

The genome-wide dynamics of purging during selfing in maize

Kyria Roessler^{1,8}, Aline Muyle^{1,8}, Concepcion M. Diez², Garren R. J. Gaut³, Alexandros Bousios⁴, Michelle C. Stitzer⁵, Danelle K. Seymour¹, John F. Doebley⁶, Qingpo Liu^{7*} and Brandon S. Gaut^{1*}

Self-fertilization (also known as selfing) is an important reproductive strategy in plants and a widely applied tool for plant genetics and plant breeding. Selfing can lead to inbreeding depression by uncovering recessive deleterious variants, unless these variants are purged by selection. Here we investigated the dynamics of purging in a set of eleven maize lines that were selfed for six generations. We show that heterozygous, putatively deleterious single nucleotide polymorphisms are preferentially lost from the genome during selfing. Deleterious single nucleotide polymorphisms were lost more rapidly in regions of high recombination, presumably because recombination increases the efficacy of selection by uncoupling linked variants. Overall, heterozygosity decreased more slowly than expected, by an estimated 35% to 40% per generation instead of the expected 50%, perhaps reflecting pervasive associative overdominance. Finally, three lines exhibited marked decreases in genome size due to the purging of transposable elements. Genome loss was more likely to occur for lineages that began with larger genomes with more transposable elements and chromosomal knobs. These three lines purged an average of 398 Mb from their genomes, an amount equivalent to three *Arabidopsis thaliana* genomes per lineage, in only a few generations.

Darwin showed that the self-fertilization of plants leads to reduced vigour and fertility—that is, inbreeding depression¹. His work supported the hypothesis that self-fertilization is strongly disadvantageous and provided a rationale for the prevalence of outcrossing in nature². He did not, however, know the genetic basis of inbreeding depression. It is now thought to be caused by increased homozygosity, which increases the genetic load by uncovering recessive deleterious alleles and/or eliminating heterozygosity at loci with an overdominant advantage³. The decrease of heterozygosity (H) is expected to occur at a regular rate; in a selfed lineage, H is expected to be halved each generation. However, the actual rate of H decline is likely to be slowed by various factors, such as interference due to linkage, epistatic interactions⁴ and selective pressure to retain heterozygosity at overdominant and associative-overdominant loci^{5,6}. These factors presumably contribute to the fact that inbred lines of maize and *Caenorhabditis* species retain heterozygosity even after many generations of selfing^{7–10}.

One way to combat the increased load caused by inbreeding is the removal, or ‘purging’, of recessive deleterious alleles. When purging is effective, there may be no inbreeding depression¹¹. Purging is expected to occur rapidly when recessive alleles have lethal effects^{12,13}, but should be less efficient for non-lethal recessive alleles⁶. The existence of purging is supported by experiments, theory and forward simulations^{3,14,15}, but it is expected to vary across species on the basis of features such as population history, mating system and the distribution of fitness effects. Given this variation, one meta-analysis has concluded that purging is an ‘inconsistent force’ in the evolution of inbreeding plant populations⁶.

Recently, researchers have argued that genomic data provide more precise insights into inbreeding effects than previous

approaches^{4,16}. Here we extend this argument to the phenomenon of purging, beginning with three simple predictions. The first is that selfed offspring will exhibit a bias against the retention of putatively deleterious single-nucleotide polymorphism (SNP) variants, because these SNPs become uncovered in a homozygous state. The second is that purging of SNP variants will be inconsistent across genomic regions, on the basis of the amount of recombination. All else being equal, regions of high recombination should purge deleterious variants more efficiently, because recombination reduces interference among selected sites^{17,18}. The third prediction is that purging reduces genome size (GS). We make this prediction because GS correlates strongly with transposable element (TE) content^{19–22} and because plant TE insertions are thought to be predominantly deleterious²³. As a consequence, inbreeding should purge TE insertions by favouring the retention of haplotypes with fewer TEs. This may be especially true for TE insertions near genes, which may be deleterious through their effects on gene expression^{24–26}. Consistent with these predictions, selfing species tend to have smaller genomes than outcrossers in both plants^{27,28} and animals²⁹.

In this study, we use an experimental evolution approach to investigate the dynamics of purging on a genome-wide scale. The experiment mimics an immediate transition to selfing, because it consists of 11 outcrossed maize parental lines that were self-fertilized for six or more generations. Given these selfed lineages, we gathered flow cytometric and whole-genome-resequencing data from a subset of the lines to address three sets of questions. First, does GS decrease rapidly in selfed lineages? If so, are TEs the primary component of loss? Second, are putatively deleterious alleles purged more rapidly than putatively neutral alleles, and if so, does purging vary with recombination rate? Finally, does H decline at expected rates over time?

¹Department of Ecology and Evolutionary Biology, UC Irvine, Irvine, CA, USA. ²Department of Agronomy, University of Cordoba, Cordoba, Spain.

³Department of Cognitive Science, UC Irvine, Irvine, CA, USA. ⁴School of Life Science, University of Sussex, Brighton, UK. ⁵Department of Plant Sciences, UC Davis, Davis, CA, USA. ⁶Department of Genetics, University of Wisconsin, Madison, WI, USA. ⁷The Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, College of Agriculture and Food Sciences, Zhejiang A&F University, Lin'an, Hangzhou, China.

⁸These authors contributed equally: Kyria Roessler, Aline Muyle. *e-mail: liuqp@zafu.edu.cn; bgaut@uci.edu

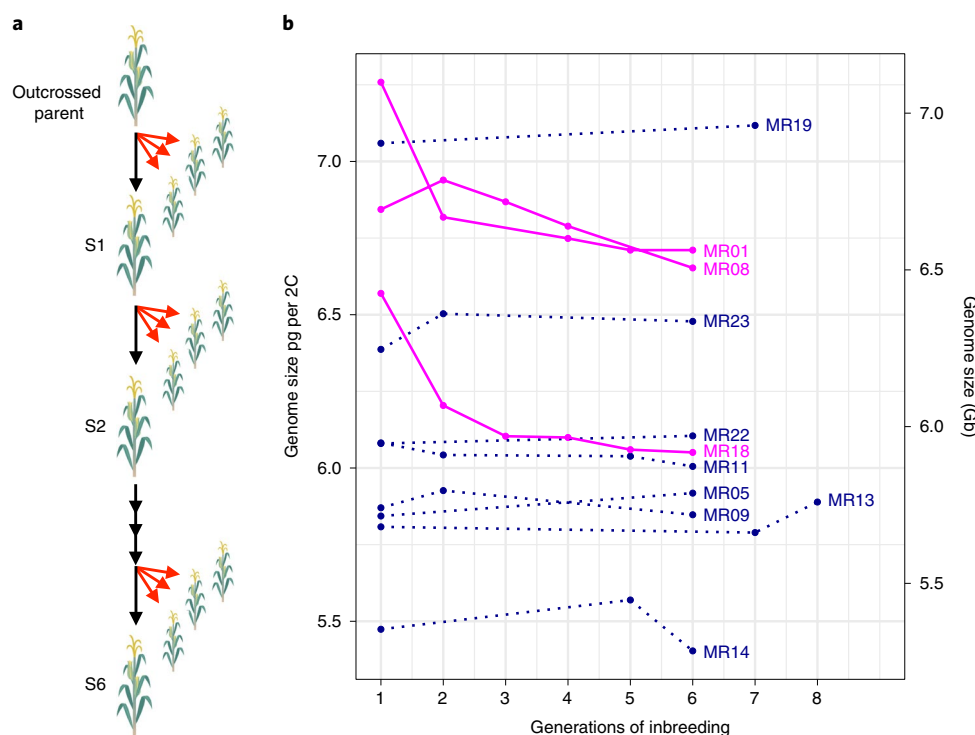


Fig. 1 | Study design and estimates of genome size. a, Schematic of the study design. An outcrossing parent was selfed to make the S1 generation and then subsequently selfed until S6 or higher. The selfed, single-seed descent lineages are represented by black arrows. Our study used sibling seed sampled from each generation, represented by red arrows. **b**, Estimates of genome size, in pg per 2C content, across generations of selfing. Each of the 11 lines is represented. Dark lines represent significant decreases of GS. Dotted lines show no detectable changes in GS over time. Sample size varied between one and three for each line and generation (see Supplementary Table 1 for details) with a total of $n=96$ plants sampled. See Supplementary Table 2 for raw values, Supplementary Tables 3 and 4 for statistics, and Supplementary Fig. 3 for a detailed plot of the raw data per line.

Results

Plants, phenotypes and genome sizes. The plant material came from a previous experiment in which 11 heterozygous maize landraces were self-fertilized to create homozygous lines³⁰. For each landrace, the experiment began with a single, outcrossed parent of unknown genotype, and selfing was continued for at least six generations by single-seed descent. For this study, we germinated seeds from intervening generations—that is, from S1 to at least S6. Each of our seeds was a sibling to the seed that was used to propagate the ensuing generation (Fig. 1a). Following germination, we sowed three plants per line per generation. The plants did not flower under our growth conditions, but we measured growth rate and mortality (proxies for fitness) over a 45d period. Growth rate and mortality varied among the eleven lines (Supplementary Figs. 1 and 2).

To test for GS change, we gathered flow cytometry estimates for 96 plants and five B73 controls. Plant choice was restricted by mortality, but the 96 plants were chosen to represent a time series for each of the 11 lines, with more than one plant per generation where possible (Supplementary Table 1). We included three technical replicates per plant, for a total of 303 assays (Supplementary Table 2). We then investigated our prediction of GS loss in two ways. First, we contrasted GS between the S1 generation and the latest (at least S4) generation with at least two siblings. By this measure, three lines (MR01, MR08 and MR18) exhibited significant decreases in GS (Wilcoxon rank-sum test; $P<0.05$), with no detectable GS shifts for the remaining eight lines (Wilcoxon rank-sum test; $P>0.5$; Supplementary Table 3). Second, we plotted flow cytometry data as a function of time, including data from intermediate generations (Fig. 1b and Supplementary Fig. 3). The results again indicated that MR01, MR08 and MR18 exhibited significant decreases in GS and that the other lines showed no detectable loss (Supplementary Table 4).

