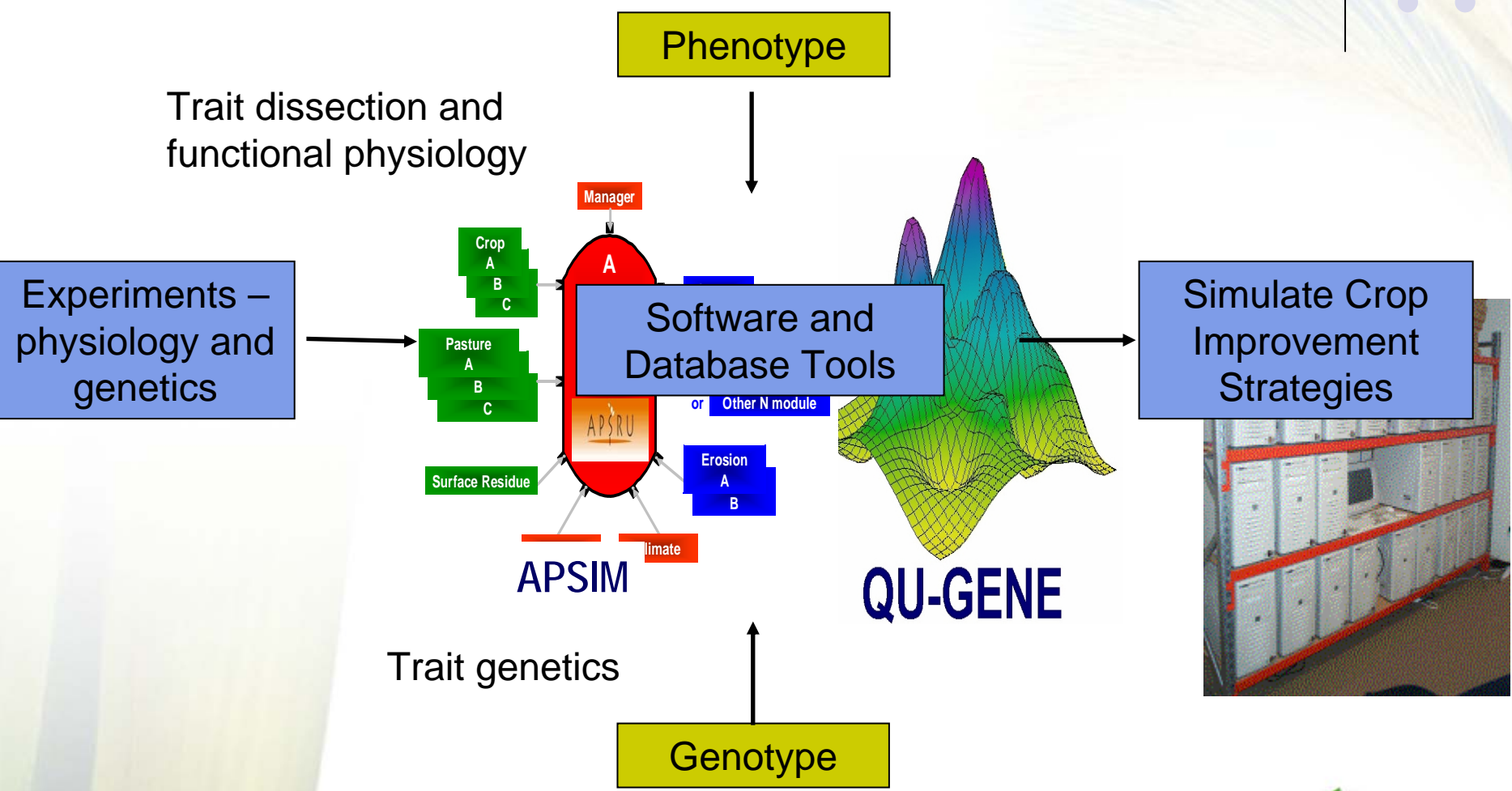
## Aim 2: Enhance selection response...



