Supplements

Data preparation and quality control

Both datasets were combined based on reference genome V2 positions into the hapmap format. For SNPs where the opposite strand was targeted by the two array platforms the corresponding alleles were converted to their complementary basepair and compared to the landrace population reference. We removed insertions, unmapable SNPs (`chr0` and duplicated SNPs), non-polymorphic sites, and SNPs with quality classes 'off-target variant' and 'call rate below threshold'. We further removed SNPs that violated the Hardy-Weinberg equilibrium ($\chi^2, mid - p < 0.05$) in at least one LR accession from all LRs using plink 1.9 (Chang *et al.* (2015), see Table S2). A vcf file for the whole dataset was constructed using TASSEL 5 (Bradbury *et al.* 2007). Accession filtered datasets were written using custom R scripts with various packages.

Imputation allows haplotype analysis

In our study we relied on published genotyping data based on two different genotyping arrays (Ganal *et al.* 2011; Unterseer *et al.* 2014). While the data was highly consistent and reconstructed the population structure correctly (Figure 1A and S5), these platforms come with several limitations, reducing the ability to detect rare and potentially deleterious alleles and reductions of diversity. By imputing the DH dataset we combined phase information of both the LR populations and the DH lines and were able increased the SNP density for the DH data. The imputed dataset enabled us to identify major reductions of mean haplotype diversity in polymorphic windows and large regions of complete fixation in the DH populations. The extent of this loss of diversity could only be detected using imputed genotypes. While imputation came with an error rate of 10.6 % to 15.9 %, we were able to increase the number of sites in the DH lines from 37 thousand to over 530 thousand. Imputational error rates depend highly on the reference panel used, MAF, SNP density and chromosome sample size; error rates in the literature range from 1 % to 15 % (Browning and Browning 2009; Howie *et al.* 2009; Khatkar *et al.* 2012). We show that the estimated imputational error rate is randomly distributed across the genome (Figure S2) and the correlation with the genetic distance of neighboring SNPs is low (Figure S1). Furthermore, the mean haplotype diversity of the unimputed 50k dataset (Figure S11) showed significant reductions of DH diversity compared to the LR in every accession and we found corresponding outliers using imputed and unimputed data (e.g., chromosome 3, near centromere). Additionally, we reanalyzed the fate haplotypes of the unimputed 50k and used windows based on LD (window size 0.2 cM). This resulted a reduction of fixed major haplotypes, but in similar amounts of lost major haplotypes confirm our findings of the 600k data run. Hence, we conclude that the trade-off between maker density and imputation error is justified for the haplotype analysis, and the information gain associated with imputation overcomes the loss in statistical power due to undetected genetic diversity in the DH. While genotyping arrays have high genotyping accuracy for called SNPs, future studies should use genome-wide sequencing to avoid imputation and ascertainment issues. This would allow harnessing the full potential of DH lines from landraces to study the causes of inbreeding depression in maize.

Table S1 Sample sizes for DH and LR

Population	DH	LR	Sum
BU	36	22	58
GB	59	46	105
RT	44	23	67
SC	58	23	81
SF	69	23	92
Sum	**266**	**137**	**403**

Table S2 Number of SNPs removed during quality control

stage	removed
duplicated SNPs	389
Chromosome 0	310
non polymorphic sites	77798
Insertions	107
quality tag: `CallRateThresh`	846
quality tag: off-target variant (`OTV`)	2747
violated HW	814
SUM removed SNPs	**83011**

Table S3 Datasets used in this study

Dataset	Populations	SNPs	Individuals
50k DH	BU, GB, RT, SC, SF	37,967	266
50k LR	BU, GB, RT, SC, SF	37,967	137
600k DH	BU, GB, RT, SC, SF	533,190	266
600k LR	BU, GB, RT, SC, SF	533,190	137
GWAS DH	CG, EF, GB, RT, SF, SM, WA	37,884	404

Table S4 Number of outlier SNPs identified in the aSFS test

Overlap	BU	GB	RT	SC	SF		SUM unique
1	2516	849	4301	1193	446		9305
2	915	566	1176	695	402		1877
3	252	340	320	372	333		539
4	161	276	260	290	281		317
5	307	307	307	307	307		307
							12345
SUM	4151	2338	6364	2857	1769	**17479**	**SUM**
Outlier %	11.15 %	6.16 %	16.78 %	7.53 %	4.66 %		

Table S5 Fate of most common haplotypes in a total of 34,833 50kb windows in the 600k data.

Accession	fixed	lost	segregating	Sum (windows)	Sum (haplotypes)	fixed %	lost %	segregating %
BU	3774	6723	13186	23683	209982	15.94	28.39	55.68
GB	3222	3917	16684	23823	244164	13.52	16.44	70.03
RT	5999	3682	12339	22020	113418	27.24	16.72	56.04
SC	5199	5138	12993	23330	160030	22.28	22.02	55.69
SF	4806	6106	12837	23749	194009	20.24	25.71	54.05

Table S6 ANOVA tables of the outlier characterization using GWAS effect sizes. The term 'outlier' refers to the SNPs classified as 'outlier' or 'non-outlier' in the aSFS.

Trait	Term	df	SumSq	MeanSq	F-value	p-value
shoot vigor	outlier	1	1.65e-07	1.65e-07	1.57	0.21
	frequency bin	9	4.89e-05	5.44e-06	51.8	1.45e-94
	outlier:frequency bin	9	1.34e-05	1.48e-06	14.2	4.26e-23
	Residuals	113000	0.0118	1.05e-07		
female flowering	outlier	1	9.94e-08	9.94e-08	0.0264	0.871
	frequency bin	9	0.00197	0.000218	58.1	1.35e-106
	outlier:frequency bin	9	0.000572	6.35e-05	16.9	3.48e-28
	Residuals	113000	0.424	3.76e-06		
fusarium	outlier	1	1.11e-07	1.11e-07	0.0798	0.778
	frequency bin	9	0.000266	2.95e-05	21.1	4.13e-36
	outlier:frequency bin	9	3.61e-05	4.01e-06	2.87	0.00217
	Residuals	113000	0.157	1.4e-06		
grain yield	outlier	1	5.07e-05	5.07e-05	1.48	0.224
	frequency bin	9	0.0097	0.00108	31.5	1.1e-55
	outlier:frequency bin	9	0.00344	0.000382	11.2	1.31e-17
	Residuals	113000	3.86	3.42e-05		
oil content	outlier	1	9.52e-06	9.52e-06	23.7	1.12e-06
	frequency bin	9	2.29e-05	2.55e-06	6.35	4.75e-09
	outlier:frequency bin	9	3.49e-05	3.87e-06	9.65	6.95e-15
	Residuals	113000	0.0452	4.01e-07		

(continued next page)

Table S6 (continued) ANOVA tables of the outlier characterization using GWAS effect sizes.

