Breeding Inducer Development

Different approaches that can be used for maize maternal haploid inducer development are explained in this entry. For a complete review on maternal haploid inducers, see: https://www.mdpi.com/2223-7747/9/5/614/htm (https://www.mdpi.com/2223-7747/9/5/614/htm)

1. Introduction

Independent of the breeding strategy, a certain sequence of selection steps is recommended in haploid inducer development. To describe a typical breeding approach for inducer development, we propose the scenario of a tropical breeding program interested in developing its own, adapted inducer from a licensed, exotic inducer. The exotic inducer is fixed for *R1-nj*, *Pl1*, *mtl* and *zmdmp*, but has poor agronomic performance and environmental adaptation. This inducer will be crossed to a set of elite inbred lines which possess good agronomic performance and environmental adaptation, but lack the desirable alleles at these four genes (Table 1). Since genes that hamper *R1-nj* expression occur at a higher frequency in tropical germplasm [1], employing a set of adapted inbreds, rather than only one, reduces the chance of developing inducers with poor *R1-nj* pigmentation. Segregation for *R1-nj* will be observed in F_2 seeds, where seeds displaying anthocyanin pigmentation in the scutellum and aleurone layer have at least one copy of this allele and can be discriminated from homozygous recessive seeds, which are uncolored. From our experience, it takes approximately 15 minutes for a person to discriminate an ear with 400 seeds using the *R1-nj* phenotype. Assuming a wage of $12.00 per hour, the cost per seed selected would be $0.0075. If available, a color sorter machine color can be used to rapidly separate colored from uncolored seed at a very low cost. If early fixation of *R1-nj* is a goal, then MAS can be applied on pre-selected F_2 seeds. The price to outsource MAS with a single simple sequence repeat (SSR) marker is roughly U$1 per plant [2][3]. DNA extraction accounts for approximately 16% of this cost [3], which reduces part of the cost per additional marker employed. In well-stablished breeding companies, in house genotyping costs are usually around $0.10 per data point per genotype. If seed chipping and high-density genotyping are available at a low cost, then MAS or GS can be applied to reduce the number of F_2 individuals to be tested in the field.

Table 1. Traits of the exotic inducer and noninducer parents assumed for the comparison of the different approaches for inducer development.

Exotic Inducer	Adapted Inbred Lines
R1-nj (purple embryo marker)	*r1-nj*
Pl1 (red root marker)	*pl1*
mtl	*MTL*
zmdmp	*ZmDMP*
Poor agronomic performance	Good agronomic performance
Poor environmental adaptation	Good environmental adaptation

F$_2$ plants that were self-pollinated and are homozygous for *R1-nj* can be easily identified and selected at the moment of harvest, since all of their seeds should show the typical anthocyanin pigmentation produced by this allele. The red root marker, conditioned by the dominant *Pl1* allele, can also be visually selected. However, its selection is more laborious than the purple embryo marker, since seeds need to be first germinated in small trays and, after selection, transplanted to either bigger containers or the field. Employing markers to select for this allele allows direct sowing in the field, simplifying operations and reducing costs. The presence of the *mtl* and *zmdmp* alleles, which disrupts maternal haploid induction [4] and increase the HIR of inducers fixed for *mtl* [5], respectively, can be verified by progeny testing. However, MAS is likely the easiest way to select for the presence of these alleles. Because strong gametic and zygotic selection against *mtl* occur [6][7][8], self-pollinating genotypes heterozygous at this locus will not produce the expected frequency of homozygous recessive individuals. From our experience, only 12–15% of F$_2$ plants are fixed for the *mtl* allele [9]. Therefore, the rapid fixation of the *mtl* allele is desirable to simplify the selection process.

Just a small fraction of the F$_2$ plants (~0.4%) will be fixed for *R1-nj*, *Pl1*, *mtl* and *zmdmp*. An extremely large population of F$_2$ plants would have to be sown if a reasonable number genotypes fixed at these loci were to be selected. Since *mtl* is required for haploid induction and *R1-nj* can be inexpensively selected, MAS in the F$_2$ generation can be used to ensure the fixation of *mtl* and elimination of genotypes homozygous for the *pl1* or *ZmDMP* alleles. This will increase the chance of obtaining F$_3$ genotypes fixed for *R1-nj*, *Pl1* and *zmdmp*. In this generation, phenotypic selection can be used to fix *R1-nj*, while MAS can used to fix *Pl1* and *zmdmp*. Table 2 lays out a breeding scheme for inducer development.

