

Rapid cycling genomic selection in maize landraces

Clara Polzer¹, Hans-Jürgen Auinger¹, Michelle Terán Pineda¹, Armin C. Hölker^{1,a}, Manfred Mayer^{1,b}, Thomas Presterl², Carolina Rivera Poulsen¹, Sofia da Silva², Milena Ouzunova², Albrecht E. Melchinger^{1,3}, Chris-Carolin Schön^{*1}

¹ Plant Breeding, TUM School of Life Sciences, Technical University of Munich, Freising, Germany, 85354

² KWS SAAT SE & Co. KGaA, Einbeck, Germany, 37574

³ Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Stuttgart, Germany, 70593

^a present address: KWS SAAT SE & Co. KGaA, 37574 Einbeck, Germany

^b present address: Bayer Crop Science, 46325 Borken, Germany

*Corresponding author

Email address: chris.schoen@tum.de

Supplemental tables

Table S1 Description of data sets used for model training, selection, and evaluation of selection response. Given are the sample size (N), the number of parents (N parents), the number of families (N families), the average number of progenies per family (N progenies) with minimum and maximum per population in brackets and the years for genotypic or phenotypic evaluation of each data set. The training set is derived from three open-pollinated landraces (OPV).

Data set	N	Origin/ N parents	N families	N progenies	Years of evaluation
Training set					2017/2018
(DH lines)					
PKL	899	Landraces Petkuser Ferdinand Rot (PE), Kemater Landmais Gelb (KE), Lalin (LL)		OPV derived	
C0	402	Landrace PE		OPV derived	
Selection					2019/2020
R1					
C1-S ₁	1003	10	44	22.80 (11, 26)	
C2-S ₀	1005	44	22	45.68 (6, 69)	
R2					
C1-S ₁	1001	8	25	40.04 (19, 42)	
C2-S ₀	906	36	18	50.33 (3, 154)	
DH lines					2022/2023
C0r	105	Landrace PE		OPV derived	
R1					
C1-DH	100	10	26	3.84 (1, 10)	
C2-DH	104	36	18	5.78 (1, 27)	
C3-DH	90	28	14	6.43 (1, 14)	
R2					
C1-DH	100	8	20	5.00 (1, 29)	
C2-DH	111	34	17	6.53 (1, 17)	
C3-DH	78	32	16	4.88 (1, 20)	

Table S2 Traits, abbreviations, description of trait collection and environments for each trait.

Trait	Abbreviation	Description	Environments	Number of Environments
Early plant height	PH_V4 and PH_V6	In cm, from ground to longest, manually vertically stretched leaf, in V4 and V6 stage. Averaged over three plants.	ROG 2022 EIN 2022/23 BBG 2022/23 OLI 2022/23	7
Final plant height	PH	In cm, from ground to first branch of the tassel. Averaged over three plants.	ROG 2022 EIN 2022/23 BBG 2022/23 OLI 2022/23	7
Days to female flowering	FF	Days after sowing when 50% of the plants have a fully developed female flower (upper 1-2 cm of the silk visible).	ROG 2022 EIN 2022/23 BBG 2022/23 OLI 2022/23	7

Table S3 Means, genotype (σ_g^2), genotype $\times$ environment (σ_{gl}^2) and error (σ_e^2) variance components, each with their standard error, and heritabilities with confidence intervals for data sets C0 and its random subset C0r evaluated in 11 environments in 2017/18 for the selection traits plant height in V4, V6 and final growth stage (PH_V4, PH_V6, PH) and for days to female flowering (FF).

	PH_V4		PH_V6	
	C0	C0r	C0	C0r
Mean $\pm$ SE	43.8 $\pm$ 0.3	43.9 $\pm$ 0.5	83.7 $\pm$ 0.5	83.8 $\pm$ 0.8
$\sigma_g^2 \pm$ SE	24.8 $\pm$ 2.0	21.5 $\pm$ 3.3	77.2 $\pm$ 6.0	62.1 $\pm$ 9.6
$\sigma_{gl}^2 \pm$ SE	18.4 $\pm$ 0.7	14 $\pm$ 1.1	38.7 $\pm$ 1.7	33.4 $\pm$ 2.9
h^2 (CI)	0.91 (0.90, 0.92)	0.92 (0.89, 0.94)	0.93 (0.92, 0.94)	0.92 (0.90, 0.94)