For MR01 and MR18, a model of exponential decay fit the data better than a linear model, suggesting that GS loss occurred more rapidly in the early generations (Supplementary Table 4).

We made three further observations on the basis of flow cytometric data. First, GS loss occurred in three of the four lines with the largest S1 genomes (Fig. 1b). These rankings were non-random by permutation test ($P=0.006$), illustrating an increased tendency for lines with larger genomes to lose size. Second, because none of the lines exhibited a significant GS increase, the probability of GS loss was significantly higher than GS gain (two-sided binomial test; $P=0.04$). Finally, we estimated the number of bases lost by each line, assuming a reference value of 5.64 pg per 2C (where 2C is twice the amount of DNA in an unreplicated haploid nucleus) for maize B73³¹ and a conversion rate of 1 pg=978 Mb³². Line MR01, for example, had an average GS estimate of 7.26 pg per 2C in S1 and a corresponding average of 6.75 pg per 2C in S4. The difference between generations was therefore 0.51 pg, which corresponds to a loss of 7.0%, or 499 Mb. Similarly, lines MR08 and MR18 lost 2.8% (or 186 Mb) and 7.9% (or 508 Mb), respectively, between generations 1 and 6.

Genomic components correlate with GS variation across samples. We predicted that purging would lead to GS loss, which was true for 3 of the 11 lines. We also predicted that loss would be dominated by TEs, but TEs are not the only potential genomic component that may contribute to rapid GS reduction. GS loss could also be attributed to: (1) the loss of genes, (2) variation in ribosomal DNA (rDNA) copy number^{33,34}, (3) fluctuations in the number of chromosomal knob and CentC satellite repeats^{22,35} or (4) the loss of supernumerary B chromosomes, which are small³⁶ but can be multicopy³⁷ and vary among accessions³⁸.

Table 1 | Estimates of the variance components based on ANOVA applied to read-count data

	Group	Landrace	Generation	Group $\times$ generation	Line $\times$ generation
TEs	14.72 ***	70.65***	2.82*	5.41*	0.65
Genes	1.46	21.49	4.85	0.017	18.51
Knobs	35.60 ***	56.21***	0.44	0.50	2.54
B-chromosomes	7.02	25.80*	7.27	6.76	25.53*
rDNA	2.00	40.49*	2.27	1.53	13.47

Each of the five genomic components (TEs, genes, knob repeats, B-chromosome specific repeats and rDNA) was tested individually. Sample sizes were $n=2$ for groups, $n=6$ for landraces and $n=2$ for generations. * $P < 0.05$; 0.05 $> P > 0.001$ and *** $P < 0.001$. P -values were FDR-corrected on the basis of all tests in Table 1. Exact P -values are provided in Supplementary Table 5.

To investigate the genomic regions responsible for GS change, we resequenced genomes of 33 plants, including data from S1 and at least S5 for the three lines that exhibited GS loss (MR01, MR08 and MR18; the GSD group) and from three control lines (MR09, MR19 and MR22; the GS_{con} group) (Supplementary Table 1). The data were mapped to the maize B73 AGPv4 genome with four annotated genomic components—genes, rDNA, TEs and knob-specific repeats—and to B-chromosome repeats (see Methods). Total read counts varied among individuals; thus comparison across individuals and generations required normalization. Similar to previous studies^{21,22}, we normalized across libraries on the basis of the ratio of read counts to genes, but in this case we focused on benchmarking universal single-copy orthologues (BUSCO³⁹; Methods). Our reasoning was that BUSCO genes were unlikely to contribute to short-term GS change, because they are conserved across the kingdom Plantae. Simulations demonstrated that this normalization approach leads to accurate inferences of relative read counts in genomic components (such as TEs) that may vary across generations, even with low (2 $\times$) coverage (Supplementary Fig. 4).

Given normalized read-count data, we examined the relationship between GS (as measured by flow cytometry) and sequence counts across the entire sample of 33 plants. Regressing each component separately, there was no significant relationship to GS for genic content (linear regression, $r^2 = -0.027$, $P = 0.63$) or B-chromosome content ($r^2 = -0.015$, $P = 0.45$). There was borderline significance for rDNA ($r^2 = 0.079$, $P = 0.07$), but strongly positive relationships between GS and both knob-repeat content ($r^2 = 0.662$, $P = 4.5 \times 10^{-8}$) and TE content ($r^2 = 0.901$, $P < 10^{-15}$; Supplementary Fig. 5). When all of the components were combined into a single linear model, only TE counts remained significant (linear model, $t = 9.18$, $P = 2.55 \times 10^{-9}$), but knobs were again significant after TE counts were removed from the model (linear model, $t = 5.78$, $P = 5.02 \times 10^{-6}$). Hence, GS correlates most strongly with TE content, but there is a suggestion that knobs also contribute to GS variation.

Genomic components that contribute to temporal loss. TEs and knobs contribute to GS variation, but which among the five components varied over time and contributed to GS change? To address this question, we applied analysis of variance (ANOVA) to read-count data from each of the five genomic components separately. The ANOVA tested for significant differences between groups (GSD versus GS_{con}), among landraces (for example, MR01 to MR22) and between generations (S1 to S6). It also tested for group $\times$ generation and landrace $\times$ generation interactions. We were particularly interested in group $\times$ generation interactions, because they identify components that differentiate the GSD versus GS_{con} groups over time.

We applied ANOVA to each of the six genomic components separately (Table 1 and Supplementary Table 5) and plotted normalized counts for groups (Fig. 2) and landraces (Supplementary Fig. 6). Focusing first on genes, the ANOVA had no significant terms

(F -tests; $P > 0.05$; all P -values corrected for false discovery rate (FDR) for all of the tests in Table 1). The lack of significance was reflected in plots of read counts, because there were only moderate differences between groups and among landraces, with no consistent trend over time (Fig. 2). For rDNA, the ANOVA detected differences among landraces ($F = 5.28$, $P = 0.004$), with 41% of the variance explained but with no other significant terms. By comparing GS estimates to read counts (see Methods), we estimated the average amount of DNA loss in Mb attributable to rDNA repeats in each line and each generation. No line had more than 8 Mb of estimated rDNA, and the temporal difference between S1 and S6 was less than 0.7 Mb for most lines (Supplementary Table 6). A third component was B chromosomes. Only one line (M18) had substantial hits (number of mapping reads) to B-chromosome repeats, representing an average of 10.7 Mb of DNA content across S1 individuals. By S6, counts were at background levels, indicating B-chromosome loss. Given these patterns, the ANOVA detected significant landrace ($F = 5.90$, $P = 0.021$) and landrace $\times$ generation terms ($F = 4.85$, $P < 0.022$), but no group effects.

We next turned to the two genomic components that correlated strongly with GS across the entire dataset: TE counts and knob repeats. TE counts exhibited significant terms across groups ($F = 53.94$, $P = 2.38 \times 10^{-7}$; 14% variance explained), landraces ($F = 64.71$, $P = 2.91 \times 10^{-11}$; 70% variance explained), generations ($F = 10.35$, $P = 0.018$; 2.8% variance explained) and group $\times$ generation interactions ($F = 19.84$, $P = 0.0013$; 5.4% variance explained) (Table 1). The plots of TE counts were consistent with these statistical results, because they show that: (1) the GSD group had higher overall TE counts than the GS_{con} group; (2) landraces within GSD exhibited reductions in TE counts from generation S1 to S6, but (3) landraces within GS_{con} did not (Fig. 2). By equating GS to read counts, we estimated that the Mb loss due to TEs was 481 Mb for MR01, 199 Mb for MR08 and 465 Mb for MR18, representing more than 90% of the estimated shift in GS over time for each line. By contrast, the GS_{con} lines exhibited temporal TE changes of about 10 Mb each (Supplementary Table 6).

Finally, knob counts differed between groups ($F = 158.99$, $P = 2.91 \times 10^{-10}$; 56% variance explained) and among landraces ($F = 62.75$, $P = 2.91 \times 10^{-10}$; 35% variance explained), with the GSD group having generally higher counts. However, knob counts did not exhibit significant interaction terms or variation between generations, which was surprising given the correlation between knob counts and GS across all samples (Supplementary Fig. 5). We therefore investigated the possibility that the lack of significance reflected reference bias by repeating analyses with the W22 reference⁴⁰. The results largely corroborated the B73 results but did produce a significant group $\times$ generation interaction for knobs ($F = 10.88$, $P = 0.0128$) (Supplementary Table 7). Based on the W22 reference, the average Mb loss over generations due to knobs was 136 Mb in MR01, 59.4 Mb in MR08 and 77.0 in MR189, but TEs explained more temporal variation in every case (341 Mb, 130 Mb and 413 Mb, respectively; Supplementary Table 8).

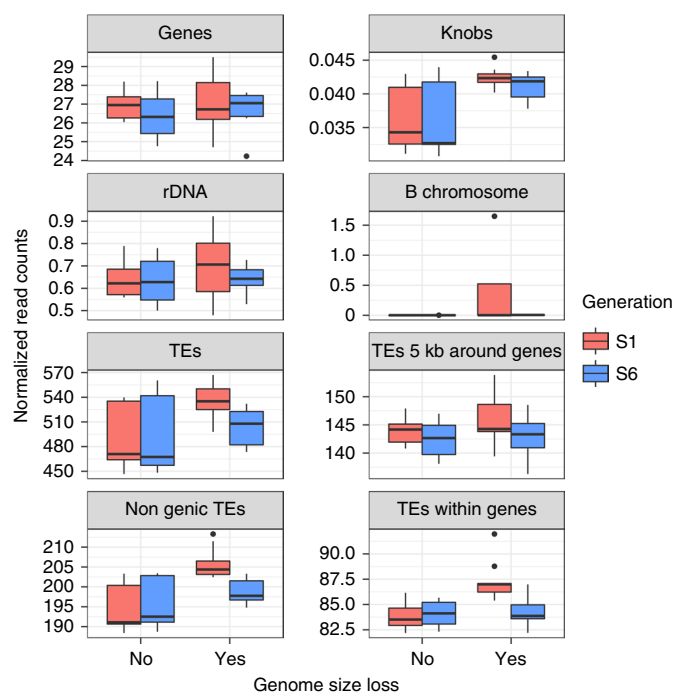


Fig. 2 | Various components of the genome compared between GSD and GS_{con} groups and between S1 and S6. Sample sizes are shown in Supplementary Table 1, *P*-values are provided in Supplementary Table 5 and Supplementary Fig. 6 reports this information for each of the lines separately. The box shows the median, lower and upper quartiles. The whiskers extend to the largest or lowest value no further than 1.5 × interquartile range from the hinge. Outliers are plotted as dots above the whiskers.