....by exploiting 'unexplained' variance



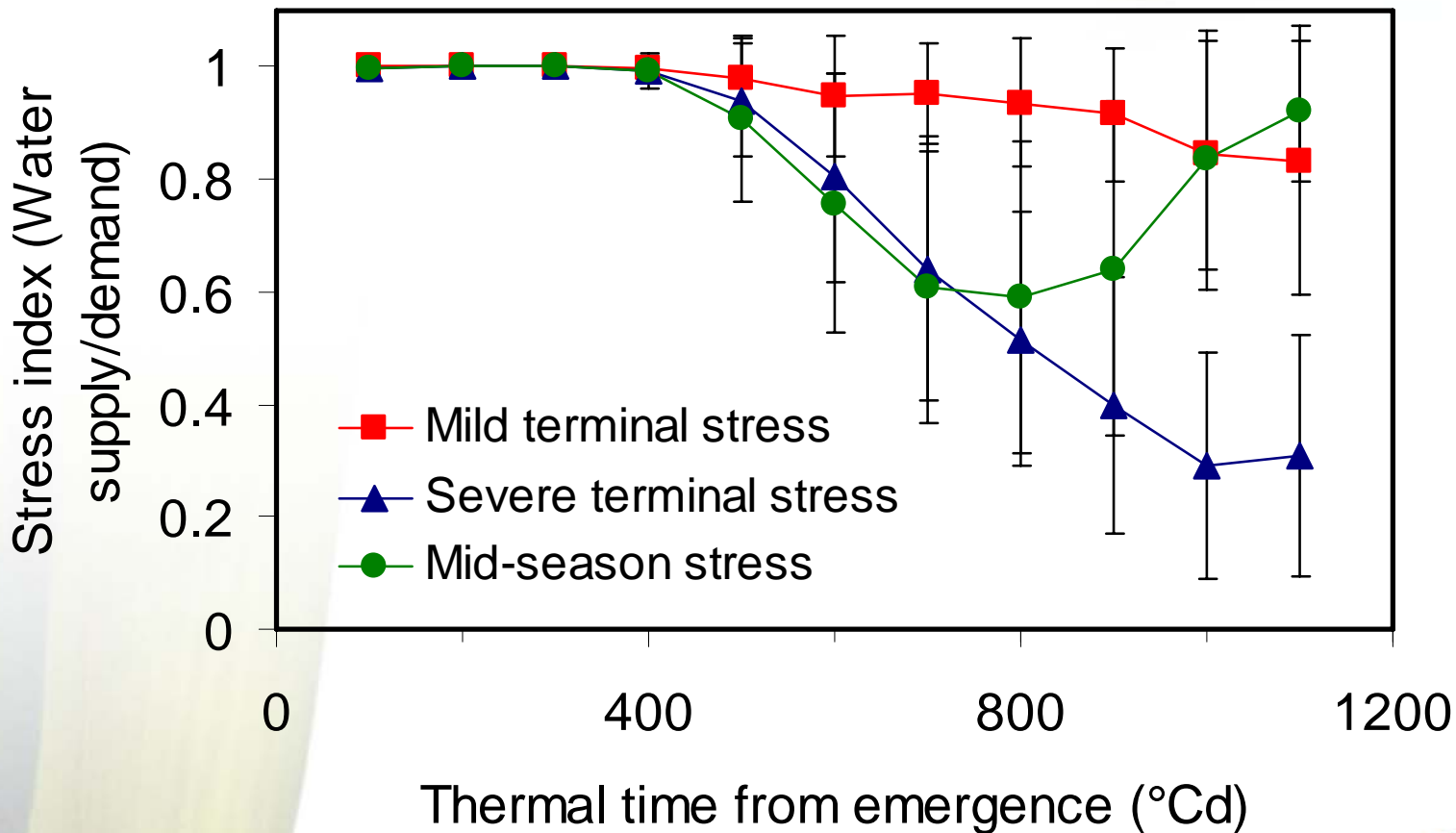
# Predictive Research framework



# Marker-assisted Selection: Environment Classification



## Seasonal patterns of water limitation



# Investigating Breeding Strategies

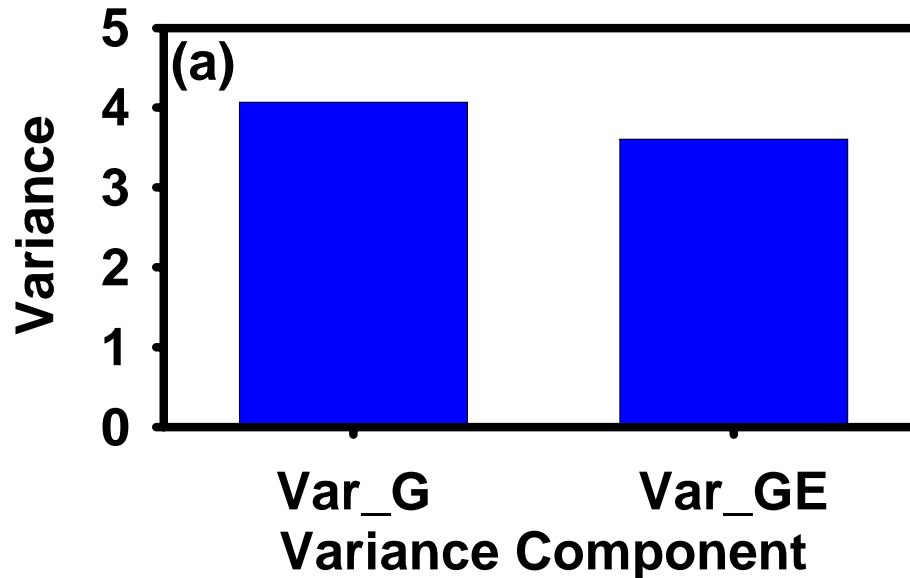


- Simulate marker-assisted selection (MAS) using analysis results in QU-GENE
  - assumed markers close to genes