Table 2. Selection steps in haploid inducer development.

Generation	Selection Applied	Resulting Breeding Lines
Parental	Selection of elite inbred lines for crossing with exotic inducer	F$_1$
F$_1$	Discard of F$_1$ families with undesirable characteristics	F$_2$
F$_2$	Discard F$_2$ seeds lacking the purple embryo pigmentation. With MAS, fix *mtl* and discard *pl1/pl1* or *ZmDMP/ZmDMP* genotypes.	F$_3$ fixed for *mtl*
F$_3$	Fix *R1-nj* by harvesting ears of F$_2$ plants that only have colored seeds. Fix *Pl1* and *zmdmp* with MAS. The HIR of the selected genotypes can be evaluated through crosses with one donor	F$_4$ fixed for *mtl*, *R1-nj*, *Pl1* and *zmdmp*
F$_{4,5,6...}$	Phenotypic and/or genotypic selection for polygenic traits of importance to inducers	F$_{5,6,7...}$

Plant height, lodging tolerance, tassel size, pollen production, length of pollen shed, HIR, seed set and tolerance to ear rots are traits that affect inducer performance and are under polygenic control. Different breeding methods, such as phenotypic selection, MAS, and more recently, genomic selection, have been used to improve quantitatively inherited traits. Genomic selection requires genotyping and phenotyping a panel of individuals to establish a prediction model with which to estimate the performance of individuals that were only genotyped [10]. Despite requiring an additional initial effort for model construction, its value for the improvement of quantitively inherited characteristics has already been demonstrated [11][12][13][14], and is now widely adopted in commercial breeding programs. Employing genomic selection for the simultaneous improvement of all polygenic traits important to inducers might be economically interesting, since the cost associated with phenotyping a large sample of individuals is presumably high. In addition, genotyping costs continue to decrease [15][16].

The value of genomic prediction for the improvement of traits important to haploid inducers was already assessed, and moderate to high prediction accuracies were observed in most of them [11]. For polygenic traits affected by large effect QTL, such as plant height, flowering time and HIR, a combination of *de novo* GWAS and GS might result in higher prediction accuracies than those observed in GS [17]. In the *de novo* GWAS + GS approach, markers with lowest p-values are treated as fixed effects in the prediction model, whereas in GS, all markers are treated as random effects [17].

Evaluating HIR is time-consuming and labor-intensive, since usually many cross-pollinations are performed to produce enough seeds with which to obtain a reliable estimate of a genotype's induction ability. Visual discrimination of haploid and diploid seeds is also very time-consuming, besides being quite susceptible to human errors, which affect the quality of the data collected. For instance, multiple factors can affect *R1-nj* expression and its visibility in the scutellum, such as environmental conditions [18][19][20][21], seed morphology and the presence of inhibitors genes in the donor germplasm [22][23][24]. These factors complicate haploid discrimination, which when coupled with a natural variation in selection stringency among evaluators, might affect the reliability of the estimates of HIR. Therefore, when evaluating the HIR of different inducers for the purpose of selection, a good balance needs to be found between sample size and the number of people involved in this task. Because of the high cost, labor-intensive and complicated task of obtaining reliable estimates of HIR, genomic selection might be helpful to improve this trait.

Tandem selection, independent culling, and index selection are common approaches used for the selection of multiple traits. Tandem selection proposes improving one trait at a time until the desired level for a given trait is achieved. This strategy is useful when selection for different traits occurs during different stages of a breeding cycle. For instance, in the first generations of an inducer breeding cycle (e.g., F_2–F_3), selection for qualitative traits, such as *R1-nj*, *Pl1* and *mtl*, or traits with moderate to high heritability, such as flowering time and plant height, can be effectively performed. At later stages of a breeding cycle (e.g., after F_4) selection for low heritable traits, such as yield and resistance to some diseases, is more effective.

2. Strategy

The strategy of independent culling levels proposes establishment of thresholds for each trait, and only genotypes that reach those thresholds are selected. For instance, if a breeder decides that an inducer must produce at least of 200 seeds in cross-pollinations and that it must have an HIR above 5.0% to ensure the economic viability of the DH technique, then all inducers that do not meet both criteria are discarded. The disadvantage of this selection strategy is that genotypes with very good performance for one trait, but with a below-threshold performance for another, are discarded, even though they might still have value as a source of favorable alleles for a trait.