TE locations, types and mechanism. We predicted that GS loss could reflect purging of TEs near genes due to their deleterious effects on gene expression^{24,25}. To address this prediction, we separated TEs from B73 into three bins: non-genic TEs, which mapped to TEs more than 5 kb away from genes; near-genic TEs that were within 5 kb of a gene; and the subset of near-genic genes that overlapped with annotated genes—that is, they fell within introns or UTRs. Both non-genic and overlapping TEs exhibited significant group × generation interactions ($F = 18.46$ and 13.97 , $P \leq 0.001$; all *P*-values FDR-corrected; Supplementary Table 9), explaining 9.1% and 12.4% of the total variance for non-genic and overlapping TEs, respectively. Despite our prediction, none of the ANOVA terms were significant for TEs near (within 5 kb) genes, but three components were borderline significant ($F = 3.34$, $P < 0.10$; Supplementary Table 10), including both the generation and landrace × generation terms. Of note, the latter reflects the fact that five of the six lines lost near-genic TEs through the course of the experiment (Supplementary Fig. 6), suggesting that the loss of TEs near genes was a general phenomenon across all lines. We repeated these analyses for the W22 reference, finding that all three TE locations exhibited group × generation effects (Supplementary Table 10 and Supplementary Fig. 7). Overall, then, these data suggest that TEs were lost throughout the genome, but it is unclear whether near-genic TEs were lost across all lines or only from the GSD group.

We also investigated potential biases by TE order, focusing on six TE types in the B73 reference: helitrons, long terminal repeats (LTR) retrotransposons, solo LTRs, terminal inverted repeats, short interspersed nuclear elements and long interspersed nuclear elements. All but solo LTRs exhibited significant variation between the GSD and GS_{con} groups ($F > 39.10$, $P < 1.2 \times 10^{-5}$). Four of the six also exhibited a significant group × generation interaction, which

explained > 5% of the variance for LTRs, solo LTRs and helitrons (Supplementary Fig. 8 and Supplementary Table 11). Thus, GS loss encompassed an array of TE types.

Finally, we addressed a question related to a potential mechanism of TE loss. In some plant species, TE loss is driven by unequal recombination between LTR elements⁴¹. These recombination events are expected to increase the ratio of solo LTR elements to intact LTR elements. If this mechanism operated during our experiment, the ratio of reads mapping to LTRs versus the internal regions of elements should increase over time, especially in the GSD lines. To test this idea, we independently annotated 22,530 full-length LTR elements of the *Sirevirus* genus, based on the B73 reference. We focused on *Sirevirus* for three reasons: (1) they represent a substantial proportion (around 20%) of the maize genome⁴², (2) they can be accurately annotated on the basis of numerous internal features, including the boundary between LTRs and internal regions⁴³, and (3) they provide a set of LTR elements that were annotated independently of the existing B73 v4 genome annotation. We found that both solo and intact *Sirevirus* exhibited losses over time in the GSD group (Supplementary Table 12 and Supplementary Fig. 9), consistent with our LTR analyses based on the v4 annotations. However, the ratio of mapping to LTRs versus internal regions did not exhibit an obviously increasing trend through time or a significant group × generation effect ($F = 0.27$, $P = 0.73$), as would be predicted if TE loss were driven by numerous unequal recombination events.

The fate of deleterious variants. We now turn to a second prediction about purging: over time, there should be a bias against the retention of deleterious SNP variants. We tested this prediction by first calling SNPs for each of the six lines from the GSD and GS_{con} groups and then by focusing only on biallelic SNPs that were inferred to be heterozygous ($H = 1$) in the resynthesized parent (see Methods). For each of these heterozygous sites, we predicted derived deleterious variants using SIFT⁴⁴ and noted the fate of variants in four functional classes (non-coding, synonymous, tolerated nonsynonymous and putatively deleterious nonsynonymous variants). In total, we examined 1,914,845 SNPs across the six lines (Supplementary Table 13).

As a signal of purging, we expected deleterious, derived SNP variants to exhibit biased rates of loss over time. To characterize this potential bias, we identified derived alleles by comparison to a *Sorghum* outgroup and estimated the proportion of derived allele (P_d) across sites. We expected P_d to be 50% in the parent and to remain 50% in the absence of perturbing factors like selection. To test this prediction, we combined results across the six lines and plotted P_d for each functional class in S1 and S6 (Fig. 3a). In S1, for example, average P_d estimates for non-coding and synonymous sites were below 0.5, potentially reflecting biases in ancestral inference and/or selection against a subset of these putatively neutral derived alleles between parents and S1. Consistent with the latter interpretation, P_d declined slightly from S1 to S6 for both site classes (linear model contrast $Z = 14.92$, $P < 0.001$).

These effects were greatly amplified for nonsynonymous mutations (Fig. 3a). For example, putatively deleterious, derived nonsynonymous SNPs had a P_d of 0.384 in S1, representing a significant decrease relative to that of synonymous and non-coding variants (linear model contrast $Z = 44.89$, $P < 0.001$; Supplementary Table 14). Between S1 and S6, P_d fell even further, from an average of 0.384 to 0.334 (linear model contrast $Z = 20.83$, $P < 0.001$). Overall, putatively deleterious SNPs demonstrated accelerated rates of loss over time relative to other variant classes.

Recombination is expected to mediate the effects of selection, because it uncouples interference between linked variants. Therefore, deleterious variants should be purged more rapidly in regions of high recombination. To explore this prediction, we contrasted genomic regions that encompass the highest and lowest

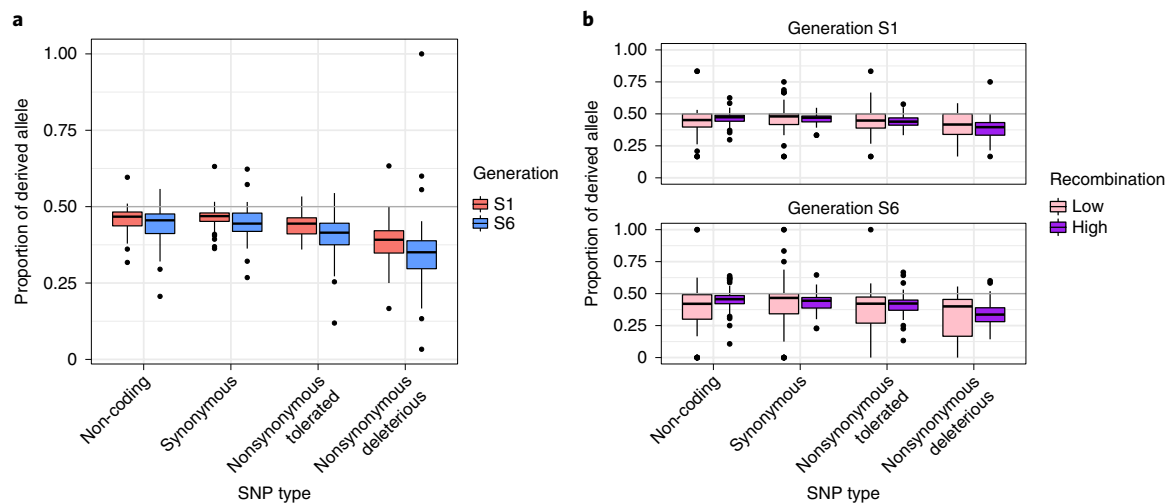


Fig. 3 | Proportion of derived allele across SNP types, generations and recombination categories. **a**, The proportion of the derived allele for the four mutational classes predicted by SIFT—that is, non-coding, synonymous, nonsynonymous tolerated and nonsynonymous deleterious. The graph reports the proportion for generations S1 and S6 across six lines (MR01, MR08, MR09, MR18, MR19 and MR22). P_d was averaged across individuals for each of the ten chromosomes and six lines separately ($n = 60$ biologically independent chromosomes for each bar of the plot, $n = 480$ total samples across generations and categories). **b**, As in **a**, except the genome was separated into high- and low-recombination quartiles, illustrating that purging occurs more rapidly in high-recombination regions. As in **a**, $n = 60$ for each bar of the plot. Box plot as described in Fig. 2.

quartiles of recombination rates, as defined by crossover events (r) (ref. 7). The results showed the expected pattern: in S6, P_d was lower in high- compared with low-recombining regions for both classes of nonsynonymous variants (tolerated: $Z = -3.37$, $P = 0.006$; deleterious: $Z = -4.95$, $P = 5.98 \times 10^{-6}$ based on linear model contrasts; Supplementary Table 15). Recombination did not have an effect on P_d for nonsynonymous SNPs in S1, consistent with the fact that time is required for recombination to break down linkage between loci.