  - random environments sampled for each cycle of testing and selection
  - equal weighting given to genotypic and phenotypic information
- Contrasted strategies using differing levels of physiology/modelling evidence



# Analysing Genetic Effects

## I. Conventional variance components

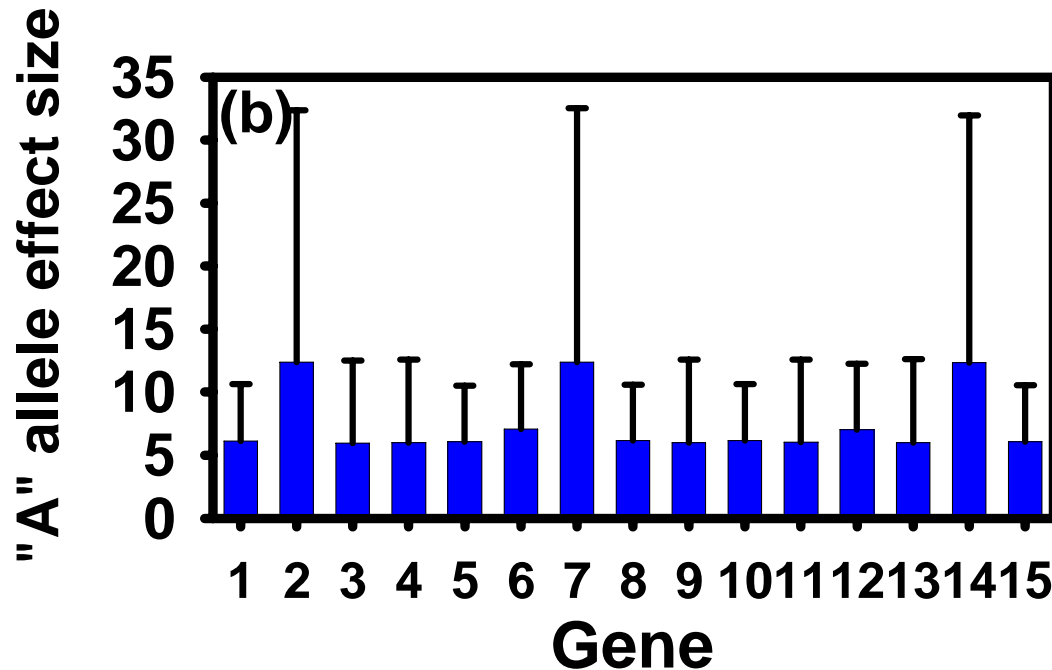


- GxE for yield arising from additive allelic effects at trait level as modified through crop model

