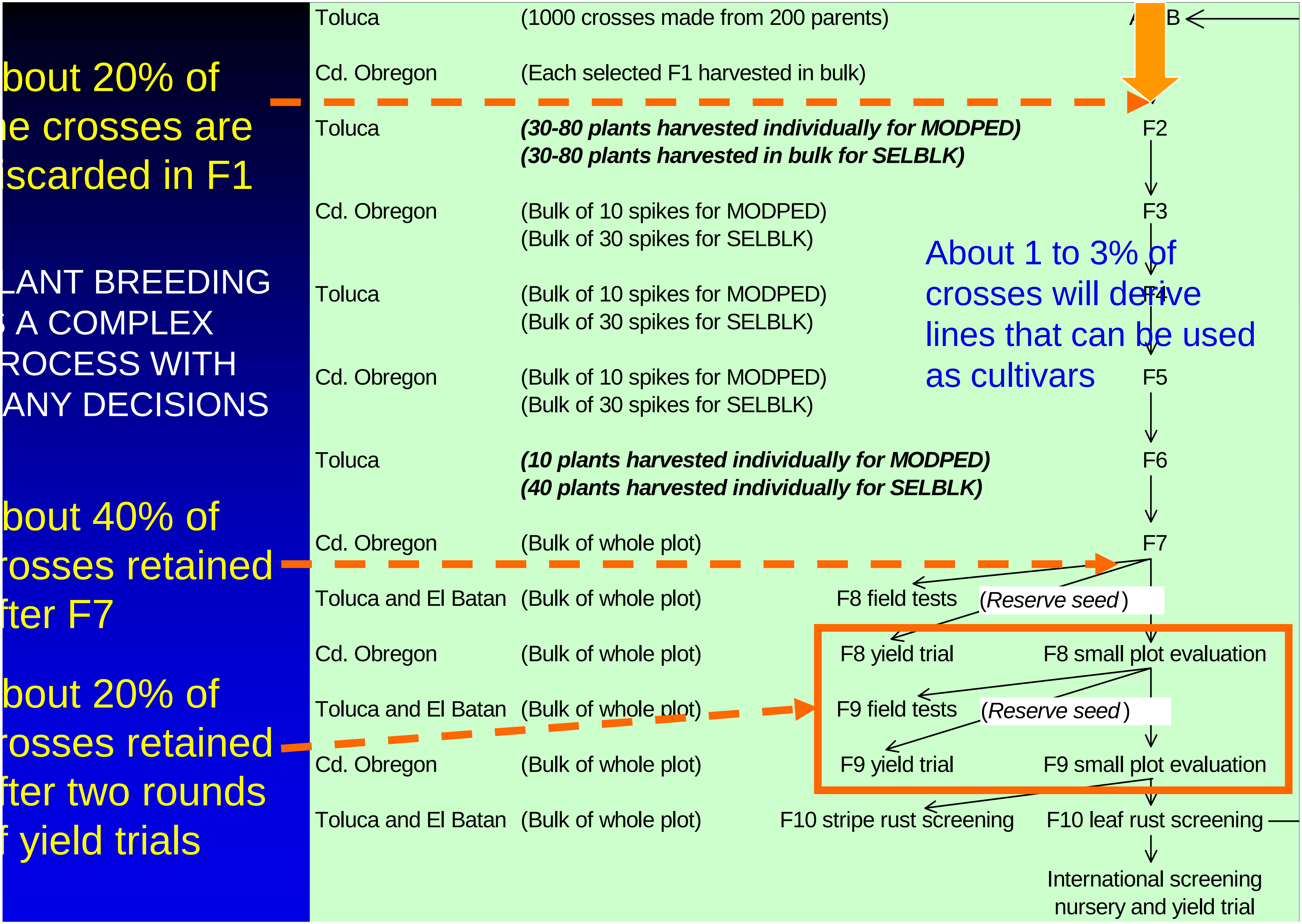
Current Status and Future Plans for Molecular Breeding Simulation Tools

