Generation
Cultivating Plant Diversity for the Resource-Poor

HarvestPlus
Breeding Crops for Better Nutrition

Simulation Approaches in Plant Breeding

Jiankang Wang
http://www.isbreeding.net

CIMMYT

THE UNIVERSITY OF QUEENSLAND

GRDC
Grains Research & Development Corporation

Outlines for presentation

- Conventional quantitative genetics
- Plant breeding and quantitative genetics
- QTL mapping
- Background of breeding simulation
- QuLine tool for genetics and breeding
- QuLine applications
- Demonstration and Hands-on

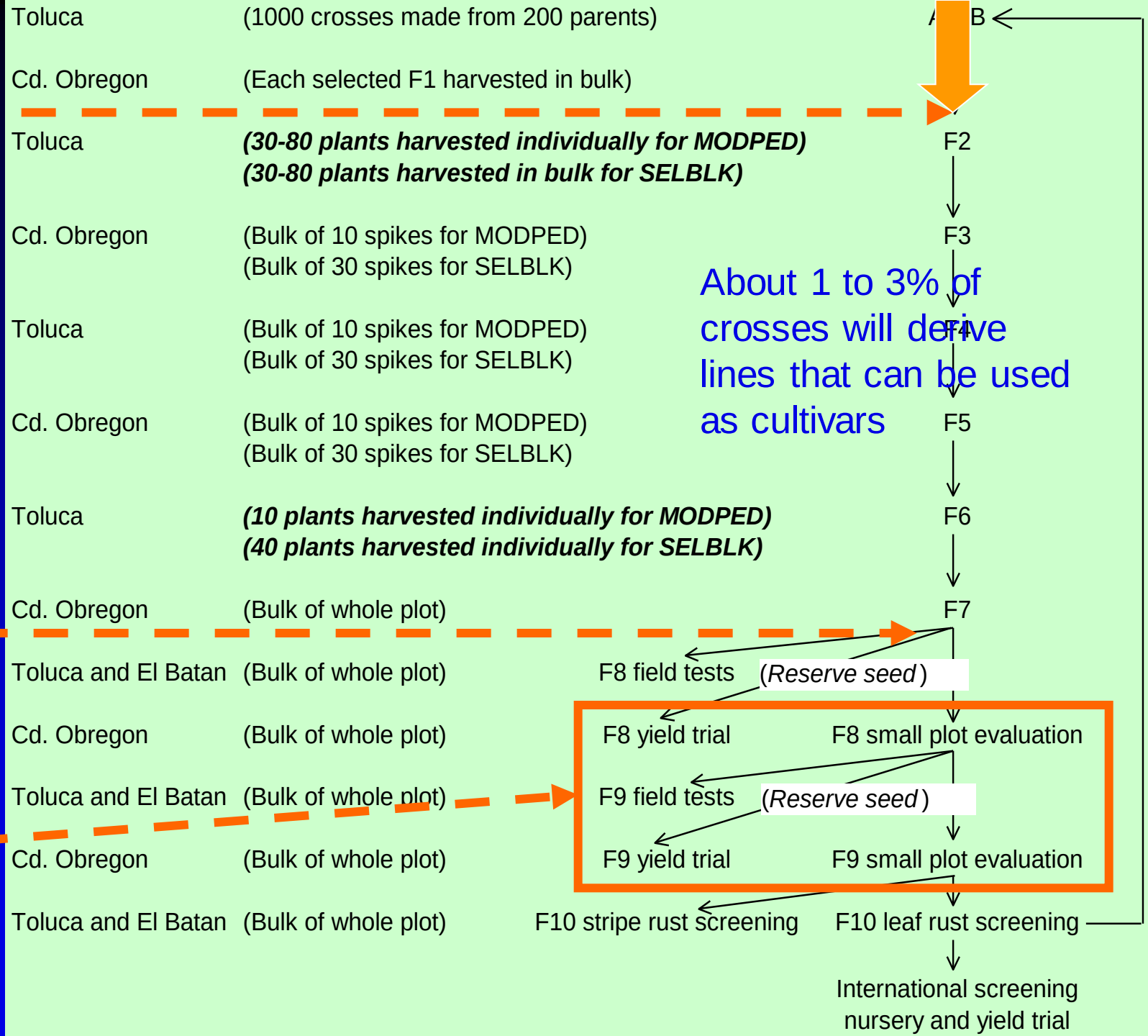
Background on breeding simulation

About 30% of the crosses are discarded in F1

PLANT BREEDING IS A COMPLEX PROCESS WITH MANY DECISIONS

About 40% of crosses retained after F7

About 20% of crosses retained after two rounds of yield trials



Why do we need tools in breeding?

- To improve the efficiency of traditional phenotypic selection through exploring various options
- To avoid the simplified assumptions made in classical quantitative genetic theory
- To better use the 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?

- To build a bridge between the biological data and breeders' requirements
- To combine all these sources of data into "knowledge" that breeders can use in their breeding programs

Questions that can be studied by QuLine: A genetic and breeding simulation tool

1. Comparison of breeding efficiencies from different selection strategies and their modifications. Which breeding method should be adopted?
2. Balance between the number of crosses and population size of segregating generations. What shall the breeder do if the available resources increase or reduce?
3. Evaluation of marker-assisted selection (MAS). When and how should MAS be used?
4. Comparative value of single, top, back, and double crosses in breeding?

Questions that can be studied by QuLine (mainly for inbred line development)

5. The correlation between parents and their offspring. Can F1 or F2 hybrids predict the performance of their advanced lines? For early generation selection, how early is too early?
6. When to use DH (doubled haploid)?
   - Early generation: many individuals need to be tested
   - Advanced line stage: lines will be good for other traits, but the desired genotype may be lost during selection due to population size, trait associations and genetic drift

QuLine tool for genetics and breeding

Available breeding simulation tools

- **QuLine**, a computer software that simulates breeding programs for developing inbred lines
- **QuHybrid**, a computer software that simulates breeding programs for developing hybrids
- **QuMARS**, a computer software that simulates marker-assisted recurrent selection and genome-wide selection

Test cross implemented in QuHybrid

Family	PopBefore(1)	PopBefore(2)	PopBefore(3)	…	Testers
Family size	5	3	6	…	3
	F11 F12 F13 F14 F15	F21 F22 F23	F31 F32 F33 F34 F35 F36	…	T1 T2 T3

Making testcrosses

Testcross 1	F11×**T1**	F12×**T1**	F13×**T1**	F14×**T1**	F15×**T1**	F21×**T1**	F22×**T1**	F23×**T1**	F31×**T1**	F32×**T1**	F33×**T1**	F34×**T1**	F35×**T1**	F36×**T1**
Testcross 2	F11×**T2**	F12×**T2**	F13×**T2**	F14×**T2**	F15×**T2**	F21×**T2**	F22×**T2**	F23×**T2**	F31×**T2**	F32×**T2**	F33×**T2**	F34×**T2**	F35×**T2**	F36×**T2**
Testcross 3	F11×**T3**	F12×**T3**	F13×**T3**	F14×**T3**	F15×**T3**	F21×**T3**	F22×**T3**	F23×**T3**	F31×**T3**	F32×**T3**	F33×**T3**	F34×**T3**	F35×**T3**	F36×**T3**

Among family selection

Testcross 1	F11×**T1**	F12×**T1**	F13×**T1**	F14×**T1**	F15×**T1**	F21×**T1**	F22×**T1**	F23×**T1**	F31×**T1**	F32×**T1**	F33×**T1**	F34×**T1**	F35×**T1**	F36×**T1**
Testcross 2	F11×**T2**	F12×**T2**	F13×**T2**	F14×**T2**	F15×**T2**	F21×**T2**	F22×**T2**	F23×**T2**	F31×**T2**	F32×**T2**	F33×**T2**	F34×**T2**	F35×**T2**	F36×**T2**
Testcross 3	F11×**T3**	F12×**T3**	F13×**T3**	F14×**T3**	F15×**T3**	F21×**T3**	F22×**T3**	F23×**T3**	F31×**T3**	F32×**T3**	F33×**T3**	F34×**T3**	F35×**T3**	F36×**T3**

Within family selection

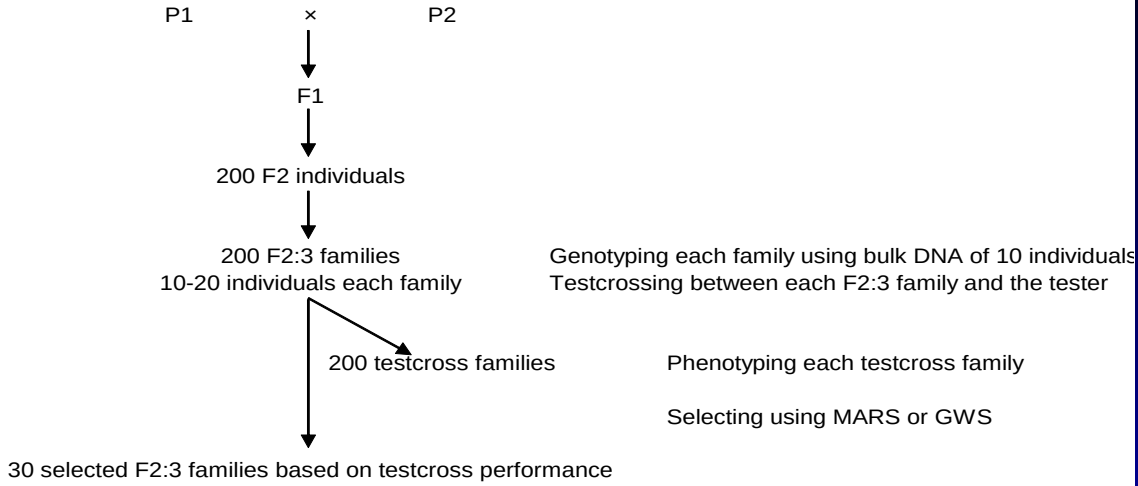
						F21×**T1**	F22×**T1**	F23×**T1**	F31×**T1**	F32×**T1**	F33×**T1**	F34×**T1**	F35×**T1**	F36×**T1**
						F21×**T2**	F22×**T2**	F23×**T2**	F31×**T2**	F32×**T2**	F33×**T2**	F34×**T2**	F35×**T2**	F36×**T2**
						F21×**T3**	F22×**T3**	F23×**T3**	F31×**T3**	F32×**T3**	F33×**T3**	F34×**T3**	F35×**T3**	F36×**T3**

Population after selection, this population will be used as the starting for the next generation

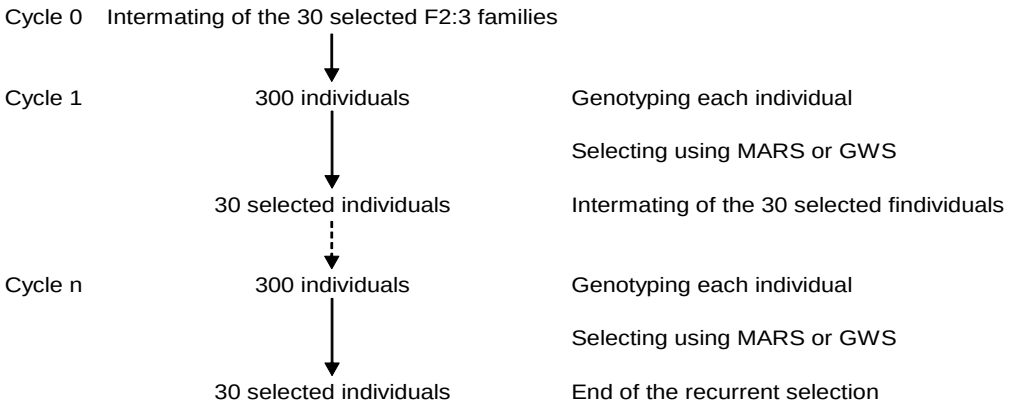
Family	PopAfter(1)	PopAfter(2)	PopAfter(3)
Family size	0	2	4
		F22 F23	F31 F33 F34 F36

The flowchart of QuMARS

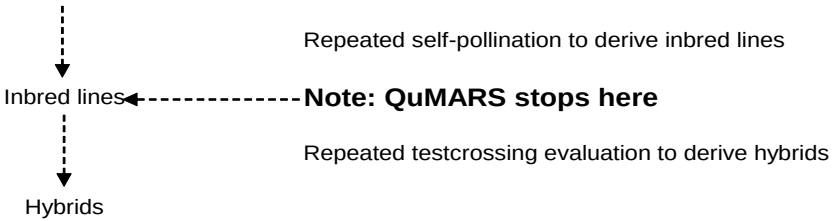
Stage I: Testcrossing: Starting from a single cross, To generate the initial population for recurrent selection; To build the prediction equation of breeding traits on markers



Stage II: Recurrent selection: To intermate the selected families (S0); To grow and genotype randomly mated progenies; To select individuals based on MARS or GWS



Stage III: Inbred development: To generate inbred lines through pedigree breeding



What can QuLine do?