Declining heterozygosity. Finally, we measured a phenomenon of empirical interest, which is the rate of loss of heterozygosity over generations. This is a difficult task, given our low coverage data, but we took advantage of the fact that SNPs inferred to be heterozygous in the parental generation can be in only one of two states within S1 and S6: heterozygous or homozygous. Moreover, these two states are expected to fall into blocks, with the transition between blocks defined by recombination events. To identify these blocks, we examined windows of 100 SNPs in size, focusing on genic SNPs, and used a Bayesian clustering method to assign windows as either heterozygous or homozygous for each individual (see Methods). The proportion of heterozygous blocks across the genome (H_b) can be compared directly to the null expectation that $H = 0.50$ in S1 and 0.015 in S6.

We applied this approach successfully to the two lines with highest coverage (MR09 and MR22) (Fig. 4) and offer five observations about heterozygosity. First, H_b exceeded 60% in both MR09 (65.7%) and MR22 (63.7%) for generation S1, representing a significant deviation from the null expectation (one-sided Wilcoxon test, $P = 0.0019$ and $P = 0.019$, respectively). Second, H_b significantly exceeded the expected value of 1.5% in S6, at 14.2% for MR22 and 4.8% for MR09 (one-sided Wilcoxon test, $P = 0.00098$ and $P = 0.019$, respectively). Third, for reasons that are not immediately apparent, the difference between the two lines in S6 was also significant (one-sided Wilcoxon test, $P = 0.00036$). Fourth, heterozygous blocks had a significantly higher proportion of nonsynonymous SNPs (7.19%) compared to homozygous blocks (6.14%, one-sided $\chi^2 = 27.72$, $P = 1.4 \times 10^{-7}$); the same was true for putatively deleterious SNPs (one-sided $\chi^2 = 4.2969$, $df = 1$, $P = 0.038$). Finally, heterozygosity was also related to recombination, because heterozygosity and r were

modestly but significantly correlated across windows in S6 (linear regression adjusted $r^2 = 0.016$; $P = 1.5 \times 10^{-4}$).

Discussion

In this study, we took an experimental approach to assess the genomic effects of selfing, with a focus on the dynamics of purging. Previous studies have investigated the effects of selfing by, for example, contrasting selfing and outcrossing plants in flowering phenology, population structure, genomic diversity²³ and evolutionary fate⁴⁵. Yet, most of these effects probably accrue after, not during, the transition to selfing. A smaller number of studies have found evidence of purging by comparing inbreeding depression between naturally inbreeding and naturally outcrossing species^{3,11,46}. By contrast, the immediate genomic effects of purging have gone largely undocumented.

Rapid genome flux. Our experiment has documented rapid GS loss in 3 of 11 selfed lineages (Fig. 1). These observations add to a growing consensus that GS can change rapidly in plant species. Other examples include GS changes in flax over a single generation³⁴, GS shifts on experimental time-scales in *Festuca*⁴⁷ and GS reductions in maize after six generations of selection for early flowering⁴⁸. To our knowledge, however, the magnitude of GS loss that we observed in this study is unprecedented. On the basis of estimates using flow cytometry, the three lines lost about 6% of their genome or 398 Mb, on average, from S1 to S6. To put these changes in context, the GS of two fully sequenced maize inbred lines (Mo17 and B73) differ by only ~25 Mb⁴⁹.

Following precedence^{20–22,50,51}, we used read counts to infer the size of genomic components, focusing on genes, TEs, knob repeats, rDNA and B chromosomes. Among these five components, it is clear that TEs are the major source of loss, which is not surprising given that DNA derived from TEs constitutes more than 85% of the maize genome⁵², and that previous studies have shown that TEs contribute to plant GS variation^{20–22,50,51}. GS shifts are not always caused by TE content, however. In flax and *A. thaliana*, GS shifts are fuelled primarily by variation in rDNA repeats^{33,34}, and GS differences between selfing and outcrossing *Caenorhabditis* species are roughly equally apportioned among genes and TEs^{29,53}.

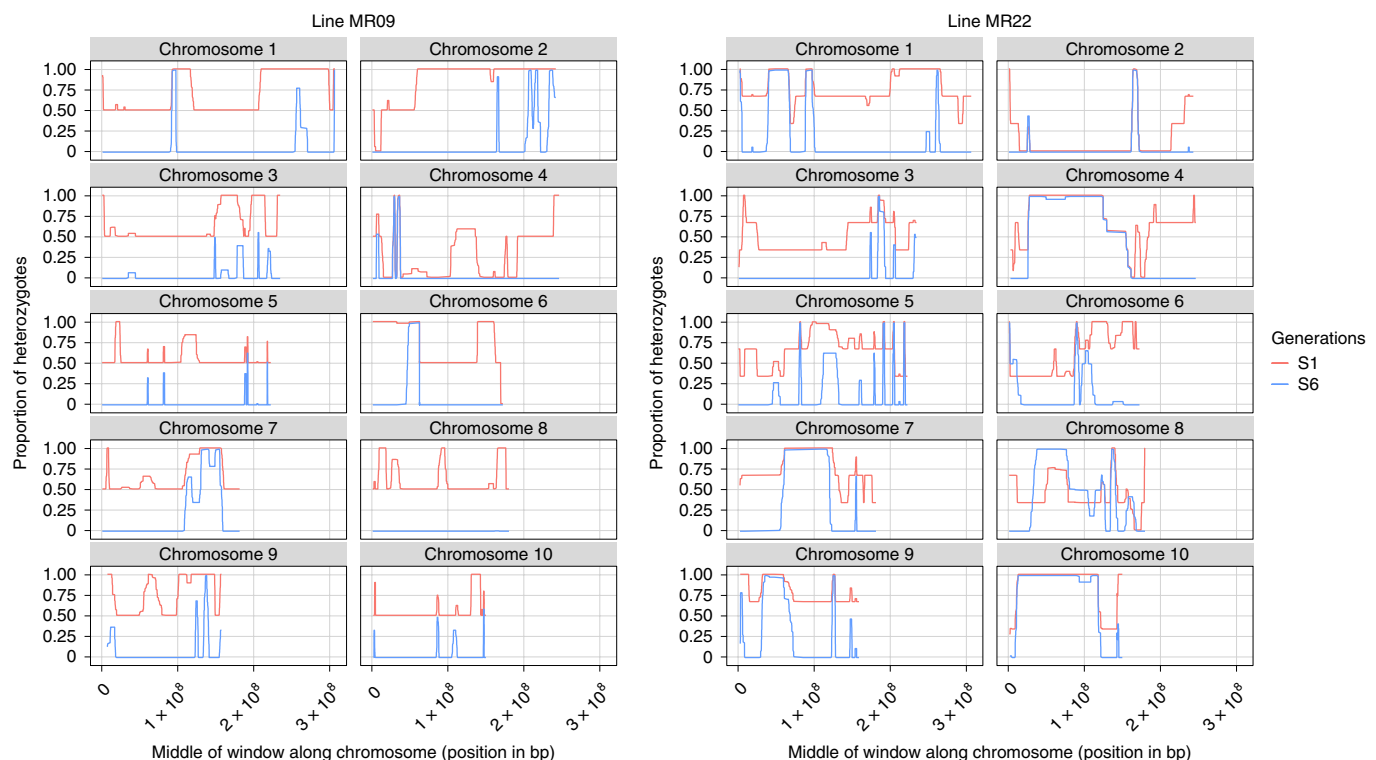


Fig. 4 | Inference of heterozygous and homozygous genomic regions, based on SNPs inferred to be heterozygous in the parent. Each of the ten chromosomes for two lines (MR22 and MR19). Heterozygosity was averaged across individuals for each line and generation separately. For each chromosome, the x axis represents length along the chromosome and the y axis is the proportion of heterozygous sites within 100-SNP sliding windows. Red and blue lines represent the S1 and S6 generations. Both lines have more regions of heterozygosity than expected (see text for statistics). Sample sizes are shown in Supplementary Table 1 ($n = 2$ or 3 plants depending on the line and generation).

Given that TEs are the major source of GS loss, we examined loss according to both TE type and location. Within the GSD group, loss occurred for all six TE orders we tested (Supplementary Fig. 8 and Supplementary Table 6). This finding mirrors previous studies that have compared TE content among *Zea* genomes^{20,21}. For example, Tenaillon et al.²⁰ estimated that 70% of the GS difference between two species (*Zea luxurians* and maize) was due to TE losses and gains, but the relative abundance of TE families was conserved between species²⁰. We predicted that GS loss should be especially evident for TEs that are near genes, because they may have deleterious effects on gene expression^{24,25}. The results varied depending on the reference. With B73, the landrace $\times$ generation effect for near-genic TEs was borderline significant ($P = 0.058$) because five of the six resequenced lines lost these TEs over time, irrespective of their inclusion in the GSD or GS_{con} group (Supplementary Fig. 5). This result implies that the loss of near-genic TEs may be a general property of selfing. However, the W22 results do not fully support this claim, because they suggest that the pattern of loss in near-genic TEs varied between groups. Given these results, we cannot yet conclude that the loss of near-genic TEs is a general outcome of selfing. As the resolution of genome assemblies improves, we advocate further investigation of this question while also recognizing that TE families vary in both their tendency to insert near genes and their epigenetic profiles.

In this context, it is important to emphasize the limitations of the read-count approach for estimating genomic components. The approach is better suited for broad-scale inferences about genome content than for inferences about the fate of specific genes, TE insertions or chromosomal regions. Here our inferences about location are based on the reference genome and may not accurately reflect the genome of our sample. We investigated reference biases

by applying our read-count approach to two references (B73 and W22). With either reference, there was little evidence that genes, rDNA and B chromosomes contributed substantively to GS loss, but the magnitude of the TE and knob effects did vary by reference sequence. With B73, TEs explained more than 90% of loss from S1 to S6 and as much as 481 Mb. With W22, the estimated TE loss was more modest, explaining around 75% of GS shift on average, with the remainder of loss assigned to knob repeats. The difference in results probably reflects annotation and assembly differences between references, because we disregarded counts from regions where annotated features overlapped. In B73, TEs often overlapped with putative knob regions, but overlaps occurred less frequently in W22. Our results therefore contain a cautionary tale about annotation biases, but we also suspect that the implication of knobs as a component of GS loss is reasonable, given our own (Supplementary Fig. 5) and previous evidence that knobs contribute to maize GS variation^{22,50}. Notably, the total Mb loss explained by TEs and knobs was consistent, regardless of the reference used.

