

The Development of Drought Tolerant Maize Germplasm in sub-Saharan Africa using Marker-Assisted Recurrent Selection and Genomic Selection

Yoseph Beyene¹, Kassa Semagn², Jose Crossa³, Stephen Mugo¹, Barbara Meisel⁴, Sylvester Oikeh⁵, Amsal Tarekegne⁶, Michael Olsen¹ and Prasanna B

Presentation 7th International Crop Science Congress ,August 14–19, 2016
Beijing, China.

Outline

- Proof of concept GS in tropical maize
- Results of genetic gain studies from GS
- Results of MARS
- Lesson learnt from GS/MARS
- Conclusions



Effect of GS models and marker density on prediction accuracy using 284 maize lines

<i>Trait–environment (heritability)</i>	<i>BL</i>	<i>RKHS</i>	<i>RBFNN</i>
<u>55,000 SNPs</u>			
FFL-WW (0.89)	<u>0.814</u>	<u>0.836</u>	<u>0.834</u>
FFL-SS (0.81)	0.754	0.763	0.757
MFL-WW (0.88)	0.817	0.841	0.832
MFL-SS (0.83)	0.776	0.782	0.780
ASI-WW (0.79)	0.582	0.586	0.594
ASI-SS (0.77)	0.612	0.621	0.605
GY-WW (0.49)	0.557	0.548	0.529
GY-SS (0.38)	<u>0.326</u>	<u>0.330</u>	<u>0.288</u>
<u>1,148 SNPs</u>			
FFL-WW	0.781	0.588	—
FFL-SS	0.774	0.648	—
MFL-WW	<u>0.790</u>	0.607	—
MFL-SS	0.778	<u>0.674</u>	—
ASI-WW	0.513	0.547	—
ASI-SS	0.517	0.572	—
GY-WW	0.525	0.514	—
GY-SS	<u>0.415</u>	<u>0.453</u>	—

N=284 inbred lines

Crossa et al. (2014)
Heredity **112**:48–60

Effect of GS models and imputations on prediction accuracy using 504 DH lines

Grain yield	
Pedigree	0.5181
RKHS	0.5274
RKHS _{IM}	0.5405
RKHS _H	<u>0.5657</u>
GBLUP	0.4997
GBLUP _{IM}	0.5065
GBLUP _H	0.4927

Grain Yield	
Pedigree + RKHS	0.5718
Pedigree + RKHS _{IM}	0.5790
Pedigree + RKHS _H	<u>0.5821</u>

Values are correlations between predicted (based on markers) and actual phenotypes

Crossa et al. (2013): 3G 3:1-24

GBLUP = Genomic Best Linear Unbiased Predictors; RKHS = Reproducing Kernel Hilbert Spaces ; IM=imputed; H= haplotype

Effect of GS models and imputations on prediction accuracy using 504 DH lines

Anthesis day	
Pedigree	0.5888
RKHS	0.6993
RKHS _{IM}	0.7104
RKHS _H	<u>0.7184</u>
GBLUP	0.6968
GBLUP _{IM}	0.7031
GBLUP _H	0.6866

Anthesis day	
Pedigree + RKHS	0.7183
Pedigree + RKHS _{IM}	0.7262
Pedigree + RKHS _H	<u>0.7313</u>

Values are correlations between predicted and actual phenotypes



Effect of marker density on GS in biparental populations

- 19 biparental maize populations
- Marker density
 - Low marker density (162-212 SNPs)
 - high marker density (48,662-78,005 GBS SNPs)
- Managements
 - 3 to 4 drought environments
 - 3 to 4 optimum environments



Average prediction accuracy by marker density, trait & water management for 19 maize pops

	<i>Grain yield</i>	<i>Anthesis date</i>	<i>Plant height</i>
Low density, optimum	0.48	0.49	0.60
High density, optimum	0.49	0.49	0.60
Low density, drought	0.36	0.38	0.34
High density, drought	0.39	0.38	0.40

- High marker density improved prediction accuracy for both grain yield and plant height under drought but not anthesis date
- Marker density did not show any difference in prediction accuracy for all 3 traits

Zhang et al. (2015) Heredity 114: 291–299

Lessons learned from GS work at CIMMYT

1. Choice of models showed some level of differences in prediction accuracies
2. Relative to pedigree or marker only models, there were consistent gains in prediction accuracy by combining pedigree and marker data
3. About 200 SNP markers were found to sufficient to get good prediction in biparental maize populations for simple traits with moderate-to-high heritability, but GBS outperformed low-density SNPs for complex traits with low-to-moderate heritability
4. The prediction accuracies of grain yield were consistently lower than those of simple traits (such as anthesis date)
5. Prediction accuracies under stress conditions were consistently lower and more variable than optimum environments



GS/MARS in CIMMYT Africa maize breeding

- We used GS in rapid cycle of recombination in bi-parental populations
- Total 34 bi-parental populations (total =6252 F_{2:3}, each with ca. 184 progenies)
- **Objective:** To improve drought tolerance using GS/MARS
- Phenotyped in 2-4 managed water-stressed and 3-4 well-watered environments in SSA
- Genetic gain studies completed for 18 populations
 - 10 populations using MARS
 - Selection at C₀ based on phenotype and marker data, followed by one recombination and two selfing
 - 8 populations using GS
 - Selection at C₀ based on phenotype only, followed by three cycles of marker only recombination