Trait	Term	df	SumSq	MeanSq	F-value	p-value
plant height	outlier	1	0.000502	0.000502	10.5	0.00119
	frequency bin	9	0.033	0.00367	76.8	1.44e-142
	outlier:frequency bin	9	0.0124	0.00138	28.9	9.98e-51
	Residuals	113000	5.38	4.78e-05		
protein content	outlier	1	8.13e-07	8.13e-07	2.36	0.124
	frequency bin	9	0.000184	2.04e-05	59.3	5.19e-109
	outlier:frequency bin	9	3.04e-05	3.37e-06	9.8	3.76e-15
	Residuals	113000	0.0388	3.44e-07		

Additional figures

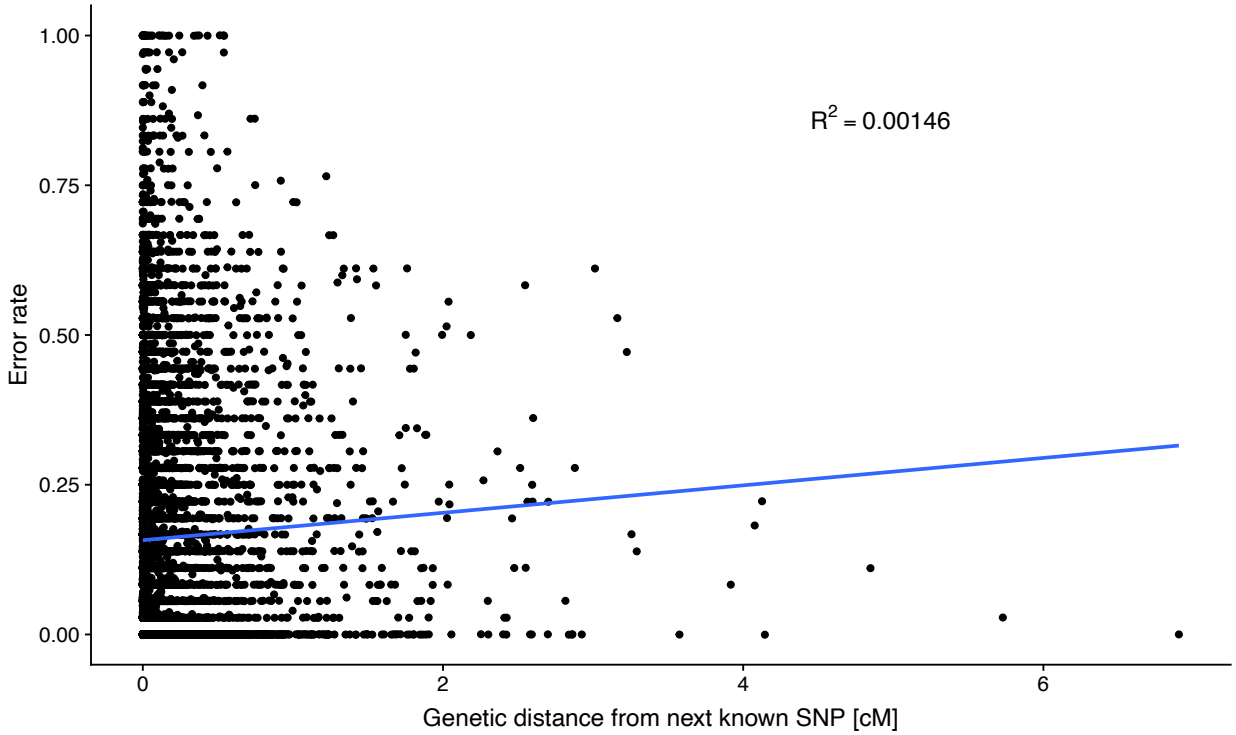

Figure S1 Correlation of marker density and imputation error rate shows low R^2.

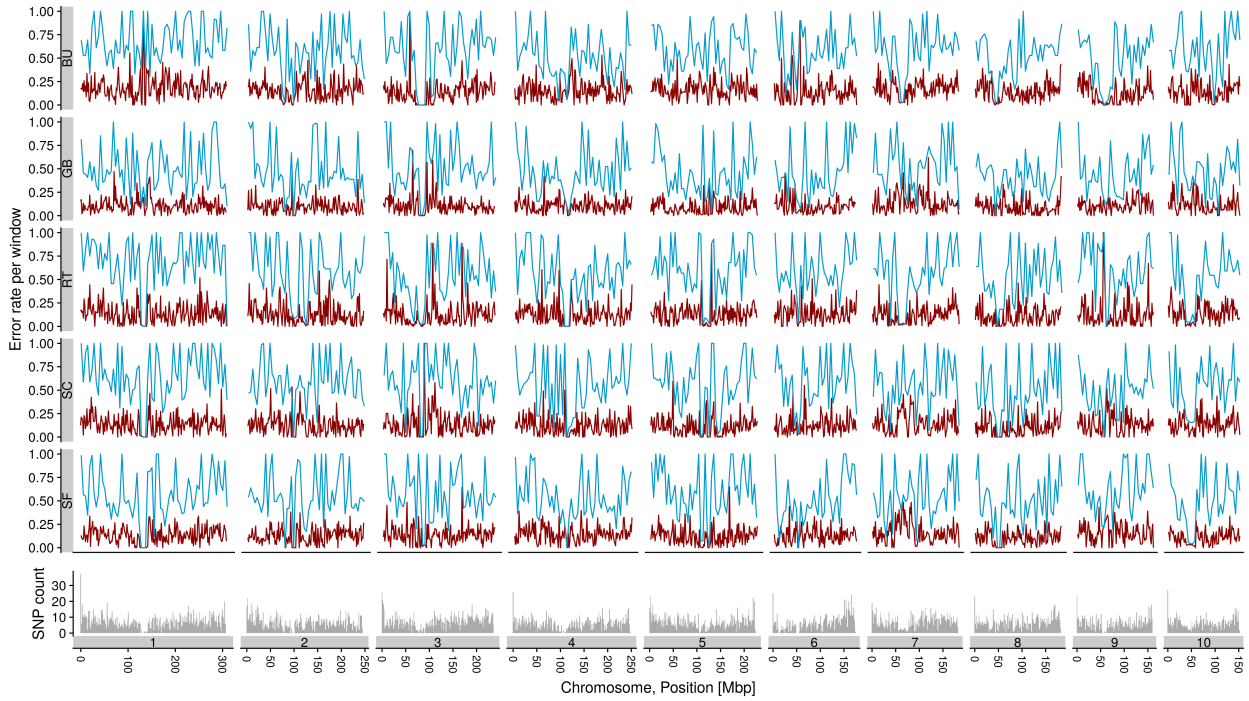

Figure S2 Imputation error rate for DH lines in five accessions. 10,000 known random SNPs were dropped and imputed to compute the error rate represented by mean error in 1.5 Mbp window (red line) and maximum error in 4.5 Mbp window (blue line). SNP density of the Illumina chip in 1.5 Mbp windows shown in bottom panel.

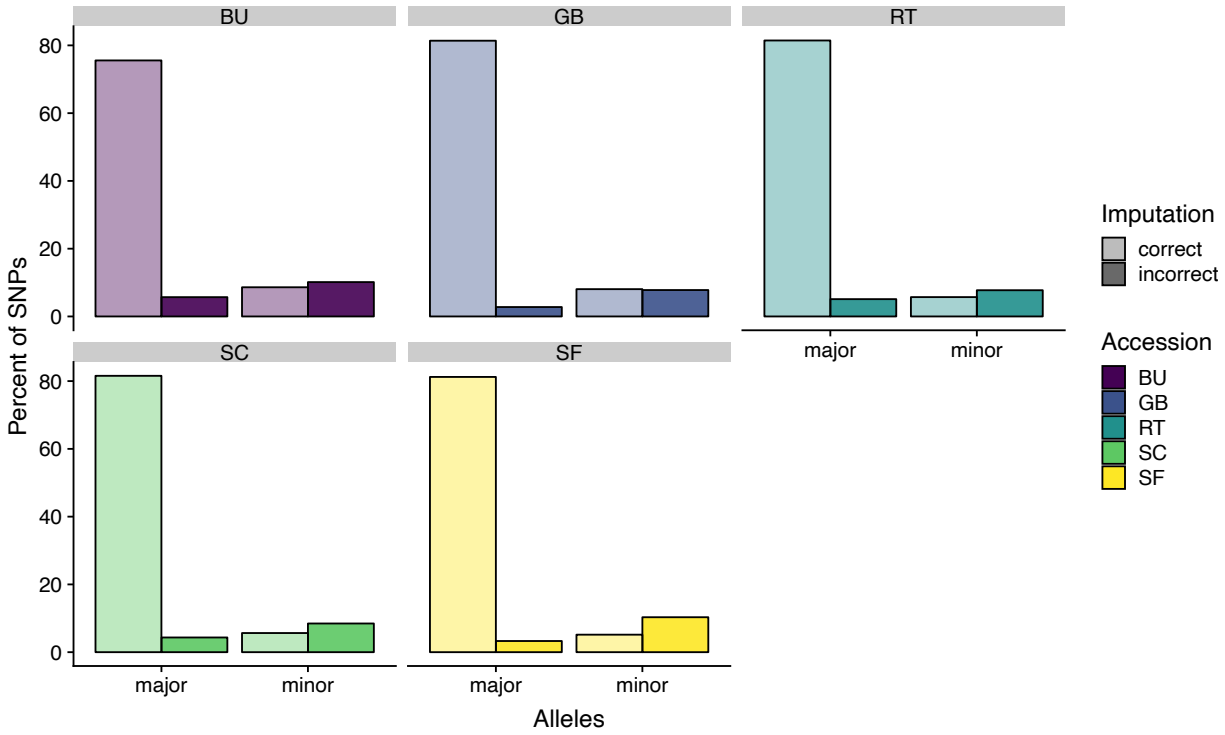

Figure S3 Percentages of correctly and incorrectly imputed SNPs in the imputation test runs.