- Comparison of genetic gains from different selection methods
  - Change in population mean
  - Change in gene frequency
  - Change in Hamming distance (distance of a selected genotype to the target genotype)
- Comparison of cross performance
  - Selection history
  - Rogers' genetic distance
  - Number of lines retained from each cross
- Comparison of cost efficiency
  - Number of families
  - Individual plants per generation
- Validation of theories

In genetics
(implemented by the QU-GENE engine)

- Most genetic phenomena, if not all, can be defined in the QU-GENE engine input file (QUG).
- Among them are:
  - Multiple alleles (e.g. Glutenin genes in wheat)
  - Linkage (between gene and marker, between genes, between markers)
  - Additive, dominance and epistasis
  - Pleiotropy (one gene effects multiple traits)
  - Genotype by environment interaction
  - Molecular markers (dominant, or co-dominant)

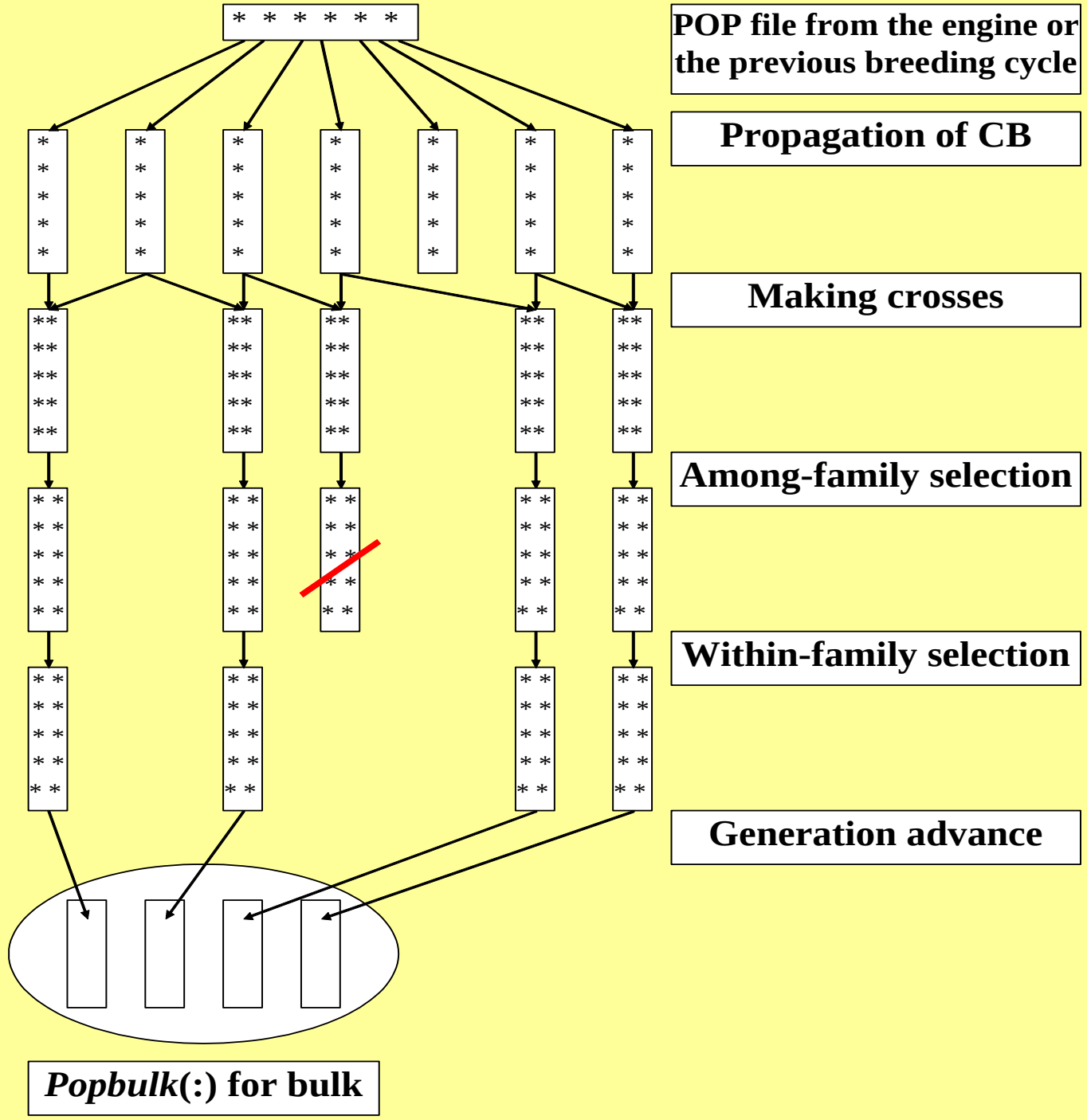
In breeding
(implemented by the QuLine module)

- Most, if not all, breeding methods for self-pollinated crops, can be defined and then simulated in QuLine.
- Among them are:
  - Pedigree system (including SSD)
  - Bulk-population
  - Doubled haploid
  - Marker-assisted selection (include marker-based selection)
  - Recurrent selection within one population
  - Many modifications and combinations

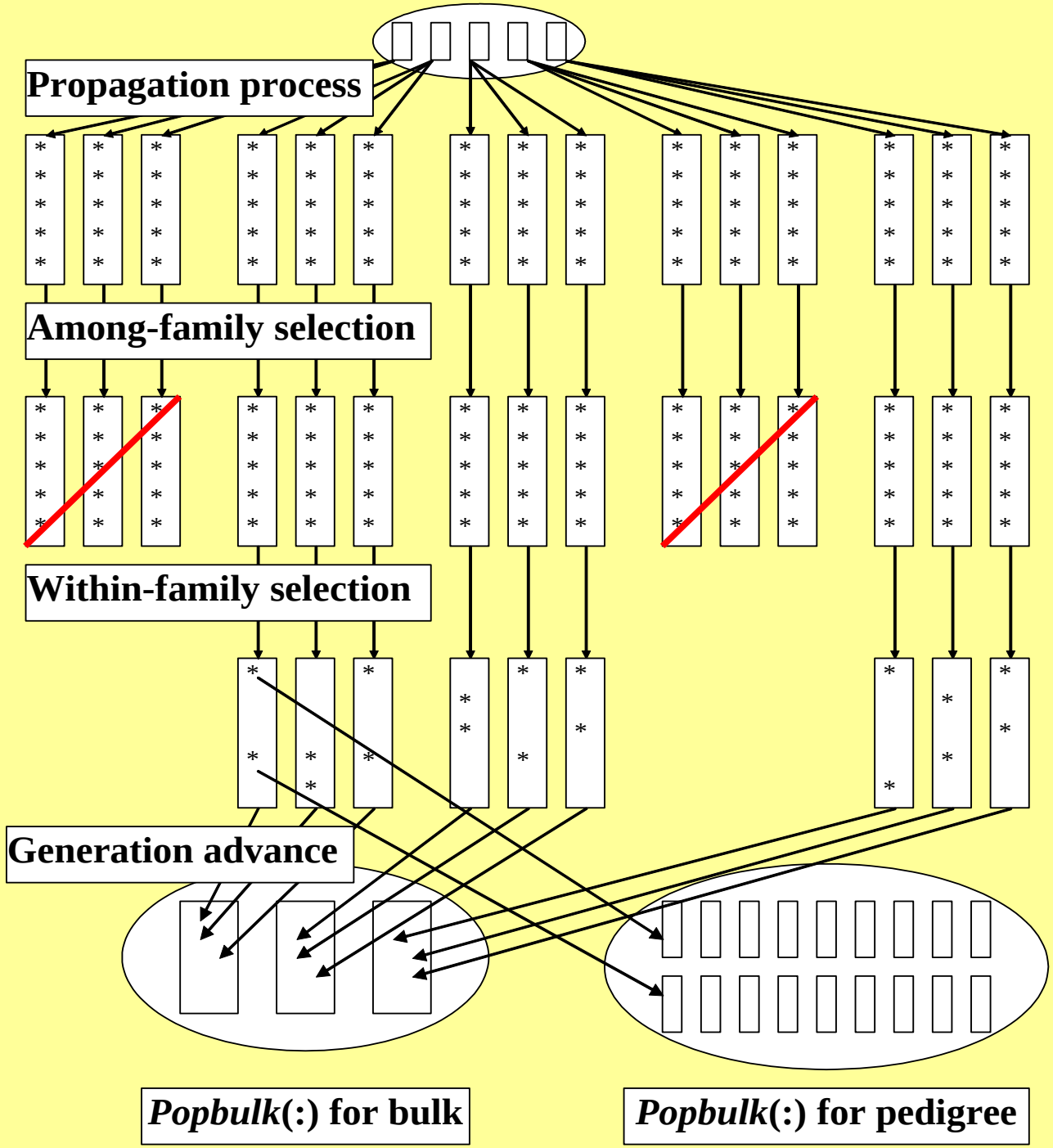
How does QuLine work?

- Two input files are needed
  - ***QUG file*** containing the necessary information for a genotype and environment (GE) system and initial population(s) of genotypes. It is the input for the QU-GENE engine. Two kinds of output files will be generated from the engine.
    - GES file for defining a GE system (= input for QuLine)
    - POP file for defining the initial population (= input for QuLine)
  - ***QMP file*** containing the necessary information for the breeding strategies to be simulated (e.g. pedigree, bulk, SSD, DH, etc.) (= input for QuLine)

General procedure for crossing and selection in F1



General procedure for propagation and selection in F2 and onwards



Define the QMP file for the selected bulk selection method: an example

Define the breeding method in a way that the computer can understand

General simulation parameters

- Number of runs: any integer
- Number of breeding cycles: any integer
- Number of crosses in F1 generation: any integer
- Indicator for crossing block update, 0 or 1.
  - 0: All individuals after a breeding cycle are used as the parents for the next cycle;
  - 1: The best individuals in the final selected population and the initial crossing block are selected as the parents for next cycle

General simulation parameters

- Indicator for outputting GE system details, 0 for no output, and 1 for output
- Indicator for outputting population details, 0 for no output, and 1 for output
- Indicator for outputting selection history, 0 for no output, and 1 for output
- Indicator for outputting Rogers distance for each cross and the lines retained from each cross, 0 for no output, and 1 for output
- Indicator for outputting correlation coefficient, 0 for no output, and 1 for output

The number of models in the GE system and the number of runs for breeding strategy

- Four nested loops in QuLine
  - Loop for all models

    All random effects in the GE system will be assigned values

    Determine the parents for all crosses for the first cycle
    - Loop for all runs
      - ✓Loop for all strategies
        - »Loop for all cycles
- All strategies start from the same point (same crossing block and same crosses), so that they can be properly compared.

Parameters to describe a set of breeding strategies

- Number of strategies
- For each strategy
  - Strategy name: any character
  - Number of generations in the strategy: any integer more than 0
  - Definition for each generation

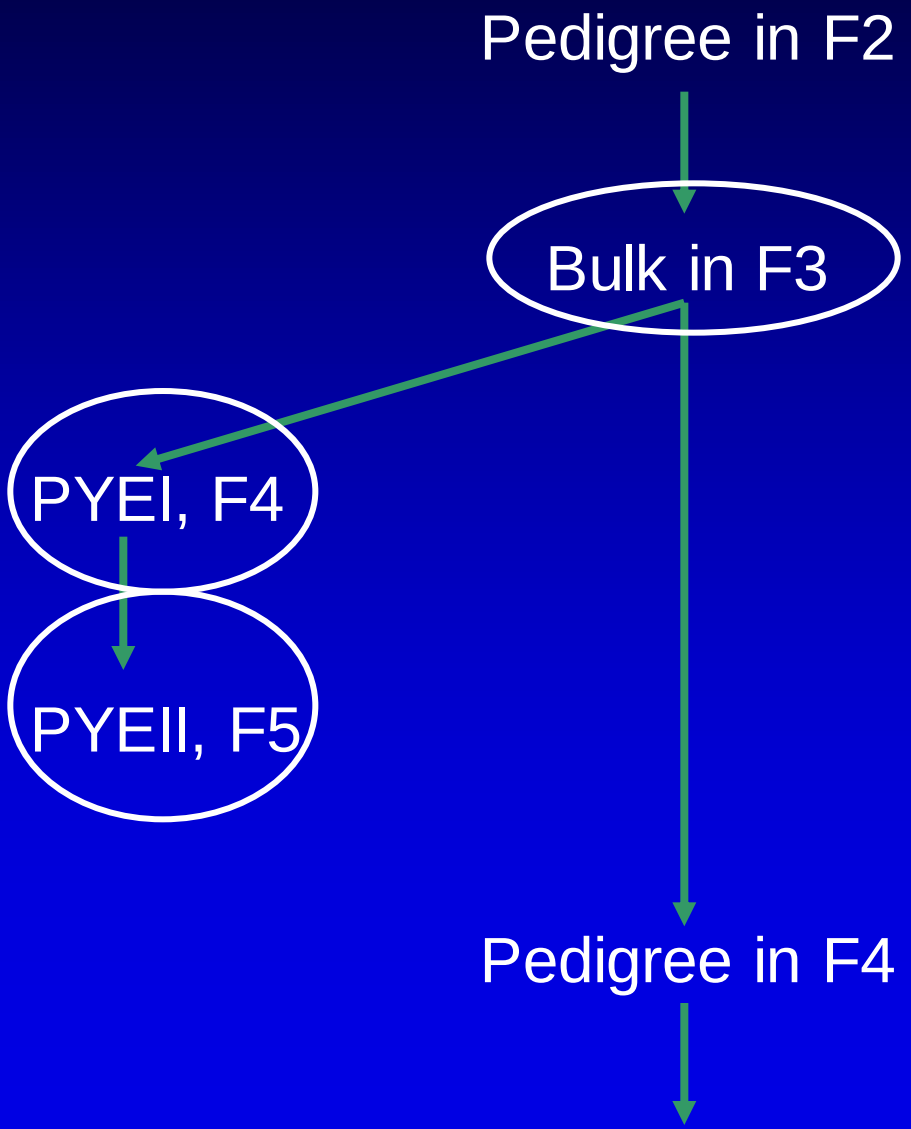
Definition of a generation

- Number of selection rounds in the generation: any integer more than 1
- Seed source indicator
  - 0: Seed for selection round *i* (*i* > 1) come from round 1
  - 1: Seed for selection round *i* (*i* > 1) come from round *i*-1
- Definition of each selection round

