

The 9th Workshop on QTL Mapping and Breeding Simulation
The University of Sydney, Cobbitty NSW, 7-9 March 2012

Strategic Applications of Breeding Simulation

Jiankang Wang
CIMMYT China and CAAS

E-mail: wangjk@caas.net.cn or jkwang@cgiar.org
<http://www.isbreeding.net>

Outline of the presentation

- Comparison of modified pedigree (MODPED) and selected bulk (SELBLK)
- Modeling of the Single Backcrossing Breeding Strategy (SBBS)
- Comparison of direct, indirect, index, and independent selections

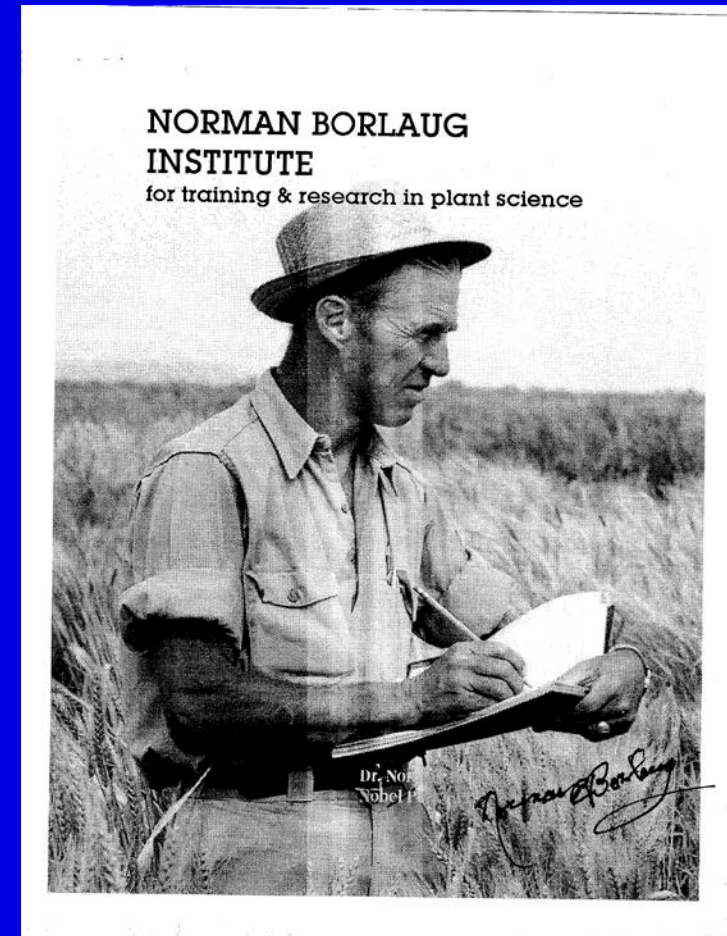
Comparison of two breeding strategies: modified pedigree (MODPED) and selected bulk (SELBLK)