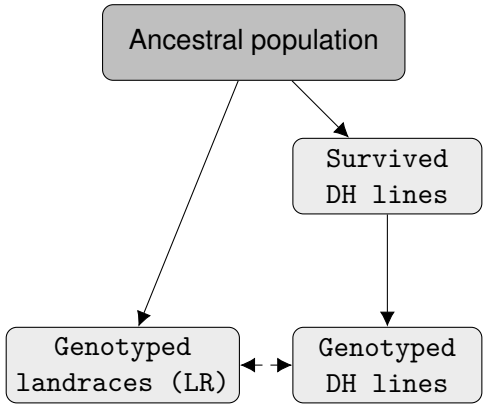

Figure S4 Assumed, simplified sampling structure for DH and LR we used to calculate ancestral frequencies and *p*-values.

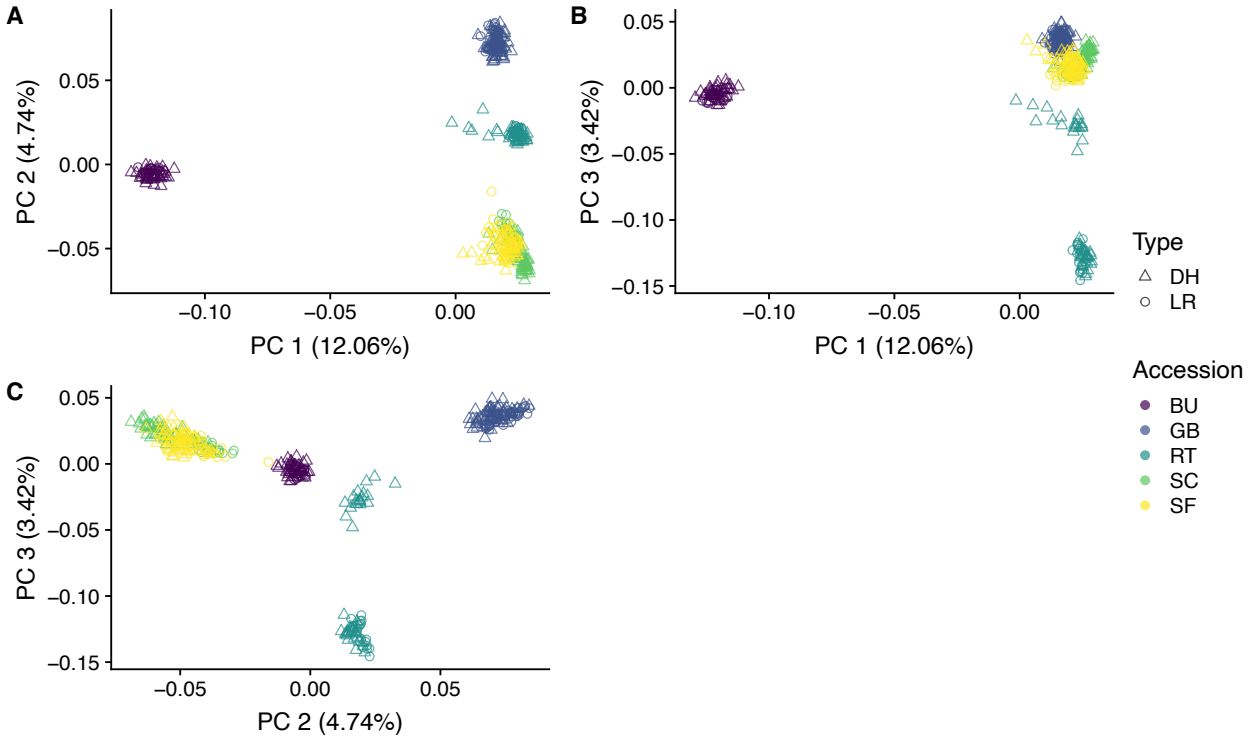

Figure S5 Principal component analysis for DH and LR of the 50k dataset, plot of principal component 1 and 2 (A) shows common clusters for LR and DH in respective accessions. However, principal component 3 separates the DH set of accession RT (B, C).