An example for seed source indicator 0

Practical breeding	Virtual breeding	Families selected
F6 field test at Toluca	F6	452 → 408 → 14,760
F7 field test at Cd. Obregon	F7	*Round 1 for F7* 14,760 → 3,868
F8 field test at Toluca and El Batan	F8 (T), F8 (B)	*Rounds 2 and 3* 3,868 → 2,163 → 1,974
F8 yield trial and F8 small plot evaluation at Cd. Obregon	F8 (YT)	*Round 4* 1,974 → 779
	F8 small plot	**779**

An example (LRC, Toowoomba, Australia) for seed source indicator 1



Families selected

Round 1 for F3

4000 → 1200

Round 2

1200 → 600

Round 3

600 → 100

100 → 1000

Definition of each selection round

- Title for the generation
- Seed propagation type (within a family)
  - *clone*, asexual reproduction
  - *DH*, doubled haploid
  - *self*, self-pollination
  - *backcross*, backcrossed to one parent
  - *topcross*, crossed with a third parent (three-way cross)
  - *doublecross*, crossed with another F1
  - *random*, random mating
  - *noself*, random mating but self-pollination is eliminated

Definition of each selection round

- Generation advance method (or harvest method): management of the selected individual plants in a family
  - *pedigree*: the selected plants in each family are harvested individually, resulting a few families in the next generation
  - *bulk*: the selected plants in each family are harvested in bulk, resulting one family in the next generation

Definition of each selection round

- Field experiment design
  - Number of replications for each family
  - Number of plants in each replication
  - Number of test locations
  - Environment type for each test location
    - defined in the GE system
    - if 0, randomly determined based on environment frequency

Definition of each selection round

- Among family and within family selection
  - Number of traits used for selection
  - Definition of each trait

Definition of each trait used in selection

- Trait number, for the trait in selection (0 when marker score is used in selection)
- Selection mode
  - **T for top, e.g. yield, tillering, grains per spike and 1000-kernel weight**
  - **B for bottom, e.g. lodging and rusts**
  - **M for middle, e.g. height and heading**
  - **R for random, for some special studies**
  - **TV for top value**
  - **BV for bottom value**
  - **TN for a number of individuals/families with top phenotypic values**
  - **BN for a number of individuals/families with bottom phenotypic values**
  - **RN for a number of individuals/families to be selected randomly**
- Selected proportion or value: the proportion or value of individual plants in a family (for within family selection) or of families (for among family selection) to be selected

Proportion selection

Threshold selection

Number selection

An example of generation definition

Rounds of selection	Seed source indicator	Generation title	Seed propagation type	Generation advance method	Replications	Plot size	Test locations	Environment type
1	0	F6	self	pedigree	1	750	1	2, Toluca
4	0	F7	self	bulk	1	70	1	1, Obregon
		F8(T)	self	bulk	1	70	1	2, Toluca
		F8(B)	self	bulk	1	70	1	3, El Batan
		F8(YT)	self	bulk	1	100	1	1, Obregon
1	0	F8(SP)	self	bulk	1	30	1	1, Obregon

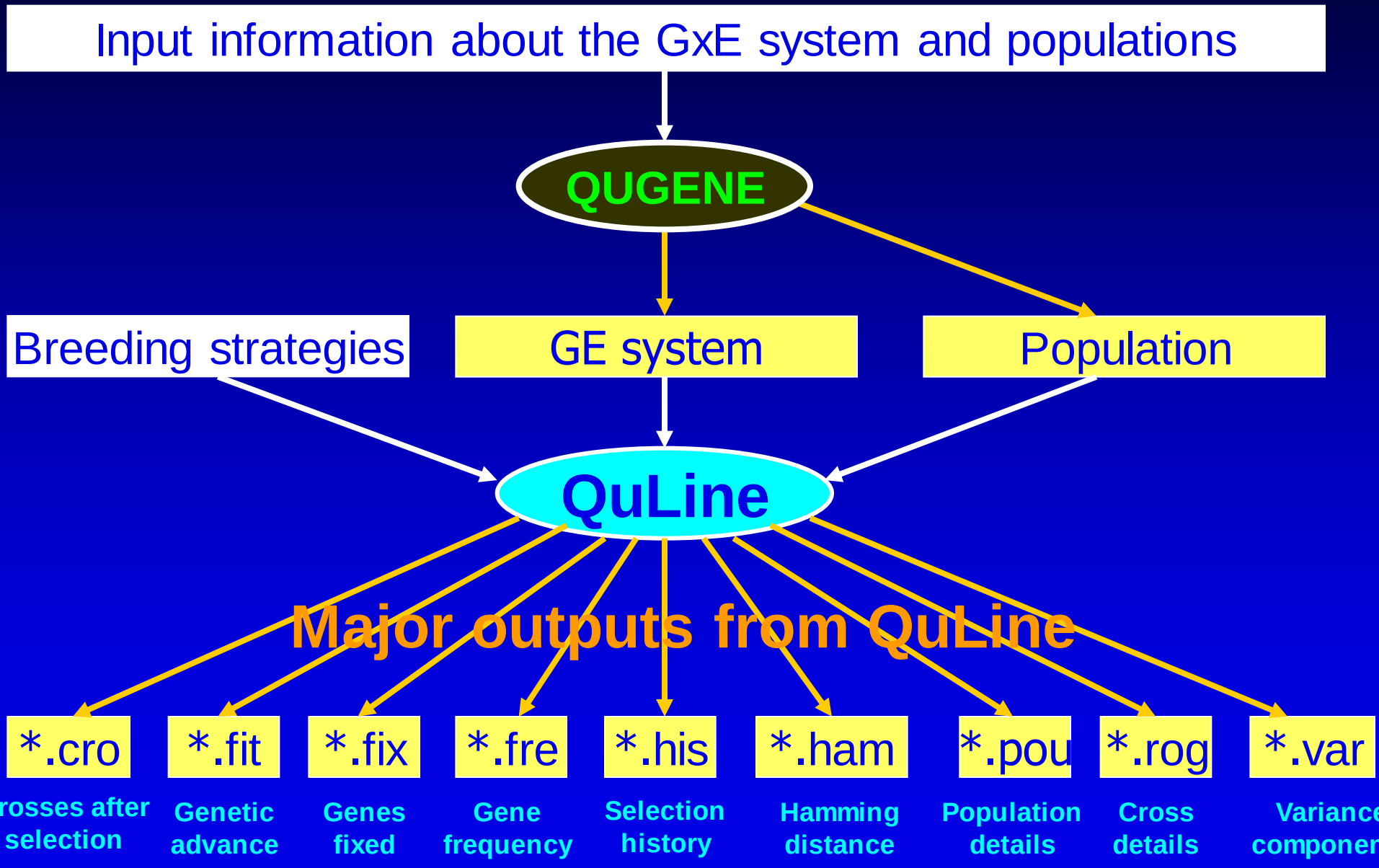
Traits, their selection modes and selected proportions

Trait	Yield	Lodg-ing	Stem rust	Leaf rust	Stripe rust	Height	Tillering	Heading	Grains per spike	1000 kernel weight	**Total**
Selection mode	**T**	**B**	**B**	**B**	**B**	**M**	**T**	**M**	**T**	**T**	
F6, among		**0.99**			**0.96**		**0.95**				**0.90**
F6, within		**0.90**			**0.70**	**0.90**	**0.95**	**0.98**	**0.10**		**0.05**
F7, among		**0.85**		**0.70**		**0.98**	**0.85**	**0.96**	**0.70**	**0.75**	**0.25**
F8(T), among		**0.55**			**0.70**	**0.99**	**0.98**	**0.99**	**0.90**		**0.55**
F8(B), among				**0.90**							**0.90**
F8(YT), among	**0.40**										**0.40**
F8(SP), among											**1.00**

In QuLine, a breeding program looks like …

```
!NR SS GT        PT         GA         RP PS    NL  ET...              Row 1
!                             AT (ID SP  SM)...                       Row 2
!                             WT (ID SP  SM)...                       Row 3
1   0  CB        self       bulk       1  10    1   2
                              0
                              0
1   0  F1     singlecross bulk        1  20    1   1
                              7   2 B 0.98  3 B 0.99  4 B 0.85  6 M 0.99  7 T 0.90  8 B 0.98  9 T 0.97
                              0
1   0  F2        self       pedigree   1  1000  1   2
                              7   2 B 0.99  3 B 0.99  5 B 0.90  6 M 0.99  7 T 0.99  8 B 0.99  9 T 0.99
                              8   2 B 0.95  4 B 0.99  5 B 0.40  6 M 0.85  7 T 0.60  8 B 0.90  9 T 0.50  10 T 0.60
1   0  F3        self       bulk       1  70    1   1
                              7   2 B 0.90  3 B 0.99  4 B 0.70  6 M 0.97  7 T 0.75  8 B 0.95  9 T 0.80
                              5   4 B 0.90  6 M 0.95  8 B 0.95  9 T 0.30 10 T 0.60
1   0  F4        self       bulk       1  70    1   2
                              6   2 B 0.90  5 B 0.65  6 M 0.97  7 T 0.85  8 B 0.97  9 T 0.85
                              5   5 B 0.90  6 M 0.95  8 B 0.95  9 T 0.30 10 T 0.60
1   0  F5        self       bulk       1  70    1   1
                              6   2 B 0.90  4 B 0.75  6 M 0.97  7 T 0.85  8 B 0.95  9 T 0.85
                              5   4 B 0.90  6 M 0.95  8 B 0.95  9 T 0.30 10 T 0.60
1   0  F6        self       pedigree   1  140   1   2
                              6   2 B 0.90  5 B 0.75  6 M 0.97  7 T 0.85  8 B 0.97  9 T 0.85
                              5   5 B 0.90  6 B 0.90  7 T 0.95  8 B 0.95  9 T 0.10
4   0  F7        self       bulk       1  70    1   1
                              7   2 B 0.90  4 B 0.75  6 M 0.97  7 T 0.90  8 B 0.95  9 T 0.85 10 T 0.75
                              0
       AL(T)     self       bulk       1  70    1   2
                              6   2 B 0.95  5 B 0.90  6 M 0.99  7 T 0.98  8 B 0.99  9 T 0.85
                              0
       AL(B)     self       bulk       1  70    1   3
                              1   4 B 0.90
                              0
       PYT       self       bulk       1  100   1   1
                              1   1 T 0.40
                              0
```

Steps to run QuLine

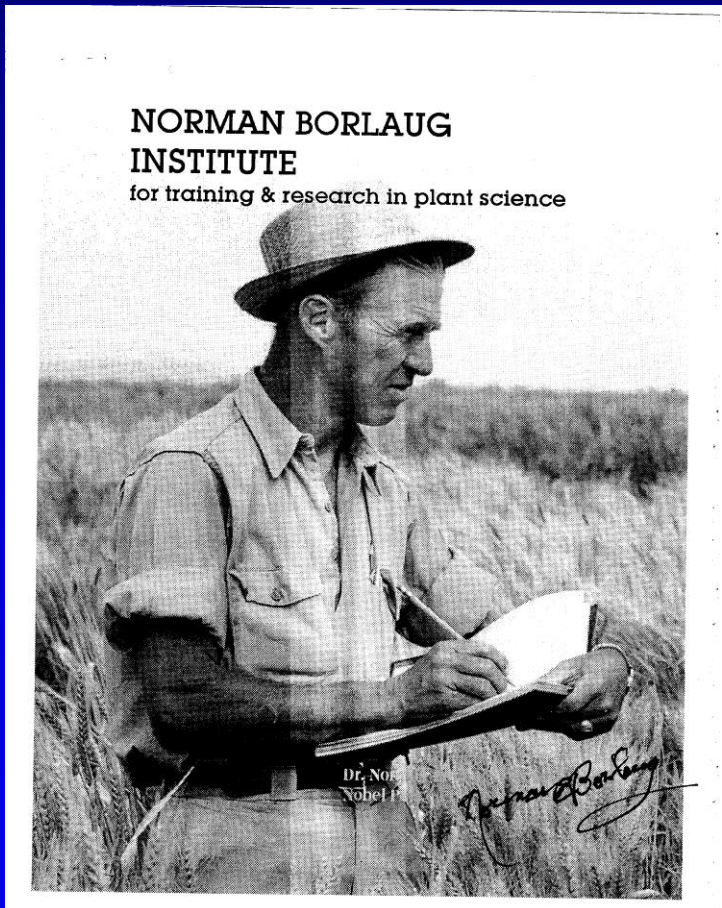


Comparison of two breeding strategies: modified pedigree (MODPED) and selected bulk (SELBLK)

Crop Science, 2003, 43: 1764-1773
Crop Science, 2004, 44: 2006-2018

Dr. Borlaug and the Green Revolution

- Mexican-Rockefeller Foundation Agricultural Program, 1944
- Mexico became self-sufficient in most food commodities by the late 1950s
- Young scientists were trained in Mexico
- Semidwarf wheats (*Rht1* gene and Norin 10)
- 1965-1975, wheat and rice production had increased by 50%
- Awarded Nobel Peace Prize in 1970



CIMMYT's Shuttle Breeding

Breeding methods with self-pollinated crops

Allard, R.W. 1960. Principles of plant breeding. John Wiley & Sons, Inc.