Jiankang Wang, CIMMYT/CAAS

Outlines of presentation

- Some backgrounds
- QuCim tool for genetics and breeding
- What have been done using QuCim?
- Future plans



bout 20% of
ne crosses are
iscarded in F1
LANT BREEDING
; A COMPLEX
ROCESS WITH
ANY DECISIONS
bout 40% of
rosses retained
fter F7
bout 20% of
rosses retained
fter two rounds
f yield trials
Toluca
(1000 crosses made from 200 parents)
B
Cd. Obregon
(Each selected F1 harvested in bulk)
Toluca
(30-80 plants harvested individually for MODPED)
(30-80 plants harvested in bulk for SELBLK)
F2
Cd. Obregon
(Bulk of 10 spikes for MODPED)
(Bulk of 30 spikes for SELBLK)
F3
Toluca
(Bulk of 10 spikes for MODPED)
(Bulk of 30 spikes for SELBLK)
F4
Cd. Obregon
(Bulk of 10 spikes for MODPED)
(Bulk of 30 spikes for SELBLK)
F5
About 1 to 3% of crosses will derive lines that can be used as cultivars
Toluca
(10 plants harvested individually for MODPED)
(40 plants harvested individually for SELBLK)
F6
Cd. Obregon
(Bulk of whole plot)
F7
Toluca and El Batan
(Bulk of whole plot)
F8 field tests
(Reserve seed)
Cd. Obregon
(Bulk of whole plot)
F8 yield trial
F8 small plot evaluation
Toluca and El Batan
(Bulk of whole plot)
F9 field tests
(Reserve seed)
Cd. Obregon
(Bulk of whole plot)
F9 yield trial
F9 small plot evaluation
Toluca and El Batan
(Bulk of whole plot)
F10 stripe rust screening
F10 leaf rust screening
International screening nursery and yield trial

Why do we need tools in breeding?

- To improve efficiency of traditional phenotypic selection through exploring options
- Limitations in quantitative genetic theory
- Large amount of gene information available from
  - Genomics research
    - January 1992: 59,317 datasets, 77,805,556 bp
    - March 2005: 43,118,204 datasets, 47,009,081,750 bp
  - QTL mapping (CAB, April 2005)
    - 3497 publications on QTL mapping
    - 1581 publications in plants

Why do we need tools in breeding?

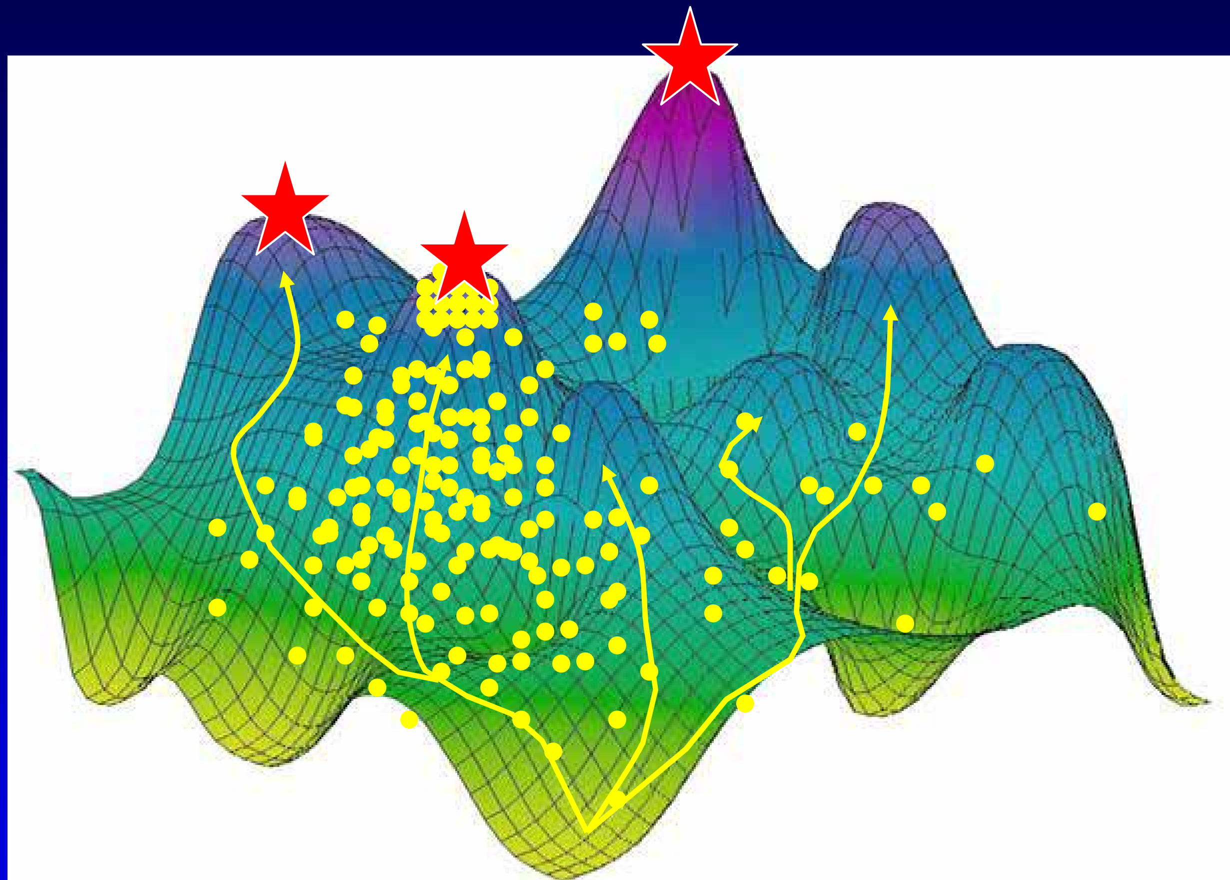
- The low efficiency of traditional phenotypic selection
- Limitations in theory
- Large amount of gene information available from genomics research and QTL mapping