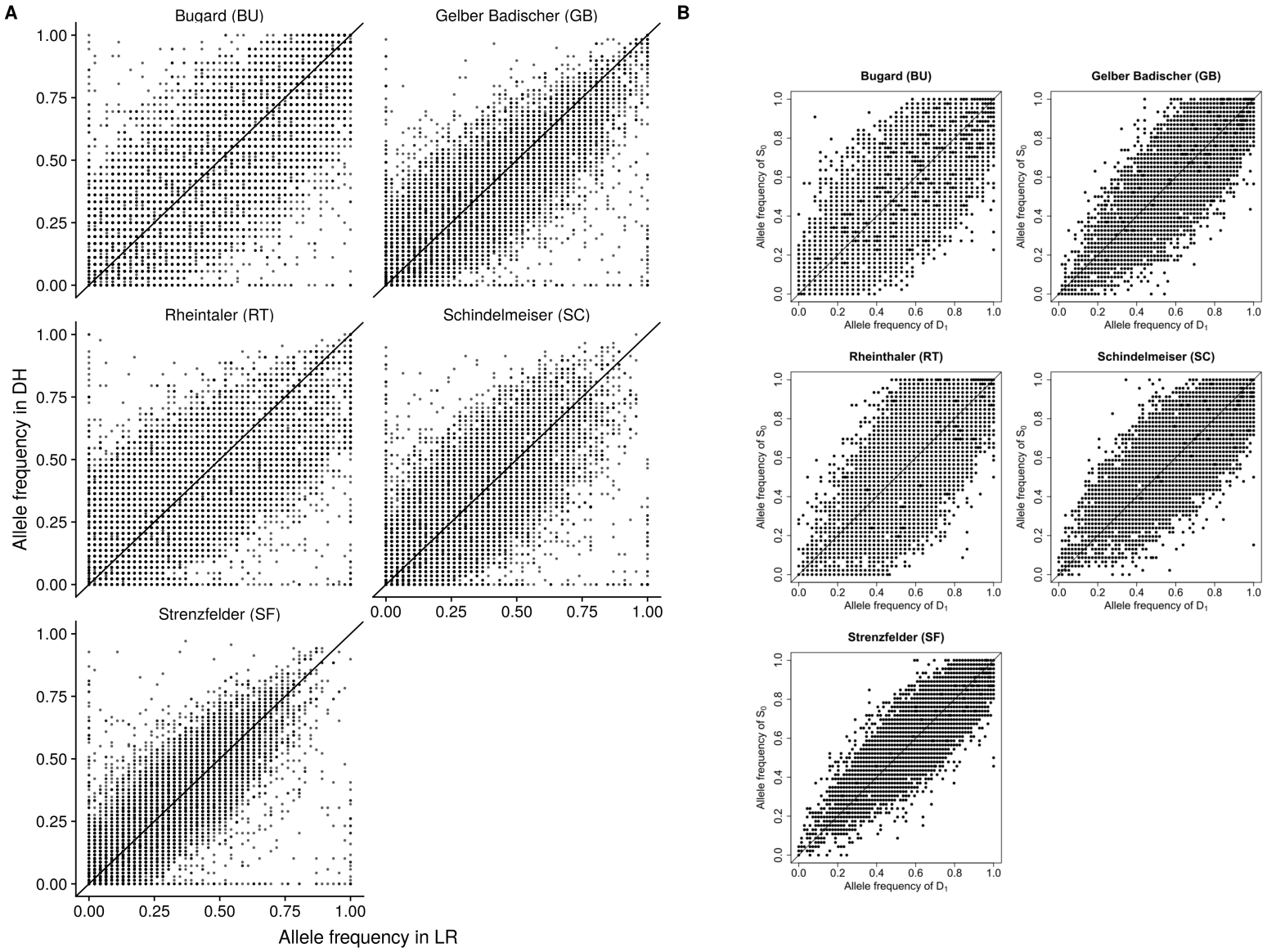

Figure S6 Comparison of joint frequency spectra with published data. (A) Landrace population and DH line alternative allele frequencies estimated from filtered supplementary dataset from Melchinger *et al.* (2017). (B) For comparison published figure (Melchinger *et al.* 2017) with reference allele frequencies depicted on the axes.

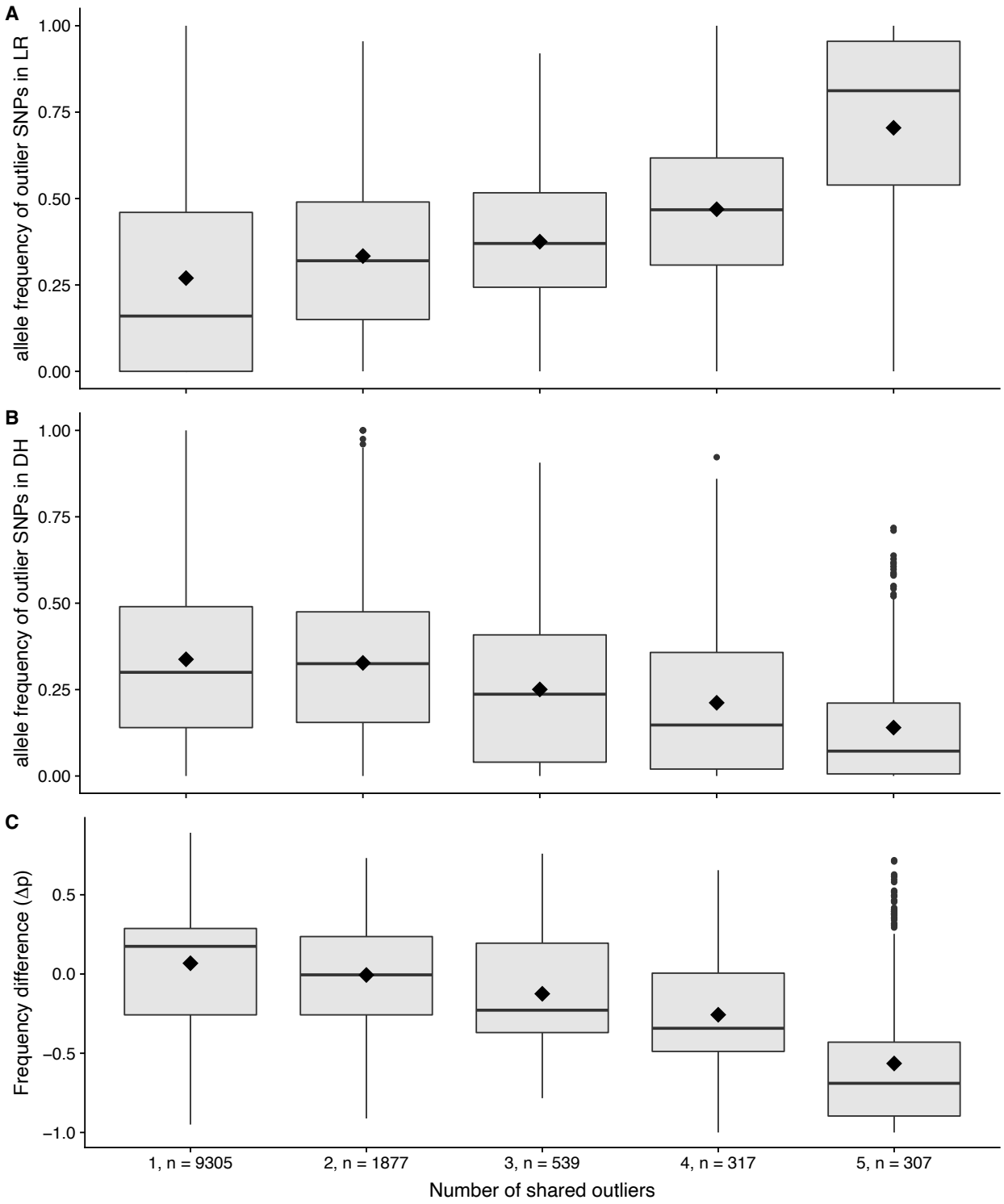

Figure S7 Mean allele frequencies in LR populations (A) and DH lines (B) calculated per number of shared outlier alleles. Outlier alleles, that are shared more often across populations are more likely to have low frequencies and to be lost, while unique outliers change only little in frequency (C). Numbers on x axis ticks correspond to number of shares and number of alleles in this column.

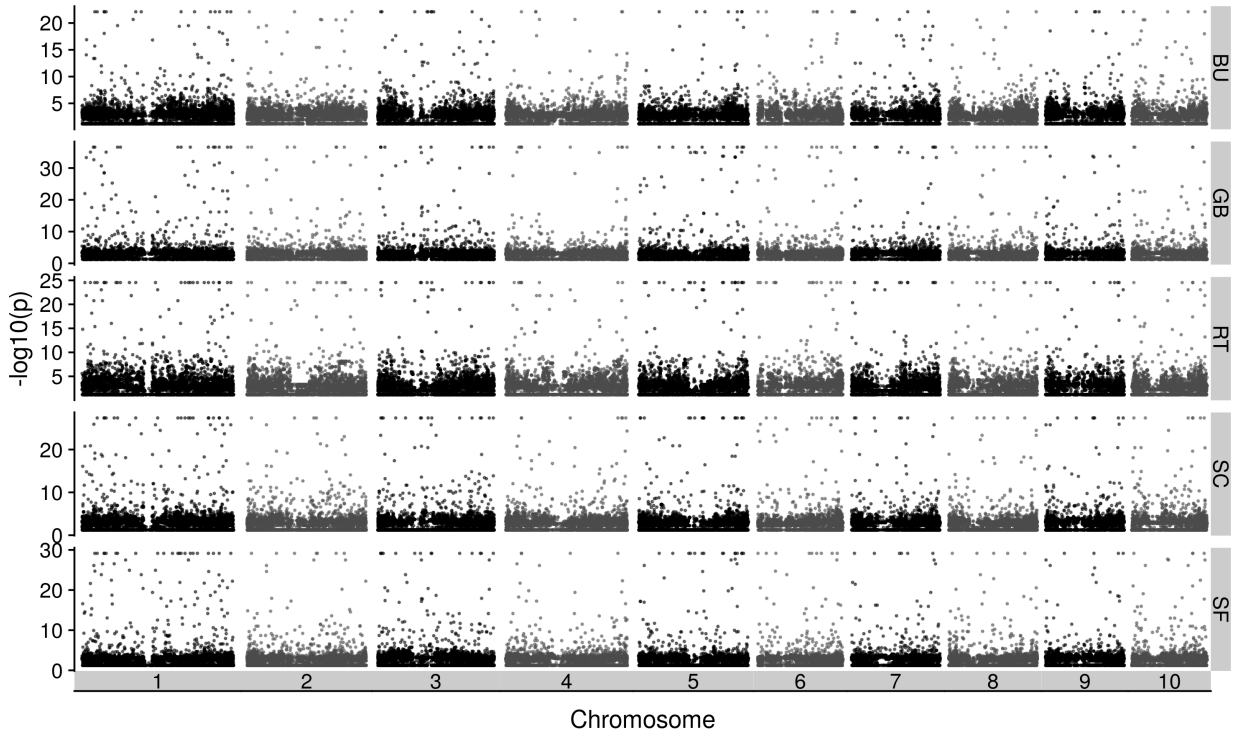

Figure S8 Joint probabilities of DH and LR allele frequency for all accessions and chromosomes.