MARS TC evaluated under Optimum, Disease and Drought

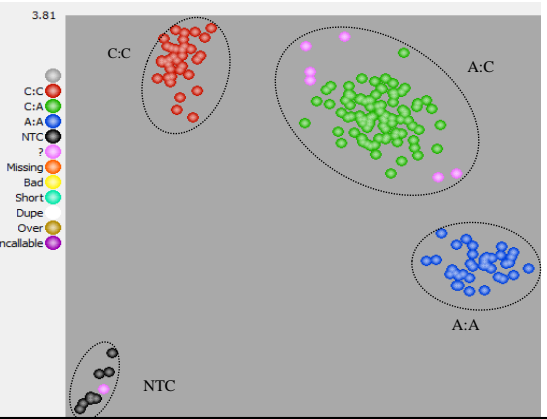
KTI

Kiboko



Kakamega

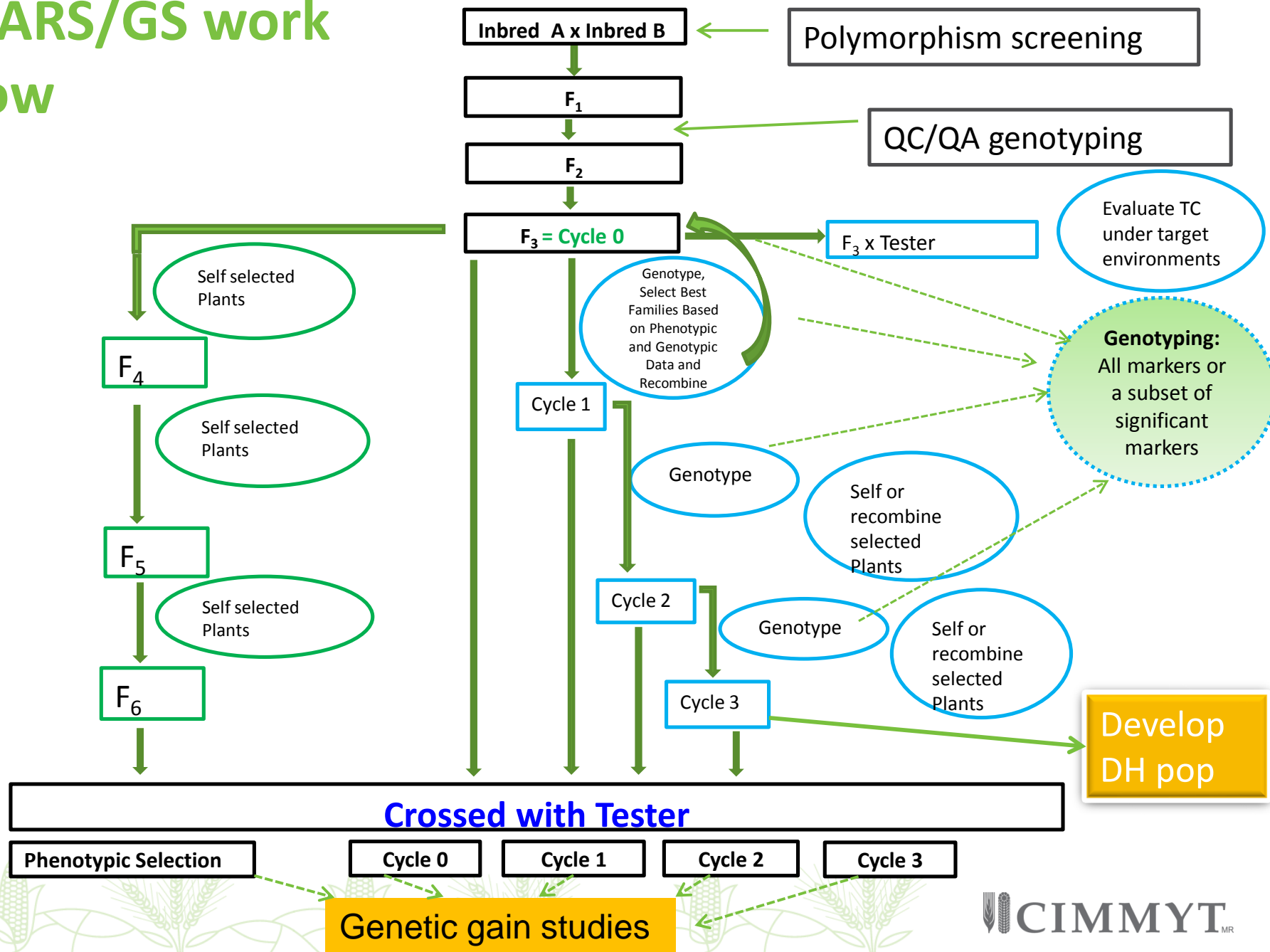




- Critical steps:**
1. Planning and coordination
 2. Data turnaround time from leaf sampling to genotyping)
 3. Data analysis and selection
 4. Synchrony and pollination



MARS/GS work flow



Genetic Gain Studies on C0-C3 from MARS

Cycle	Design
C0	Made balanced bulk from each F2:3 family within a population
C1	Made balanced bulk from the selected individuals/families
C2	Made balanced bulk from the selected individuals /families
C3	Made balanced bulk from selected individuals /families
Lines from Pedigree	Five random F6 lines from each population
Reference entries	2 parents and F1