# Analysing Genetic Effects



## II. Average gene effect/QTL analysis

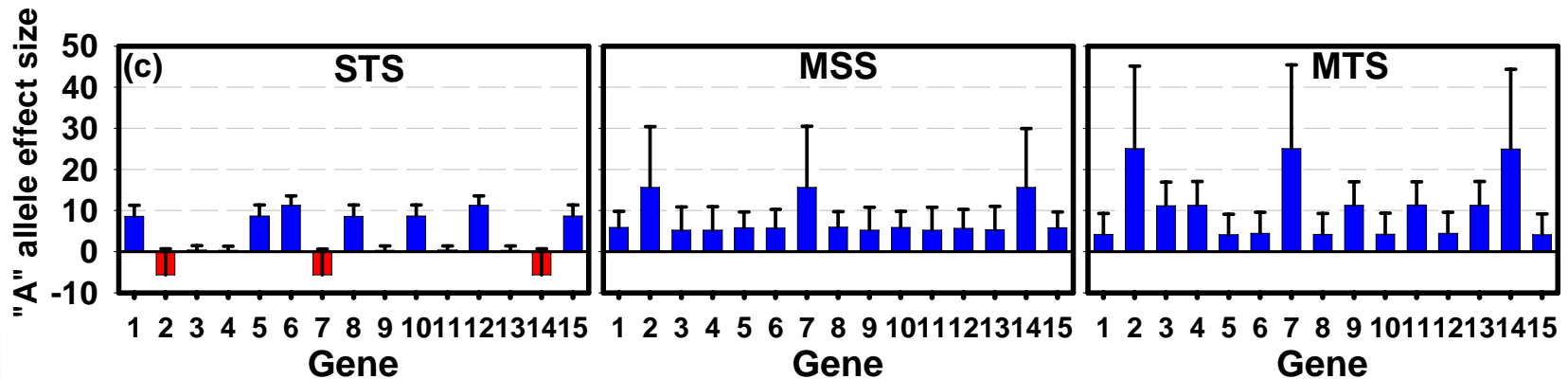


- Averaged over all environments
- Equivalent to massive QTL study for main effects