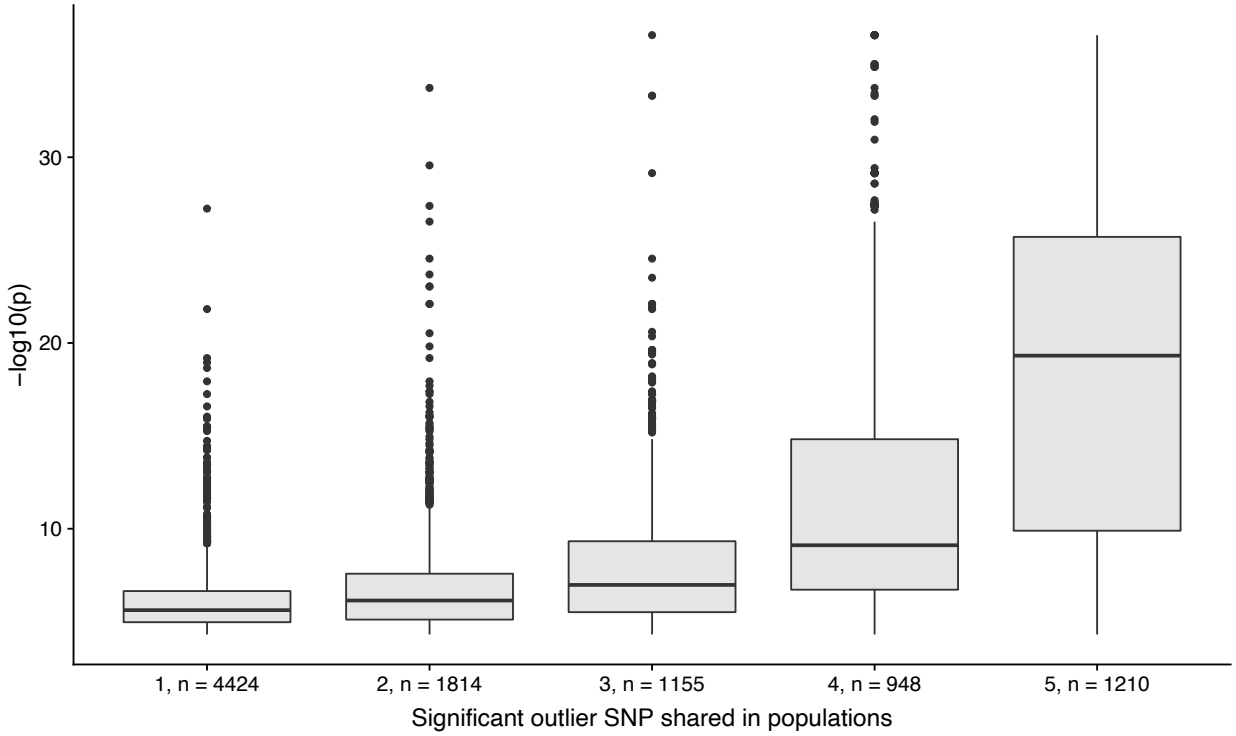

Figure S9 Shared significant $-log_{10}(p)$ values of the probability test and their overlaps among accessions show that high values are found primarily in frequently shared outlier SNPs. Numbers on x-axis correspond to the shared populations and number of outlier-SNPs in this class.

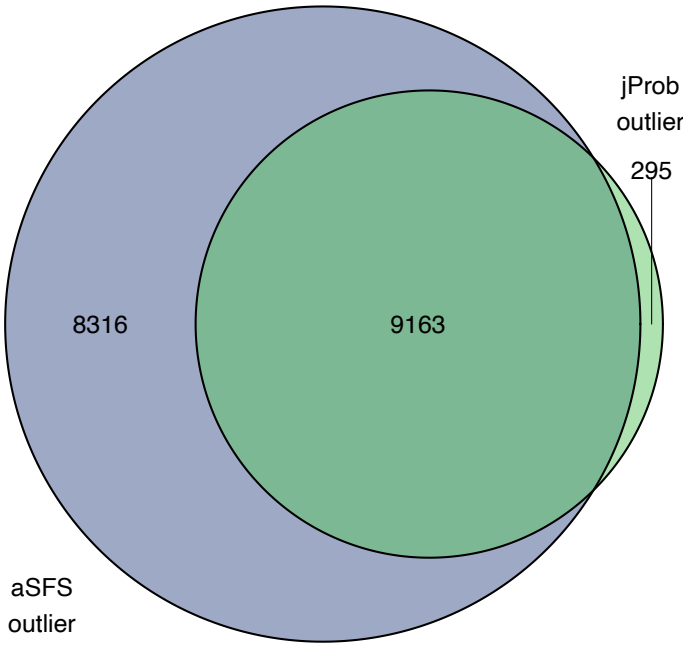

Figure S10 Shared outliers of aSFS and joint probability (jProb) tests, numbers in circles refer to summarized number of outlier in category.