Crop Science, 2003, 43: 1764-1773

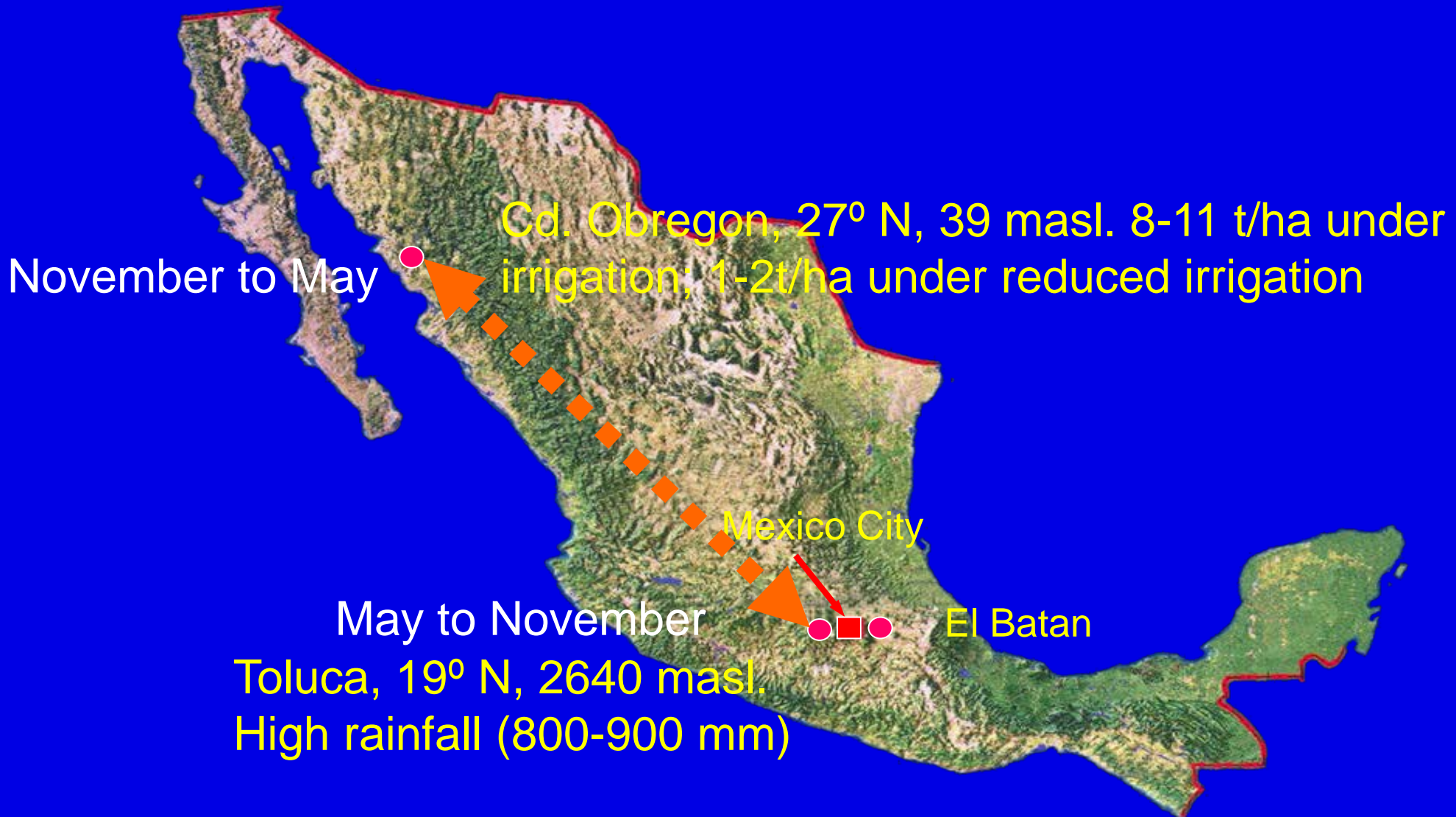
Crop Science, 2004, 44: 2006-2018

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
- Criticism of the Green Revolution
