Prospective Targeted Recombination and Genetic Gains for Quantitative Traits in Maize

Rex Bernardo*

Abstract

Advances in clustered regularly interspaced short palindromic repeats (CRISPR) technology have allowed targeted recombination in specific DNA sequences in yeast (*Saccharomyces cerevisiae*). My objective was to determine if the selection gains from targeted recombination are large enough to warrant the development of targeted recombination technology in plants. Genomewide marker effects for quantitative traits in two maize (*Zea mays* L.) experiments were used to identify targeted recombination points that would maximize the per-chromosome genetic gains in a given cross. With nontargeted recombination in the intermated B73 × Mo17 population, selecting the best out of 180 recombinant inbreds led to a 7.1% gain for testcross yield. Having one targeted recombination on each of the 10 maize chromosomes led to a predicted gain of 15.3% for yield. Targeted recombination therefore led to a predicted relative efficiency ($RE_{Targeted}$) of (0.153 ÷ 0.071) = 212% of targeted recombination compared with nontargeted recombination. For the five other traits in the intermated B73 × Mo17 population and for four traits in 45 other maize crosses, the $RE_{Targeted}$ values ranged from 105 to 600%. The targeted recombination points differed among traits and crosses. Predicted gains increased when the number of targeted recombinations per chromosome increased from one to two. Overall, the results suggested that targeted recombination could double the selection gains for quantitative traits in maize and that the development of targeted recombination technology is worthwhile. Empirical experiments with current marker-assisted breeding procedures are needed to validate the per-chromosome predicted gains.

Published in Plant Genome
Volume 10. doi: 10.3835/plantgenome2016.11.0118

TARGETED RECOMBINATION refers to the possibility of inducing homologous recombination (in mitosis) or crossing over (in meiosis) at a specific point along a chromosome. In this study, I investigated whether targeted recombination in maize would lead to large genetic gains for quantitative traits. Exploiting targeted recombination in crops has two prerequisites: (i) targeted recombination is technically feasible, and (ii) ideal recombination points on each chromosome can be determined.

Genome editing technologies have enabled the deletion, addition, or substitution of DNA bases at known genes in plants (Gaj et al., 2013; Pennisi, 2013; Belhaj et al., 2013; Bortesi and Fischer, 2015). A different application of CRISPR technology involves its use not for genome editing but for inducing mitotic recombination at precise locations in the genome. Recently, Sadhu et al. (2016) demonstrated that homologous repair of CRISPR-induced double-strand breaks leads to chromosomes with precise recombination breakpoints in yeast. Repair of the double-strand break via homologous recombination led to yeast cells that exhibited a loss-of-heterozygosity event [see figure 1 in Sadhu et al. (2016)]. By creating a loss-of-heterozygosity mapping panel of yeast strains that differed in their CRISPR-induced recombination points, Sadhu et al. were then able to fine-map Mn sensitivity on chromosome 7 to a single polymorphism.

Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, 411 Borlaug Hall, 1991 Upper Buford Circle, Saint Paul, MN 55108. Received 22 Nov. 2016. Accepted 23 Feb. 2017. *Corresponding author (bernardo@umn.edu). Assigned to Associate Editor Shawn Kaeppler.

Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; QTL, quantitative trait loci; RR-BLUP, ridge regression-best linear unbiased prediction; $R_{Nontargeted}$, response to selection with nontargeted recombination; $R_{Targeted}$, response to selection with targeted recombination; $RE_{Targeted}$, relative efficiency of selection with targeted recombination versus selection without targeted recombination; SNP, single nucleotide polymorphism

It seems plausible, at least in concept, that the technology found useful in yeast by Sadhu et al. (2016) can be adapted to eventually obtain F_1–derived doubled haploids that have recombination events precisely where a plant breeder wants them. A doubled haploid with targeted recombination may be obtained if the following protocols are developed: (i) a multiplex CRISPR system induces mitotic double-strand breaks at multiple target DNA sequences, (ii) cells with the desired loss-of-heterozygosity events are screened, and (iii) the desired cells are regenerated into whole plants that carry the targeted recombinations. From the regenerated plants, doubled haploids can be developed via standard procedures that have already been found routinely useful in breeding programs. Doubled haploids that exhibit the targeted recombinations can be identified via markers. In addition to the possible use of CRISPR technology (Sadhu et al., 2016), targeted meiotic recombination via the *Spo11* endonuclease is being investigated [Bastianelli and Nicolas, 2015; http://www.meiogenix.com, (accessed 7 Apr. 2017)].

Of the two prerequisites mentioned above, determining where to induce recombinations is currently more straightforward. Many studies have shown that genome-wide selection (or genomic selection; Meuwissen et al., 2001) is effective in plants (Lorenzana and Bernardo, 2009; Albrecht et al., 2011; Massman et al., 2013b; Asoro et al., 2013; Combs and Bernardo, 2013; Jacobson et al., 2014; Lian et al., 2014; Beyene et al., 2015; Rutkoski et al., 2015). Estimated marker effects can be used to predict the recombination points to target on each chromosome. Suppose that eight single nucleotide polymorphism (SNP) markers are found on a chromosome (Fig. 1). Suppose further that for a given trait, the effects of the SNP alleles carried by two parental inbreds range from −0.9 to 0.9. The parental genotypic values (0.2 and −0.2) are obtained as the sum of the genotypic values associated with the SNP alleles carried by each parent. To identify where recombination should be targeted, recombination is assumed at each of the marker intervals, and the recombination point and specific homolog that lead (in doubled haploid form) to the highest predicted genotypic value are determined. In Fig. 1, a targeted recombination between the second and third SNP loci leads to a doubled haploid with a predicted genotypic value of 2.2. A second targeted recombination, between the two rightmost markers on the chromosome, would increase the predicted genotypic value to 2.8.