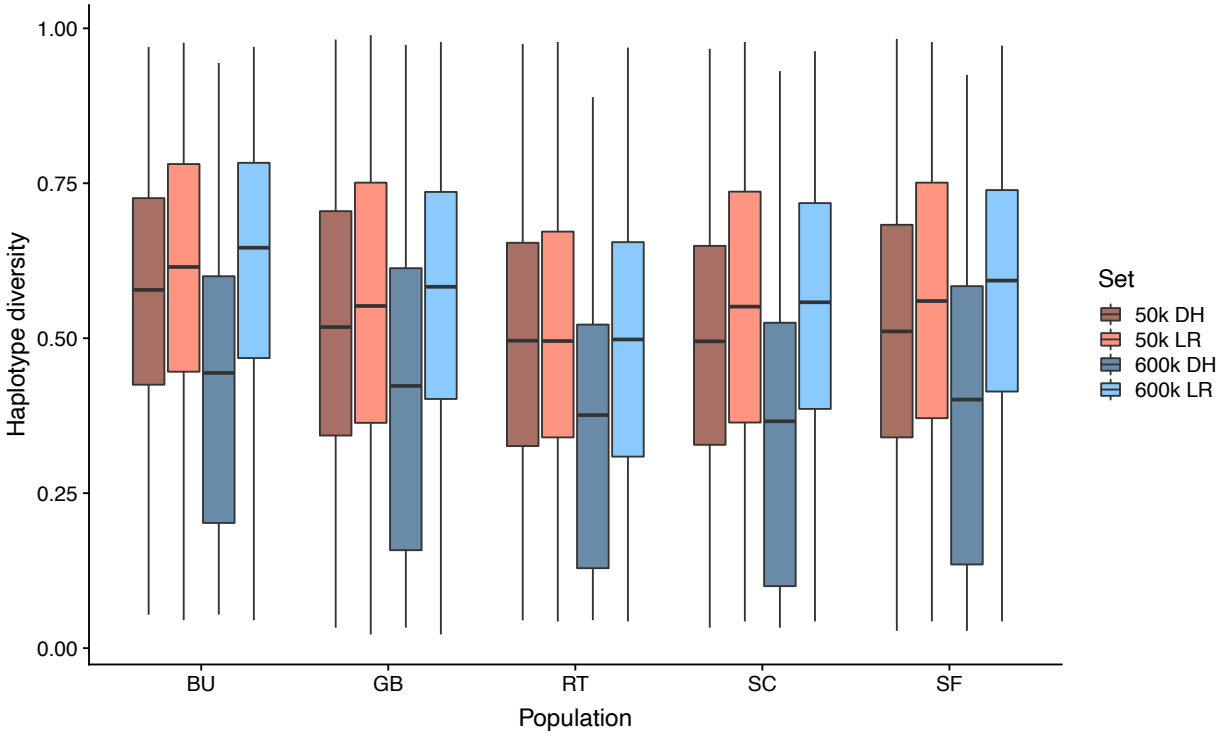

Figure S11 Comparison of haplotype diversity in 50k and 600k datasets in 50kb windows with more than one haplotype shows increase in haplotype diversity in the 600k LR dataset compared to the 600k DH, as well as a reduction of 50k LR compared to 600k DH. This difference is not visible in the DH, indicating that some diversity is missed during the imputation.

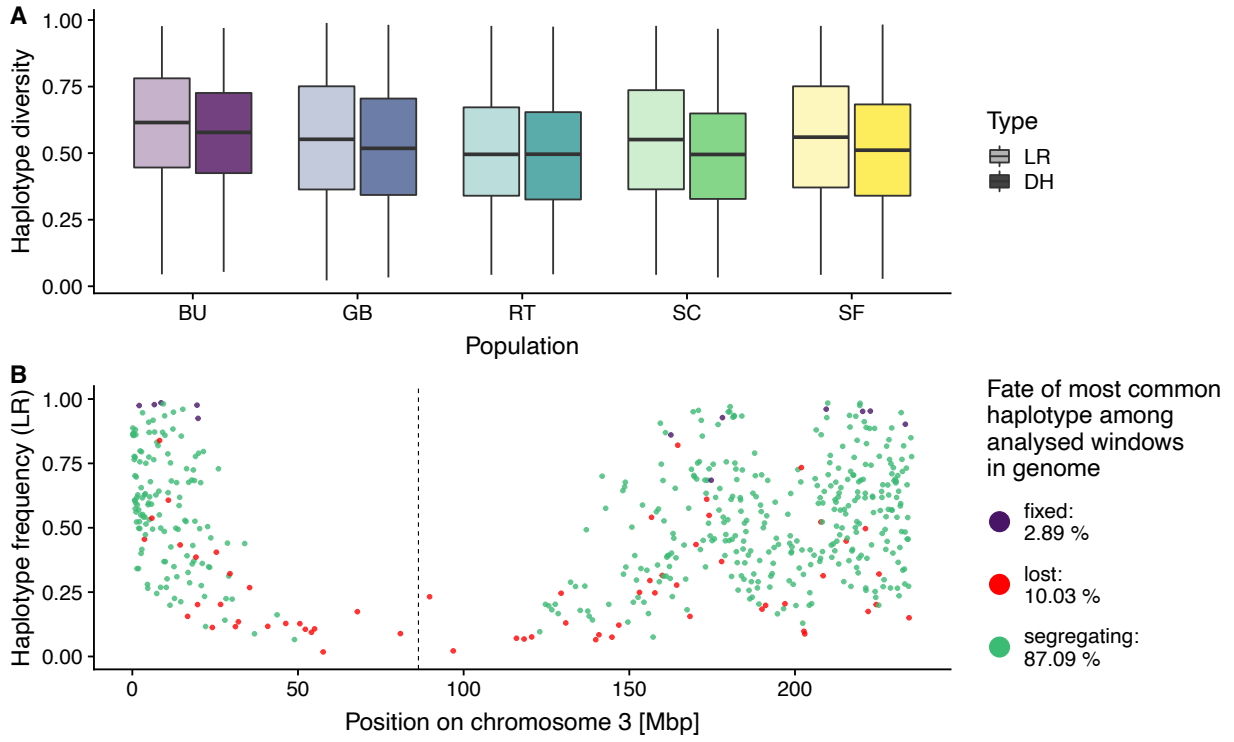

Figure S12 Similar to Figure 3 we analysed haplotypes of the 50k data using windows based on genetic distance (0.2 cM windows). (A) The reduction in haplotype diversity is less pronounced due to higher ascertainment bias and lower SNP density of the 50k chip and imputational error of the 600k DH. (B) Fixed major haplotypes occur very rarely, lost haplotype, however, quite frequently. Reduced SNP density and differences in ascertainment panel sizes between 50k and 600k arrays remove the signal of the putative inversion.

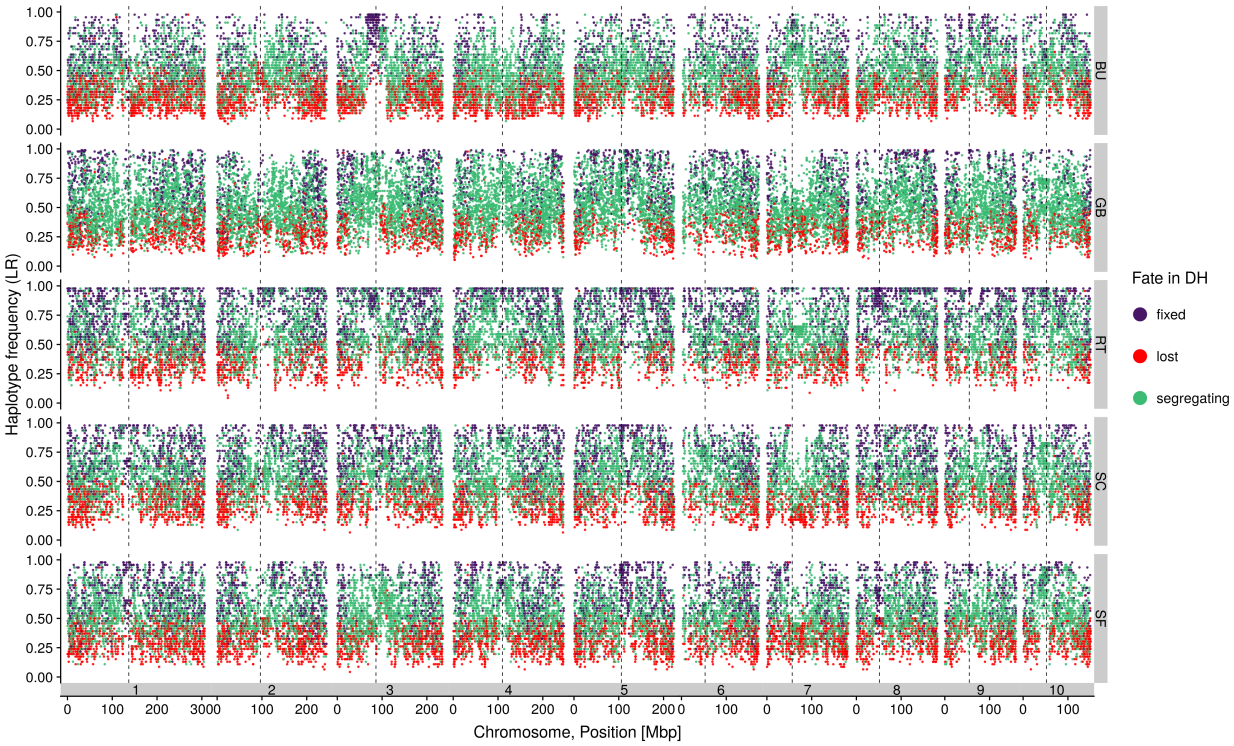

Figure S13 Fate of the most common haplotypes in all accessions. Centromeres are shown as dashed lines.