 - The rich got richer, the poor got poorer
 - Environment damage
 - Loss of biodiversity



CIMMYT's Shuttle Breeding



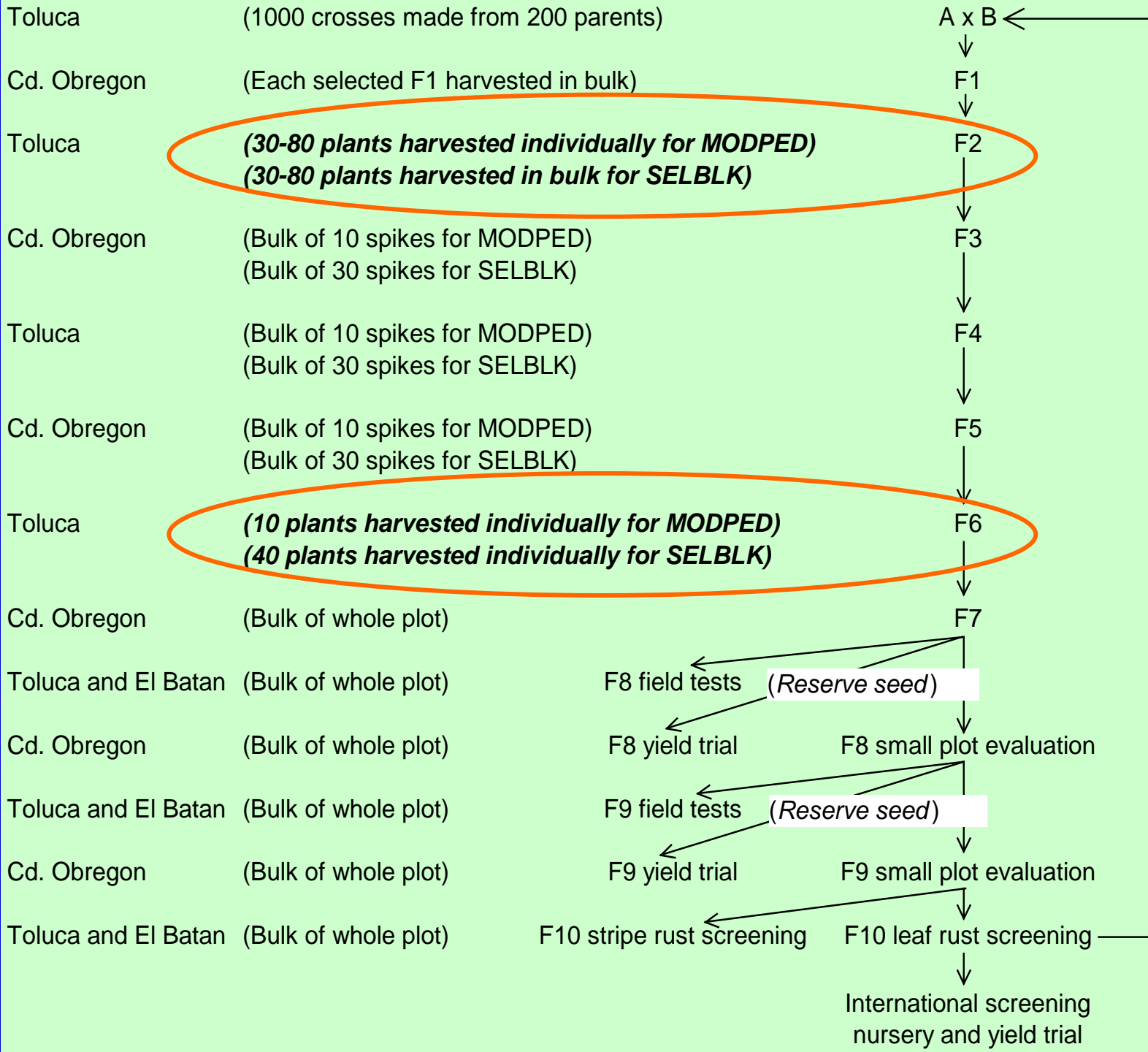
Breeding methods in CIMMYT's wheat breeding program