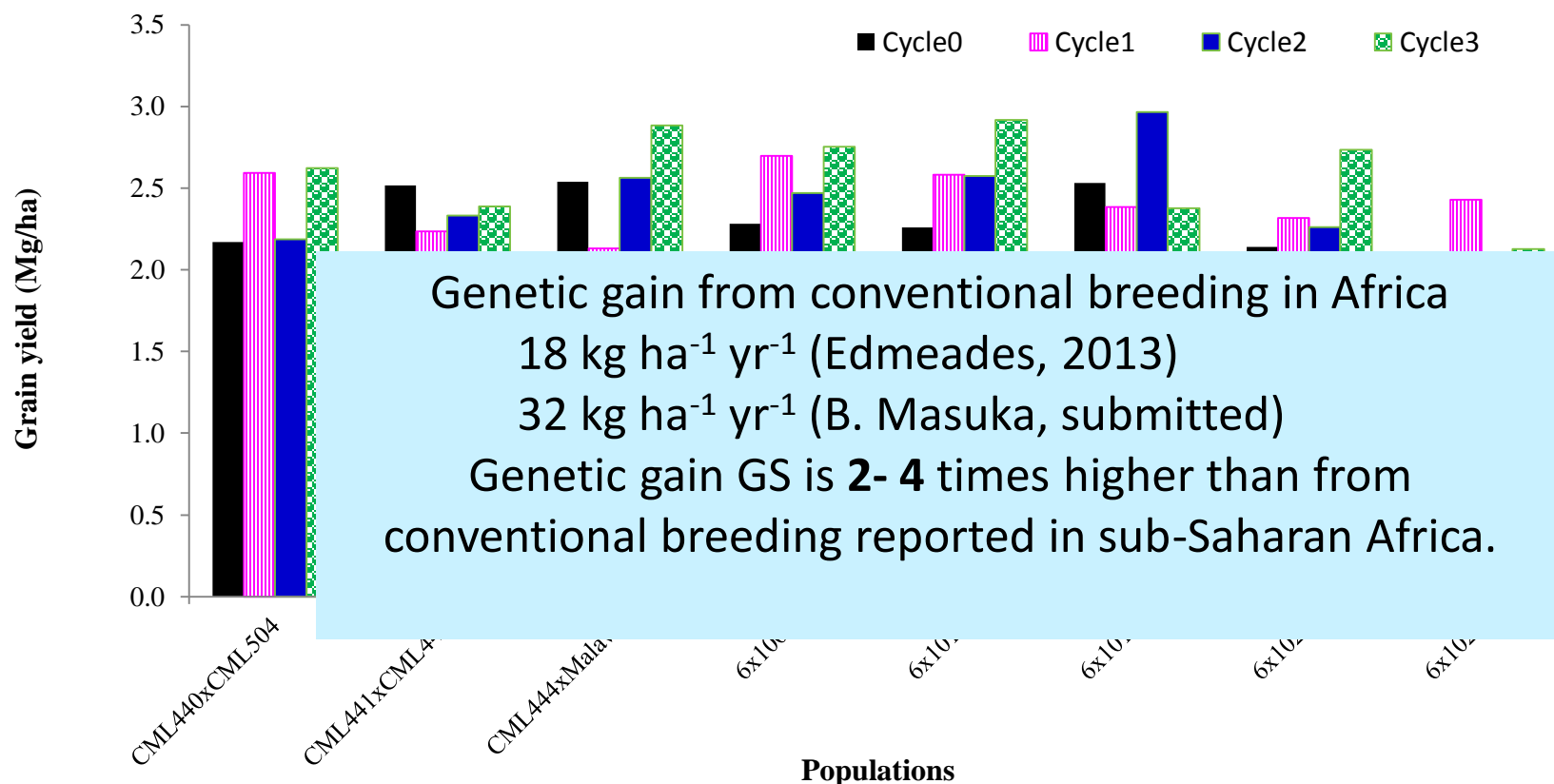
x

Single cross tester

+ Best currently available commercial hybrids (checks)