Much work is obviously needed to transform targeted recombination from possibility to reality in plants. Investment into both public and private research on targeted recombination technology will depend on the prospective genetic gains for important traits. If targeted recombination in maize is expected to increase the genetic gain for grain yield by 5 to 10%, then targeted recombination does not hold much promise because such gains can be achieved by using a larger population size and more stringent selection intensity. However, if targeted recombination can increase the yield gains by a substantial amount, then developing targeted recombination technology in plants may be worthwhile. To date, there have been no estimates of the genetic gains from prospective targeted recombination for quantitative traits in crops.

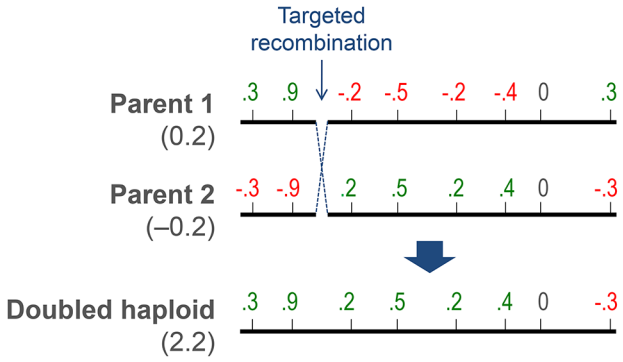

Fig. 1. Targeted recombination on the basis of genome-wide marker effects. The genotypic values of two parental inbreds and a doubled haploid developed via targeted recombination are in parentheses.

This study in maize was an attempt to begin to determine whether there is sufficient downstream impetus to develop targeted recombination technology in crops. My objectives were to assess (i) the predicted responses to selection if one or two targeted recombinations were to occur on each maize chromosome, (ii) the extent to which the predicted responses with targeted recombination vary among maize traits and populations, and (iii) the consistency of the ideal recombination points among maize traits and populations.

MATERIALS AND METHODS

Maize Populations and Marker and Phenotypic Data

Phenotypic and marker data from two previous experiments from the University of Minnesota were used. Experiment 1 included 180 recombinant inbreds from the intermated B73 × Mo17 population (Lee et al., 2002). The recombinant inbreds were testcrossed to a proprietary Monsanto inbred that performed well when crossed to both B73 and Mo17, and the testcrosses were evaluated for grain yield (kg ha^{-1} at 150 g H_2O kg^{-1}), moisture (g kg^{-1}), plant height (cm), stalk lodging (%), root lodging (%), and stover quality traits that are important for producing cellulosic ethanol (Lewis et al., 2010). The stover quality traits included the amount of cell wall glucose released from the stover on thermochemical pretreatment and enzymatic saccharification (“glucose release”) (Dien et al., 2006) and the concentration of Klason lignin on a cell wall basis (“lignin”). The field trials were conducted at four Minnesota locations in 2007. Information on yield, moisture, stalk lodging, and root lodging were combined by Massman et al. (2013b) into a yield index. Retrospective weights applied by a team of maize

breeders for these four traits (Bernardo, 1991) were used to calculate the yield index.

Marker data for the intermated B73 × Mo17 recombinant inbreds were obtained from Liu et al. (2010). The 892 SNP loci considered in this study were a subset of the 1016 SNP loci on the Iowa State University linkage map (Liu et al., 2010). To reduce multicollinearity among the SNP markers, linkage disequilibrium between adjacent SNP markers was calculated with reference to the 180 recombinant inbreds, and the 892 SNP loci were obtained after discarding one of a pair of adjacent SNP loci that had $r^2 > 0.95$. The SNP genotypes were coded as 1 (homozygous for the B73 allele), −1 (homozygous for the Mo17 allele), and 0 (heterozygous). Missing marker data were projected as follows. If the two flanking loci had the same genotype (1 and 1, −1 and −1, or 0 and 0), then the middle SNP with missing data was assigned the flanking marker genotypes. In all other cases, the marker data remained missing. Among the 160,560 data points between the 180 recombinant inbreds and 892 SNP loci, only 546 (0.34%) were heterozygous and only 679 (0.42%) had missing data after projection. For each trait, genomewide marker effects were obtained via ridge regression-best linear unbiased prediction (RR-BLUP) as implemented in RRBLUP2 in-house software (available at http://bernardo-group.org, accessed 7 Apr. 2017).

Experiment 2 comprised a subset of 10 representative inbreds (Bernardo, 2015) chosen from a larger panel of 271 inbreds (Schaefer and Bernardo, 2013). These 10 inbreds were A321, A632, A641, F2, LH82, LH127, ND203, PH207, PHG80, and PHH93. The 271 inbreds were evaluated for anthesis date (growing degree-days from planting to when 50% of the plants were shedding pollen), plant height (cm), kernel starch concentration (g kg^{-1}), and kernel protein concentration (g kg^{-1}) at five Minnesota locations in 2011 and one Minnesota location in 2012. Genomewide marker effects at 28,826 SNP loci were previously calculated by Schaefer and Bernardo (2013) using rrBLUP software in R (Endelman, 2011).

Predicted Gains from Targeted Recombination

Targeted recombination was studied for each cross and for each trait. Selection in Experiment 1 was for higher yield, yield index, and glucose release and for lower moisture, plant height, and lignin. Selection in Experiment 2 was for higher starch and for lower values of protein (which is negatively correlated with starch), anthesis date, and plant height.

In Experiment 1, genomewide marker effects in the intermated B73 × Mo17 population were used to predict the performance of a doubled haploid that has a recombination event at the first marker interval, the second marker interval, and so on within each chromosome (as described in the Introduction). Each recombination event leads to two possible doubled haploids, depending on which chromosome in a homologous pair is eventually inherited by the homozygous line. The better of the two was recorded. Furthermore, double recombination events were considered between all possible pairs of marker intervals within a chromosome. The predicted performance with double recombinants was recorded in the same manner as that for single recombinants. The same procedures were used for each of the 45 pairwise combinations among the 10 inbreds in Experiment 2. For simplicity, particularly given that the technology has not yet been developed, the number of targeted recombinations in this study was limited to a maximum of two per chromosome.