# Analysing Genetic Effects



## III. Average gene effects within environment types

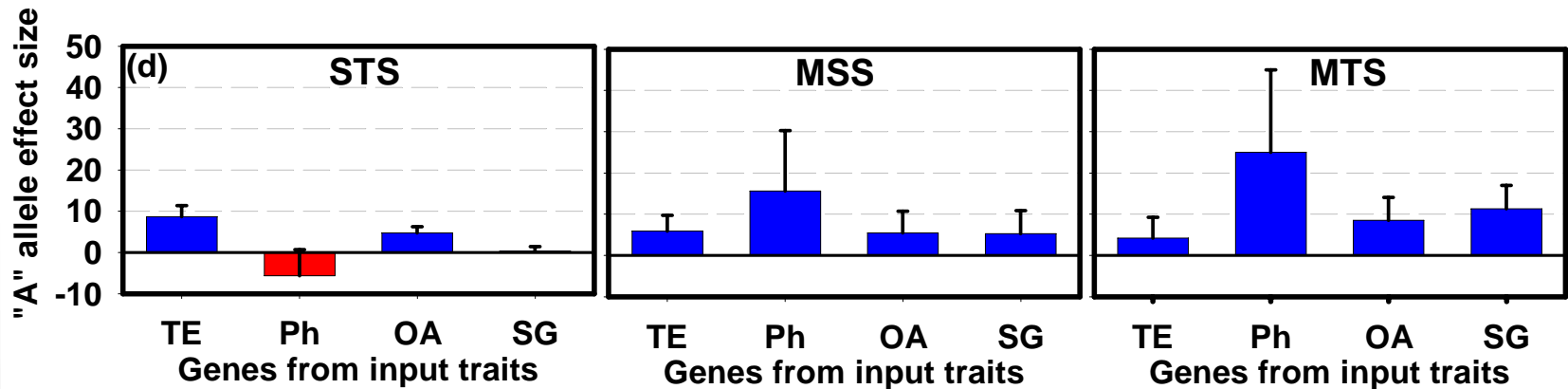


- Environment context dependency
- QTL x E