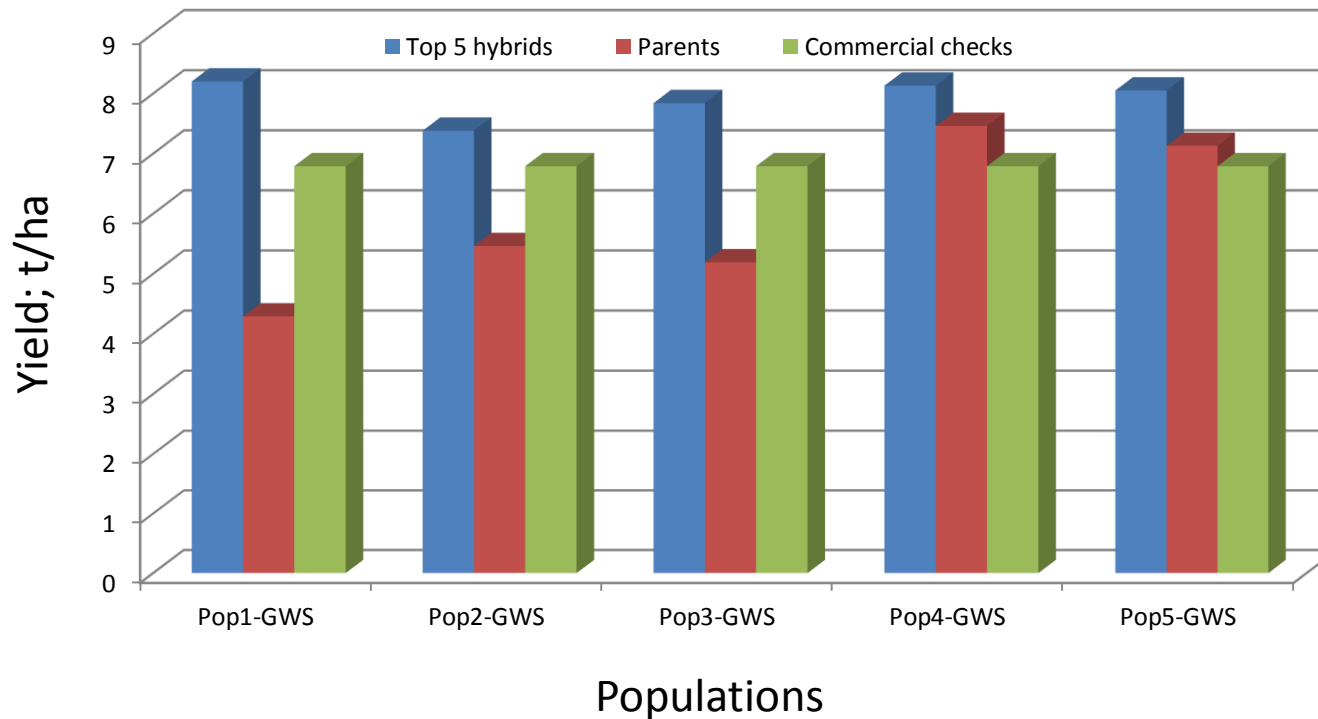


Gain in grain yield using genomewide SNPs under drought environments in SSA



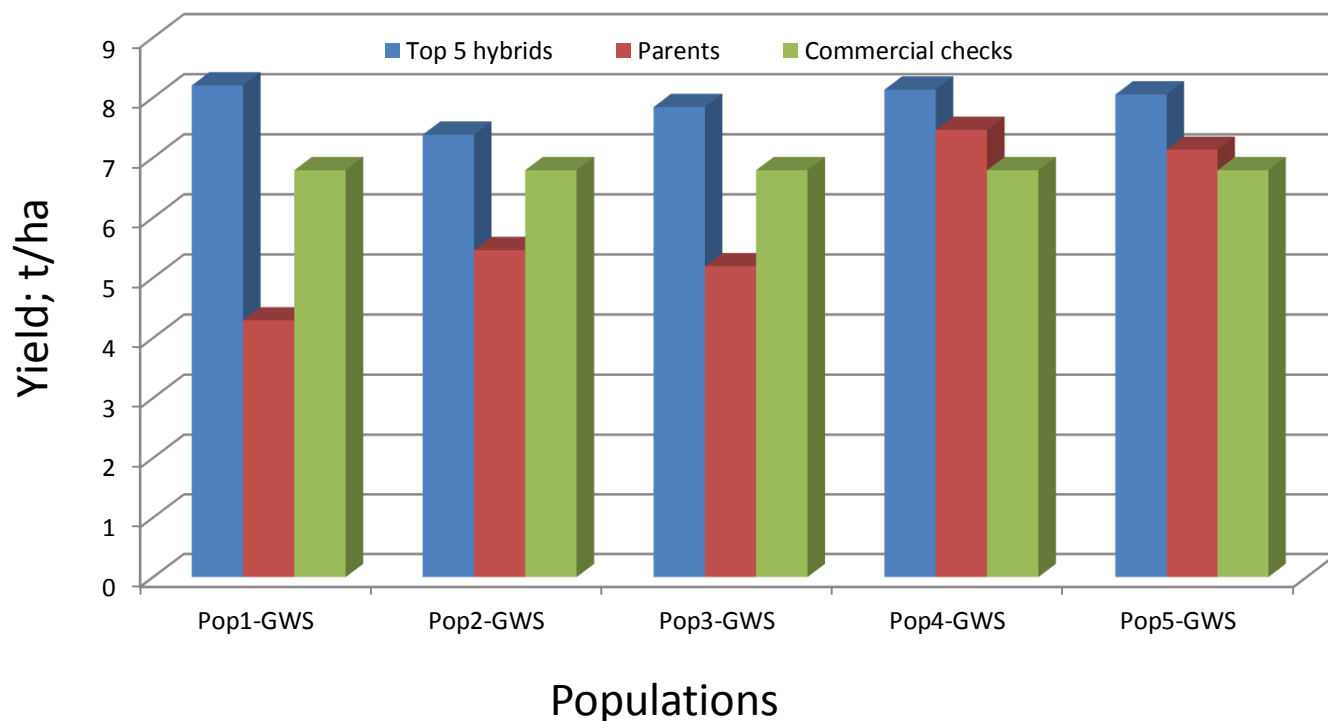
Overall gain in GY : 70.5 kg ha⁻¹ year⁻¹

Performance of hybrids developed from GS-C3-DH lines, founder parents and commercial checks under optimum