	PH		FF	
	C0	C0r	C0	C0r
Mean $\pm$ SE	125 $\pm$ 1	127 $\pm$ 2	79.2 $\pm$ 0.2	78.3 $\pm$ 0.3
$\sigma_g^2 \pm$ SE	285 $\pm$ 21	221 $\pm$ 33	14.8 $\pm$ 1.1	10.2 $\pm$ 1.6
$\sigma_{gl}^2 \pm$ SE	91 $\pm$ 4	95 $\pm$ 8	4.2 $\pm$ 0.2	3.7 $\pm$ 0.4
h^2 (CI)	0.96 (0.95, 0.96)	0.94 (0.92, 0.96)	0.95 (0.95, 0.96)	0.94 (0.92, 0.95)

Table S4 Phenotypic correlations between the selection traits calculated in DH populations C0r to C3. The selection criterion applied in each cycle aimed at directional selection for early plant height in growth stage V4 and V6 (PH_V4, PH_V6) and stabilizing selection for final plant height (PH).

	PH_V4 - PH_V6		PH_V4 - PH		PH_V6 - PH	
	R1	R2	R1	R2	R1	R2
C0r	0.87*		0.43*		0.52*	
C1	0.90*	0.88*	0.35*	0.34*	0.42*	0.33*
C2	0.82*	0.90*	0.27	0.40*	0.49*	0.49*
C3	0.80*	0.76*	0.24	0.26	0.47*	0.40*

* Significant at the 0.05 probability level

Table S5 Bootstrap results for differences in prediction ability between pairs of DH populations obtained with the training set PKL\C0r in replication R1 and R2 (Rep). Q2.5 and Q97.5 are the 2.5% and 97.5% quantiles of the bootstrap distribution. Population pairs with a significant difference in prediction ability (95% bootstrap interval), are marked in bold.

Rep	Population pair	PH_V4			PH_V6			PH			FF		
		Q2.5	Mean	Q97.5	Q2.5	Mean	Q97.5	Q2.5	Mean	Q97.5	Q2.5	Mean	Q97.5
R1	C0r vs. C1	-0.28	-0.03	0.22	-0.27	-0.03	0.20	-0.12	0.07	0.27	-0.19	0.02	0.23
	C0r vs. C2	-0.19	0.07	0.32	-0.10	0.16	0.42	-0.13	0.06	0.25	-0.22	-0.02	0.17
	C0r vs. C3	-0.17	0.08	0.34	-0.14	0.12	0.39	-0.05	0.16	0.37	-0.19	0.01	0.21
	C1 vs. C2	-0.15	0.10	0.34	-0.05	0.20	0.44	-0.22	-0.01	0.19	-0.24	-0.04	0.17
	C1 vs. C3	-0.14	0.12	0.37	-0.10	0.16	0.41	-0.14	0.09	0.30	-0.21	0.00	0.20
	C2 vs. C3	-0.24	0.02	0.28	-0.32	-0.04	0.23	-0.12	0.10	0.31	-0.16	0.03	0.22
R2	C0r vs. C1	-0.16	0.12	0.42	-0.16	0.12	0.40	-0.17	0.04	0.24	-0.23	-0.03	0.17
	C0r vs. C2	0.09	0.35	0.59	0.18	0.44	0.70	-0.08	0.12	0.31	-0.19	0.01	0.20
	C0r vs. C3	-0.28	-0.04	0.20	-0.24	0.02	0.27	-0.16	0.03	0.22	-0.15	0.07	0.29
	C1 vs. C2	-0.07	0.23	0.51	0.03	0.32	0.60	-0.14	0.08	0.30	-0.16	0.04	0.24
	C1 vs. C3	-0.45	-0.16	0.12	-0.39	-0.10	0.18	-0.22	-0.01	0.20	-0.12	0.10	0.33
	C2 vs. C3	-0.63	-0.39	-0.14	-0.68	-0.42	-0.16	-0.29	-0.09	0.11	-0.15	0.06	0.28

Table S6 Adjusted coefficients of determination ($R^2_{adj.}$) for the multiple regression analyses of the selection criterion SC (see Eq. 2) on the underlying components for three populations. C0 denotes the DH population from landrace PE and C1-S₁ and C2-S₀ the populations comprising the selection candidates in cycles C1 and C2. Predictor variables were the genomic estimated breeding values (GEBVs) for PH_V4, PH_V6, and PH_trans as described in Materials and Methods.