# Analysing Genetic Effects

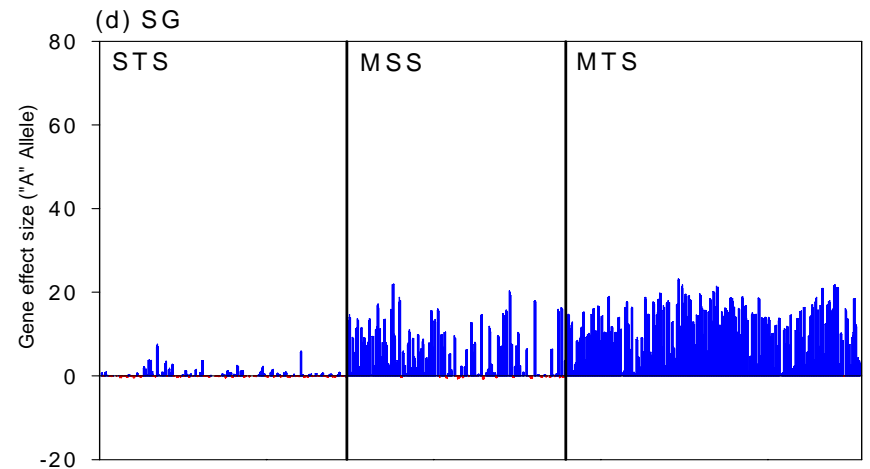
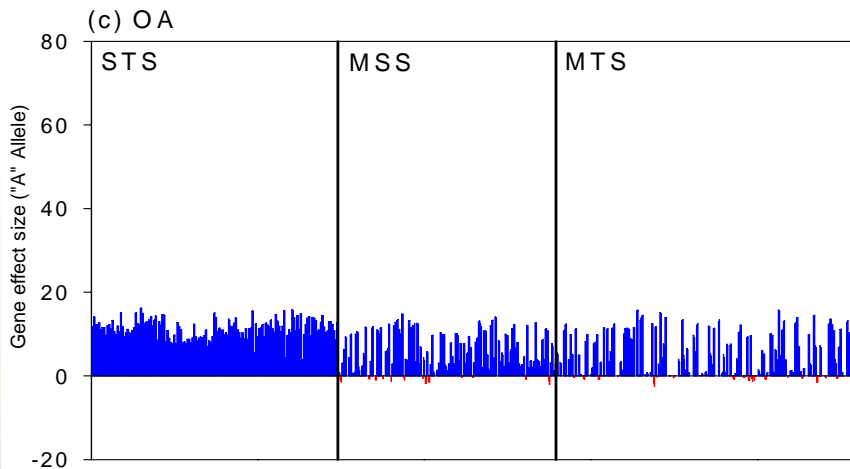
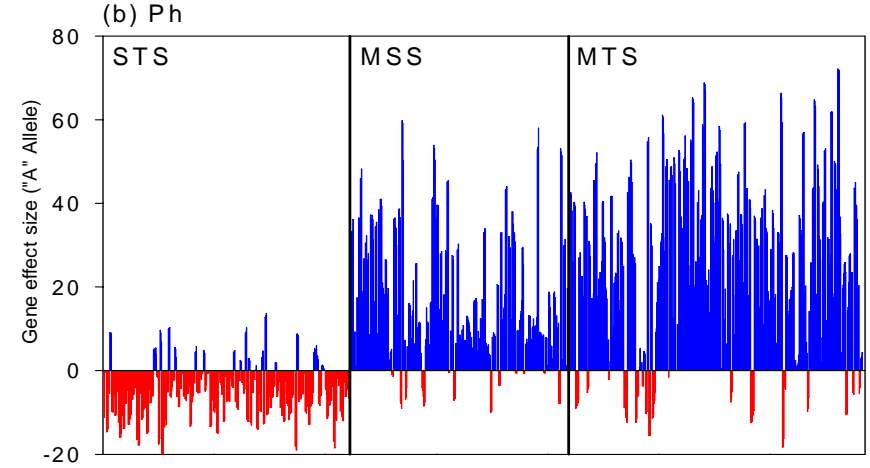
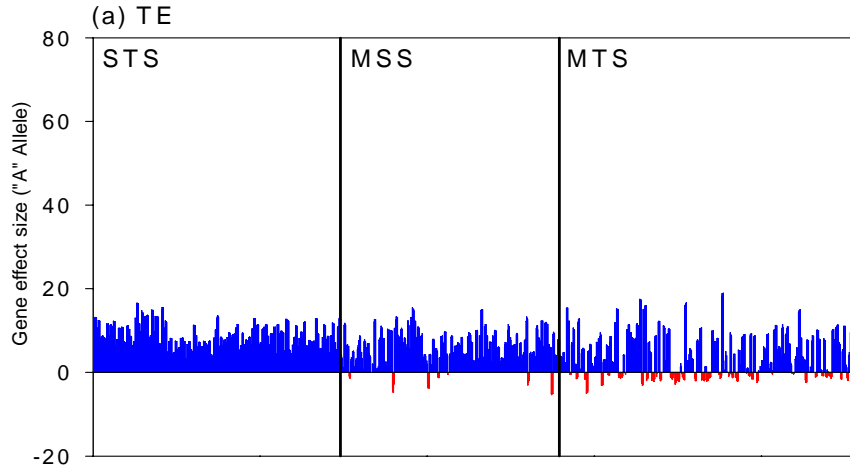