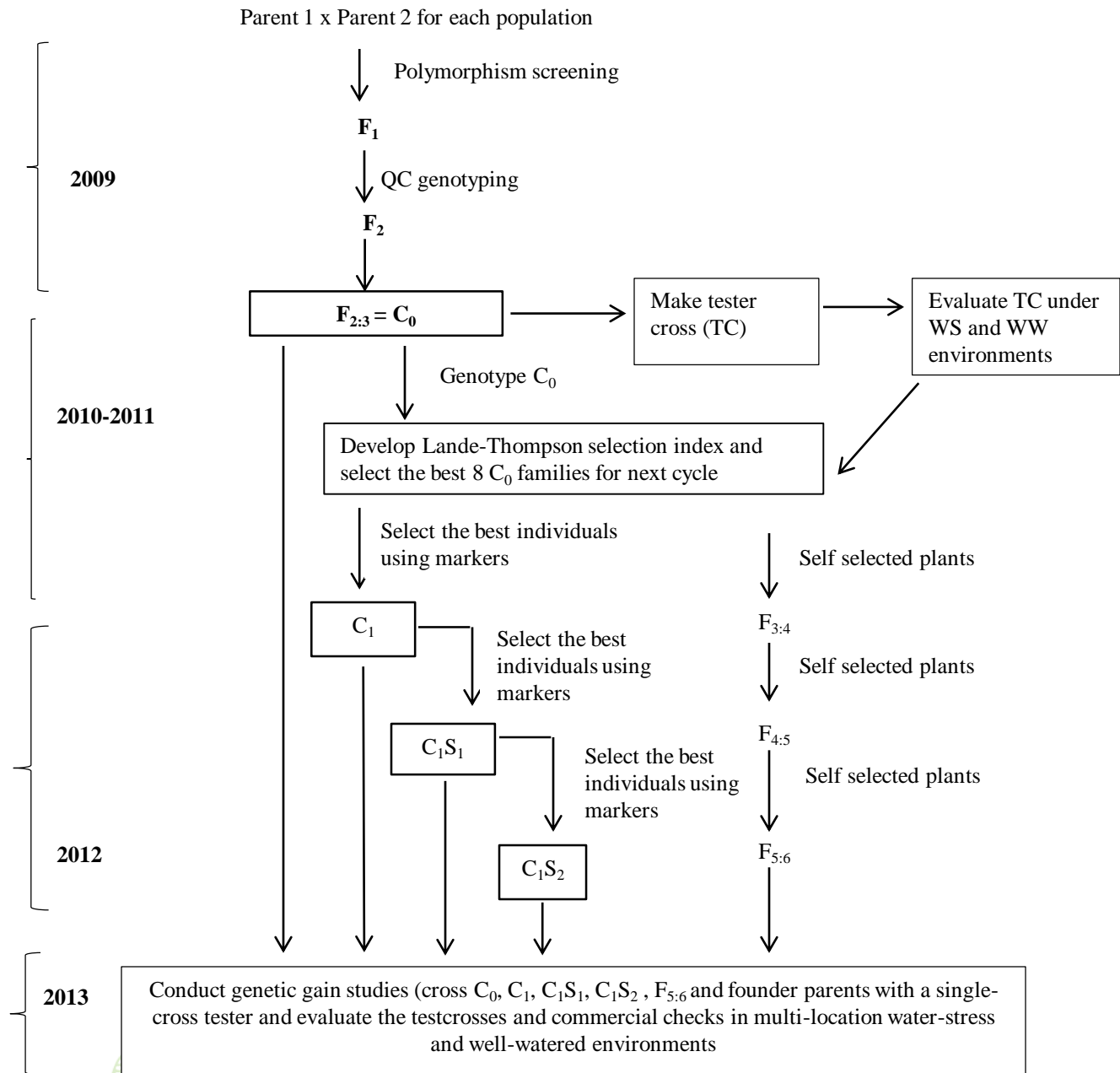


Phenotyped at 3 optimum locations
Gain over the commercial checks= 8.7 to 20.8%
Gain over the parents = 9.0 to 91.1%

Performance of hybrids developed from GS-C3-DH lines, founder parents and commercial checks under drought



- Phenotyped at one managed drought location
- Gain over the commercial checks= 47.3 to 98.2%
- Gain over the parents = 17.2 to 57.3%



Number of families and individual plants planted, selected and advanced to the next selection cycle during recombination in the 10 MARS populations.

Cycle	Form C_1 from C_0	Form C_1S_1 from C_1	Form C_1S_2 from C_1S_1
Total number of individuals per population	88	184	184
Number of families	8	8	23
Number of individuals per family	11	23	8
Number of crosses or selfing	24	40	92
Crossing operation	recombine	self	self
Number of individuals or families selected for next cycle	8	23	92
Number of nursery rows	8	8	23



Genetic gain per population - MARS

Summary for grain yield under 4 managed drought (Mg/ha)



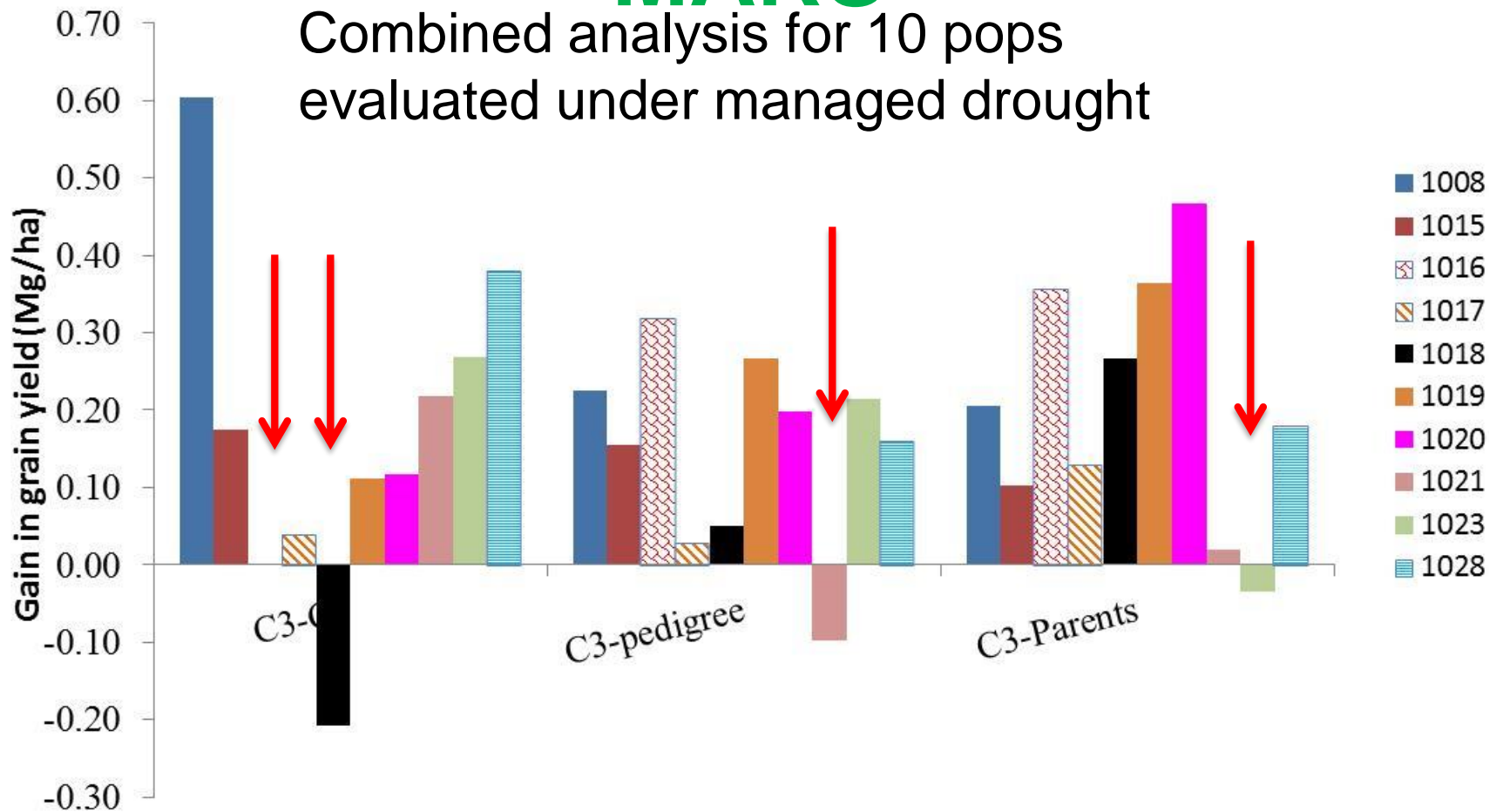
	1008	1015	1016	1017	1018	1019	1020	1021	1023	1028
Cycle0	2.361	2.739	2.965	2.834	2.894	2.362	2.519	2.151	2.265	1.993
Cycle1	2.582	2.701	2.784	3.191	2.529	2.073	2.798	2.621	2.356	2.134
Cycle2	2.775	2.466	2.341	2.952	2.301	2.246	2.389	3.008	2.378	2.332
Cycle3	2.965	2.914	2.965	2.873	2.687	2.473	2.636	2.369	2.533	2.373
Pedigree	2.74	2.759	2.647	2.845	2.637	2.206	2.437	2.466	2.318	2.213
Parents	2.76	2.812	2.608	2.744	2.42	2.109	2.169	2.35	2.567	2.194
LSD _{0.05}	0.345	0.401	0.517	0.503	0.52	0.387	0.353	0.549	0.636	0.459