- 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									
Heading	0.60									
Grains/spike	0.09						-0.17			
Seed weight	-0.07						-0.30		-0.07	

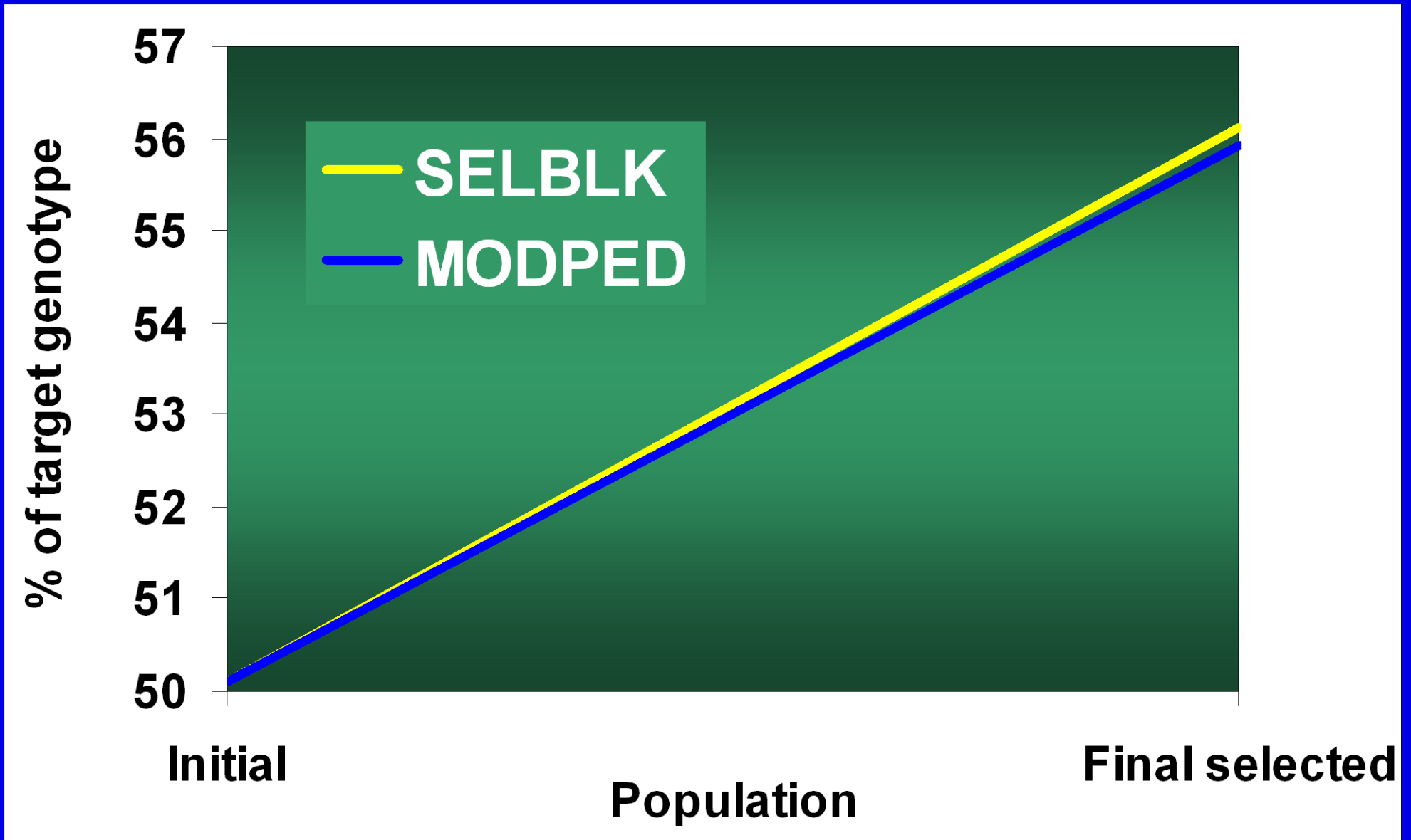
Estimated by CIMMYT breeders

Estimated from the defined genetic model