Altogether, our results support the hypothesis that GS loss is a potential outcome of selfing. This hypothesis is based on the observations that genomes are smaller in selfers compared to their outcrossing sister taxa in *Caenorhabditis*^{29,53} and across plant taxa²⁸, where it is likely that other factors, such as the reduced spread of transposable elements, also contribute to differences. Assuming that GS loss is common during selfing, one must ask why 8 of our 11 lines exhibited no detectable loss. The lack of loss is probably not a question of statistical power, because five lines were estimated to have slightly larger GS, on average, in S6 relative to S1 (Fig. 1). Here our lack of the parental genome could be misleading, because our experimental design could not monitor loss from the parent to S1. The greatest loss is expected to occur within this first generation,

given that two of three G Δ lines lost GS exponentially over time. We can nevertheless provide some predictive insights by contrasting data between G Δ and G S_{con} groups. Neither group exhibited particularly low growth rates or high mortality (Supplementary Figs. 1 and 2), so GS loss did not obviously relate to these fitness proxies. However, the three lines with GS loss did have larger S1 genomes (Fig. 1b), with significantly more TEs and knobs than the G S_{con} group (Fig. 2 and Table 1). Hence, to a first approximation, genomes with high TE and knob content are more prone to loss.

Heterozygosity, recombination and the fate of deleterious SNPs.

Several previous studies have shown that H declines at lower rates than expected under selfing^{4,9}. In S1 eucalyptus trees, for example, average H was 65.5%, compared with the expectation of 50% (ref. ⁴). We also found elevated heterozygosity in our lines. In S1, H_b was ~65% for MR09 and MR22 (Fig. 4). By S6, both lines retained significantly more heterozygosity than the expected value of 1.5%. Observed values of H_b in S6 imply that, assuming constancy across generations, the rate of heterozygosity retention was 0.60 per generation ($=e^{(\log(0.048)/6)}$) for MR09 and 0.72 per generation ($=e^{(\log(0.142)/6)}$) for MR22.

What can account for this retention of heterozygosity? One explanation is genotyping error. Such errors are not only possible but are likely to be prevalent at individual sites with our low coverage data. For this reason, we focused on a window-based method that assigned blocks of 100 SNP sites into one of two states—heterozygous or homozygous. This approach should mitigate the effect of miscalls at individual sites, and we also employed the method using conservative assumptions; for example, blocks with uncertain assignments were not counted as heterozygous (see Methods). Nevertheless, there is a region on chromosome 8 of MR22 that has higher heterozygosity in S6 than S1 (Fig. 4); such a pattern could be real, given our sampling strategy (Fig. 1a), or may hint to some underlying error in assignments. Towards that end, we also examined obvious potential sources of error by, for example, testing for correlations between the location of heterozygous windows in MR09 and MR22. No correlation was found ($r^2=0.03846$, linear regression, $P=0.5871$), suggesting that underlying genomic features (for example, sets of paralogs that can cause SNP miscalls¹⁰) did not consistently inflate heterozygosity across lines. Altogether, we believe our heterozygosity estimates to be reasonable and probably conservative; together with previous work^{4,9}, they suggest that heterozygosity generally declines more slowly than expected. Nevertheless, more studies are needed to characterize this important dynamic, perhaps by incorporating more intervening generations.

Biological explanations for slower-than-expected rates of heterozygosity decline usually invoke either overdominance or associative overdominance, with the latter thought to be the prevailing force maintaining heterozygosity in selfed lineages^{3,4,9,54}. If higher-than-expected levels of heterozygosity are caused in part by linkage to deleterious variants, then heterozygosity should be higher in regions of low recombination, where selection against deleterious variants is inefficient because loci are coupled. Consistent with this prediction, heterozygosity is elevated in regions of low recombination in the maize nested association mapping population^{8,55}. Here, over the short-term timescale of our experiment, we find that heterozygosity is lower in regions of low recombination, probably reflecting linked selection⁵⁶ against strongly deleterious variants.

Another feature of recombination is that it has the capacity to uncouple linked variants, making selection more efficacious. Putatively deleterious variants are purged from our lines more rapidly than presumably neutral variants (Fig. 3a), and they are purged more rapidly from high versus low recombination regions in S6 (Fig. 3b). We infer that recombination separates deleterious variants from linked variation, permitting the independent loss of the deleterious variant and allowing neutral diversity to remain⁵⁷. A

similar relationship between heterozygosity and recombination was discovered recently within hybrid genomes of swordtail fish⁵⁸. In these hybrids, high recombination regions retained heterozygosity because recombination breaks up incompatibilities that otherwise contribute to hybrid load.

Outstanding questions. At least three questions remain. First, what is the mechanism of TE (and knob) removal? One potential explanation is ectopic and/or unequal recombination, which can leave a signature of an increased ratio of solo to intact LTR elements⁴¹, but we found no evidence for this effect. It is possible, of course, that unequal recombination caused a small number of large deletion events, with only minor effects on the ratio of solo:intact elements. We nonetheless favour a non-exclusive mechanism for GS loss in this experiment, which is that selection tends to act against the larger haplotype when there is a size difference in a heterozygote. Under this scenario, selfed plants with the best collection of small(er) haplotypes are favoured by the selfing process, leading to GS reductions. If true, we expect the resolution of selfing to be a contest between haplotypes, with recombination occasionally reducing interference and combining linked structural variants from different haplotypes onto a single chromosome. Under this model, we can make two predictions: (1) parental plants of higher heterozygosity and larger differences in size between haplotypes are more likely to lose GS, and (2) regions of higher recombination will tend to lose more Mb, due to more efficient selection against large(r) haplotypes. These predictions remain to be tested, underscoring how little we know about selfing, purging and its effects on genomic variants.

Second, what is the proximal cause of GS loss? Our results suggest that the primary effect of selfing is to uncover deleterious recessive mutations, leading to selection against homozygous recessives. But is there a phenotype that drives this selection? GS is known to correlate with several traits, including reproductive rates, growth rates, flowering time, cell sizes and other factors^{22,48,59–62}. Selection on one or several of these diverse characteristics may have occurred during the formation of the inbred lines. However, we cannot find any pattern among our lines that suggest selection was more pronounced on the G Δ versus G S_{con} groups. For example, each of the members of G Δ group (MR01, MR08 and MR18) originated from landraces in the tropical lowlands and were bred in lowland tropical nurseries, but the same is true of MR05, MR09, MR11, M22 and MR23, none of which exhibited obvious GS loss.

Finally, what bearing do these results have on broader questions about plant evolution? First, they inform on processes of genome evolution and show that selection can have several effects even over the very-short term. This includes purging deleterious alleles in higher recombination regions more efficaciously and removing linked variation in regions of low recombination. The data also hint that interference between deleterious variants contributes to the retention of heterozygosity, because regions of high heterozygosity are enriched for deleterious variants in S6. Second, this work relates to the finding that indirect selection for recombination modifiers is favoured under selfing^{63,64}. Our results demonstrate that high recombination rates are advantageous for quickly purging genetic load, which in theory could drive the observed trend toward higher chiasmata frequencies in selfing plants compared to outcrossers.

Methods

Plant materials and phenotypic analyses. Our experiment was based on 11 maize landraces (Supplementary Table 1) that were inbred by J. Doebley (University of Wisconsin) and maintained through single-seed descent for several generations³⁰. The parents represented outcrossed landraces of unknown genotype. For each line and generation, one seed was grown and selfed, and the remaining sibling seeds were stored. We grew the sibling seeds in the UC Irvine greenhouses after germination on Petri dishes. Ten seeds per cultivar were sown in individual pots on 22 July 2014 and grown in a growth chamber under controlled conditions of 12 h

light at 26 °C, 12 h dark at 20 °C, a relative humidity of 70% and 500–600 cal cm⁻² of radiation per d. The third and fourth leaves of each plant were harvested when 12–13 cm long and then frozen in liquid nitrogen and stored at –80 °C. The 11 cultivars, with a subset of 6 plants per cultivar per generation, were grown in 4 completely randomized blocks, with B73 as the control across blocks. Measures for height were taken 9, 17, 30 and 45 d after sowing; mortality was also noted throughout the duration. Mortality and growth rates were compared among lines. We estimated the exponential growth rate for each individual and used a one-way ANOVA to test whether the estimated growth rates differed between lines. A logistic regression model was applied to mortality, and a likelihood-ratio test was used to compare mortality between lines. We did not measure fitness via fecundity, because none of the lines produced seed under our experimental conditions.

Flow cytometric data and analyses. To estimate GS, leaf samples were sent to Plant Cytometry Services. Following a previous study³¹, flow cytometry used 4',6-diamidino-2-phenylindole (DAPI) staining. *Ilex crenata* Fastigiate (2C = 2.2 pg) and maize B73 (2C = 5.64 pg)³¹ were used as internal standards. Three technical replicates were performed for each plant (Supplementary Table 2). To assess whether GS changed as a consequence of selfing, we performed linear regressions, exponential decay analyses and Wilcoxon rank-sum tests in R, combining biological and technical replicates for rank-sum tests. Flow cytometric data were converted to picograms assuming that the maize B73 reference had a value of 5.64 pg per 2C (ref. ³¹); picograms were translated to Mb assuming 1 pg = 978 Mb³². To infer a significant trend toward genome loss, we estimated that the probability of loss was 3 lines out of 11 trials ($P = 0.273$) and calculated the probability of observing zero GS increases over 11 trials with a two-sided binomial.