Grain yield of hybrids derived from 6 pops was the highest at C₃ than all others

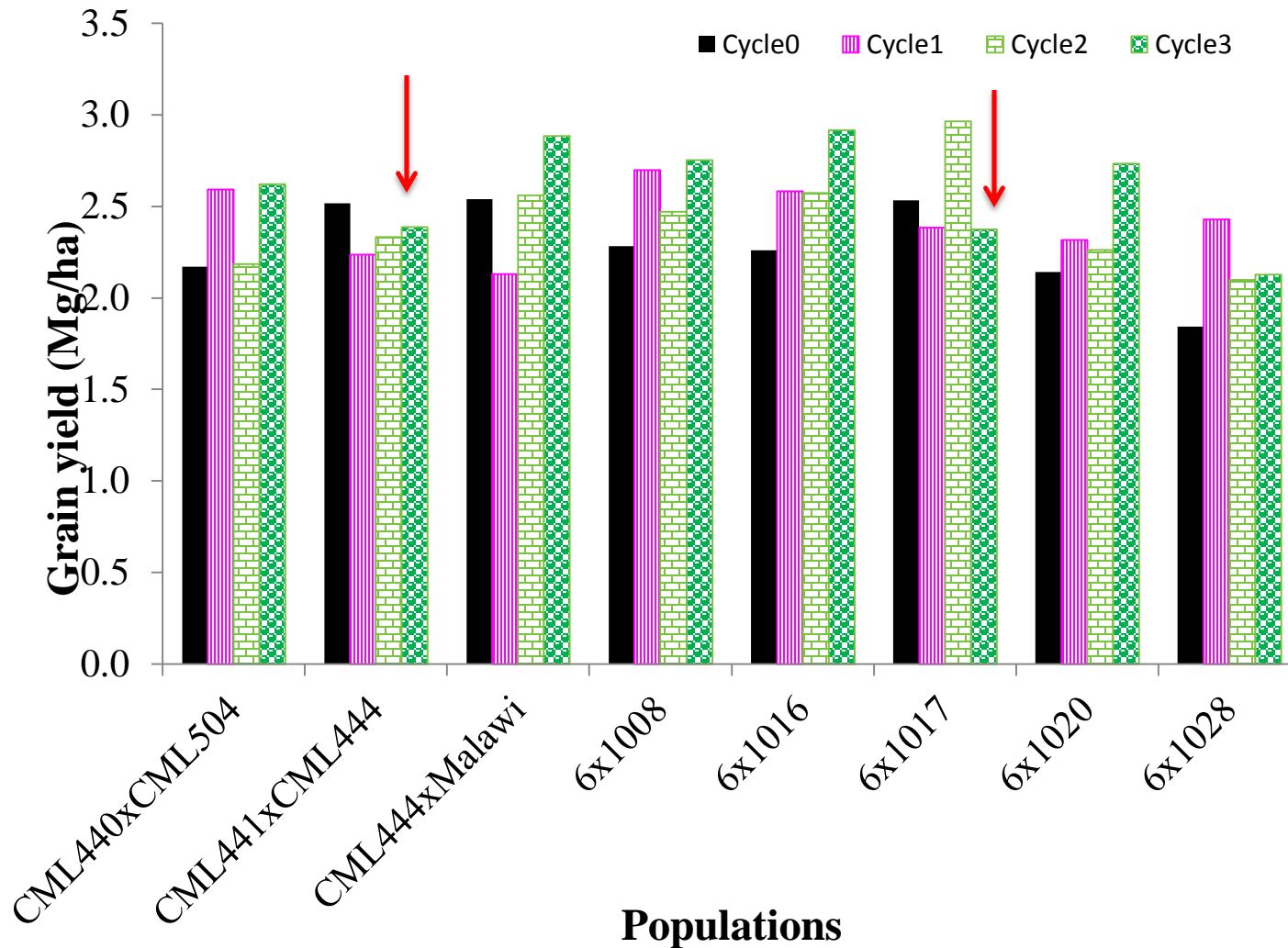


... Genetic gain per population - MARS

Combined analysis for 10 pops evaluated under managed drought

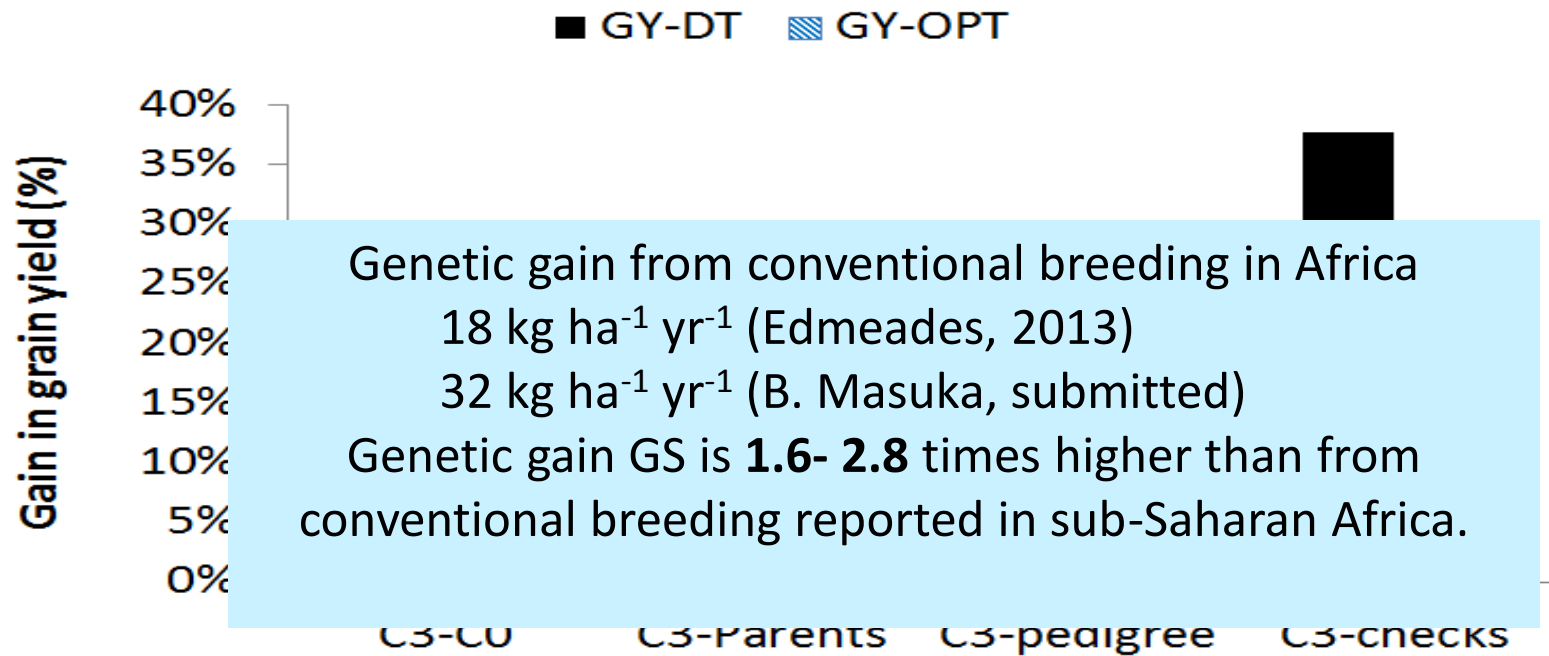


Genetic gain per population - GS



Grain yield (Mg/ha) for 8 populations advanced using genomewide markers and evaluated under 4 managed drought sites 

Gain in grain yield under drought and optimum across 10 MARS pops



Overall gain in GY : 51 kg ha⁻¹ year⁻¹