The single-recombination and double-recombination points that led to the best performance were identified on each chromosome. Given that the different chromosomes assort independently, the performance attributed to each chromosome was summed across all chromosomes. In both Experiments 1 and 2, the response to selection with targeted recombination ($R_{Targeted}$) was calculated as the predicted performance of a doubled haploid that had the ideal recombinations on each of the 10 chromosomes, minus the mean of the cross as obtained from RR-BLUP analysis.

In Experiment 1, the $R_{Targeted}$ in the B73 × Mo17 population was compared with the response when the best recombinant inbred out of the 180 recombinant inbreds was selected, the latter selection response being denoted by $R_{Nontargeted}$. Genomewide prediction involves changes in scale between the predicted versus the observed performance. For example, the SD of the observed yields (kg ha^{-1}) of the 180 recombinant inbreds was 681, with the observed yields ranging from 7099 to 10,352. Because of shrinkage of the genomewide marker effects, the SD of the predicted yields of the 180 recombinant inbreds was 271, with the predicted yields ranging from 8279 to 9618. To avoid the effects of scale, $R_{Nontargeted}$ was assessed on the basis of the marker-predicted performance rather than the observed performance of the 180 lines. This procedure for assessing $R_{Nontargeted}$ was facilitated by the high correlations between marker-predicted and observed performance. These correlations in Experiment 1 (without cross-validation) were 0.80 for yield, 0.85 for lignin, 0.86 for the yield index, and >0.89 for the three other traits. As mentioned later, the ratio between $R_{Targeted}$ and $R_{Nontargeted}$ was calculated; this ratio was unaffected by scale. On the other hand, targeted recombination would involve predicting the performance of a line that is not part of the training population (i.e., with cross-validation). The estimates of prediction accuracy (Dekkers, 2007) with cross-validation in the intermated B73 × Mo17 population were 0.61 for yield, 0.58 for moisture, 0.73 for yield index, 0.70 for plant height, 0.57 for glucose release, and 0.54 for lignin (Lorenzana and Bernardo, 2009).

For each of the 45 pairs of parents in Experiment 2, the performance of the best 1% of lines (assuming random recombination) was obtained via virtual populations as described by Bernardo (2015). In particular, each virtual population consisted of 1000 doubled haploids simulated from the F_1 cross between the two parents. The number of recombination events per chromosome was simulated as a truncated Poisson process, with a maximum of three recombination events per chromosome.

Table 1. Prospective targeted recombination and genetic gains in the intermated B73 × Mo17 maize population (Experiment 1; Lewis et al., 2010; Liu et al., 2010) and in a maize diversity panel (Experiment 2; Schaefer and Bernardo, 2013). Predicted performance assumes one (x = 1) or two (x = 2) targeted recombinations on each of the 10 maize chromosomes.

	Experiment 1						Experiment 2,	
			Predicted performance		$RE_{Targeted}$§		mean¶ of $RE_{Targeted}$	
Trait	Mean†	Best 1 out of 180‡	x = 1	x = 2	x = 1	x = 2	x = 1	x = 2
					%			
Grain yield (kg ha^{-1})	8982	9618	10,330	10,599	212	254	–	–
Moisture (g kg^{-1})	201	170	125	93	240	340	–	–
Yield index	3.35	4.5	6.15	6.58	242	280	–	–
Lignin (g kg^{-1})	211	207	195	187	403	631	–	–
Glucose release (g kg^{-1})	516	539	573	598	243	352	–	–
Plant height (cm)	282	259	224	204	252	337	208 (108–600)	263 (126–815)
Anthesis date (GDD#)	–	–	–	–	–	–	214 (105–536)	261 (119–637)
Starch (g kg^{-1})	–	–	–	–	–	–	213 (108–556)	274 (125–754)
Protein (g kg^{-1})	–	–	–	–	–	–	189 (109–577)	232 (126–692)

† Mean of the 180 intermated B73 × Mo17 recombinant inbreds.

‡ Marker-predicted performance of the best recombinant inbred out of 180.

§ RE_{Target}, relative efficiency of selection with targeted recombination versus selection without targeted recombination. All RE_{Target} values in Experiment 1 and all mean RE_{Target} values in Experiment 2 were significantly greater (P = 0.05) than 100%.

¶ Among 45 pairwise crosses in Experiment 2. The range is shown in parentheses.

Growing degree-days

The mean number of recombinations was one per 100 cM and the chromosome sizes were those in a published maize linkage map (Senior et al., 1996). The $R_{Nontargeted}$ was obtained from the predicted mean of the best 10 out of 1000 simulated doubled haploids (best 1% selected). The correlations between predicted and observed performance without cross-validation in Experiment 2 were $\geq$ 0.96 (Bernardo and Thompson, 2016). These high correlations were consistent with the very high coefficients of determination attributed to inbreds (0.92–0.99) in the Schaefer and Bernardo (2013) study. The estimates of predictive ability from delete-one cross-validation among the 271 inbreds in Experiment 2 were 0.55 for starch, 0.61 for protein, 0.64 for plant height, and 0.74 for anthesis.

Significance Tests

The relative efficiency of selection with targeted recombination versus selection with nontargeted recombination was calculated as: $RE_{Targeted} = 100 \times (R_{Targeted} \div R_{Nontargeted})$.

In Experiment 1, the significance of $RE_{Targeted}$ (for the null hypothesis of $RE_{Targeted}$ = 100%) was tested by (i) obtaining values of $RE_{Targeted}$ within each of the 10 chromosomes, and (ii) bootstrapping on the within-chromosome values of $RE_{Targeted}$. Bootstrap confidence intervals on $RE_{Targeted}$ were obtained for a significance level of P = 0.05 with 10,000 bootstrap samples. In Experiment 2, the mean $RE_{Targeted}$ was calculated across the 45 pairs of parents, and t-tests (P = 0.05) were conducted on the basis of the variance of the estimates of $RE_{Targeted}$ among the 45 crosses.