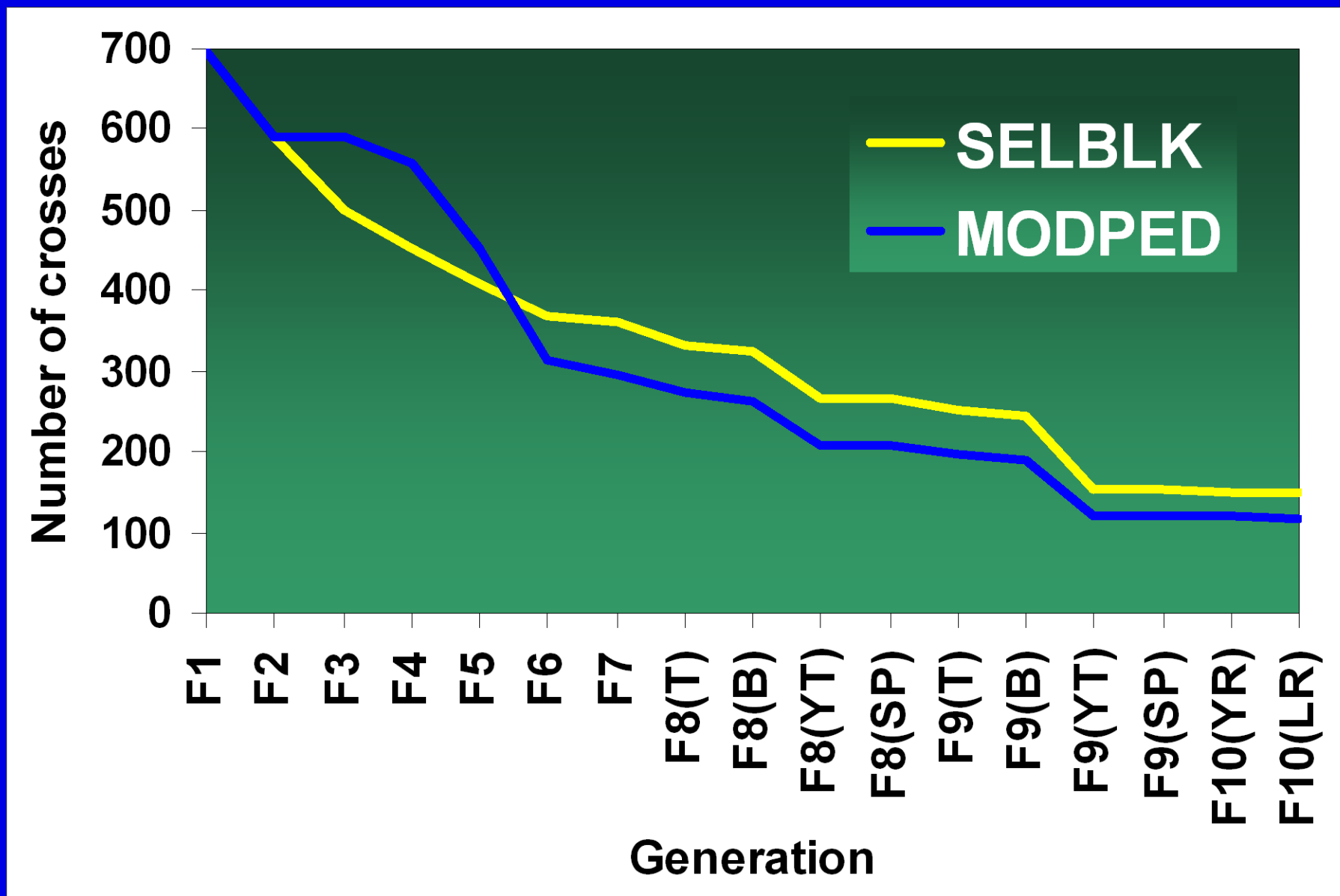
Experimental design

- 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

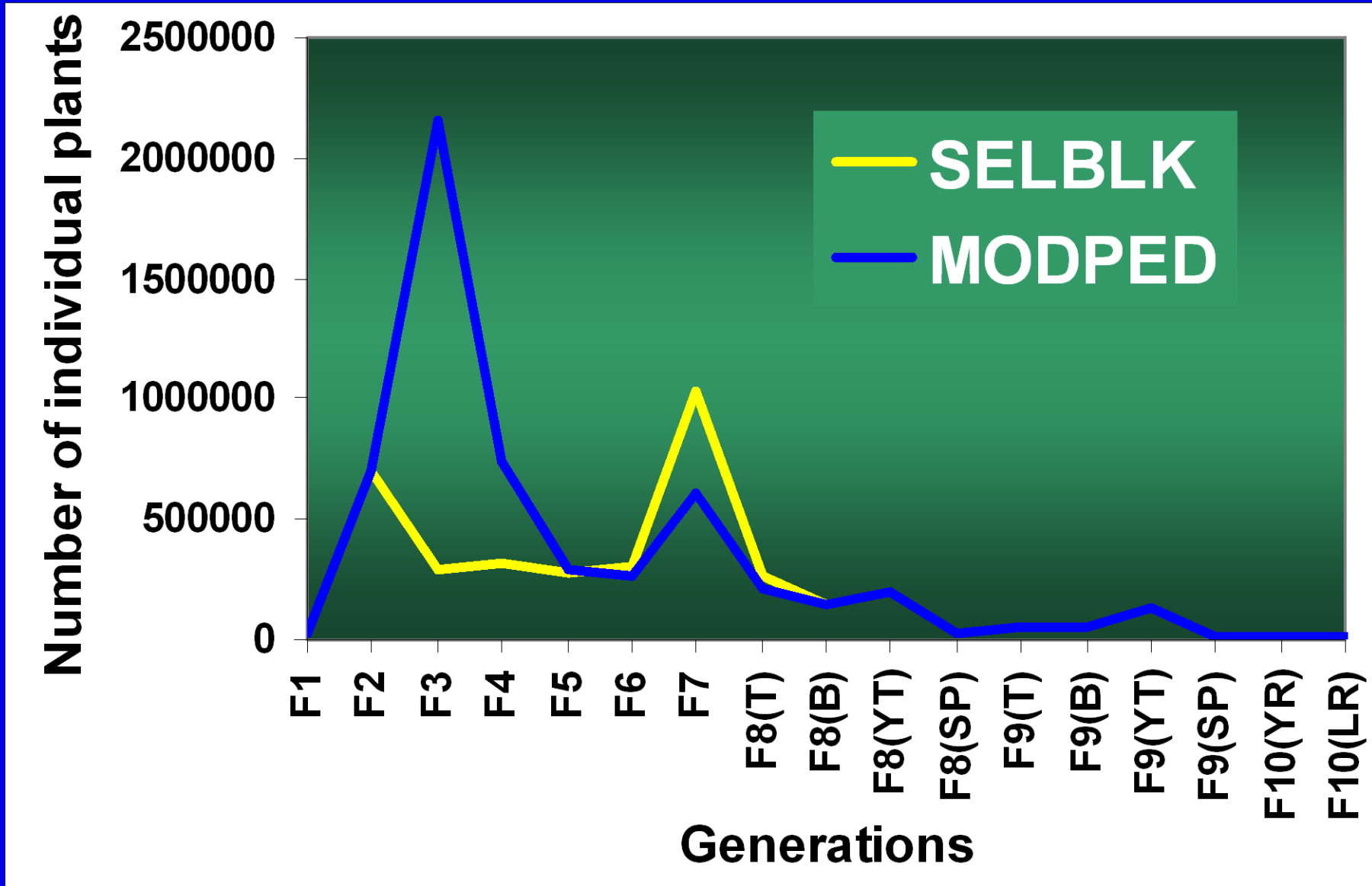
Result 1: Genetic gain in yield from SELBLK is 3.3% higher than MODPED. SELBLK is slightly more efficient.