Stoskopf, N.C. 1993. Plant Breeding — Theory and Practice. Westview Press.

- Mass and pure-line selection
- The pedigree system
- The bulk population method
- The backcross breeding method
- Single seed descent (SSD), a special case of pedigree system
- Recurrent selection breeding method
- Mutation breeding
- Haploid breeding system (doubled haploid)

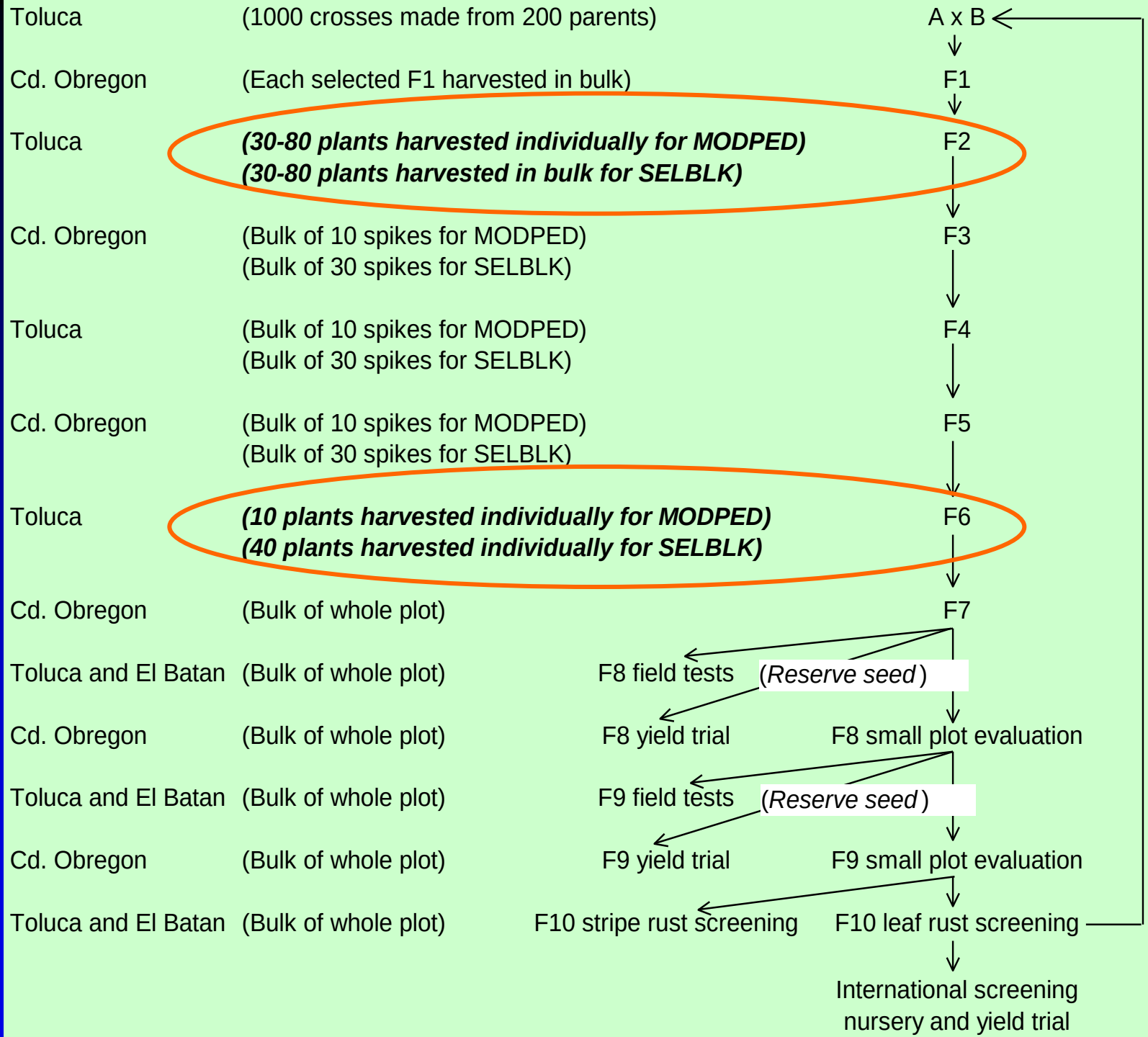
Breeding methods in CIMMYT's wheat breeding program

- Pedigree system: before 1984.
  - "Pedigree selection" is used from F2 to F6.
- Modified pedigree/bulk (MODPED): in 1985-1989/94.
  - "Pedigree selection" is used in F2 and F6, and "bulk selection" is used in other generations.
- Selected bulk (SELBLK): after 1995.
  - "Pedigree selection" is used only in F6, and "bulk selection" is used in other generations.

Germplasm flow for simple crosses made in Toluca and targeted to ME1

MODPED: modified pedigree/bulk

SELBLK: selected bulk methods



Trait, segregating gene number, gene effects and trait heritability

Trait	Genes	Gene effect type	AA	Aa	aa	Trait range	h_b^2 (Indiv. plant)
Yield	20, 40	E0, E1, E2	Random value from UD (0, 1)				0.05
Lodging	3	additive	0	5	10	0-30	0.10
Stem rust	5	additive	0	0.5	1	0-5	0.30
Leaf rust	5	additive	0	5	10	0-50	0.30
Yellow rust	5	additive	0	5	10	0-50	0.30
Height	3	additive	40	30	20	120-60	0.45
Tillers/plant	3	additive	5	3	1	15-3	0.35
Heading	5	additive	20	16	12	100-60	0.30
Grains/spike	5	additive	14	10	6	70-30	0.35
Seed weight	5	additive	12	8.5	5	60-25	0.35

Trait correlation and pleiotropy

Trait	Yield	Lodging	Stem rust	Leaf rust	Yellow rust	Height	Tillers/ plant	Heading	Grains/ spike	Seed weight
Yield		-0.50	-0.20	-0.10	-0.10	-0.50	0.40	0.30	0.50	0.40
Lodging	***-0.56***									
Stem rust	***-0.25***									
Leaf rust	***-0.05***									
Yellow rust	***-0.09***									
Height	***-0.62***									
Tillers/plant	***-0.08***								-0.20	-0.40
Heading	***0.60***									
Grains/spike	***0.09***						***-0.17***			-0.30
Seed weight	***-0.07***						***-0.30***		***-0.07***	

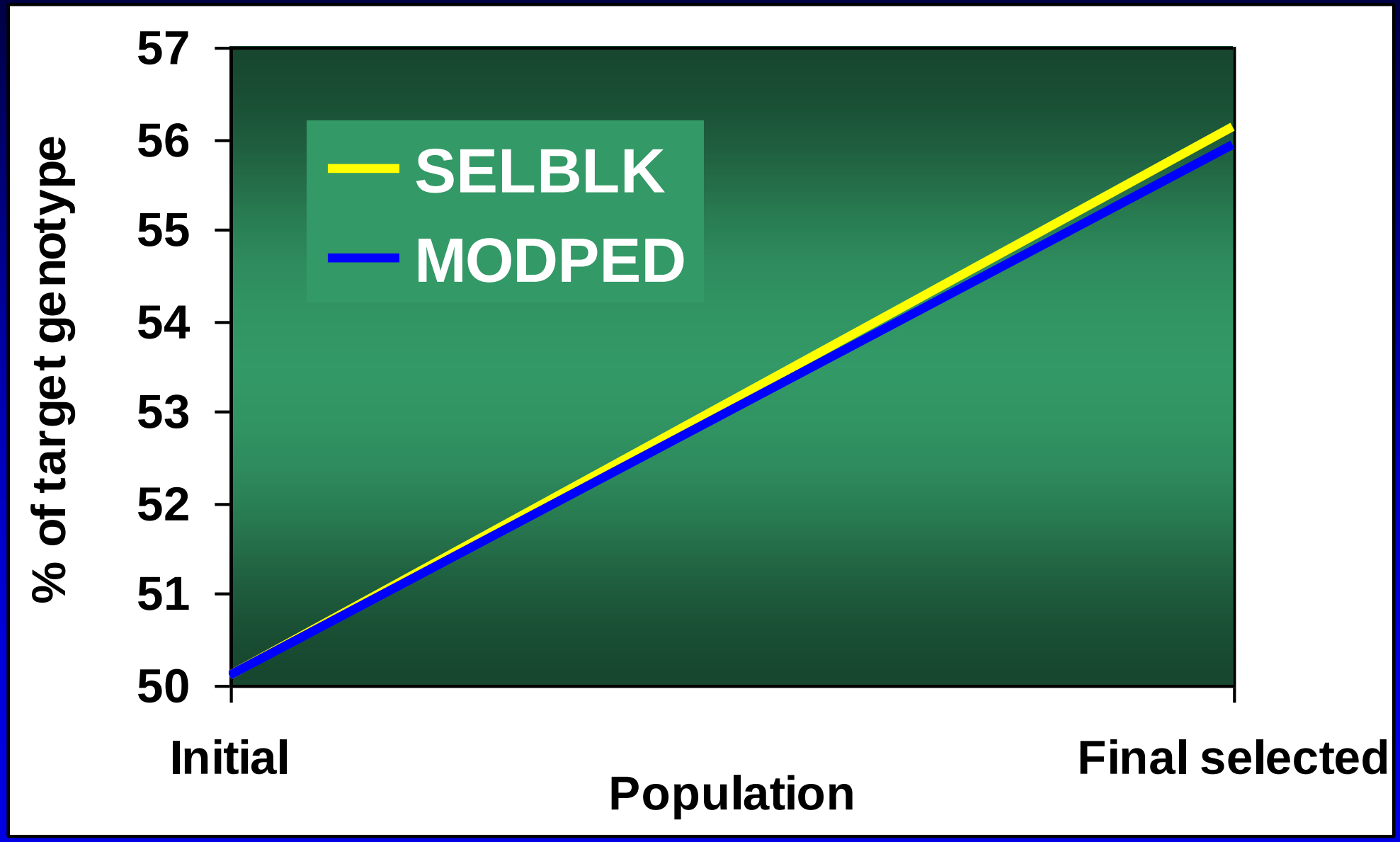
Estimated by CIMMYT breeders

Estimated from the defined genetic model

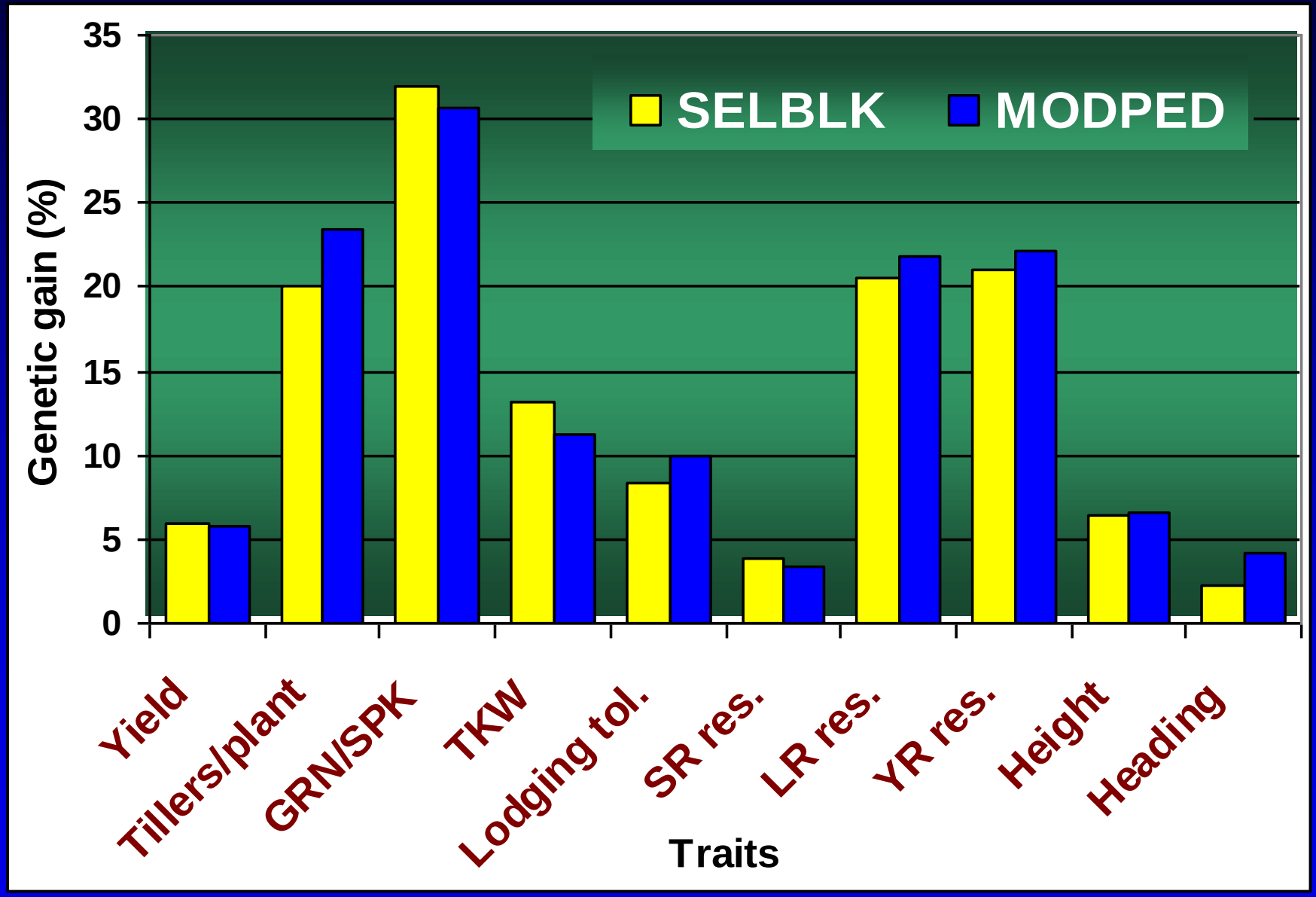
Experiment design

- 12 Genotype and environment (GE) systems
  - Two yield gene numbers: 20 and 40, two alleles for each gene
  - Pleiotropy (same gene effects various traits): absent and present
  - Epistasis (multiple gene interaction): no epistasis, digenic-epistasis, and tri-genic epistasis
  - Linkage: no linkage (independent gene segregation)
- Initial population
  - 200 parents, gene (allele) frequencies of 0.5 for all genes
- 1000 crosses were made
- 258 lines were selected after 10 generations of selection