## IV. Average gene effects by traits and environ types



- precise phenotyping to inform genetic analysis
- QTL x E x trait

# Analysing Genetic Effects



- Context dependency for environment by trait
- QTLxTrialxTrait effects

# Investigating Breeding Strategies

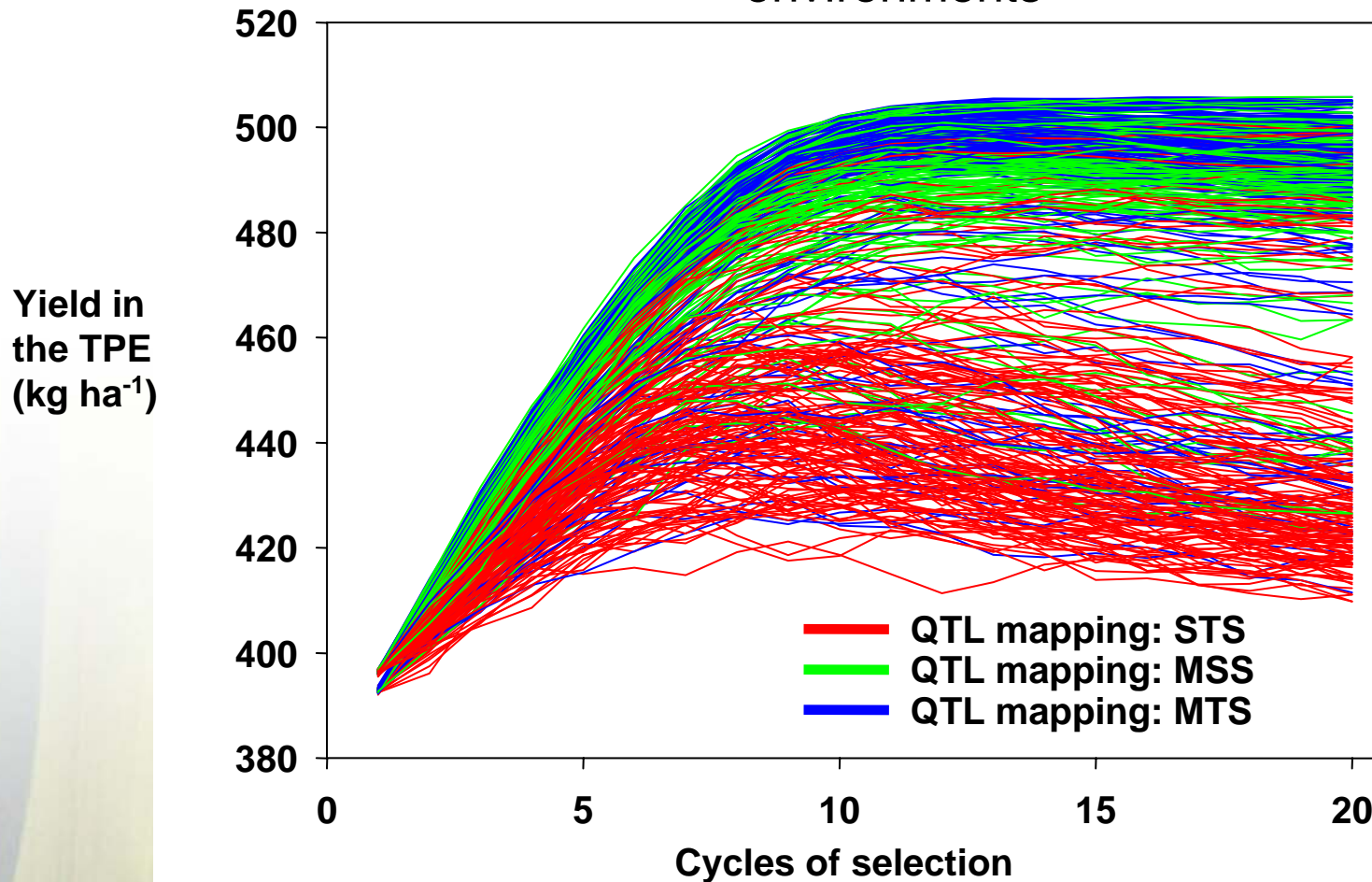


- Simulate marker-assisted selection (MAS) using analysis results in QU-GENE
  - assumed markers close to genes
  - random environments sampled for each cycle of testing and selection
  - equal weighting given to genotypic and phenotypic information
- Contrasted strategies using differing levels of physiology/modelling evidence

# Investigating Breeding Strategies



Effectiveness of MAS based on QTL analysis in each of the 547 environments





## Marker-based selection – next steps

- Using marker-based selection
  - Adding in QTL for drought related traits (coleoptile length, TE)
  - Examining effects of linkage and ‘noise’
- Using marker-assisted selection
  - Effects of GxE
  - Scaling of physiological models
    - E.g. leaf elongation rate to yield



## Use of physiological models in genetics

- Drought is a complex trait
- Where are the quantitative geneticists?
- Simple vs complex traits
  - Simple genetics OR
  - More direct relationship to traits of interest
- Present
  - Stress index to classify trial
  - Models as 'virtual genotypes' in trial
  - Models to simulate trials of many genotypes and environments in order to evaluate methods for
    - Screening genotypes
    - Designing crossing and selection schemes



## Use of physiological models in genetics

- Future
  - Many, many QTL identified for many traits
  - How much do we need to know for QTL or genetic knowledge to be useful?
  - This can be answered and we can use this information to FOCUS research on traits that have the best promise in terms of
    - effectiveness of physiological mechanism
    - sufficiently resolved genetic understanding
  - Meeting the aims of
    - Reducing trait complexity to
    - Enhancing selection response