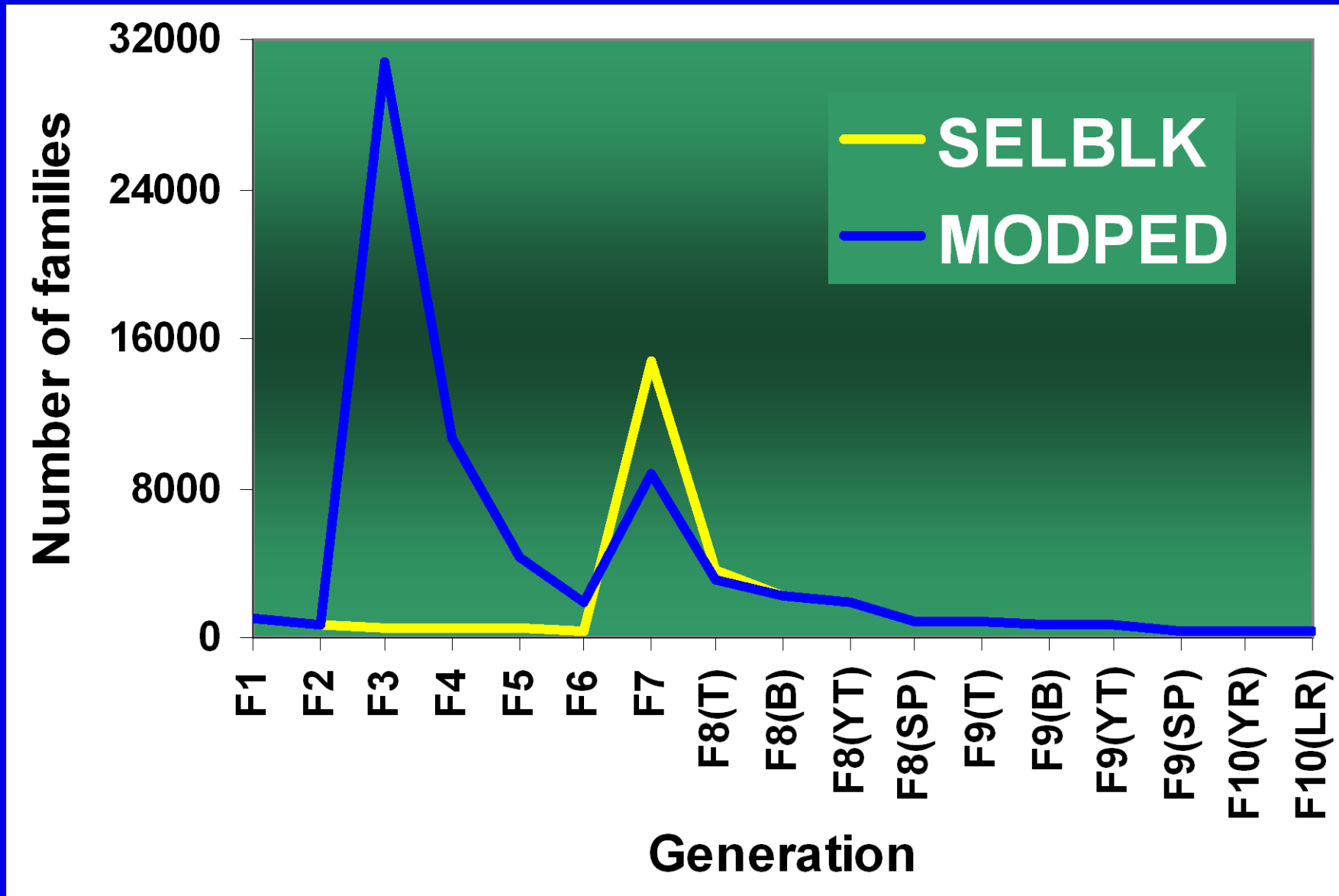
Result 2: SELBLK retained 25% more crosses in the final selected population (more genetic diversity retained)



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)

Wang, J., * R. P. Singh, H.-J. Braun, and W. H. Pfeiffer. 2009. Investigating the efficiency of the single backcrossing breeding strategy through computer simulation. **Theor. Appl. Genet.** 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 parental 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 <i>Triticum durum</i> and <i>Aegilops tauschii</i>	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

➤ ***“All models are wrong, but some are useful!”***

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 ₁ F ₁	Backcrossing to the adapted parents	100	400	100	50
B ₁ F ₂	Selfing	100	1200	100	30
B ₁ F ₃	Selfing	100	400	100	10
B ₁ F ₄	Selfing	100	400	100	10
B ₁ F ₅	Selfing	100	400	100	10
B ₁ F ₆	Selfing	1000	200	30	200
Final selected advanced lined		10			