RESULTS AND DISCUSSION

Selection Gains from Targeted Recombination

The results indicated that only 10 targeted recombinations, one on each of the 10 maize chromosomes, could double the selection gains for quantitative traits in maize. In the intermated B73 × Mo17 population, the mean grain yield (in testcross combination) of the 180 recombinant inbreds was 8982 kg ha^{-1} (Table 1). Among the 180 recombinant inbreds, the marker-predicted mean of the best inbred (M0014) was 9618 kg ha^{-1}. Selecting the best inbred out of 180 therefore led to a (9618 – 8982) ÷ 8982 = 7.1% gain in yield. In contrast, having one targeted recombination (x = 1) on each of the 10 maize chromosomes led to a predicted yield of 10,330 kg ha^{-1}. This predicted yield corresponded to a 15.3% gain in yield compared to the mean of the population, and to a relative efficiency of $RE_{Targeted}$ = (10,330 – 8962) ÷ (9618 – 8982) = 212% compared with nontargeted recombination. For the five other traits in the intermated B73 × Mo17 population, the $RE_{Targeted}$ values with x = 1 ranged from 240 to 403%.

Among the 45 maize biparental crosses in Experiment 2, the mean $RE_{Targeted}$ values with x = 1 ranged from 189% for kernel protein concentration to 214% for anthesis date (Table 1). For each trait, the $RE_{Targeted}$ values differed among the crosses. For plant height, for example, the $RE_{Targeted}$ with x = 1 ranged from 108% in A632 × ND203 to 600% in F2 × ND203. Caution is needed in interpreting these results because the highest $RE_{Targeted}$ value tended to correspond to the smallest actual gains. To illustrate, the $RE_{Targeted}$ of 108% for plant height in

A632 × ND203 corresponded to responses of $R_{Nontargeted}$ = −21.1 cm with nontargeted recombination and $R_{Targeted}$ = −23.0 cm with targeted recombination (x = 1). In the opposite extreme, the $RE_{Targeted}$ of 600% in F2 × ND203 corresponded to $R_{Nontargeted}$ = −1.7 cm and $R_{Targeted}$ = −10.3 cm. Similar results were obtained for the other traits in Experiment 2. Both F2 and ND203 are short inbreds [106 cm in the Schaefer and Bernardo (2013) field trials], whereas A632 is a taller inbred (147 cm). The differences in height between the parents were reflected in the estimated genetic variance (V_G) in virtual crosses (Bernardo, 2015), which was V_G = 367 cm^2 in A632 × ND203 and V_G = 0.43 cm^2 in F2 × ND203. These results suggest that targeted recombination may be most advantageous when little genetic variance is expected in the biparental cross. On the other hand, empirical validation of the usefulness of virtual populations is still needed (Bernardo, 2015), particularly because genomewide marker effects estimated from diverse germplasm panels might not readily apply to predicting the performance of progeny within a much narrower biparental cross (Massman et al., 2013a).

The $RE_{Targeted}$ values were always larger with two targeted recombinations (x = 2) per chromosome than with x = 1 (Table 1). In Experiment 1, the $RE_{Targeted}$ values increased from ranging between 212 and 403% with x = 1 to ranging between 254 and 631% with x = 2. In Experiment 2, the mean $RE_{Targeted}$ values increased from ranging between 189 and 214% with x = 1 to ranging between 232 and 274% with x = 2. The predicted performance with x = 2 in the B73 × Mo17 cross was 10,599 kg ha^{-1} for yield and 93 g kg^{-1} for moisture. It remains to be seen whether these predicted gains are realistic. Based on their field breeding experience, maize breeders may conclude that developing an inbred with a moisture content of 9.3% from the B73 × Mo17 cross is unrealistic. On the other hand, an inbred with the 20 specific targeted recombinations in the B73 × Mo17 genome has probably never been obtained either.

For each trait and chromosome in the intermated B73 × Mo17 population, x = 1 was always superior to having a parental chromosome passed intact from a parent to a doubled haploid offspring (no recombination). Likewise, x = 2 was always superior to x = 1. In contrast to the intermated B73 × Mo17 population, the 45 crosses in Experiment 2 had instances in which having no recombination was superior to x = 1. For example, out of the 450 combinations between the 10 chromosomes and 45 crosses, having no recombination on a specific chromosome was superior to x = 1 on that chromosome in 7 out of 450 cases (1.5%) for plant height and in 6 out of 450 cases (1.3%) for anthesis date. For plant height, chromosome 4 was involved in three out of the seven instances in which having no recombination was superior. These three instances involving chromosome 4 corresponded to A641 × PHH93, F2 × PH207, and F2 × PHG80. These three crosses involved a cross between a short and tall inbred; although F2 was the shortest inbred, PHH93, PH207, and PHG80 were not the tallest inbreds among the 10 lines. Even though having no recombination was sometimes superior to x = 1 in Experiment 2, x = 2 was always superior to having no recombination or to x = 1 in this experiment. This result indicated that when two parents differed widely in their performance, more than one targeted recombination per chromosome was sometimes needed to obtain a progeny line that outperformed the better parent in Experiment 2.

The $RE_{Targeted}$ values will decrease if $R_{Nontargeted}$ increases because of a higher selection differential. However, increases in $R_{Nontargeted}$ will be minor when selection is already stringent. To illustrate, the standardized selection differential with the best out of 180 inbreds selected in Experiment 1 was 2.710 (Becker, 1984). If the population size is doubled to 360 and the best inbred is selected, the standardized selection differential increases to 2.935. Increasing the population size by 100% will therefore increase the selection response by only (2.935– 2.710) ÷ 2.710 = 8%. The nonlinear relationship between the proportion selected and the standardized selection differential therefore limits the extent to which a more stringent selection process (with nontargeted recombination) will decrease the $RE_{Targeted}$ values.