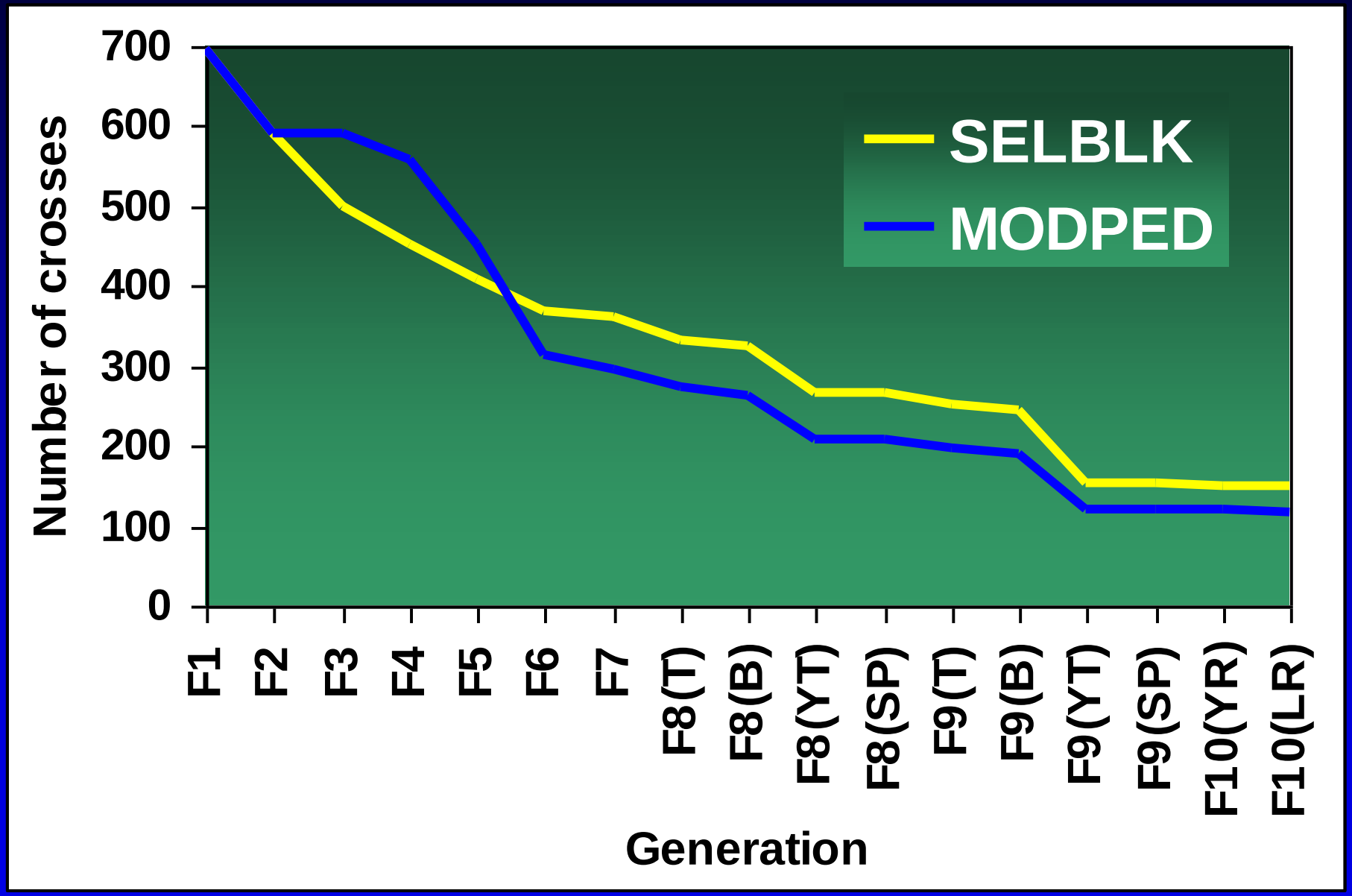
Result 1: Genetic gain in yield from SELBLK is 3.3% higher than MODPED. SELBLK is slightly more efficient.



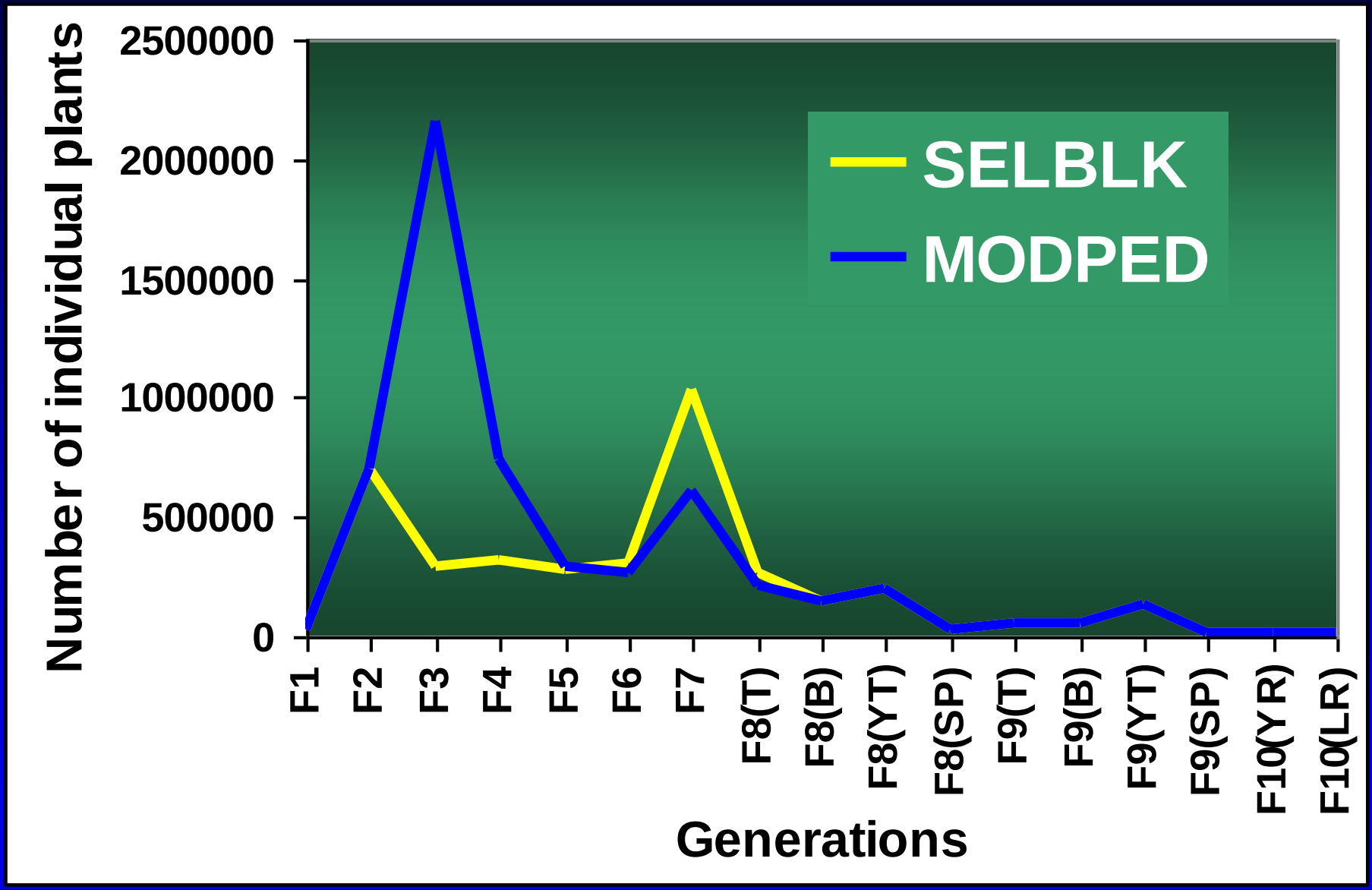
For gains per spike and 1000-kernel weight, SELBLK has a faster genetic gain. For tillers/plant, MODPED has a faster genetic gain.



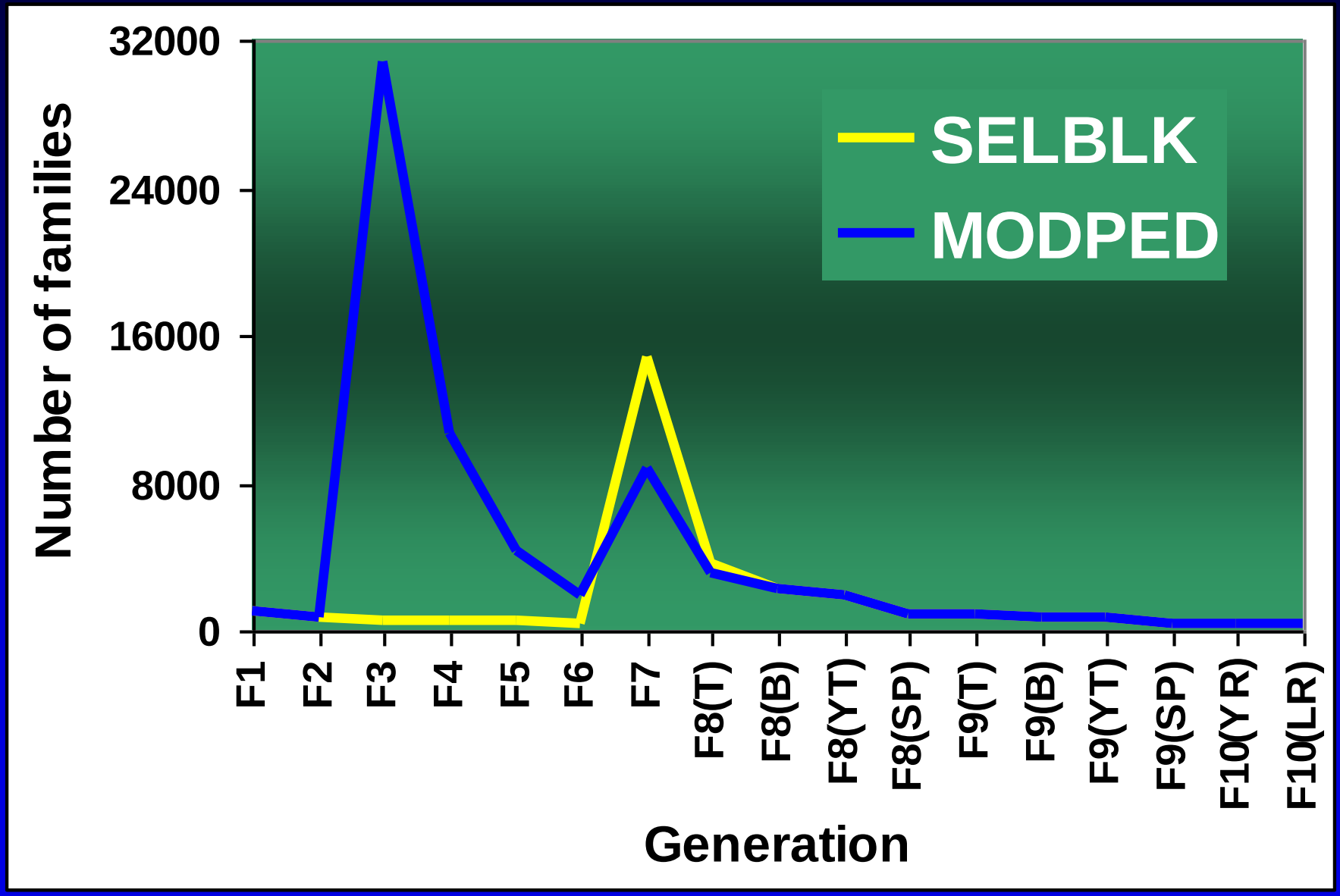
Result 2: SELBLK retained 25% more crosses in the final selected population (more genetic diversity retained)



Result 3: SELBLK required 1/3 less land from F1 to F8 than MODPED. SELBLK is more cost-effective.