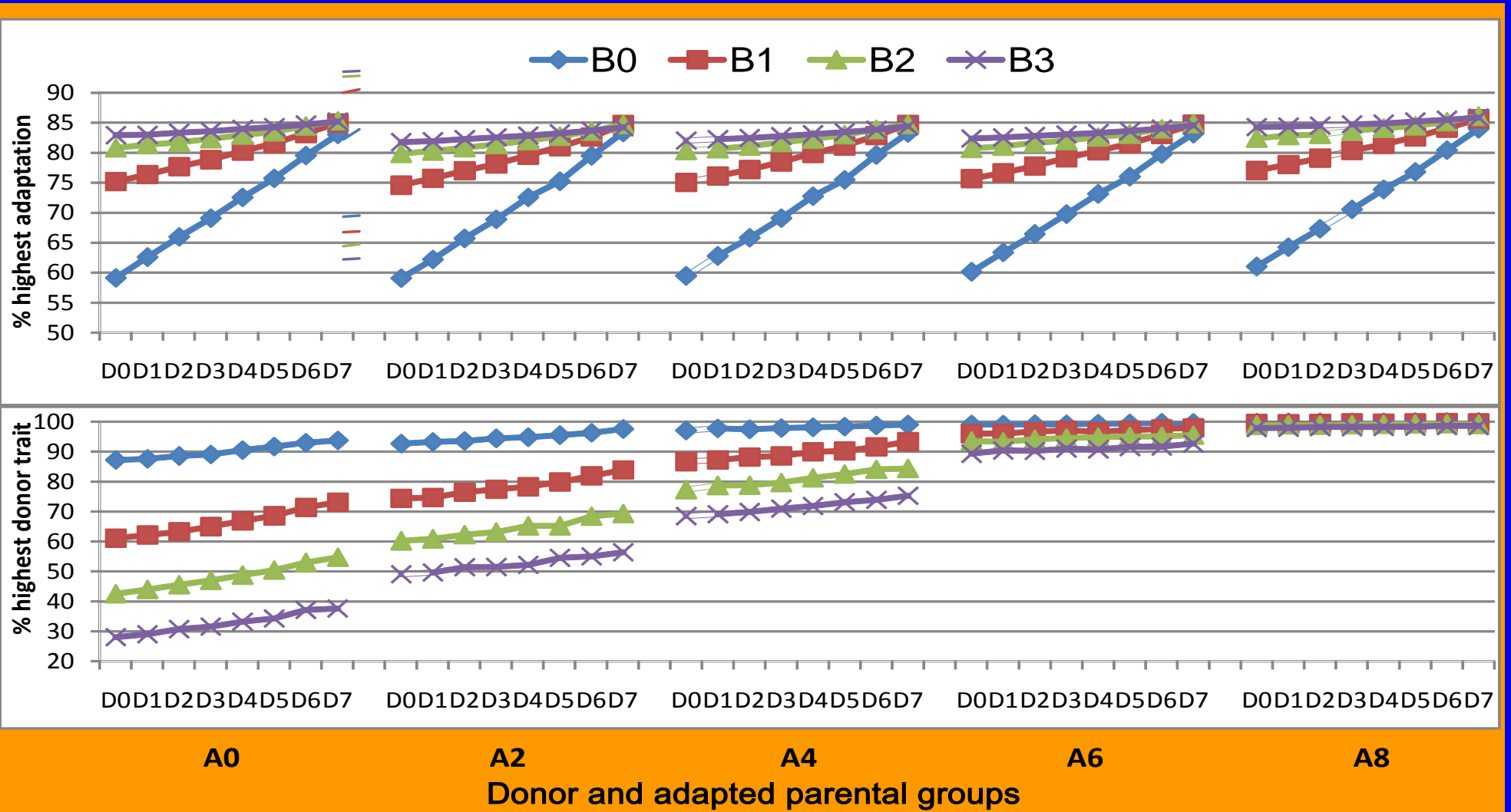
Four crossing strategies defined in QuLine

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 ₁ F ₁	BC ₁ F ₁	BC ₁ F ₁	Bulk
F ₃	BC ₁ F ₂	BC ₂ F ₁	BC ₂ F ₁	Bulk
F ₄	BC ₁ F ₃	BC ₂ F ₂	BC ₃ F ₁	Bulk
F ₅	BC ₁ F ₄	BC ₂ F ₃	BC ₃ F ₂	Bulk
F ₆	BC ₁ F ₅	BC ₂ F ₄	BC ₃ F ₃	Pedigree
F ₇	BC ₁ F ₆	BC ₂ F ₅	BC ₃ F 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