- Bridge between the biological data and breeders' requirements
- Combine all these sources of data into "knowledge" for the breeding program

QuCim: A simulation tool for genetics and breeding

- QU-GENE (QUantitative GENEtics)
  - A simulation platform for quantitative analysis of genetic models, developed by The University of Queensland, Australia
- QuCim (funded by GRDC 2000-2005)
  - A QU-GENE application breeding simulation module, specifically designed for CIMMYT's wheat breeding programs
  - Simulate most, if not all, breeding programs for developing inbred lines
  - Version 1.1 released on July, 2003 (Workshop in Brisbane, Australia)
  - More than 100 global requests for QuCim 1.1

Landscape representation of a complex GE system (the real GE system is multi-dimensional)

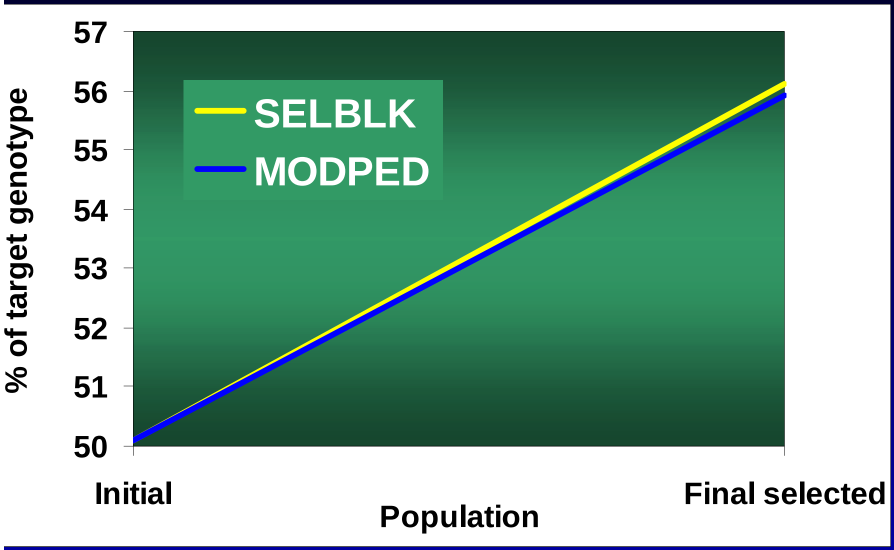


The breeding process can be viewed as a searching process for peaks on the landscape