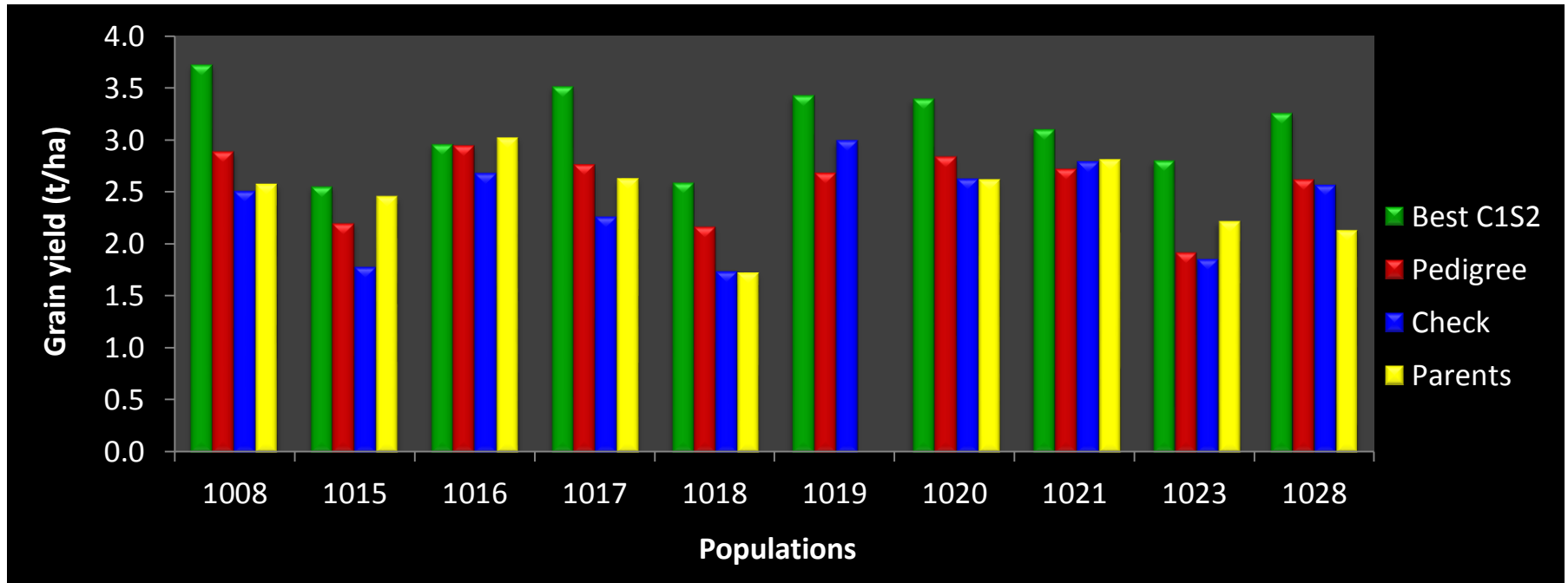
Beyene et al. (2016) Crop Sci. 56:1–10

Performance of hybrids developed from individual C₁S₂ of MARS

- For each of the 10 populations,
 - 47-74 C₁S₂ lines were extracted after 3 cycles of MARS
 - The best five S₅ lines developed through pedigree selection,
 - The founder parents for each of the population
- The above lines were crossed with a single-cross tester from the opposite heterotic group.
- Five commercial checks
- All the hybrids evaluated in 1-3 managed stress and 3-5 well watered conditions with two replications