Result 4: SELBLK produced 40% less families (plots) to be planted from F1 to F8 (less labor required)



Modeling of the Single Backcrossing Breeding Strategy (SBBS)

Theor. Appl. Genet., 2009, 118: 683-694

Ravi Singh, wheat breeder and pathologist in CIMMYT

- Adapted wheats X durable disease resistance donors
- 1-2 times of backcrossing with the adapted
- Large BCF2 population
- Select adapted lines combined with durable resistances

Estimated percentages of favourable alleles or gene combinations in different parental lines in wheat breeding at CIMMYT

Category	favorable genes (%)	Example	% total lines
Elite adapted lines (EAL)	80-85	Major released cultivars in targeted mega-environments (MEs) either developed by CIMMYT or by partners	10
Adapted lines (AL)	75-80	Elite advanced lines from CIMMYT's International Nursery and Yield Trials	60
Intermediate adapted lines (IAL)	65-75	Advanced lines from CIMMYT's Yield Trials in Ciudad Obregón and Toluca, Mexico	10
Un-adapted (or non-adapted) lines (UAL)	20-40	Land races	2
Second generation of re-synthesized wheat (SYNII)	40-60	Derived lines between the first generation of re-synthesized wheat derivatives and adapted lines	10
First generation of re-synthesized wheat (SYNI)	20-40	Derived lines between primary re-synthesized wheat and adapted lines	5
Primary re-synthesized wheat (SYN0)	0-30	Inter-specific crosses between *Triticum durum* and *Aegilops tauschii*	3

Two traits defined in QU-GENE

- Adaptation
  - 200 genes on the 21 wheat chromosomes
  - Lowest adaptation with no favorable alleles: 0
  - Highest adaptation with all favorable alleles: 100
  - Heritability: 0.5
- Donor traits (DT) to be transferred
  - 10 genes governing the donor traits
  - Lowest DT with no favorable alleles: 0
  - Highest DT with all favorable alleles: 10
  - Heritability: 0.5

Adapted parental groups in simulation

Gene frequency of favorable adaptation alleles fixed at 0.8

- A0: the frequency of favorable DT alleles is 0
- A2: the frequency of favorable DT alleles is 0.2
- A4: the frequency of favorable DT alleles is 0.4
- A6: the frequency of favorable DT alleles is 0.6
- A8: the frequency of favorable DT alleles is 0.8

Donor parental groups in simulation

Gene frequency of favorable DT alleles fixed at 1.0

- D0: the frequency of favorable adaptation alleles is 0
- D1: the frequency of favorable adaptation alleles is 0.1
- D2: the frequency of favorable adaptation alleles is 0.2
- D3: the frequency of favorable adaptation alleles is 0.3
- D4: the frequency of favorable adaptation alleles is 0.4
- D5: the frequency of favorable adaptation alleles is 0.5
- D6: the frequency of favorable adaptation alleles is 0.6
- D7: the frequency of favorable adaptation alleles is 0.7

Crosses made between different parental groups for wheat breeding at CIMMYT

Category	Percentage of total crosses	Similarity to defined parental groups
(EAL+AL) × (EAL+AL)	65	(A0+A2+A4+A6+A8) ×D7
(EAL+AL) × IAL	10	(A0+A2+A4+A6+A8) ×D5+D6)
(EAL+AL) × UAL	5	(A0+A2+A4+A6+A8) ×(D2+D3+D4)
(EAL+AL) × SYNII	10	(A0+A2+A4+A6+A8) ×(D6+D7)
(EAL+AL) × SYNI	7	(A0+A2+A4+A6+A8) ×(D4+D5)
(EAL+AL) × SYN0	3	(A0+A2+A4+A6+A8) ×(D0+D1+D2)

The single backcrossing breeding strategy (SBBS)

Generation	Seed propagation method	No. crosses or families grown	Individuals per cross or family	No. selected crosses or families	No. selected individuals in each cross or family
F_1	Hand pollination between adapted and donor lines	100	20	100	20
B_1F_1	Backcrossing to the adapted parents	100	400	100	50
B_1F_2	Selfing	100	1200	100	30
B_1F_3	Selfing	100	400	100	10
B_1F_4	Selfing	100	400	100	10
B_1F_5	Selfing	100	400	100	10
B_1F_6	Selfing	1000	200	30	200
Final selected advanced lined		10			

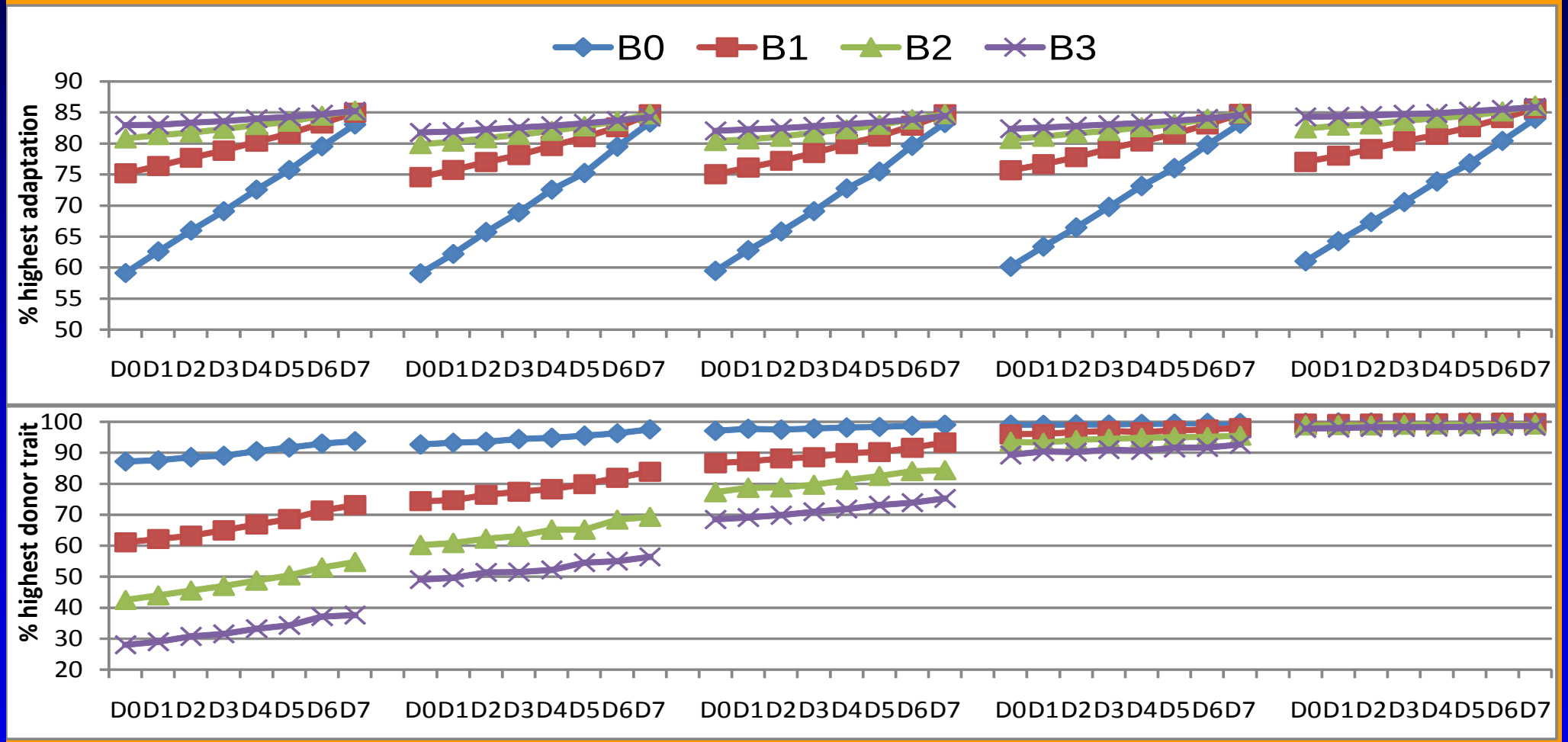
Four crossing strategies defined in QuLine

No backcross	One time of backcross	Two times of backcross	Three times of backcross	Generation advance method
B0	B1	B2	B3	
F_1	F_1	F_1	F_1	Bulk
F_2	BC_1F_1	BC_1F_1	BC_1F_1	Bulk
F_3	BC_1F_2	BC_2F_1	BC_2F_1	Bulk
F_4	BC_1F_3	BC_2F_2	BC_3F_1	Bulk
F_5	BC_1F_4	BC_2F_3	BC_3F_2	Bulk
F_6	BC_1F_5	BC_2F_4	BC_3F_3	Pedigree
F_7	BC_1F_6	BC_2F_5	BC_3F_4	Bulk

Six selection schemes

- **AD:** Adaptation is selected first, followed by the selection for DT
- **DA:** DT is selected first, followed by the selection for Adaptation
- **ADA:** Adaptation is selected first, followed by the selection for DT, and adaptation is selected again
- **DAD:** DT is selected first, followed by the selection for Adaptation, and DT is selected again
- **ADAD:** Adaptation and DT are selected two times in each generation, and adaptation is selected first
- **DADA:** Adaptation and DT are selected two times in each generation, and DT is selected first

Genetic advance of selection scheme AD



Conclusions

- We recommend the use of SBBS based on three assumptions:
  - multiple genes governing the phenotypic traits to be transferred from donor parents to adapted parents
  - donor parents still have some favorable genes that may contribute to the improvement of adaptation in the recipient parents even under low adaptation
  - the conventional phenotypic selection is applied or the individual genotypes cannot be precisely indentified