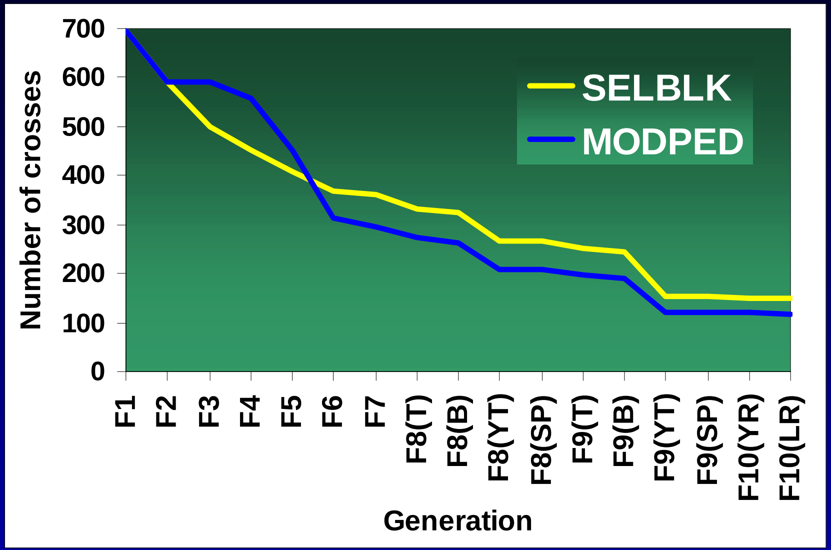
What has been done using QuCim?

- Comparison of two breeding strategies: modified pedigree (MODPED) and selected bulk (SELBLK), ***Crop Science*** (2003)
- Effects on selection of dominance and epistasis, ***Crop Science*** (2004)
- Cross performance prediction using known gene information, ***Aust. J. Agri. Sci.*** (2005)
- Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy, ***Molecular Breeding*** (2005)
- More to be submitted

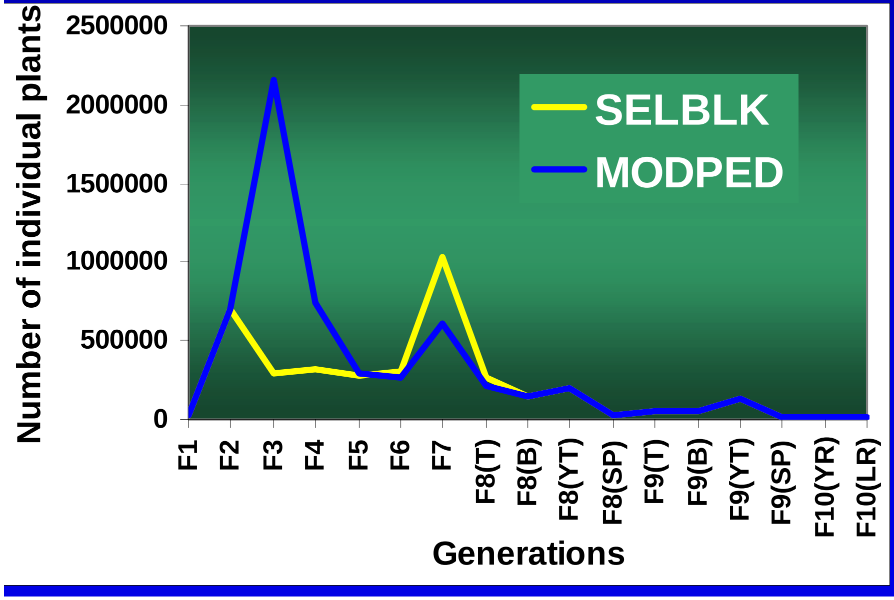
Genetic gain in yield from SELBLK is 3.3% higher than MODPED