Positions of Targeted Recombination

Different positions of recombination on the same chromosome typically led to peaks and valleys in the predicted response to selection for a given trait. Consider the contribution of chromosome 1 to yield in the intermated B73 × Mo17 population (Fig. 2). Yield contributed by chromosome 1 was predicted to be the highest (306 kg ha^{-1}) if a doubled haploid inherited the B73 genome at SNPs 1 to 143 and the Mo17 genome at SNPs 144 to 158. The responses for yield remained high in the marker intervals surrounding the SNP 143 to 144 peak: the yield contribution was 284 kg ha^{-1} for a targeted recombination in the SNP 139 to 140 interval, it increased to the peak in the SNP 143 to 144 interval, and it decreased to 282 kg ha^{-1} in the SNP 146 to 147 interval. Peaks and valleys were likewise observed for the other traits. The profile of per-chromosome responses attributed to targeted recombination was relatively smooth for yield but not for moisture in the intermated B73 × Mo17 population (Fig. 2). This result suggests that in a biparental cross, the favorable alleles from one parent may be concentrated in larger chromosome segments for one trait (yield) but in smaller chromosome segments for another trait (moisture).

The targeted recombination points (i.e., those which led to the largest gains) differed among the traits (Fig. 2). With x = 1 in the intermated B73 × Mo17 population, none of the 10 targeted recombinations for yield (indicated by green triangles in Fig. 2) were in the same locations as the targeted recombinations for moisture. For moisture and for the yield index (which included moisture as a component), the targeted recombination on chromosome 10 was the only one in common between the two traits. This result indicated that if

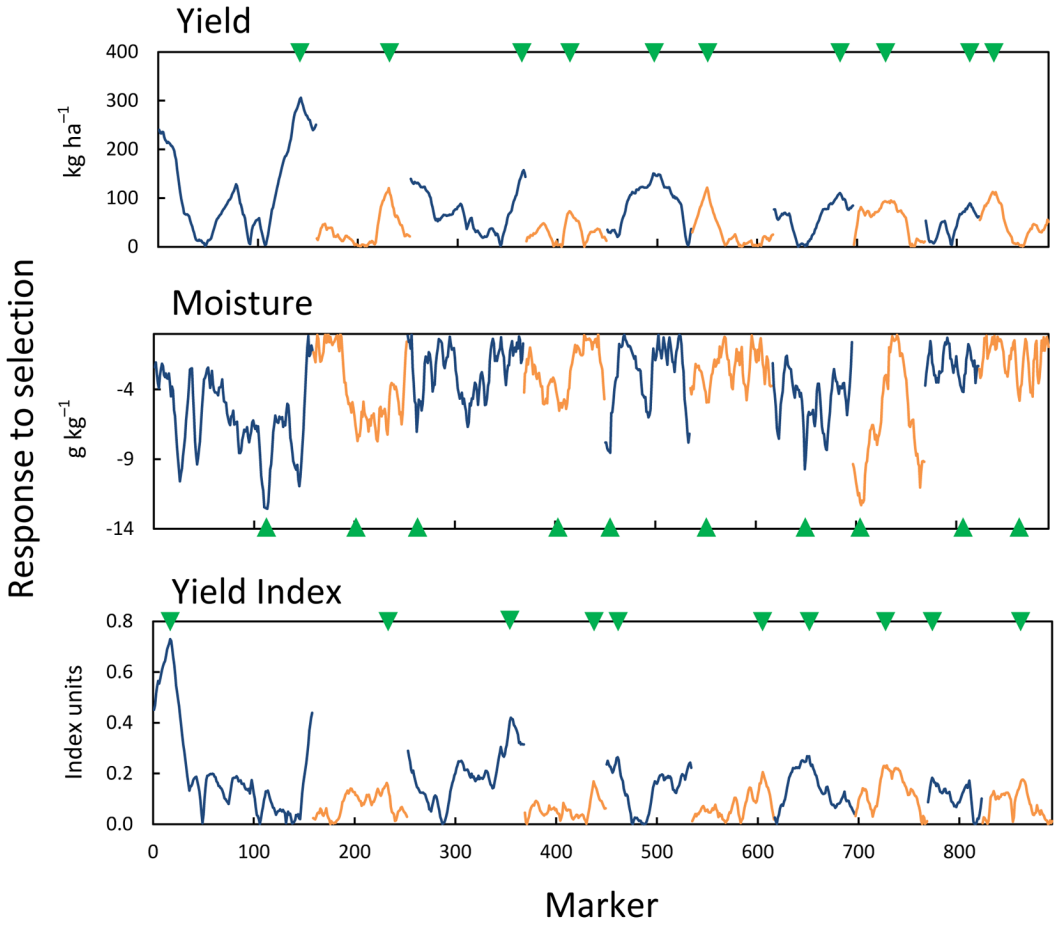

Fig. 2. Predicted changes in performance when recombination occurs at different marker intervals in the intermated B73 × Mo17 maize population. Different chromosomes are indicated by the alternating line colors. Targeted recombination points that led to the highest yield and yield index and the lowest moisture are shown by the green triangles.

targeted recombination becomes a reality in plants, a breeder would need to specify *a priori* how information on different traits is to be combined in a selection index. Currently, plant breeders use four approaches to distill information on multiple traits into a yes-or-no selection decision (Bernardo, 2010): (i) independent culling, in which minimum levels of performance are assigned for each trait; (ii) a formal selection index; (iii) an intuitive selection index (Hallauer et al., 1988); or (iv) some combination of these three aforementioned approaches. With nontargeted recombination, a breeder can wait until the progeny are developed before choosing a method to combine information on multiple traits. With targeted recombination, a formal selection index needs to be specified beforehand because different weights placed on each trait will result in different targeted recombination points among the progeny.