Predictor variable(s)	Population				
	C0	C1-S ₁		C2-S ₀	
		R1	R2	R1	R2
PH_V4	0.70	0.65	0.71	0.68	0.50
PH_V6	0.74	0.65	0.71	0.66	0.57
PH	0.38	0.21	0.43	0.17	0.21
PH_trans	0.29	0.20	0.20	0.20	0.31
PH_V4 + PH_V6	0.74	0.69	0.73	0.70	0.58
PH_V4 + PH	0.71	0.65	0.71	0.68	0.50
PH_V4 + PH_trans	0.98	0.95	0.98	0.96	0.95
PH_V6 + PH	0.74	0.65	0.72	0.67	0.57
PH_V6 + PH_trans	0.98	0.95	0.98	0.96	0.93
PH_V4 + PH_V6 + PH	0.74	0.69	0.73	0.70	0.59

Supplemental figures

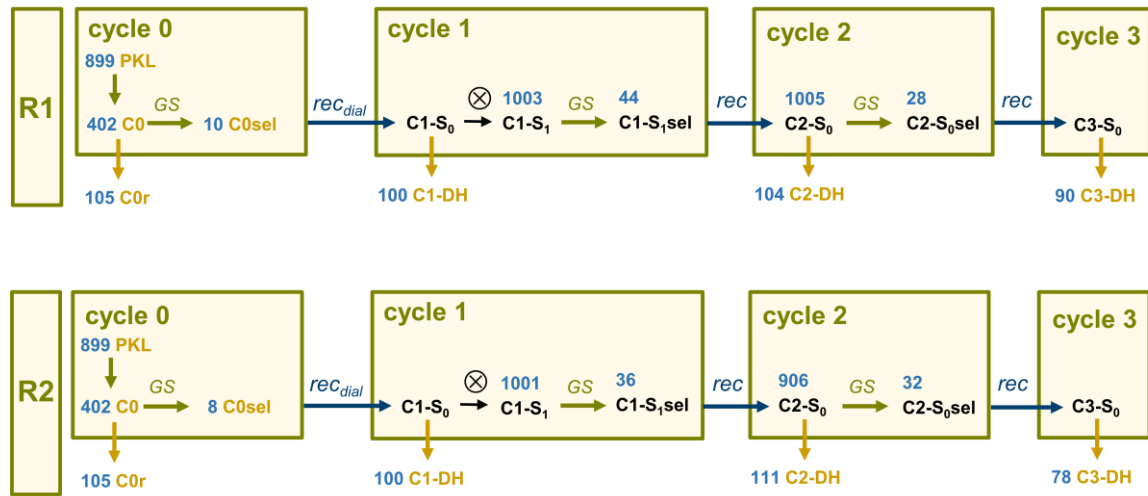


Fig. S1 Rapid cycling genomic selection (GS) scheme for early and final plant height traits, replicated twice (R1, R2). Following model training with a set of 899 DH lines, the best C0sel DH lines from population PE (C0) were selected based on a multi-trait selection criterion, recombined in a diallel crossing scheme and selfed once to produce the C1-S₁ population. Subsequently, two cycles of GS and mating of the selected candidates in pairs were conducted. From each selection cycle, DH lines were sampled (C0r) or produced (C1-DH, C2-DH, C3-DH). Blue digits indicate the number of selection units, selected fractions and phenotyped and genotyped DH lines