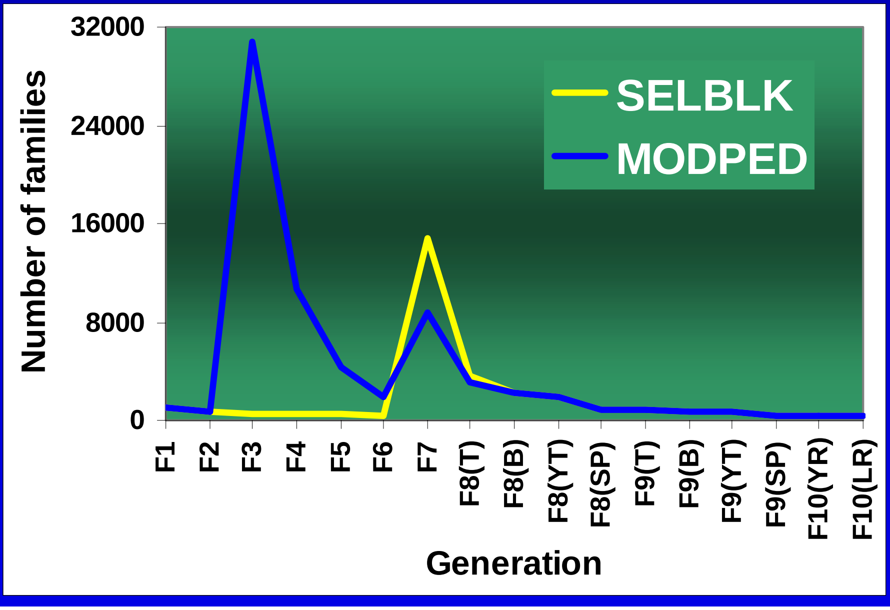
SELBLK retained 25% more crosses in the final selected population



SELBLK required 1/3 less land from F1 to F8 than MODPED



SELBLK produced 40% less families to be planted from F1 to F8



Glutenin genes and wheat quality

- Six glutenin genes
  - *Glu-A1* (1AL), *Glu-B1* (1BL), *Glu-D1* (1DL) for HMW
  - *Glu-A3* (1AS), *Glu-B3* (1BS), *Glu-D3* (1DS) for LMW
- Two end-use quality traits commonly used by wheat breeders
  - Rmax (BU), for maximum dough resistance
  - Extensibility (cm), for dough extensibility
- Multiple alleles on each gene locus
  - *Glu-A1*: *1*, *2**, and *Null*
  - *Glu-B1*: *7*, *7+8*, *7+9*, *6+8*, *20*, *13+16*, *14+15*, *17+18*, and *23+24*
  - *Glu-D1*: *2+12*, *4+12*, *5+10*, and *2+T2*

Four selection schemes

Parent 1 × Parent 2

↓

F1

↓

F2: 1000 individuals

⋮

F8: 1000 lines through SSD

	R0.04		**R0.2E0.2**		**E0.2R0.2**		**E0.04**	
	Trait to be selected	Lines selected	Trait to be selected	Lines selected	Trait to be selected	Lines selected	Trait to be selected	Lines selecte
tep 1	Rmax	**40**	Rmax	200	Extensibility	200	Extensibity	**40**
tep 2	n.a.		Extensibility	**40**	Rmax	**40**	n.a.	

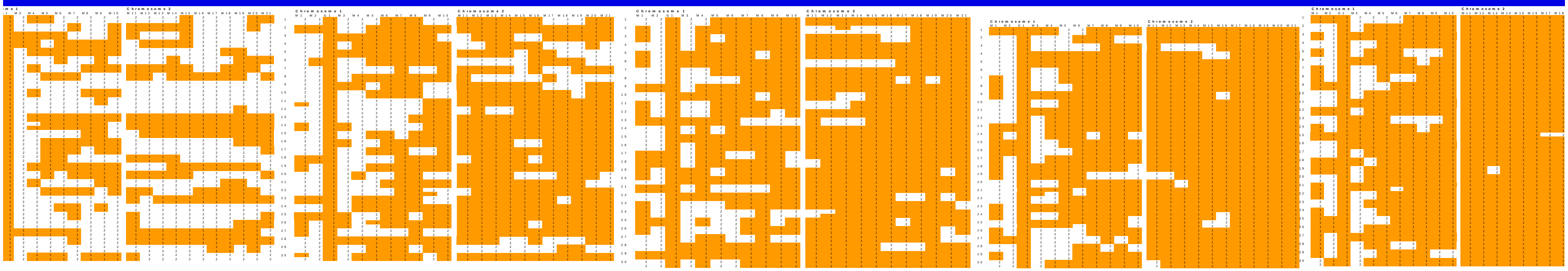
The best sister lines under each breeding objective and selection scheme