For the same trait, the targeted recombination points differed among crosses (Fig. 3). Consider the contribution of chromosome 1 to plant height when ND203 is crossed with five other inbreds, including those with the lowest RE_{Targeted} (105% for A632 × ND203) and the highest RE_{Targeted} (600% for F2 × ND203). With $x = 1$ and with ND203 as a common parent, the predicted plant height resulting from chromosome 1 was lowest at the following marker intervals: SNPs 1407 to 1408 with A641, SNPs 2885 to 2886 with A321, SNPs 2895 to 2896 with F2, SNPs 3964 to 3965 with LH82, and SNPs 4142 to 4143 with A632. These results indicated that generalizations regarding targeted recombination points cannot be made even among crosses that share a common parent.

The double recombination points were sometimes only several markers apart. For yield in the intermated B73 × Mo17 population, the two targeted recombinations on chromosome 7 were nine markers (19.6 cM) apart. For moisture, the two targeted recombinations on chromosome 10 were six markers (30.2 cM) apart. There often was correspondence between the targeted recombination point with $x = 1$ and one of the two recombination points with $x = 2$. For yield in the intermated B73 × Mo17 population, the targeted recombination point with $x = 1$ was the same as one of the two recombination points with $x = 2$ on chromosomes 2, 4, 5, 6, 7, 8, and 10. For the other traits in the intermated B73 × Mo17 population, the number of targeted recombination points common between $x = 1$ and $x = 2$ was five for the yield index, six for moisture, and eight each for plant height, lignin, and glucose release.

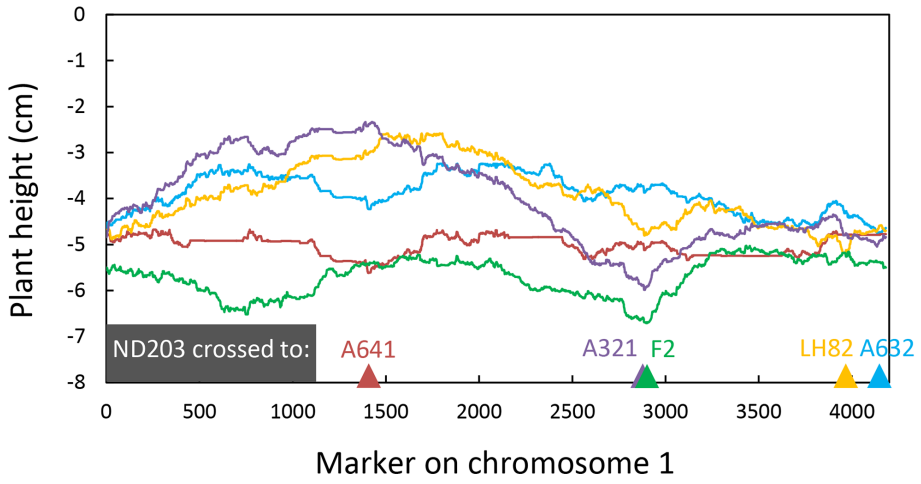

Fig. 3. Predicted contribution of chromosome 1 to maize plant height, when recombination occurs at different marker intervals, in crosses of ND203 to five other inbreds. Targeted recombination points for each cross are shown by the colored triangles beside the inbred names.

Caveats to Using Genomewide Marker Effects to Identify Targeted Recombination Points

Identifying targeted recombination points and predicting the resulting genetic gains are predicated on having good estimates of genomewide marker effects. Theoretical results (Daetwyler et al., 2008; Lian et al., 2014) and empirical studies (Lorenzana and Bernardo, 2009; Lian et al., 2014) have shown that prediction accuracy increases as the size of the training population, trait heritability, and number of markers increase. A larger number of markers can increase the resolution of the targeted recombination points, but increasing the number of SNP loci may lead to severe multicollinearity between adjacent markers and to poor estimates of marker effects. Furthermore, major quantitative trait loci (QTL) were not detected in either Experiment 1 (Lorenzana et al., 2010) or Experiment 2 (Schaefer and Bernardo, 2013). However, if major genes or QTL are known to exist in the population, a Bayesian model (Resende et al., 2012) or a fixed-effects model for major genes or QTL (Bernardo, 2014a) is probably preferable to the RR-BLUP procedure used in this study.

In phenotypic selection, realized responses are often smaller than the predicted responses because of genotype × environment interaction and error in the heritability estimates (Hallauer and Miranda, 1989). Likewise, the realized gains from targeted recombination may be less than the predicted gains because of genotype × environment interaction and error in the estimates of genomewide marker effects. Nevertheless, the large values of RE_{Targeted} in Table 1 suggest that the realized gains from targeted recombination will remain substantial.

The use of genomewide marker effects for identifying targeted recombination points relies on the natural, meiotic recombination that has occurred and that allowed estimation of the marker effects in the first place. In other words, the targeted recombination points identified from genomewide marker effects will only be in genomic regions where there is sufficient recombination. In contrast, the targeted recombination points will not cover any genomic regions where natural recombination is not observed. In this situation, targeted recombination would serve a different purpose: it would allow us to discover whether recombination cold spots harbor unfavorable linkages that need to be broken or favorable linkages that need to be preserved.

Targeted Recombination as a Breeding Tool for Quantitative Traits

Plant breeders can currently control much but not all of the breeding process. Breeders can choose which parents to cross and which breeding methods, tools, and environments to use in determining which progeny to keep. However, breeders have thus far been unable to control recombination among linked loci in a given cross. In this study, the RE_{Targeted} values for maize yield and other agronomic traits suggested that an ability to induce one or two targeted recombinations per chromosome may double the current selection gains. Developing targeted recombination technology therefore might be worthwhile, particularly given the concerns that the current rates of increase in crop productivity are not enough to meet the goal of doubling global crop production by 2050 (Food and Agriculture Organization, 2009; Ray et al., 2013).