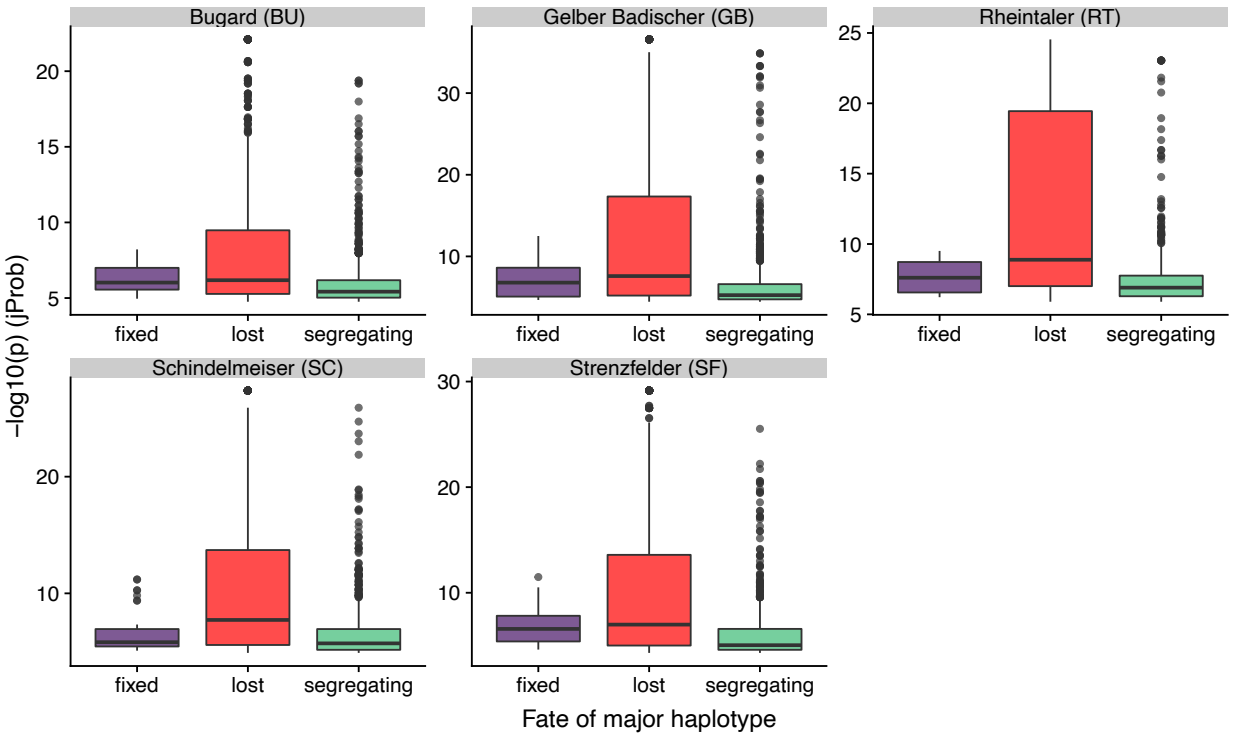

Figure S14 Fate of major haplotypes in 50kb windows outliers from the joint probability test reveal the highest significance levels for outlier in regions with large scale losses of haplotypes.

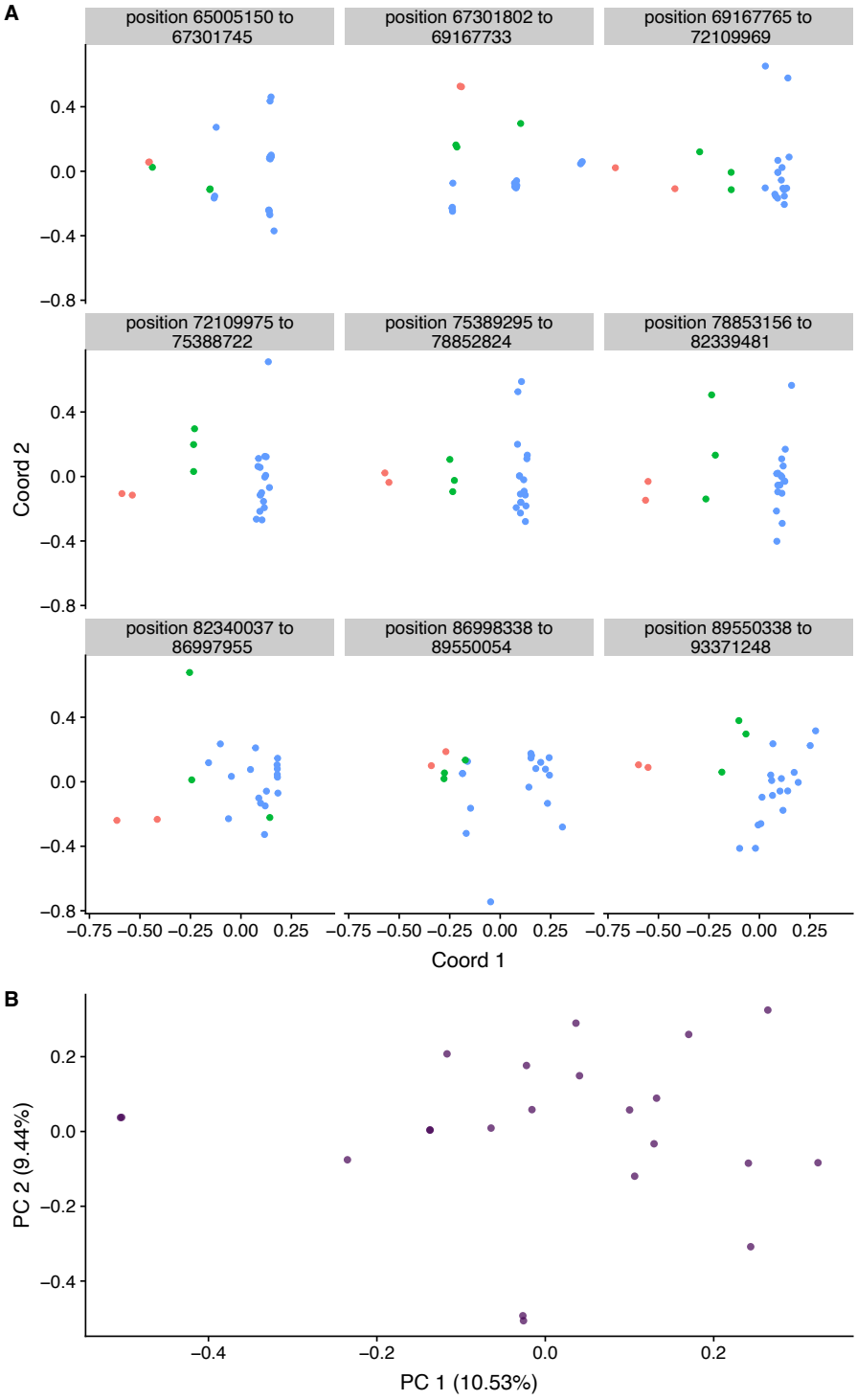

Figure S15 (A) Local PCA (500 SNPs per window) reveals structural variation in multiple consecutive windows (chromosome 3, start positions 72,109,975; 75,389,295; 78,853,156) in putative inversion region of accession BU. Facet labels correspond to window start positions. Each windowed PCA was computed using 500 SNPs of the 600k LR BU dataset. Color codes refer to 3 clusters observed in window 72,109,975 to track clustering patterns in adjacent windows. (B) No structure is observed in principle components computed for genome-wide 600k data of accession BU (LR).