Parent	Objective	R0.04	R0.2E0.2	E0.2R0.2	E0.04
Westonia	High Rmax	3, 7	3, 7	3, 7	1, 3
	High Ext.	1	1, 5	1, 3, 5	1,3,5,7
Krichauff	High Rmax	1,3,5,7	1,3,5,7	3, 7	3, 7
	High Ext.	1,3,5,7	1,3,5,7	1, 5	1, 5
Machete	High Rmax	3,4,7,8	3,4,7,8	4, 8	None
	High Ext.	1,2,5,6	1,2,5,6	1, 2, 3	1,2,3,4
Diamondbird	High Rmax	1,2,3,4	1, 3, 4	3, 4	3, 4
	High Ext.	None	None	1,2,5,6	1,2,5,6

Recovery of recurrent parent in inbred lines derived from F2 and backcross populations

No background selection

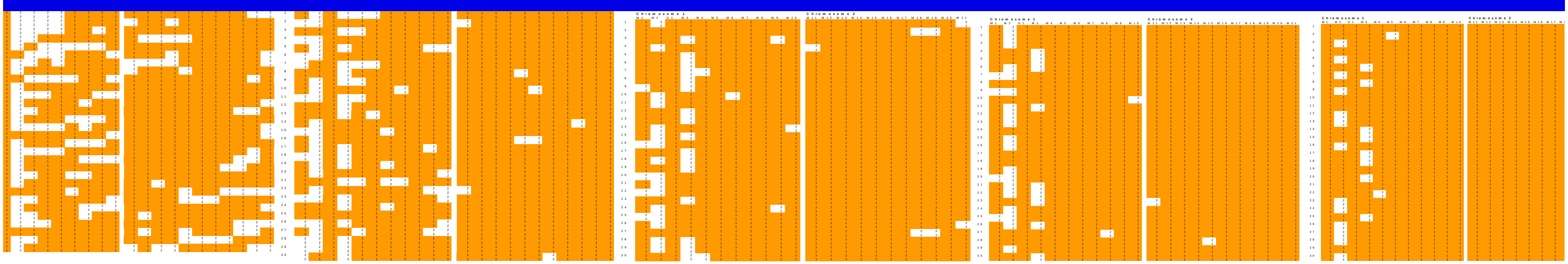
2-derived RILs | BC11-derived RILs | BC12-derived RILs | BC13-derived RILs | BC14-derived RIL



43.7% | 67.9% | 78.9% | 82.0% | 86.6%

20% selection based on background recovery

-derived RILs | BC11-derived RILs | BC12-derived RILs | BC13-derived RILs | BC14-derived RIL



70.5% | 87.0% | 92.4% | 94.5% | 96.5%

Genotype building

Gene symbol	*Rht-B1*	*Rht-D1*	*Rht8*	*Sr2*	*Cre1*	*VPM*	*Glu-B1*	*Glu-A3*	*tin*
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	0.8
Sunstate	*Rht-B1a*	***Rht-D1b***	*rht8*	***Sr2***	*cre1*	***VPM***	***Glu-B1i***	***Glu-A3b***	*Tin*
Silverstar+tin	***Rht-B1b***	*Rht-D1a*	*rht8*	*sr2*	***Cre1***	*vpm*	***Glu-B1i***	*Glu-A3c*	***tin***
HM14BS	*Rht-B1a*	*Rht-D1a*	***Rht8***	*sr2*	*cre1*	*vpm*	*Glu-B1a*	*Glu-A3e*	*Tin*
Target genotype	*Rht-B1a*	*Rht-D1a*	***Rht8***	***Sr2***	***Cre1***	***VPM***	***Glu-B1i***	***Glu-A3b***	***tin***

Tested various breeding strategies…

Related projects

- Related projects in conjunction with Graeme Hammer funded by ARC and GRDC in Australia are continuing development of simulation and analysis tools
- Already have
  - Link between QU-Gene and APSIM (Cropping system model)
  - Management system for experiment input/output and preliminary analysis
  - Computing clusters (50 to 100 processors) to undertake large experiments

Future plans in this project

- Implementing new linkage between QU-Gene and leaf elongation rate model
  - Already have prototype
  - Designing experiments to reproduce published research on QTL analysis
  - Other experiments to compare different ‘phenotypes’ that can be simulated…

Future plans in this project

- Prediction of residual genetic variation for quantitative traits that have not been selected using markers
  - Case study on coleoptile length (an important trait for drought environment) in wheat