Whole-genome sequencing and genomic composition. We selected 6 landraces and 33 individuals for whole-genome sequencing (Supplementary Table 1), focusing on the S1 and S6 generations. DNA was extracted from frozen leaf tissue using the QIAGEN DNeasy Plant Mini kit. DNA was multiplexed into libraries with Illumina TruSeq PCR Free kit. The libraries were sequenced on the HiSeq2500 (100 bp read length, paired-end, two lanes) in the UCI High-Throughput Genomics Facility in 2015 (landraces MR01, MR08, MR18 and MR19) and on the HiSeq3000 (150 bp read length, paired-end, one lane) in the UC Davis DNA Technologies Core in 2016 (landraces MR09 and MR22). Individuals were sequenced to an average coverage of ~2.5× per individual (Supplementary Table 16). Note, however, that we had > 6× coverage for each generation for each of the lines investigated given the inclusion of siblings.

Sequencing reads were processed by Trimmomatic (v.0.35) to remove barcodes and low quality reads (< 20), with a minimum read length of 36. Processed reads were mapped simultaneously onto maize genome AGP version 4.37 (AGPv4)⁶⁵ and B-specific chromosomal repeats using BWA-MEM (v.0.7.12)⁶⁶. To prevent double counts of a feature, only one of the paired reads was mapped and only the primary alignment was kept for each multi-mapping read, based on Samtools v.1.3⁶⁷.

We counted mapped reads for five annotated genomic components: genes, B-chromosome specific repeats, chromosomal knobs, rDNA and TEs. The annotation features for protein coding genes and for TEs were obtained from the Gramene database on 1 May 2017 for B73 AGPv4 (Supplementary Table 17). To annotate regions containing knob (plus CentC) regions and rDNA (plus transfer DNA) sequences, a series of fasta files (Supplementary Table 17) representing both features were mapped to the v4 genome using blat (v.36). The regions of B73 that mapped to either knobs or rDNA were then added to gff files (blatgff v.3) for read-count analyses. To count reads, all features were merged (bedtools merge v.2.25.0) to avoid double counting⁶⁸. Bedtools coverage was used to count reads that overlapped at least 90% with each feature. An identical approach was used for W22 annotations (Supplementary Table 17).

We used BUSCO genes to normalize between libraries, on the expectation that these highly conserved genes represent an invariant component of the genome. To identify a conserved set of BUSCO genes, we ran BUSCO (v.3)³⁹ on AGPv4. From the resulting set of 1,309 BUSCO genes, we eliminated any that appeared to be multicopy or that overlapped with TE annotations in B73 AGPv4, leaving a final set of 761 genes. A similar procedure in W22 yielded 918 BUSCO genes. In both references, any gene, knob or rDNA annotation that overlapped with a TE was not considered further. Within any sequencing run, normalized counts for a genomic feature were calculated as the observed number of sequence counts to that feature divided by the total number of counts that mapped to BUSCO genes. To verify that our use of BUSCO genes was accurate, we simulated datasets with BUSCO normalizations based on Chromosome 10 (see below).

Further analyses considered different families and types of TEs. These analyses were performed only in B73. For these, we first identified TEs from the AGPv4 gff file and employed their TE family designations for additional analyses. To examine the ratio of solo LTRs to complete LTRs, we de novo annotated *Sirevirus* sequences on the basis of the MASiVE algorithm⁴³. The application of MASiVE produced 22,530 full-length elements with defined boundaries between LTRs and internal regions.

To assess relationships between GS and genomic components, we used both linear regression and ANOVA, using the lm and aov modules in R (v.3.34).

ANOVA P -values were FDR-corrected. To estimate the Mb of the genome explained by various component, we: (1) translated the GS of each plant from pg per 2C to Mb using the conversion rate of 1 pg = 978 Mb³²; (2) equated Mb for each individual to the total number of reads mapped to the five genomic components, and (3) calculated the number of Mb explained per sequencing read. Finally, note that in addition to mapping to our W22 and B73 databases, for completeness we also mapped to a database consisting only of knob repeats, which avoided the complication of reference TE annotations. These analyses also detected a moderate group × generation effect (F -test; $P = 0.015$) (Supplementary Table 18), suggesting again that knob repeats contribute to GS shifts.

Testing BUSCO normalization via simulation. To compare counts among individuals, it is important to assess the accuracy of our normalization approach. We tested BUSCO normalization via simulations of TE loss and gain. For the simulations, we used the smallest chromosome—chromosome 10—for computational efficiency. We randomly removed either 10% or 20% of TEs from the chromosome, duplicated 10% of TEs, or did not change the chromosome. Each treatment was repeated five times with different random TEs removed or gained. The short-read simulator wgsim (<https://github.com/lh3/wgsim>) was used to simulate datasets with 2× and 10× coverage, mimicking the potential for different coverages among our libraries. For each simulation, reads were mapped to chromosome 10, counted across annotation features (non-BUSCO genes, TEs, knobs and rDNA) and then normalized by dividing by the total counts for BUSCO genes on chromosome 10. We simulated each set of parameters 1,000 times. On the basis of these simulations, we were able to recover the expected decrease in genomic components (Supplementary Fig. 1), but it did not recapitulate genome gain in TEs as accurately. It is likely that the inability to estimate TE gains is a feature of our simulations, because we duplicated TEs as exact, tandem copies of chromosomal TEs, which would lead to systematic undercounting of the duplicated TEs. Nevertheless, our simulations indicate that our normalization approach is sufficient to compare TE loss among datasets with different coverages and different degrees of TE loss.

Identification of SNPs and deleterious variants. To identify SNPs, paired-end sequencing reads were evaluated for quality using FastQC v.0.11.2, and were further processed to remove adapter contamination and low quality bases using Trimmomatic v.0.35⁶⁹, with the parameters LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20 and MINLEN:50. Trimmed reads were then mapped to the B73 reference genome (AGPv4.37⁶⁵; ftp://ftp.ensemblgenomes.org/pub/plants/release-37/fasta/zea_mays) using the MEM algorithm implemented in Burrows–Wheeler Aligner (BWA) v.0.7.12⁶⁶ with the parameters “-M -k 9 -T 25”. Mapping alignments from one individual were merged using Picard tools v.1.96 (<http://broadinstitute.github.io/picard/>) MergeSamFiles, and potential PCR duplicates were filtered from alignments using SAMtools v.1.1⁶⁷ rmdup. To minimize the number of mismatched bases, local realignment of reads around indels were performed using the Genome Analysis Toolkit (GATK) v.3.7⁷⁰ RealignerTargetCreator and IndelRealigner. Only uniquely mapped reads were kept for downstream SNP calling.

To detect SNPs, we used HaplotypeCaller, CombineGVCFs and GenotypeGVCFs from GATK v.3.7⁷⁰ separately on each of the six resequenced lines. Variant sites having a minimum phred-scaled confidence threshold 30 and a minimum base quality 20 were considered as SNP candidates. For the SNP set in all samples: (1) only biallelic SNPs were retained, (2) genotypes with genotype quality score < 5 were assigned as missing, and (3) the filtration “QUAL < 30.0, QD < 2.0, MQ < 10.0, DP < 3.0, ReadPosRankSum < -8.0, FS > 30.0” were set to further reduce false positives. A Python program parseVCFpy (https://github.com/simonhmartin/genomics_general) was adopted to extract the genotypes of every sample at each SNP site.

We identified putative deleterious SNPs (dSNPs) using SIFT⁷¹, which annotated SNPs as non-coding, synonymous and nonsynonymous, on the basis of the gene annotation information in Ensembl (<https://plants.ensembl.org>). The SIFT database of maize (AGPv3.22) was downloaded from SIFT 4 G (http://sift.bii.a-star.edu.sg/sift4g/public/Zea_mays/). Our SNP coordinates were converted to AGPv3 using CrossMap v.0.2.7⁷², and then SIFT 4 G⁷³ was launched to compute scores for all converted SNPs. Nonsynonymous SNPs were then predicted as deleterious or tolerated according to their computed SIFT scores. Nonsynonymous SNPs having SIFT score < 0.05 were predicted as deleterious; they were considered to be tolerated if they had a normalized probability value ≥ 0.05. For SNPs annotated by SIFT, the derived SNP was inferred using the *Sorghum* genome, on the basis of mapping the raw data from six sorghum varieties from the NCBI Short Read Archive (accession numbers DRR045087, DRR045074, DRR045075, DRR045082, DRR045083 and DRR045081) to the B73 reference. For our analyses, the derived allele was assumed to be the deleterious variant.

Recombination data. Crossover data for maize US-NAM population were retrieved from ref. ⁷. The start and end positions of crossover intervals were translated from *Z. mays* B73 AGPv2 to the AGPv4 reference, using CrossMap 0.2.7⁷². The number of crossover events in each non-overlapping, 5 Mb window was computed as in ref. ⁷; if a given crossover interval fell over > 1 window, the

proportion of the interval present in each window was added to the window crossover counts. Genomic windows were then classified into highly and lowly recombining using the crossover counts quartiles.

Examining SNP frequencies. We focused only on those SNPs for which the parent could be inferred to be heterozygous—that is, $H=1$ in the parent. Operationally, this implied that at least one heterozygote was detected in S1 or that there were two S1 homozygotes with alternative alleles. The derived allele was inferred by comparing SNPs to the *Sorghum* genome and making the hypothesis that the *Sorghum* allele is ancestral. SNPs were annotated using SIFT and classified into four categories (see main text). The proportion of the derived allele was computed for each SNP type in each chromosome separately for every line.

A generalized linear model with mixed effects was applied to the proportion of derived allele in each chromosome of every line using the R function `glmer` in the `lme4` package, using the binomial family of tests. Two fixed effects with interaction were considered in the model: the type of SNP as defined by SIFT and the inbreeding generation, see equation (1) below. The line was considered a random effect.

$$\begin{aligned} &(\text{number of derived alleles, number of ancestral alleles}) \\ &\sim \text{SNP type} * \text{Generation} + (1|\text{Line}) \end{aligned} \quad (1)$$

Both fixed effects and their interaction were significant (all $P < 2.2 \cdot 10^{-16}$) using comparison of the fit of equation (1) to simpler nested models (removing one effect at a time in equation (1)). To statistically test whether there was a significant difference between different types of SNPs and/or generations, we computed contrasts with the R package `multcomp`, which automatically corrects for multiple tests.