Breeding with known gene information

Aust. J. Agric. Res., 2005, 56: 465-473

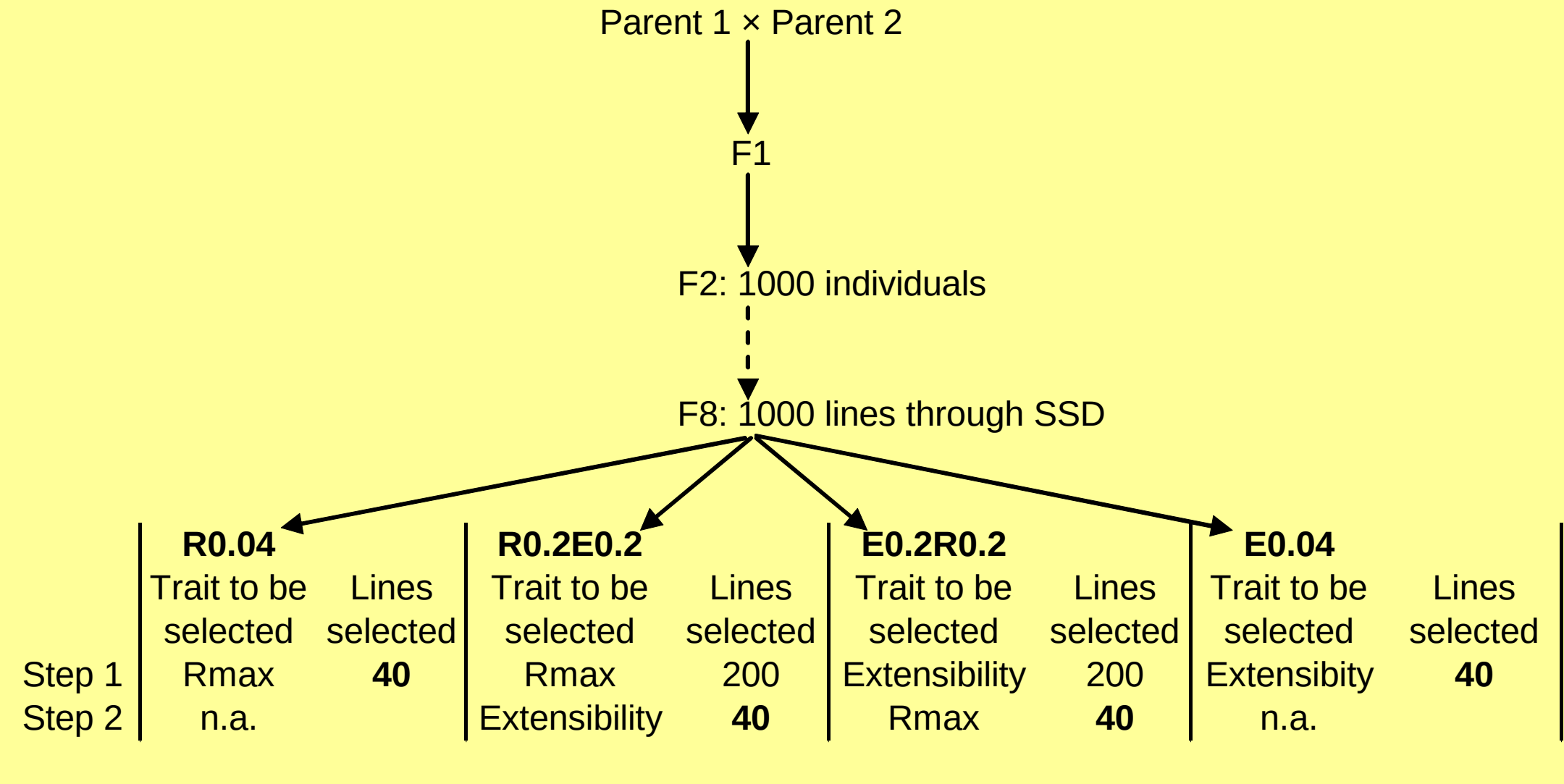
Glutenin genes and wheat quality

- 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*

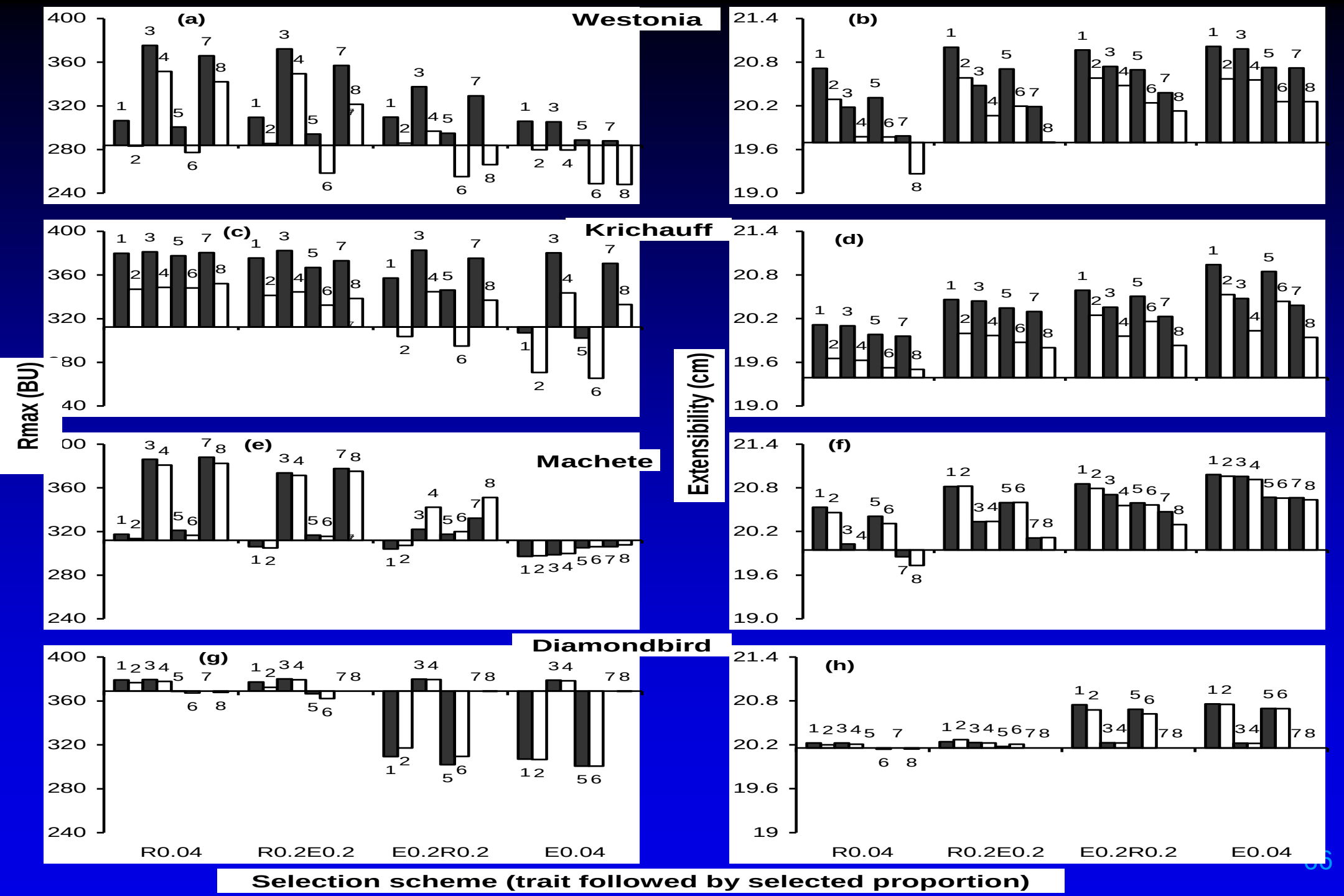
Selected parents

Parent	Rmax	Extensibility
Silverstar 1	309.80	20.78
Silverstar 2	270.20	20.31
…	…	…
Silverstar 7	368.90	20.15
Silverstar 8	329.30	19.68
Westonia	283.70	19.70
Krichauff	312.26	19.39
Machete	312.03	19.95
Diamondbird	368.88	20.16

Four selection schemes



	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 selected
Step 1	Rmax	**40**	Rmax	200	Extensibility	200	Extensibity	**40**
Step 2	n.a.		Extensibility	**40**	Rmax	**40**	n.a.	


Westonia
Krichauff
Machete
Diamondbird
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
Rmax (BU)
Extensibility (cm)
R0.04
R0.2E0.2
E0.2R0.2
E0.04
Selection scheme (trait followed by selected proportion)

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

Efficient selection of multiple genes via marker-assisted selection, an example in wheat

Wang, J.,* S.C. Chapman, D.B. Bonnett, G.J. Rebetzke, and J. Crouch. 2007. Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. **Crop Science 47**: 580-588.

Nine major genes to be pyramided in wheat

Gene	***Rht-B1***	***Rht-D1***	***Rht8***	***Sr2***	***Cre1***	***VPM***	***Glu-B1***	***Glu-A3***	***tin***
Chr.	**4BS**	**4DS**	**2DL**	**3BS**	**2BL**	**7DL**	**1BL**	**1AS**	**1AS**
Marker	Codom	Codom	Codom	Codom	Dom	Dom	Codom	Codom	Codom
MK-gene distance	0	0	0.6	1.1	0	0	0	0	0.8
HM14BS	***Rht-B1a***	***Rht-D1a***	***Rht8***	*sr2*	*cre1*	*vpm*	*Glu-B1a*	*Glu-A3e*	*Tin*
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***
Target	***Rht-B1a***	***Rht-D1a***	***Rht8***	***Sr2***	***Cre1***	***VPM***	***Glu-B1i***	***Glu-A3b***	***tin***

One strategy identified by QuLine to combine the nine genes from topcross

- Selection of **Sunstate as the final parent** (having largest number of favorable alleles) in the topcross

- Stage I: Selection for *Rht-B1a* and *Glu-B1i* homozygotes, and enrichment of *rht8*, *Cre1*, and *tin* in TCF1
- Stage II: Selection of homozygotes for one target allele, e.g. *Rht8*, and enrich remaining target alleles in TCF2
- Stage III: Selection of the target genotype in DHs/RILs

Comparison with other strategies

- For this strategy, one target genotype can be selected by screening **< 600 individuals/lines**

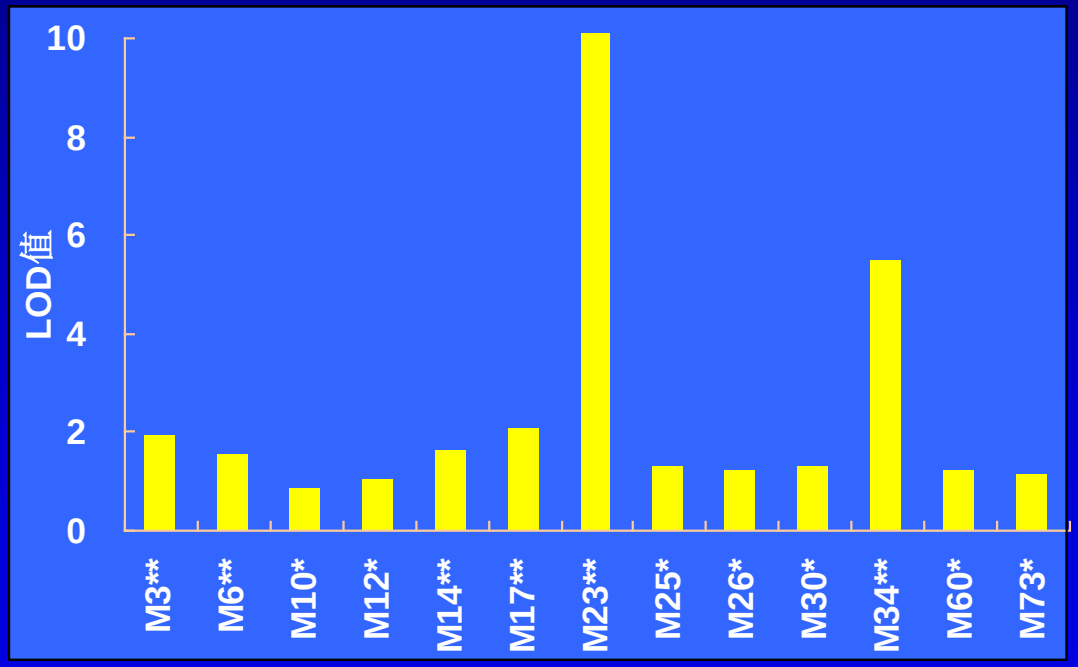
- For one-stage selection in advanced generations, one target genotype can be selected be screening **> 3500 lines**
- For one-stage selection in early generations, say TCF2, one target genotype can be selected be screening **millions of individuals**

Using QTL mapping results to design the breeding program

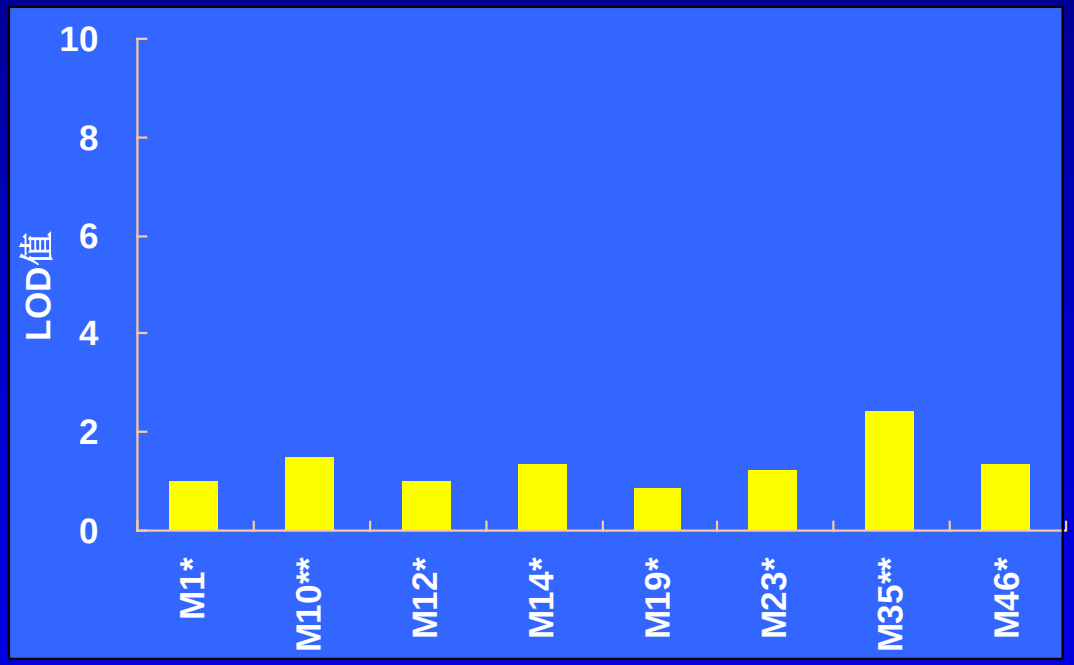
Genetical Research, 2006, 88: 93-104
Theor. Appl. Genet., 2007, 115: 87-100