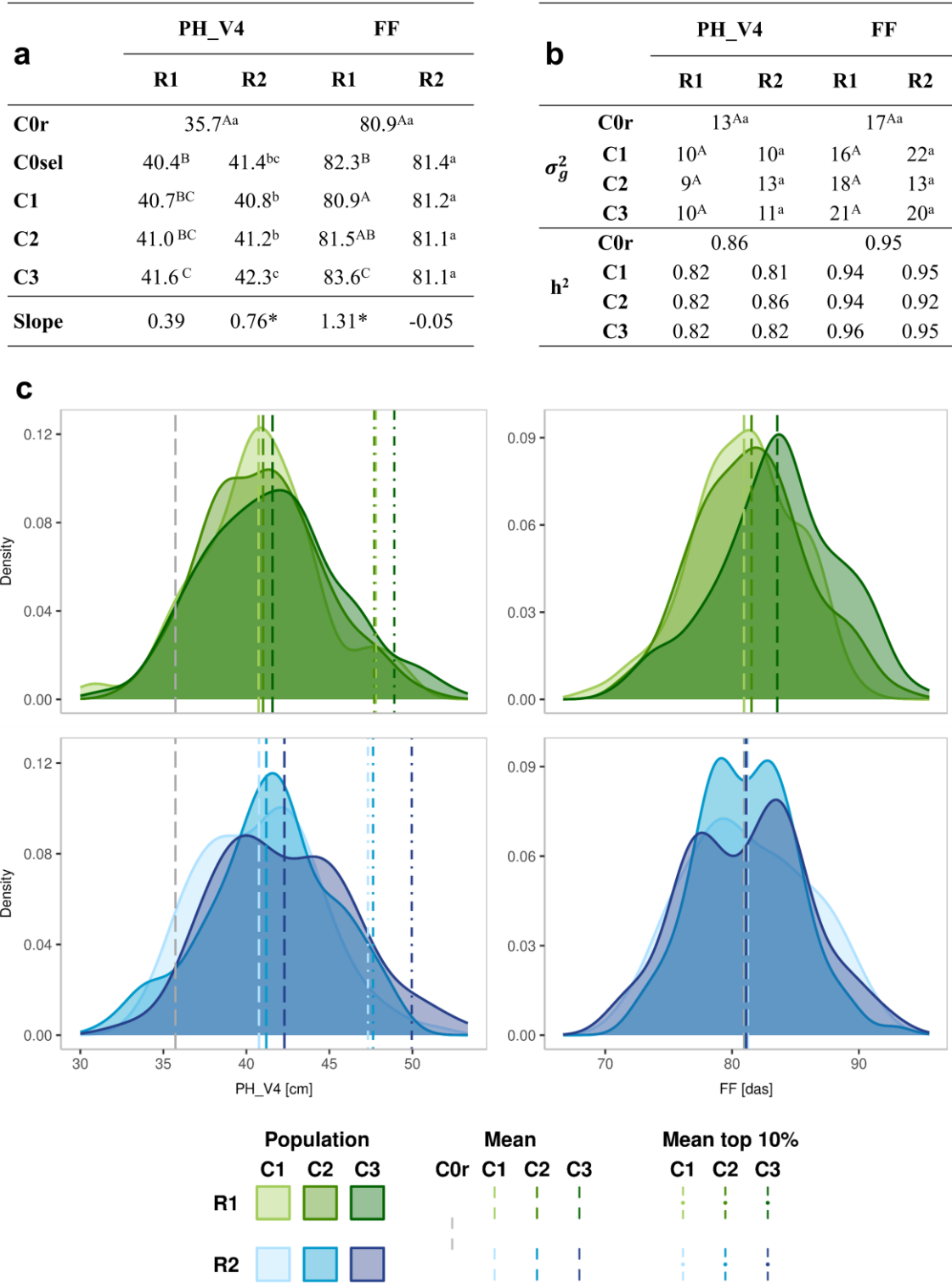


Fig. S2 Response to selection based on adjusted entry means of populations in two replications (R1 and R2) of the experiment. **a** Means of DH populations for plant height in growth stage V4 (PH_V4) and days to female flowering (FF), and the coefficient of linear regression indicating selection response per cycle from C1 to C3 (Slope). **b** Genetic variances and heritabilities. Different letters (upper case for R1; lower case for R2) in **a** and **b** indicate significant differences between populations ($P < 0.05$). **c** Density distribution of adjusted entry means of DH lines in each population. The dashed lines indicate the population means, the dashed-dotted lines the means of the best performing 10% per population for PH_V4