Targeted recombination would lead to a most fundamental change in the breeding process, as it would shift breeding for quantitative traits from largely a numbers game to more of a design approach. Targeted recombination might conjure images of being able to cherry-pick, for example, 1000 favorable alleles and quickly assemble them into a single cultivar. The approach proposed herein is more realistic and attainable in that the number of targeted recombinations per chromosome (1–2) is the same number of nontargeted recombinations typically observed in maize (Smith et al., 2008; Sleper and Bernardo, 2016).

The Sadhu et al. (2016) CRISPR system for targeted recombination worked well in yeast. There is interest in extending the Sadhu et al. system from yeast to multicellular organisms (Offord, 2016), but the feasibility and efficiency of CRISPR-induced mitotic recombination in plants remains to be seen. Repair of double-strand breaks by homologous recombination is known to occur in plant somatic cells (Schuermann et al., 2005) but the efficiency of homologous repair differs among plant species and cell types (Puchta and Fauser, 2014). Even if multiplex CRISPR systems for inducing homologous recombination become feasible, protocols need to be developed to screen for cells that have the targeted recombinations and to regenerate such cells into whole plants. Furthermore, although targeted recombination needs to be initially induced, nontargeted recombination needs to be absent when doubled haploids are developed from the regenerated plants that have the targeted recombinations. The complexity of the task is acknowledged, and the development of workflows for targeted recombination will be a major undertaking among multidisciplinary teams.

A system for targeted recombination is likely to be imperfect at the outset, and targeted recombination on all of the chromosomes might not be possible. If so, a few chromosomes could be given higher priority than others for targeted recombination. For yield in the intermated B73 × Mo17 population, the per-chromosome contribution (with $x = 1$ targeted recombination per chromosome) was lowest for chromosome 4 (73 kg ha^{-1}) and was highest for chromosomes 1 (306 kg ha^{-1}), 3 (157 kg ha^{-1}), and 5 (150 kg ha^{-1}) (Fig. 2). Given the population mean of 8982 kg ha^{-1} (Table 1), targeted recombination only on chromosomes 1, 3, and 5 leads to a predicted yield of 9595 kg ha^{-1}, which is 99.8% of yield of the best line out of the 180 lines (9618 kg ha^{-1}, Table 1) in the intermated B73 × Mo17 population. Targeted recombination on chromosomes 1, 3, and 5, coupled with selection of the best nontargeted recombinants on the seven remaining chromosomes is likely to lead to yields that exceed that of the best line with nontargeted recombination on all 10 maize chromosomes.

Because the technology for targeted recombination in plants still needs to be developed, initial validation experiments are likely to rely on the use of traditional breeding procedures to develop lines with targeted recombination on one or a few chromosomes at a time. Among the 600 combinations between the six traits and 10 chromosomes in Experiment 1, a chromosome with the desired targeted recombination was observed only once. This lone instance was for glucose release and chromosome 5 of line M0131, which had the Mo17 allele at SNPs 450 to 524 and the B73 allele at SNPs 525 to 534. Lines with targeted recombination can be developed with methods akin to marker-assisted backcrossing (Bernardo, 2014b): the procedure would involve both foreground selection [for the chromosome(s) where targeted recombination is desired] and background selection (for a uniform genetic background across the remaining chromosomes). Selection gains with targeted recombination on individual chromosomes can then be pooled across the chromosomes to assess the realized gains.

References

Albrecht, T., V. Wimmer, H.J. Auinger, M. Erbe, C. Knaak, M. Ouzunova, et al. 2011. Genome-based prediction of testcross values in maize. Theor. Appl. Genet. 123:339–350. doi:10.1007/s00122-011-1587-7

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott, N.A. Tinker, and J.-L. Jannink. 2013. Comparison of genomic, marker-assisted, and pedigree-BLUP selection methods to increase β-glucan concentration in elite oat germplasm. Crop Sci. 53:1894–1906. doi:10.2135/cropsci2012.09.0526

Bastianelli, G., and A. Nicolas. 2015. Targeting of meiotic recombination. Poster 0088 presented at: Plant and Animal Genome Conf., San Diego, CA. 10–14 Jan. 2015.

Becker, W.A. 1984. Manual of quantitative genetics. Academic Enterprises, Pullman, WA.

Bernardo, R. 1991. Retrospective index weights used in multiple trait selection in a maize breeding program. Crop Sci. 31:1174–1179. doi:10.2135/cropsci1991.0011183X003100050020x

Bernardo, R. 2010. Breeding for quantitative traits in plants. 2nd ed. Stemma Press, Woodbury, MN.

Bernardo, R. 2014a. Genomewide selection when major genes are known. Crop Sci. 54:68–75. doi:10.2135/cropsci2013.05.0315

Bernardo, R. 2014b. Essentials of plant breeding. Stemma Press, Woodbury, MN.

Bernardo, R. 2015. Genomewide selection of parental inbreds: Classes of loci and virtual biparental populations. Crop Sci. 55:2586–2595.

Bernardo, R., and A.M. Thompson. 2016. Germplasm architecture revealed through chromosomal effects for quantitative traits in maize. Plant Genome 9. doi:10.3835/plantgenome2016.03.0028

Belhaj, A., A. Chaparro-Garcia, S. Kamoun, and V. Nekrasov. 2013. Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9:39. doi:10.1186/1746-4811-9-39

Beyene, Y., K. Semagn, S. Mugo, A. Tarekegne, R. Babu, B. Meisel, et al. 2015. Genetic gains in grain yield through genomic selection in eight biparental maize populations under drought stress. Crop Sci. 55:154–163. doi:10.2135/cropsci2014.07.0460

Bortesi, L., and R. Fischer. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Adv. 33:41–52. doi:10.1016/j.biotechadv.2014.12.006

Combs, E., and R. Bernardo. 2013. Genomewide selection to introgress semidwarf corn germplasm into U.S. Corn Belt inbreds. Crop Sci. 53:1427–1436. doi:10.2135/cropsci2012.11.0666

Daetwyler, H.D., B. Villanueva, and J.A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PLoS ONE 3:e3395. doi:10.1371/journal.pone.0003395

Dekkers, J.C.M. 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. J. Anim. Breed. Genet. 124:331–341. doi:10.1111/j.1439-0388.2007.00701.x

Dien, B.S., H.-J.G. Jung, K.P. Vogel, M.D. Casler, J.F.S. Lamb, L. Iten, et al. 2006. Chemical composition and response to dilute acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. Biomass Bioenergy 30:880–891. doi:10.1016/j.biombioe.2006.02.004

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250–255. doi:10.3835/plantgenome2011.08.0024

Food and Agriculture Organization. 2009. High level expert forum—How to Feed the World in 2050, Rome, Italy. 12–13 Oct. 2009.