QTL for grain length and grain width in rice using 65 CSS lines

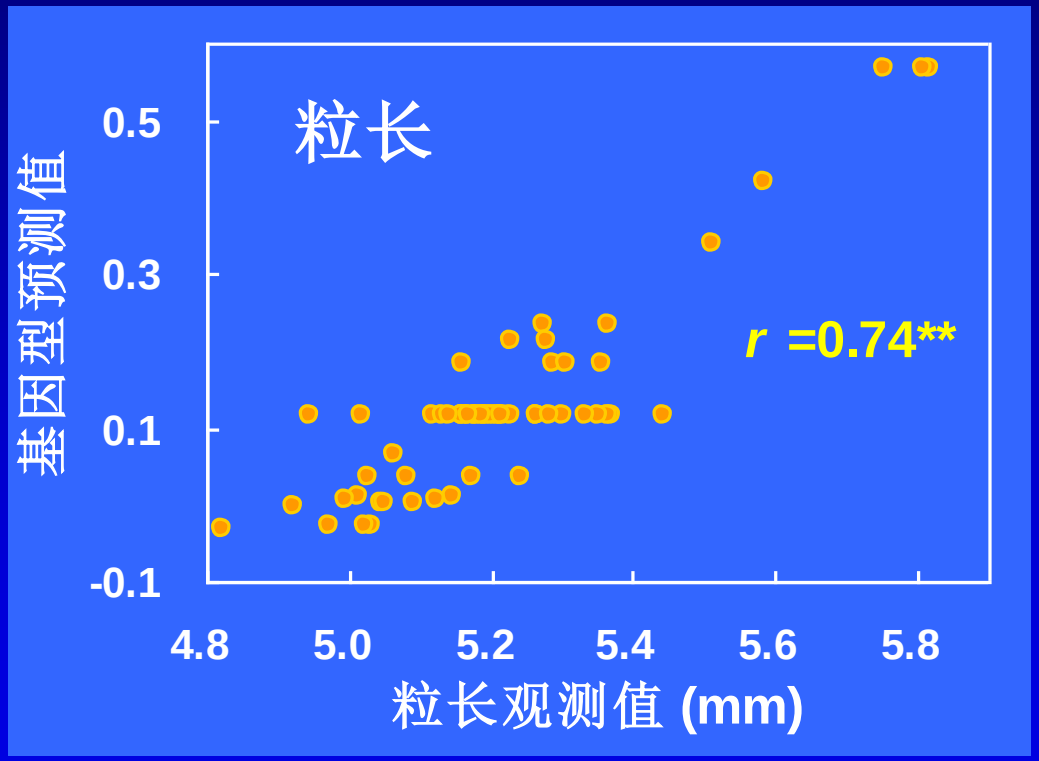
Grain length

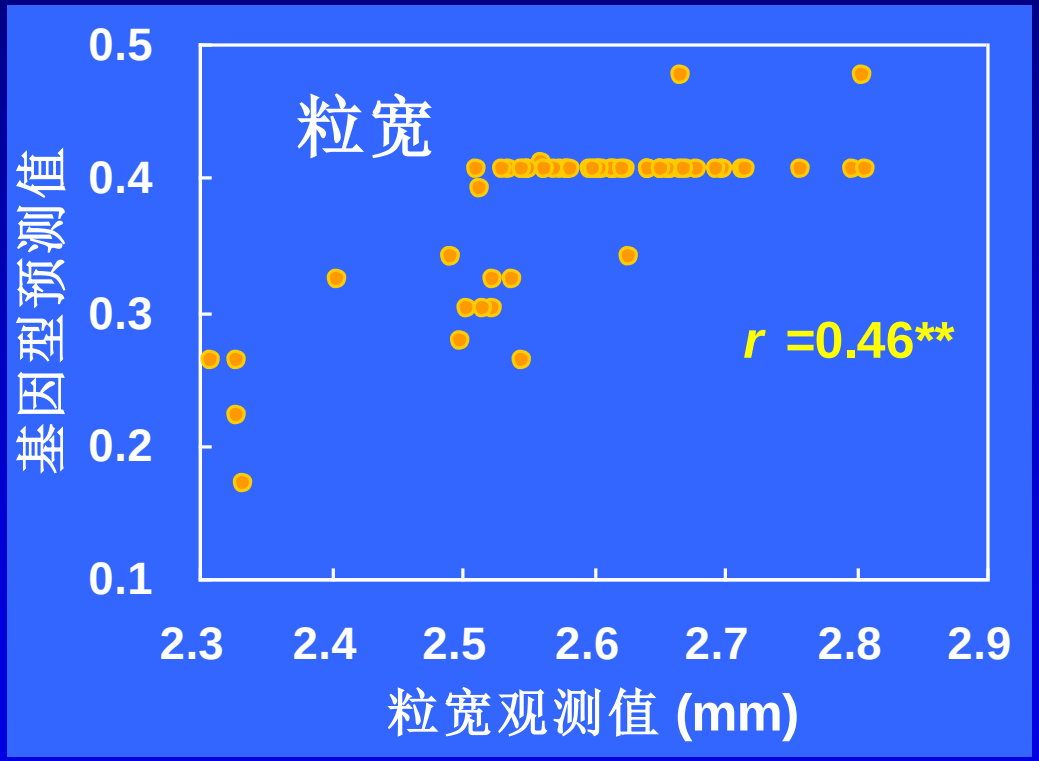


Grain width

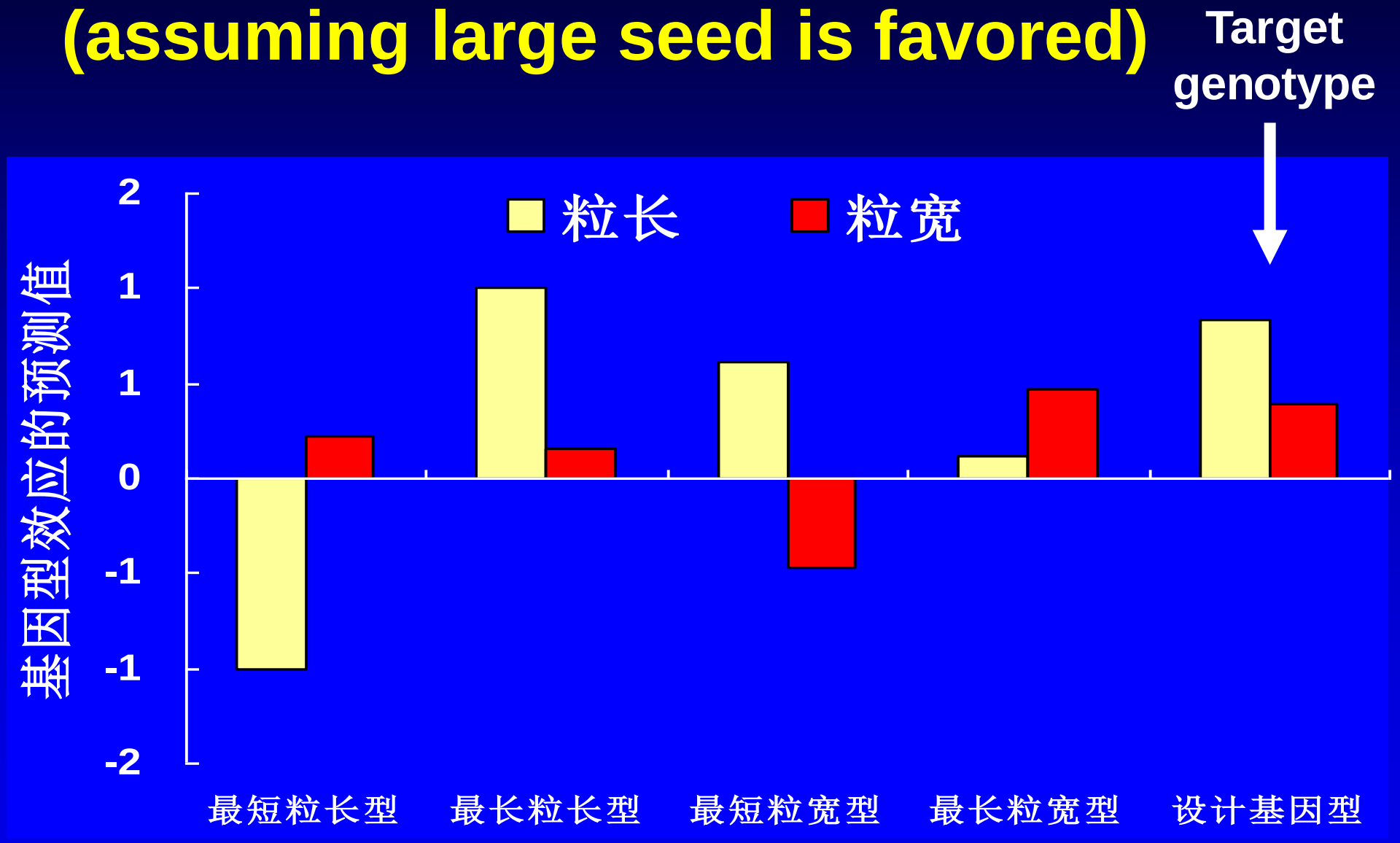


Prediction using identified QTL





Designing the target genotype
(assuming large seed is favored)



Choosing parental lines

Genotype	M1	M6	M12	M23	M25	Prediction in	
						GL (mm)	GW (mm)
Longest grain	1	2	2	2	2	6.21	2.74
Widest grain	2	1	1	1	1	5.32	3.07
Designed TG	2	2	1	2	2	6.05	2.98
CSSL5	1	2	1	1	1	5.44	3.00
CSSL16	2	1	1	2	1	5.77	2.98
CSSL19	1	1	2	1	2	5.54	2.93

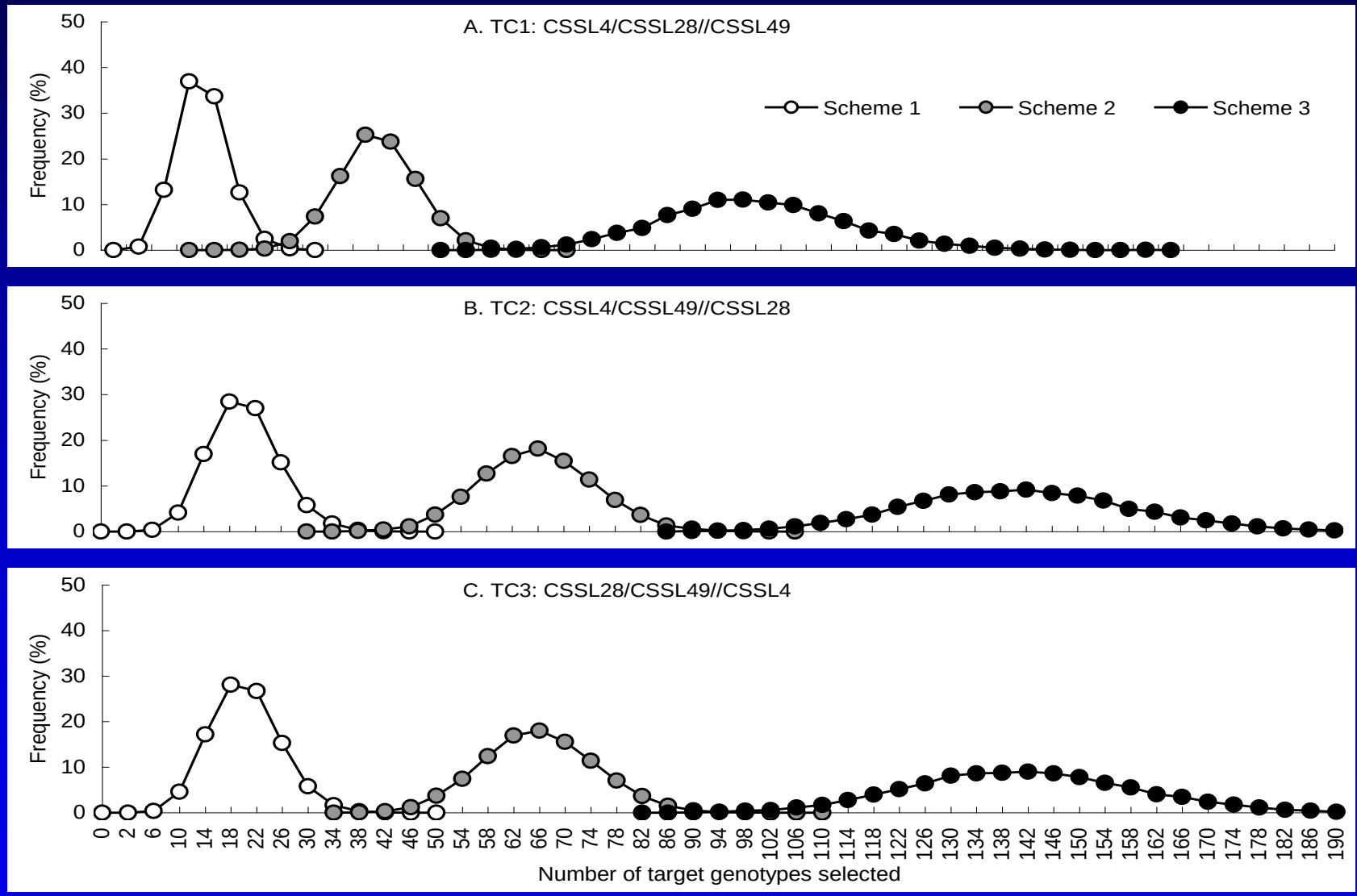
Achieving the designed TG

- Three option for crossing
  - TC1：(CSSL5×CSSL16)×CSSL19
  - TC2：(CSSL5×CSSL19)×CSSL16
  - TC3：CSSL5×(CSSL16×CSSL19)
- Two MAS schemes (just for example)
  - Scheme1：MAS in F8 only
  - Scheme2：MAS in F2 and F8

Choosing the best crossing and selection schemes

TC	MAS	Selected Ind. In F2	F8 families before MAS	F8 families after MAS (S.E.)	DNA samples	DNA samples per TG
TC1	Scheme1	**100**	**3000**	**7.6 (3.27)**	**3000**	**395**
	Scheme2	**12.0**	**359**	**7.6 (3.37)**	**459**	**60**
TC2	Scheme1	**100**	**3000**	**24.3 (7.06)**	**3000**	**123**
	Scheme2	**24.8**	**745**	**23.3 (7.16)**	**845**	**36**
TC3	Scheme1	**100**	**3000**	**11.2 (5.45)**	**3000**	**268**
	Scheme2	**7.5**	**226**	**12.3 (5.14)**	**326**	**26**

Frequency distribution of the number of DG1 from 1000 simulations (Breeding is science, art and chance!)



Breeding is a tedious and long procedure when considering the nature of complex inheritance: linkage, epistasis, pleiotropy, GbyE, etc.

Dr. Borlaug spent more than
20 years from
this tall wheat to this semi-dwarf wheat,
which triggered the Green Revolution!

中国农业科学院
The Chinese Academy of Agricultural Sciences

CIMMYT
©INTERNATIONAL MAIZE AND WHEAT IMPROVEMENT CENTER

Acknowledgements

- Grain Research and Development Corporation (GRDC), Australia (1999 – 2004)
- Generation Challenge Program (GCP) of CGIAR (2005 – 2014)
- HarvestPlus Challenge Program of CGIAR (2006 – 2014)
- National 973 Programs of China for Key Basic Scientific Research (2006 – 2015)
- National 863 Programs of China for High Technological Research (2006 – 2015)
- Natural Science Foundation of China (2008 – 2015)

GRDC
Grains Research & Development Corporation

Generation
Cultivating Plant Diversity for the Resource-Poor

Integrated Breeding

HarvestPlus

Demonstration of QuLine