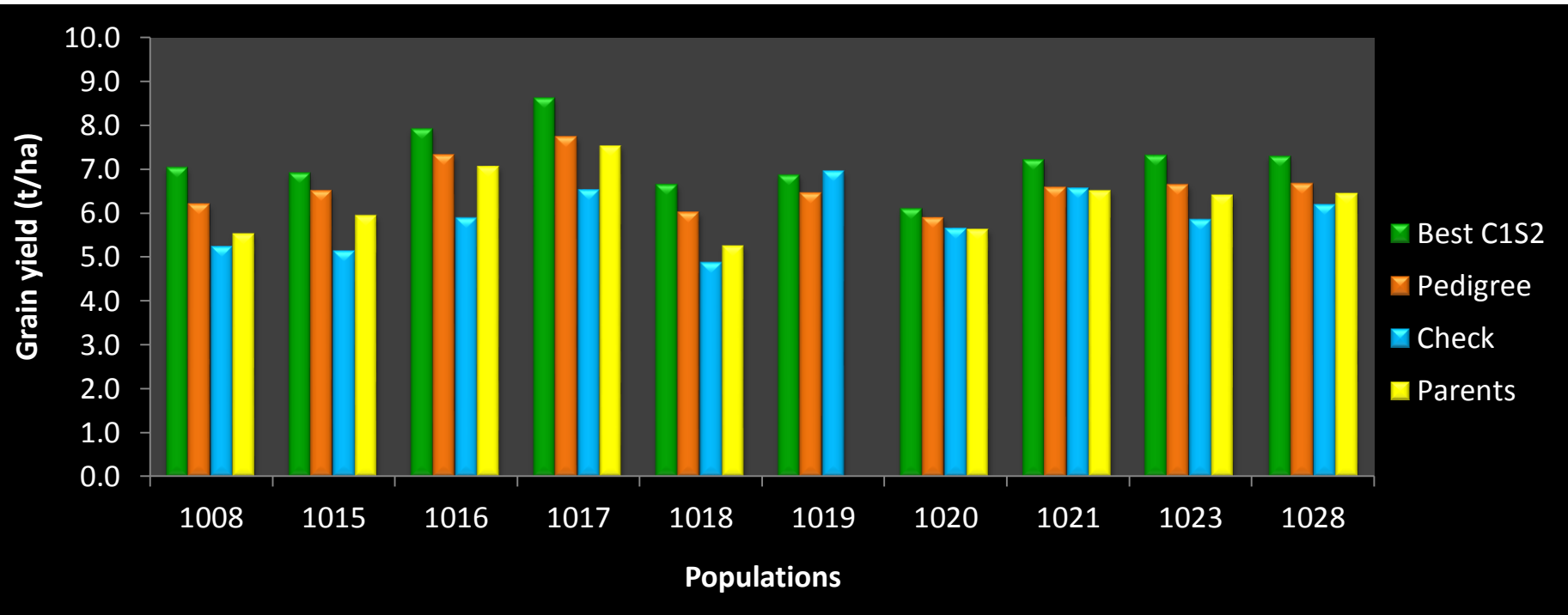
Performance of hybrids developed from individual C1S2 of MARS and pedigree breeding under drought



Heritability = 0.32 to 0.75

Beyene et al. (2015) Euphytica : online first

Performance of hybrids developed from individual C1S2 of MARS and pedigree breeding under optimum



Heritability = 0.10 to 0.49

Beyene et al. (2015) Euphytica : online first

Lines and hybrids developed through Rapid cycle of GS



Fixed lines



Drought tolerant hybrids



Lessons learnt from DTMA and WEMA MARS populations

- MARS/GS requires critical coordination among the different players (population developers, phenotypers, molecular staff, biometricians)
- MARS/GS would be advantages if the process is completed within a short period of time (time saving)
- MARS/GS provided 2-3 fold higher grain yield than pedigree selection without significantly affecting maturity and plant height **BUT**
- Although C3 derived hybrids provided higher grain yield in most populations, some populations showed either reduction or no gain compared with C0.



Challenges

1. Availability of sufficient land for timely multi-location evaluation under drought and optimum locations
2. Low heritability under drought – not every population was advanced
 - Trait selection and assigning weight for developing multi-trait index for selecting among CO families
3. Time pressure to do selection before flowering (leaf sampling, getting the marker data on time, analyzing the data and returning the selection to nursery managers for recombination or selfing)



The way forward on MARS and GS

- About 850 S5 lines extracted from C3 and over 1000 DH lines from C1 and C2 testcrossed and evaluated across locations. The best hybrids will be selected for NPT by NARS partners
- We now have confidence that both MARS and GS provided 2-3 times better grain yield than pedigree selection with early generation testing
 - Detail cost-benefit analyses on MARS, GS vs pedigree selection
 - Developing institutional/GMP-based strategy in integrating MARS and GS in our routine breeding programs for different traits



Other contributors to genetic gain studies (in alphabetical order)

- Amsal Tarekegne
- Barbara Meisel
- Bindiganavile S. Vivek
- Boddupalli M. Prasanna
- Cosmos Mogrogocho
- Dan Makumbi
- Gary N. Atlin
- Gregorio Alvarado
- Jose Crossa
- Kassa Semagn
- Lewis Machida
- Michael Olsen
- Pierre Sehabiague
- Raman Babu
- Stephen Mugo
- CIMMYT research assistants at different locations

Conclusions

- We have tested and validated the effects of different genomic selection models, marker density, genetic architecture, population structure, etc. which all are the prerequisites for integrating GS in maize breeding program
- We have conducted the largest public MARS/GS projects and demonstrated 2-3 fold higher grain yield using MARS and GS than pedigree methods in tropical maize
- Several superior hybrids developed through GS and MARS are currently under national performance trials and expect to release (commercialize) some hybrids soon in SSA



Acknowledgments

- Bill and Melinda Gates Foundation (BMGF)
- United States Agency for International Development (USAID)
- Howard G. Buffet Foundation (HGBF)





**Thank you
for your
interest!**