Gaj, T., C.A. Gersbach, and C.F. Barbas. 2013. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31:397–405. doi:10.1016/j.tibtech.2013.04.004

Hallauer, A.R., and J.B. Miranda. 1989. Quantitative genetics in maize breeding. 2nd ed. Iowa State Univ. Press, Ames, IA.

Hallauer, A.R., W.A. Russell, and K.R. Lamkey. 1988. Corn breeding. In: G.F. Sprague and J.W. Dudley, editors, Corn and Corn Improvement. 3rd ed. CSSA, Madison, WI. p. 463–564.

Jacobson, A., L. Lian, S. Zhong, and R. Bernardo. 2014. General combining ability model for genomewide selection in a biparental cross. Crop Sci. 54:895–905. doi:10.2135/cropsci2013.11.0774

Lee, M., N. Sharopova, W.D. Beavis, D. Grant, M. Katt, D. Blair, et al. 2002. Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol. Biol. 48:453–461. doi:10.1023/A:1014893521186

Lewis, M.F., R.E. Lorenzana, H.G. Jung, and R. Bernardo. 2010. Potential for simultaneous improvement of corn grain yield and stover quality for cellulosic ethanol. Crop Sci. 50:516–523. doi:10.2135/cropsci2009.03.0148

Lian, L., A. Jacobson, S. Zhong, and R. Bernardo. 2014. Genomewide prediction accuracy within 969 maize biparental populations. Crop Sci. 54:1514–1522. doi:10.2135/cropsci2013.12.0856

Liu, S., H.D. Chen, I. Makarevitch, R. Shirmer, S.J. Emrich, C.R. Dietrich, et al. 2010. High-throughput genetic mapping of mutants via quantitative single nucleotide polymorphism typing. Genetics 184:19–26. doi:10.1534/genetics.109.107557

Lorenzana, R., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor. Appl. Genet. 120:151–161. doi:10.1007/s00122-009-1166-3

Lorenzana, R.E., M.F. Lewis, H.-J.G. Jung, and R. Bernardo. 2010. Quantitative trait loci and trait correlations for maize stover cell wall composition and glucose release for cellulosic ethanol. Crop Sci. 50:541–555. doi:10.2135/cropsci2009.04.0182

Massman, J.M., A. Gordillo, R.E. Lorenzana, and R. Bernardo. 2013a. Genomewide predictions from maize single-cross data. Theor. Appl. Genet. 126:13–22. doi:10.1007/s00122-012-1955-y

Massman, J.M., H.-J.G. Jung, and R. Bernardo. 2013b. Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. Crop Sci. 53:58–66. doi:10.2135/cropsci2012.02.0112

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Offord, C. 2016. Mapping traits to genes with CRISPR. The Scientist. http://www.the-scientist.com/?articles.view/articleNo/46029/title/Mapping-Traits-to-Genes-with-CRISPR/ (accessed 7 Apr. 2017).

Pennisi, E. 2013. The CRISPR craze. Science 341:833–836. doi:10.1126/science.341.6148.833

Puchta, H., and F. Fauser. 2014. Synthetic nucleases for genome engineering in plants: Prospects for a bright future. Plant J. 78:727–741. doi:10.1111/tpj.12338

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield trends are insufficient to double global crop production by 2050. PLoS ONE 8: E66428. doi:10.1371/journal.pone.0066428.

Resende, M.F.R., Jr., P. Muñoz, M.D.V. Resende, D.J. Garrick, R.L. Fernando, J.M. Davis, et al. 2012. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). Genetics 188:695–708.

Rutkoski, J., R.P. Singh, J. Huerta-Espino, S. Bhavani, J. Poland, J.-L. Jannink, et al. 2015. Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. Plant Genome 8. doi:10.3835/plantgenome2014.10.0074

Sadhu, M.J., J.S. Bloom, L. Day, and L. Kruglyak. 2016. CRISPR-directed mitotic recombination enables genetic mapping without crosses. Science 352:1113–1116. doi:10.1126/science.aaf5124

Schuermann, D., J. Molinier, O. Fritsch, and B. Hohm. 2005. The dual nature of homologous recombination in plants. Trends Genet. 21:172–181. doi:10.1016/j.tig.2005.01.002

Schaefer, C.M., and R. Bernardo. 2013. Genomewide association mapping of flowering time, kernel composition, and disease resistance in historical Minnesota maize inbreds. Crop Sci. 53:2518–2529. doi:10.2135/cropsci2013.02.0121

Senior, M.L., E.C.L. Chin, M. Lee, J.S.C. Smith, and C.W. Stuber. 1996. Simple sequence repeat markers developed from maize sequences found in the GENBANK database: Map construction. Crop Sci. 36:1676–1683. doi:10.2135/cropsci1996.0011183X003600060043x

Sleper, J.A., and R. Bernardo. 2016. Recombination and genetic variance among maize doubled haploids induced from F_1 and F_2 plants. Theor. Appl. Genet. doi:10.1007/s00122-016-2781-4

Smith, J.S.C., T. Hussain, E.S. Jones, G. Graham, D. Podlich, S. Wall, et al. 2008. Use of doubled haploids in maize breeding: Implications for intellectual property protection and genetic diversity in hybrid crops. Mol. Breed. 22:51–59. doi:10.1007/s11032-007-9155-1