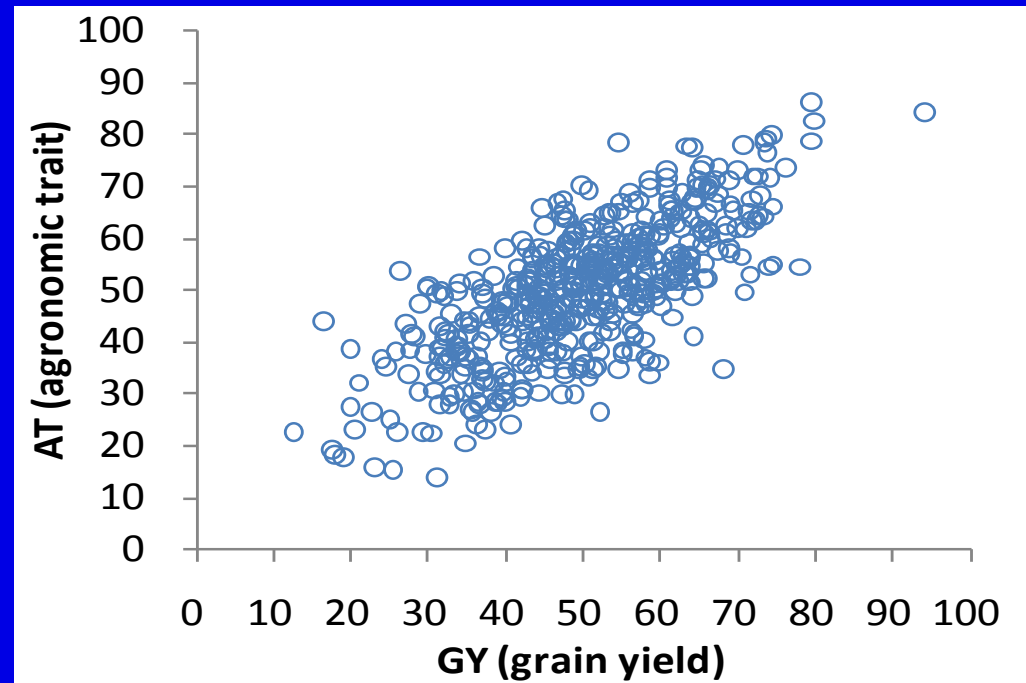


Comparison of direct, indirect, index, and independent selections

When GY and AT are both phenotyped, say in yield trials, what is the selection method resulting in the highest genetic gain in GY?

Assumptions

- Suppose heritability is 0.4 for GY, 0.6 for AT
- Correlation between GY and AT is 0.7
- 500 RIL phenotyped on both GY and AT
- Genotypic values are known in simulation
- Possible selection schemes
 - **Direct selection on GY**
 - **Index1: 0.75 GY + 0.25 AT**
 - **Index2: 0.50 GY + 0.50 AT**
 - **Index3: 0.25 GY + 0.75 AT**
 - **Direct selection on AT**



Genetic gains in GY and AT

Gain in GY			
Selected proportion	20%	5%	1%
Lines selected	100	25	5
Direct selection on GY	66.64	73.68	81.60
0.75GY + 0.25AT	66.36	73.52	80.80
0.50GY + 0.50AT	65.64	72.08	80.80
0.25GY + 0.75AT	63.94	70.24	80.80
Direct selection on AT	62.48	69.52	79.60
Gain in AT			
Selected proportion	20%	5%	1%
Lines selected	100	25	5
Direct selection on GY	62.92	69.20	76.80
0.75GY + 0.25AT	64.98	71.12	81.60
0.50GY + 0.50AT	66.34	74.40	81.60
0.25GY + 0.75AT	67.28	75.36	81.60
Direct selection on AT	67.62	75.60	82.40

Observations

- When 500 RIL were phenotyped on both GY and AT, and the same selection intensity was applied
 - The direct selection based on GY resulted in the highest gain in GY; So is the case for AT
 - An index including both GY and AT reduced the gain in GY
 - No index resulted in the same gain as the direct selection on GY
 - When GY data is available, selection on other traits is less important, when GY is the major objective

Conclusion:

There is no situation in which adding AT to an index will increase gains in GY unless H_{AT} is $\gg H_{GY}$ and r_G is very close to 1. But then, you would do better to select for AT alone!

Question:

Why would we still select for AT in breeding? Why are we still trying to identify physiological traits that can be used for the improvement of GY?

Comparison of independent selection with direct selection on GY

- Say we need to select 5 out of 500 RIL having the highest GY, phenotyping cost is \$10 for AT, and \$50 for GY (**need breeder's input here!**)
 - **Direct selection on GY:** Phenotyping the 500 RILs on GY, and the best 5 are selected; **Cost: \approx \$25,000**
 - **Independent selection (AT followed by GY):** Phenotyping the 500 RILs on AT, and the best 50 are selected based on AT; Phenotyping the retained 50 RILs, and the best 5 are selected based on GY; **Cost: \approx \$5,000 + \$2,500 = \$7,500**
- ***Is there any difference in genetic gain from direct and independent selections?***

Comparison of independent selection with the single trait selection on GY

Gain in GY	Cost (\$)			
Selected proportion	20%	5%	1%	1%
Lines selected	100	25	5	5
Direct selection on GY	66.64	73.68	81.60	25,000
0.75GY + 0.25AT	66.36	73.52	80.80	30,000
0.50GY + 0.50AT	65.64	72.08	80.80	30,000
0.25GY + 0.75AT	63.94	70.24	80.80	30,000
Direct selection on AT	62.48	69.52	79.60	5,000
Independent selection	66.26	72.88	81.20	7,500

Take-home messages

When GY and AT are both available, and GY is the main breeding target ...

- Direct selection on GY is more efficient and more important in yield trials
- Index selection including both GY and AT always resulted in reduced gain in GY
- When AT or other physiological traits have been identified to be closely correlated with GY, indirect selections based on these traits in early generations are useful if screening for AT is much less expensive