To study the effect of recombination on the proportion of the derived allele, the number of derived and ancestral alleles were summed for each chromosome of every line when considering only highly or lowly recombining genomic windows as previously defined. A similar linear model was then applied, with an additional fixed effect for recombination which interacts with the other two previous fixed effects:

$$\begin{aligned} &(\text{number of derived alleles, number of ancestral alleles}) \\ &\sim \text{SNP type} * \text{Generation} * \text{Recombination} + (1|\text{Line}) \end{aligned} \quad (2)$$

As previously, all three fixed effects and their interactions were significant when comparing model (2) to simpler nested models (all $P < 0.007$).

Heterozygosity analyses. For each individual, we used sliding windows of 100 SNPs to infer heterozygosity for genomic regions, focusing only on SNPs within genes to avoid potential misalignments due to repetitive elements. Using the set of SNPs inferred to be heterozygous in the parents, the proportion of the major allele P was calculated as follows: if a position was homozygous, then the proportion of the major allele was 1. If a position was heterozygous, then one of the two alleles was arbitrarily assigned to be the major allele and given a proportion of 0.5. The proportion P was then averaged across the 100 SNPs of each window for each individual separately to calculate P' . We assumed that the limited number of recombination events in each line over the time course of the experiment did not fully homogenize chromosomes, so that most genomic regions were either heterozygous or homozygous. Based on this approach, the genomic regions that are heterozygous should exhibit a P' close to 0.5 while genomic regions that are homozygous should have P' close to 1. Note, however, that real heterozygous loci can be misgenotyped as homozygous to make the $P' > 0.5$. Also, the maize genome contains a high number of duplicated genes, and erroneous mapping of reads from duplicated genes can cause false heterozygous SNPs in homozygous regions¹⁰, making $P' < 1$ in homozygous regions. Nevertheless, when coverage is high enough to genotype heterozygotes correctly, two peaks of $P' = 0.5$ and $P' = 1.0$ should be observed.

The distribution of P' for each line across all individuals and generations is presented in Supplementary Fig. 11. Only MR09 and MR22 exhibited the expected two peaks. These two lines have the highest coverage among the set of lines (Supplementary Table 16), and they were therefore the only lines we studied hereafter. Given the distribution of P' across genomic regions, the R package `Mclust` was used to classify each window of each individual as homozygous or heterozygous²⁴ by forcing the number of components to be 2 ($G=2$). Windows that fell between the two peaks of the P' distribution were classified as 'uncertain' if the `Mclust` classification uncertainty was > 0.1 (Supplementary Figs. 12 and 13).

For each individual, the heterozygosity status of a region was inferred from the clustering of overlapping sliding windows. The start and end of a heterozygous region were defined by (1) the start of the first window that had the given heterozygosity state and (2) the start of the closest next 'uncertain' window. All SNPs inside the region were afterwards considered to be of the inferred heterozygosity type, regardless of genotyping errors. A similar procedure was applied to homozygous regions. Although in principle the categorical status of uncertain regions could be inferred by parsimony arguments, we adopted the conservative approach to discard these blocks of uncertainty from heterozygosity calculations. Heterozygosity levels could then be averaged across individuals of the same line and generation in sliding windows containing 100 SNPs as follows:

Heterozygosity = number of inferred heterozygous SNPs / (number of inferred heterozygous SNPs + number of inferred homozygous SNPs)

Average heterozygosity levels across individuals were plotted along chromosomes for sliding windows of 100 SNPs that fall within genes (Fig. 4). For statistical tests, chromosomes were considered as biologically independent units, owing to the small number of individuals ($n=2$ or 3). The non-parametric Wilcoxon signed-rank test was used to compare the expected heterozygosity with the observed heterozygosity of the ten chromosomes averaged across individuals for each line and generation separately. As a conservative control, this analysis was repeated when considering windows with uncertain heterozygosity in the clustering method as homozygous, instead of discarding them. A similar approach with non-overlapping windows of 100 SNPs falling within genes was used to correlate heterozygosity with crossover number using the R `lm` function. The same non-overlapping windows were used to study the effect of the proportion of nonsynonymous SNPs on heterozygosity using a χ^2 contingency table test with the R function `chisq.test`.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequence data that support the findings of this study have been deposited in NCBI Short Read Archive under project code [SRP158803](https://www.ncbi.nlm.nih.gov/sra/SRP158803). The gff files used in this study, the GS flow cytometry data and the raw mapping count data are available on figshare.com (<https://doi.org/10.6084/m9.figshare.783825.v2>) or from the corresponding author. The SNP VCF files and dataset are available from the corresponding authors upon request.

Code availability

Custom code used in SNP analyses is available from the corresponding authors upon request.

Received: 4 September 2018; Accepted: 26 July 2019;

Published online: 2 September 2019

References

- Darwin, C. *The Effects of Self and Cross Fertilization in the Vegetable Kingdom* (John Murray, 1876).
- Fisher, R. A. Average excess and average effect of a gene substitution. *Ann. Hum. Genet.* **11**, 53–63 (1941).
- Charlesworth, D. & Willis, J. H. The genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–796 (2009).
- Hedrick, P. W., Hellsten, U. & Grattapaglia, D. Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytol.* **209**, 600–611 (2016).
- Schnable, P. S. & Springer, N. M. Progress toward understanding heterosis in crop plants. *Annu. Rev. Plant Biol.* **64**, 71–88 (2013).
- Byers, D. L. & Waller, D. M. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. *Annu. Rev. Ecol. Syst.* **30**, 479–513 (1999).
- Rodgers-Melnick, E. et al. Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl Acad. Sci. USA* **112**, 3823–3828 (2015).
- McMullen, M. D. et al. Genetic properties of the maize nested association mapping population. *Science* **325**, 737–740 (2009).
- Barrière, A. et al. Detecting heterozygosity in shotgun genome assemblies: lessons from obligately outcrossing nematodes. *Genome Res.* **19**, 470–480 (2009).
- Brandenburg, J. T. et al. Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. *PLoS Genet.* **13**, e1006666 (2017).
- Crnokrak, P. & Barrett, S. C. Perspective: purging the genetic load: a review of the experimental evidence. *Evolution* **56**, 2347–2358 (2002).
- Lande, R. & Schmske, D. W. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**, 24–40 (1985).
- Charlesworth, B., Charlesworth, D. & Morgan, M. T. Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature* **347**, 380–382 (1990).
- Arunkumar, R., Ness, R. W., Wright, S. I. & Barrett, S. C. The evolution of selfing is accompanied by reduced efficacy of selection and purging of deleterious mutations. *Genetics* **199**, 817–829 (2015).
- Liu, Q., Zhou, Y., Morrell, P. L. & Gaut, B. S. Deleterious variants in Asian rice and the potential cost of domestication. *Mol. Biol. Evol.* **34**, 908–924 (2017).
- Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G. & Allendorf, F. W. Genomics advances the study of inbreeding depression in the wild. *Evol. Appl.* **9**, 1205–1218 (2016).