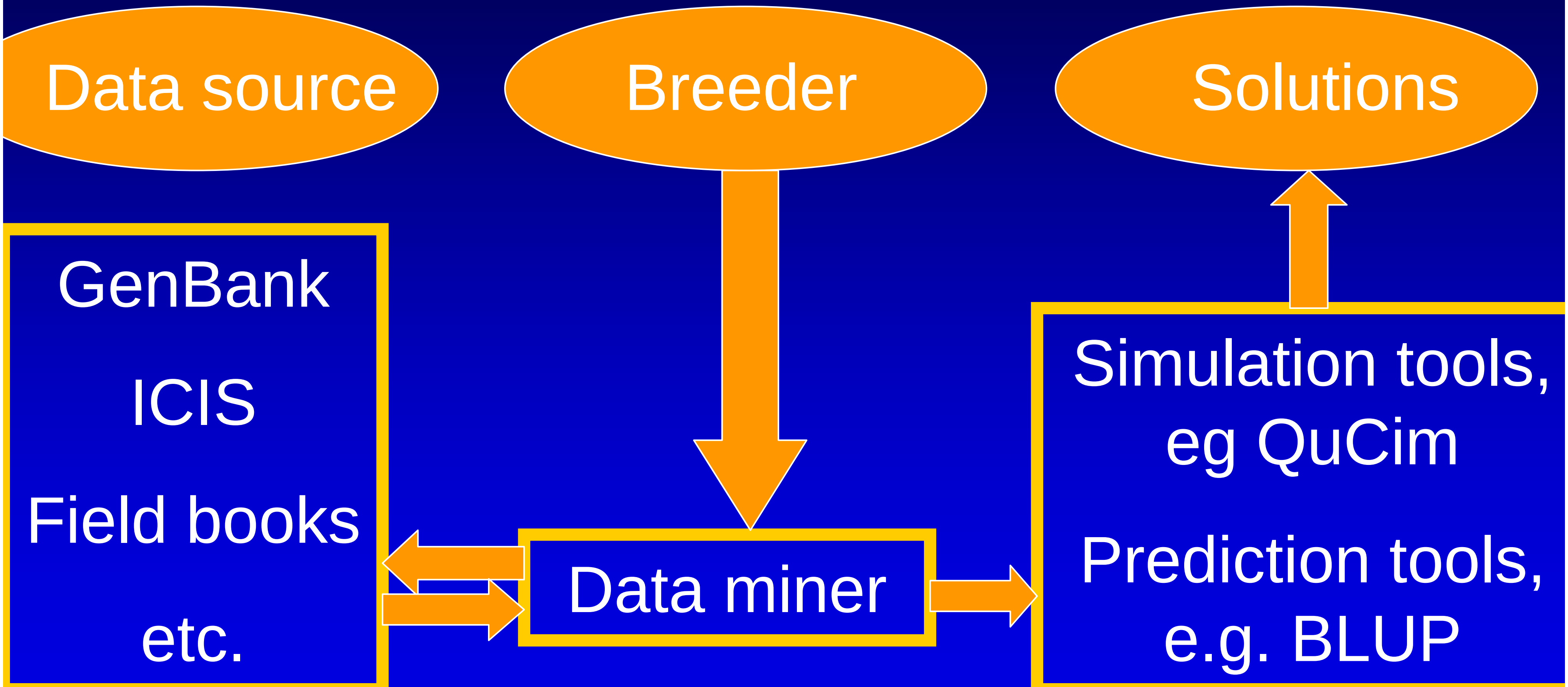
Future plans in this project

- Investigation of effects of ‘noise’ due to
  - GxE
  - Statistical error in phenotyping (low H2)
  - ‘unknown’ genes/QTLs affecting trait
  - Complexity of physiological model
    - (embedding Leaf model inside APSIM)

Future plans in this project

- Initial design of user-friendly interfaces that help
  - Build/define the genetic models
  - Build/define the crossing and selection methods

The Main Aim: An integrated decision-supported system for plant breeding



Acknowledgments

- Marker Cooper, Dean Podlich, Pioneer
- Mark Dieters, Guoyou Ye, Ian DeLacy, The University of Queensland
- Maarten van Ginkel, Richard Trethowan, Wolfgang Pfeiffer, Jose Crossa, CIMMYT
- Scott Chapman, David Bonnett, Greg Rubezke, CSIRO