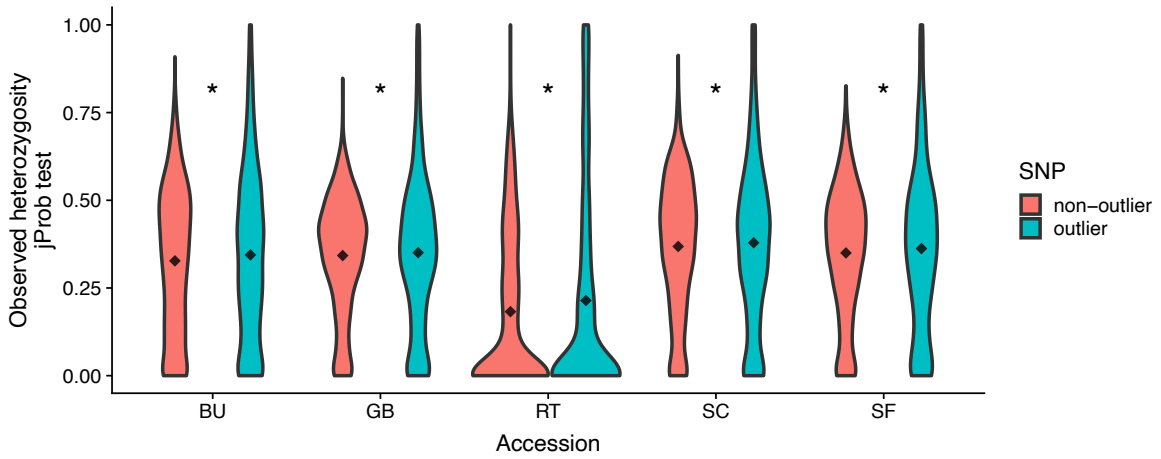

Figure S16 Violin plots for the frequencies of heterozygous genotypes of LD-pruned non-outlier SNPs and outlier SNPs in LR accessions for joint probability outlier. Diamonds indicate group means. Comparisons with asterisks have significantly different means ($p < 0.05$).

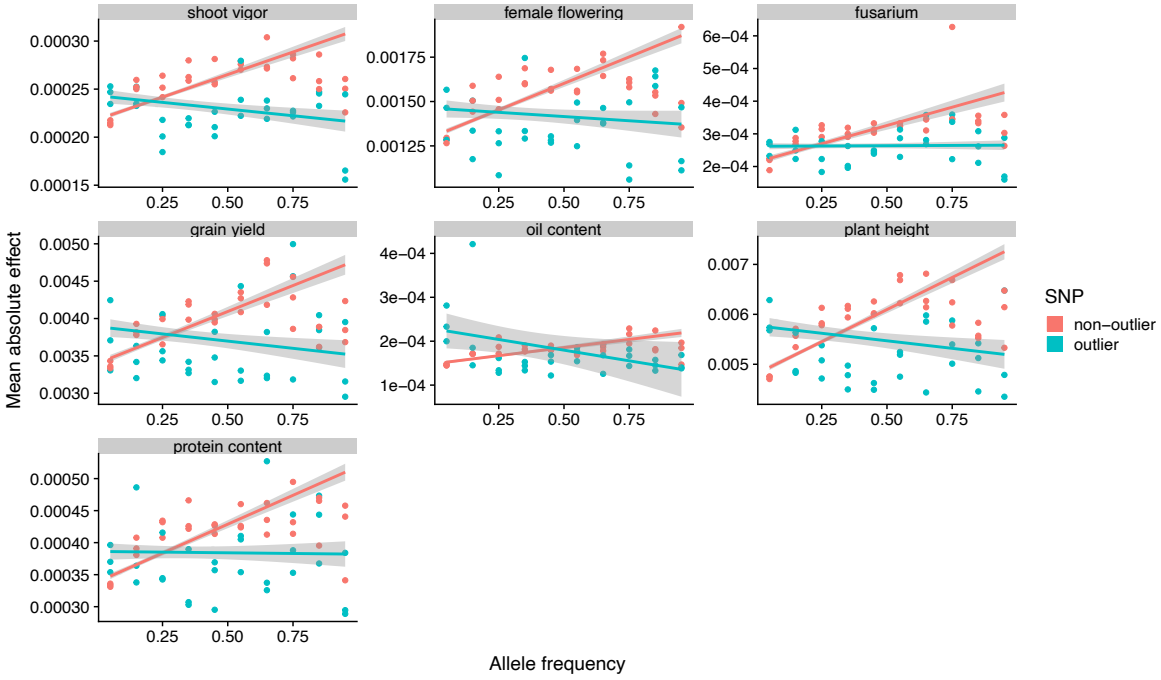

Figure S17 Comparison of mean absolute effect sizes (α) for outliers and non-outliers in different frequency bins in three populations used in GWAS. By fitting a linear regression model ($|\alpha| \sim \overline{p_{bin}}$) for fitness-related traits and outlier status we explore significant interactions in ANOVA (Table S6) between outlier status and frequency bin. Regression slopes with different signs highlight the deleterious action of outlier SNPs. Shaded areas around regression lines are 95% confidence intervals.

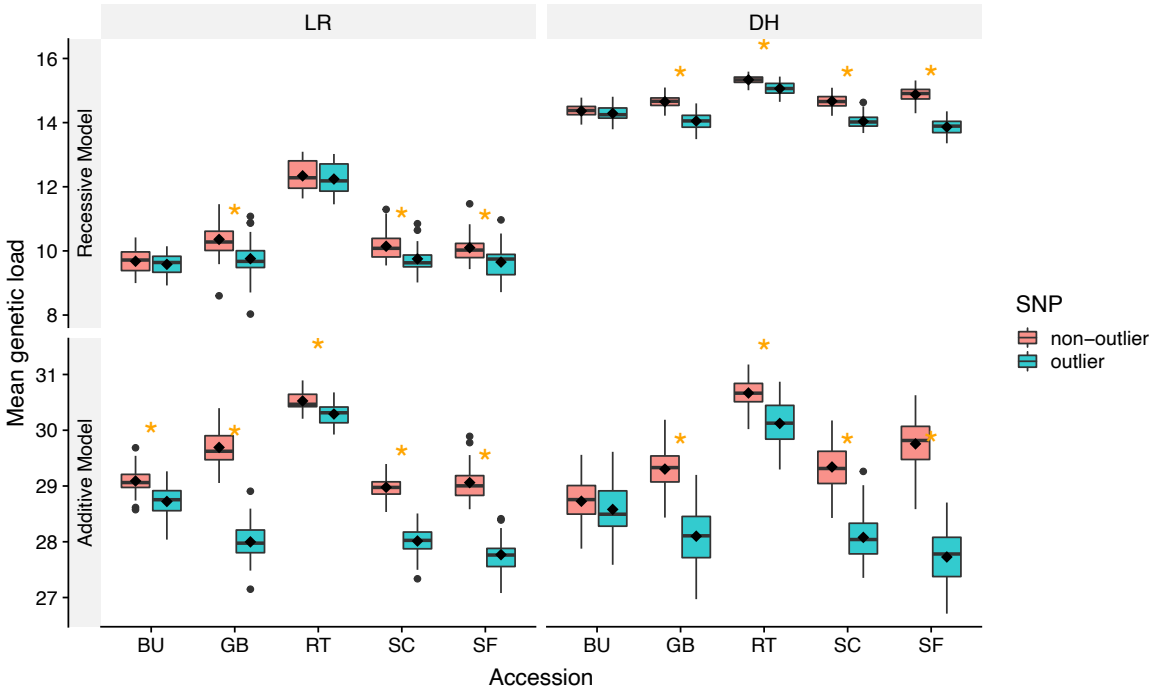

Figure S18 Mean individuals' GERP sum in 1 cM region for SNPs with GERP > 0 per DH-LR pair and SNP-type reveal differences in putative genetic load comprised within accessions and populations for the additive and recessive model. Group means are represented by diamonds. Orange asterisks mean significantly different outlier/non-outlier means within accession and population (t-test, $p < 0.05$).