17. Morran, L. T., Ohdera, A. H. & Phillips, P. C. Purging deleterious mutations under self fertilization: paradoxical recovery in fitness with increasing mutation rate in *Caenorhabditis elegans*. *PLoS ONE* **5**, e14473 (2010).
18. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
19. Tenaillon, M. I., Hollister, J. D. & Gaut, B. S. A triptych of the evolution of plant transposable elements. *Trends Plant Sci.* **15**, 471–478 (2010).
20. Tenaillon, M. I., Hufford, M. B., Gaut, B. S. & Ross-Ibarra, J. Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. *Genome Biol. Evol.* **3**, 219–229 (2011).
21. Diez, C. M., Meca, E., Tenaillon, M. I. & Gaut, B. S. Three groups of transposable elements with contrasting copy number dynamics and host responses in the maize (*Zea mays* ssp. *mays*) genome. *PLoS Genet.* **10**, e1004298 (2014).
22. Bilinski, P. et al. Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. *PLoS Genet.* **14**, e1007162 (2018).
23. Wright, S. I., Kalisz, S. & Slotte, T. Evolutionary consequences of self-fertilization in plants. *Proc. Biol. Sci.* **280**, 20130133 (2013).
24. Hollister, J. D. & Gaut, B. S. Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **19**, 1419–1428 (2009).
25. Lee, Y. C. G. & Karpen, G. H. Pervasive epigenetic effects of *Drosophila* euchromatic transposable elements impact their evolution. *eLife* **6**, e25762 (2017).
26. Quadrona, L. *The Arabidopsis thaliana* mobilome and its impact at the species level. *eLife* **5**, e15716 (2016).
27. Price, H. J. Evolution of DNA content in higher plants. *Bot. Rev.* **42**, 27 (1976).
28. Wright, S. I., Ness, R. W., Foxe, J. P. & Barrett, S. C. H. Genomic consequences of outcrossing and selfing in plants. *Int. J. Plant Sci.* **169**, 105–118 (2008).
29. First, J. L. et al. Reproductive mode and the evolution of genome size and structure in *Caenorhabditis* nematodes. *PLoS Genet.* **11**, e1005323 (2015).
30. Wills, D. M. et al. From many, one: genetic control of prolificacy during maize domestication. *PLoS Genet.* **9**, e1003604 (2013).
31. Diez, C. M. et al. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytol.* **199**, 264–276 (2013).
32. Dolezel, J., Bartos, J., Voglmayr, H. & Greilhuber, J. Nuclear DNA content and genome size of trout and human. *Cytom. A* **51**, 127–128 (2003).
33. Long, Q. et al. Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nat. Genet.* **45**, 884–890 (2013).
34. Cullis, C. A. Mechanisms and control of rapid genomic changes in flax. *Ann. Bot.* **95**, 201–206 (2005).
35. Jian, Y. et al. Maize (*Zea mays* L.) genome size indicated by 180-bp knob abundance is associated with flowering time. *Sci. Rep.* **7**, 5954 (2017).
36. Mroczek, R. J., Melo, J. R., Luce, A. C., Hiatt, E. N. & Dawe, R. K. The maize Ab10 meiotic drive system maps to supernumerary sequences in a large complex haplotype. *Genetics* **174**, 145–154 (2006).
37. Randolph, L. F. Genetic characteristics of the B chromosomes in maize. *Genetics* **26**, 608–631 (1941).
38. Yamakake, K. & Angel, T. *Cytological Studies in Maize (Zea mays L.) and Teosinte (Zea mexicana (Schrader) Kuntze) in Relation to Their Origin and Evolution* (Massachusetts Agricultural Experiment Station, 1976).
39. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
40. Springer, N. M. et al. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat. Genet.* **50**, 1282–1288 (2018).
41. Devos, K. M., Brown, J. K. & Bennetzen, J. L. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res.* **12**, 1075–1079 (2002).
42. Bousios, A. et al. The turbulent life of *Sirevirus* retrotransposons and the evolution of the maize genome: more than ten thousand elements tell the story. *Plant J.* **69**, 475–488 (2012).
43. Darzentas, N., Bousios, A., Apostolidou, V. & Tsafaris, A. S. MASiVE: Mapping and Analysis of *Sirevirus* Elements in plant genome sequences. *Bioinformatics* **26**, 2452–2454 (2010).
44. Ng, P. C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
45. Takebayashi, N. & Morrell, P. L. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.* **88**, 1143–1150 (2001).
46. Weller, S. G., Sakai, A. K., Thai, D. A., Tom, J. & Rankin, A. E. Inbreeding depression and heterosis in populations of *Schiedea viscida*, a highly selfing species. *J. Evol. Biol.* **18**, 1434–1444 (2005).
47. Smarda, P., Horova, L., Bures, P., Hralova, I. & Markova, M. Stabilizing selection on genome size in a population of *Festuca pallens* under conditions of intensive intraspecific competition. *New Phytol.* **187**, 1195–1204 (2010).
48. Rayburn, A. L., Dudley, J. W. & Biradar, D. P. Selection for early flowering results in simultaneous selection for reduced nuclear-DNA content in maize. *Plant Breed.* **112**, 318–322 (1994).
49. Sun, S. et al. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. *Nat. Genet.* **50**, 1289–1295 (2018).
50. Chia, J. M. et al. Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* **44**, 803–807 (2012).
51. Lyu, H., He, Z., Wu, C. I. & Shi, S. Convergent adaptive evolution in marginal environments: unloading transposable elements as a common strategy among mangrove genomes. *New Phytol.* **217**, 428–438 (2018).
52. Schnable, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009).
53. Yin, D. et al. Rapid genome shrinkage in a self-fertile nematode reveals sperm competition proteins. *Science* **359**, 55–61 (2018).
54. Springer, N. M. & Stupar, R. M. Allelic variation and heterosis in maize: how do two halves make more than a whole? *Genome Res.* **17**, 264–275 (2007).
55. Gore, M. A. et al. A first-generation haplotype map of maize. *Science* **326**, 1115–1117 (2009).
56. Charlesworth, B., Morgan, M. T. & Charlesworth, D. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).
57. Bersabé, D., Caballero, A., Pérez-Figueroa, A. & García-Dorado, A. On the consequences of purging and linkage on fitness and genetic diversity. *G3* **6**, 171–181 (2015).
58. Schumer, M. et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **360**, 656–660 (2018).
59. Tenaillon, M. I., Manicacci, D., Nicolas, S. D., Tardieu, F. & Welcker, C. Testing the link between genome size and growth rate in maize. *PeerJ* **4**, e2408 (2016).
60. Hu, T. T. et al. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat. Genet.* **43**, 476–481 (2011).
61. Beaulieu, J. M., Leitch, I. J., Patel, S., Pendharkar, A. & Knight, C. A. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytol.* **179**, 975–986 (2008).
62. Knight, C. A., Molinari, N. A. & Petrov, D. A. The large genome constraint hypothesis: evolution, ecology and phenotype. *Ann. Bot.* **95**, 177–190 (2005).
63. Charlesworth, D., Charlesworth, B. & Strobeck, C. Selection for recombination in partially self-fertilizing populations. *Genetics* **93**, 237–244 (1979).
64. Roze, D. & Lenormand, T. Self-fertilization and the evolution of recombination. *Genetics* **170**, 841–857 (2005).
65. Jiao, Y. et al. Improved maize reference genome with single-molecule technologies. *Nature* **546**, 524–527 (2017).
66. Li, H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**, 2843–2851 (2014).
67. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
68. Quinlan, A. R. BEDTools: The Swiss-Army tool for genome feature analysis. *Curr. Protoc. Bioinforma.* **47**, 11.12.1–11.12.34 (2014).
69. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
70. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
71. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
72. Zhao, H. et al. CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30**, 1006–1007 (2014).
73. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. & Ng, P. C. SIFT missense predictions for genomes. *Nat. Protoc.* **11**, 1–9 (2016).
74. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* **8**, 289–317 (2016).

Acknowledgements

A.M. is supported by an EMBO Postdoctoral Fellowship ALTF 775-2017 and by HFSP Fellowship LT000496/2018-L. A.B. is supported by The Royal Society (award numbers UF160222 and RGF/R1/180006). M.C.S. is supported by an NSF Graduate Research Fellowship to UC Davis (1148897). D.K.S. is supported by a Postdoctoral Fellowship from the National Science Foundation (NSF) Plant Genome Research Program (1609024). J.F.D. is supported by NSF grant IOS 1238014. Q.L. is supported by a National Natural Science Foundation of China grant (no. 31471431) and the Training Program for Outstanding Young Talents of Zhejiang A&F University to Q.L. B.S.G. is supported by NSF grants 1542703 and 1655808.

Author contributions

K.R., A.M. and B.S.G. contributed analyses, ideas and writing. G.R.J.G. and Q.L. performed analyses. K.R., C.M.D. and B.S.G. helped design the experiment, grew plants and measured phenotypes; A.B., G.R.J.G., Q.L., D.K.S., J.F.D. and M.C.S. provided materials, data and/or critical ideas. B.S.G. conceived of the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41477-019-0508-7>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Q.L. or B.S.G.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| ☐ | ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ A description of all covariates tested |
| ☐ | ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☐ | ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Collection of sequence data employed software on the Illumina HiSeq 2500 and HiSeq3000 platforms.

Data analysis

Software used for read mapping and counting: Trimmomatic (v0.35), BWA-MEM (v0.7.12), Samtools (v1.3), blat(v36), bltoggff(v3), bedtools (v.2.25.0), BUSCO(v3), MASIVE (<http://databases.bat.infspire.org/masivedb/>), wgsim(v1). Software used for SNP detection: FastQC(v0.11.2), Trimmomatic(v0.35), BWA(v0.7.12), Picard tools (V1.96), Samtools (v1.1), GATK (v3.7), parseVCF.py (github.com/simonhmartin/genomics_general), SIFT 4G, CrossMap (v0.2.7) Used throughout: Rstudio (v0.98.1103), R(v3.34)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequence data generated for this study have been deposited in NCBI Short Read Archive under project code SRP158803. The gffs used in this study, the genome size flow cytometry data and the raw mapping count data are available on figshare.com or from the corresponding author. The SNP vcfs and dataset are available upon request, as is code used in SNP analyses.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study evaluates the genome effects of inbreeding, based initially on 11 lines that were selfed for up to seven generations. The lines are described below and also in Chia et al. Nature Genetics (2012) and Wills et al. Plos Genetics (2009). With these materials, we studied genome size from the S1 to later generations (e.g., S6) in each of the 11 lines. We then focused on 3 lines that lost genome size and 3 that apparently did not lose genome size to assess genomic components that were lost and to assess purging of putatively deleterious SNPs.
Research sample	The research sample utilized a set of 11 inbred lines that had been generated by J. Doebley at the University of Wisconsin. These lines were chosen to represent diverse landraces of maize. Each began with a parent that was selfed for up to 7 generations. The lines were promulgated by single seed descent, and sibling seed were retained from each generation. Plants grown from the sibling seed were utilized in this study.
Sampling strategy	No preliminary analyses on sample size were performed. For the first analyses (genome size) we first identified lines that appeared to change genome size by utilizing materials available to us in early (S1) and late (S6) generations. For sequencing data, we chose the beginning and end point of the experiment, choosing 3 replicates of the two groups - lines that lost genome size and lines that apparently did not - for contrasts.
Data collection	Plants were grown in a randomized plot design. Samples from available plants were outsourced to get flow cytometric estimates of genome size. The sequencing data were generated by making libraries and sequencing at UC Irvine.
Timing and spatial scale	Plants were germinated and sown in Spring 2014, with plant height and mortality measured 9, 17, 30 and 45 days after sowing.
Data exclusions	We excluded sequence data from generations 2 and 4 from the analysis. These data were generated from only two lines and hence did not follow the broader study design. The data were pre-excluded because of their failure to conform to the study design.
Reproducibility	For flow cytometry, we included both technical and biological replicates. For sequence data, we used biological replicates, sequencing different sets of plants based on the their line and their number of generations of inbreeding.
Randomization	Plant growth was in a randomized block design.
Blinding	For phenotyping, plant growth and genome size assays, blinding was inherent in the randomized block design. The generation of sequencing data was blind to control vs. genome-loss groups. Sequence analysis was performed as if blinded, but members of the two groups were known.
Did the study involve field work?	☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Source tissue was leaves. Leaf punches were sent to Plant Cytometry Services (Schijndel, the Netherlands) where relative DNA measurements were performed by flow cytometry using internal standards for DAPI and PI analysis.

Instrument

DAPI: Partec CyFlow Space PI: partec cube

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Samples were analysed using Ilex crenata and maize B73 as internal standards, as stated in the manuscript.

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.