A way to avoid the strictness caused by thresholds, and to allow compensation among traits, is to employ index selection. In index selection, the phenotypic value that each genotype has for a trait is multiplied by the weight given for this trait. All weighted phenotypic values are accumulated in the index value for that genotype. The weight might be the economic value for that trait, or a value empirically determined by the breeder. Different kinds of selection indices exist, and their usefulness is context dependent. A challenge associated with selection indices is to determine appropriate weights for each trait, since they depend on the unit of measure of each trait and because they might not be stable (e.g., prices might change from season to season).

For inducer development, all methods are used and embedded in a flexible tandem selection scheme (Table 2). In the F_2 and F_3 generations, phenotypic or MAS are used to fix *R-nj*, *Pl1*, *mtl* and *zmdmp*. Then, culling is used to discard F_3 families with very low HIRs. Finally, index selection is used to identify F_4 or later generation families having the best combination of HIR and various agronomic traits. In particular for this last stage of selection, GS is a promising alternative [11], since prediction accuracies for HIR and the other traits important to inducers were found to be high, and their phenotypic evaluation is laborious and thus costly.

References

1. Chaikam, V.; Nair, S.K.; Babu, R.; Martinez, L.; Tejomurtula, J.; Boddupalli, P.M. Analysis of effectiveness of R1-nj anthocyanin marker for in vivo haploid identification in maize and molecular markers for predicting the inhibition of R1-nj expression. Theor. Appl. Genet. 2015, 128, 159–171.

2. Dreher, K.; Morris, M.; Khairallah, M.; Ribaut, J.-M.; Pandey, S.; Srinivasan, G. Is marker-assisted selection cost-effective compared to conventional plant breeding methods? The case of quality protein maize. In Proceedings of the 4th Annual Conference of the International Consortium on Agricultural Biotechnology Research (ICABR'00), Ravello, Italy, 24–28 August 2000; pp. 203–236.

3. Fox, R.L.; Hayden, M.J.; Mekuria, G.; Eglinton, J.K. Cost analysis of molecular techniques for marker assisted selection within a barley breeding program. In Proceedings of the 12th Australian Annual Barley Technical Symposium, Hobart, TAS, Australia, 12–14 September 2005.

4. Kelliher, T.; Starr, D.; Richbourg, L.; Chintamanani, S.; Delzer, B.; Nuccio, M.L.; Green, J.; Chen, Z.; McCuiston, J.; Wang, W.; et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature 2017, 542, 105–109.

5. Zhong, Y.; Liu, C.; Qi, X.; Jiao, Y.; Wang, D.; Wang, Y.; Liu, Z.; Chen, C.; Chen, B.; Tian, X.; et al. Mutation of ZmDMP enhances haploid induction in maize. Nat. Plants 2019, 5, 575–580.

6. Prigge, V.; Xu, X.; Li, L.; Babu, R.; Chen, S.; Atlin, G.N.; Melchinger, A.E. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics 2012, 190, 781–793.

7. Xu, X.; Li, L.; Dong, X.; Jin, W.; Melchinger, A.E.; Chen, S. Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J. Exp. Bot. 2013, 64, 1083–1096.

8. Barret, P.; Brinkmann, M.; Beckert, M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. Theor. Appl. Genet. 2008, 117, 581–594.

9. Frei, U.K. Personal Communication; Iowa State University: Ames, IA, USA, 2015.

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

11. Almeida, V.C.; Trentin, H.U.; Frei, U.K.; Lübberstedt, T. Genomic prediction of maternal haploid induction rate in maize. Plant Genome 2020, e20014.

12. Asoro, F.G.; Newell, M.A.; Beavis, W.D.; Scott, M.P.; Jannink, J.-L. Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. Plant Genome J. 2011, 4, 132.

13. Heffner, E.L.; Jannink, J.; Sorrells, M.E. Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. Plant Genome 2011, 4, 65–75.

14. Dos Santos, J.P.R.; Pires, L.P.M.; de Castro Vasconcellos, R.C.; Pereira, G.S.; Von Pinho, R.G.; Balestre, M. Genomic selection to resistance to Stenocarpella maydis in maize lines using DArTseq markers. BMC Genet. 2016, 17, 1–10.

15. Eathington, S.R.; Crosbie, T.M.; Edwards, M.D.; Reiter, R.S.; Bull, J.K. Molecular markers in a commercial breeding program. Crop Sci. 2007, 47, S154–S163.

16. Scheben, A.; Batley, J.; Edwards, D. Genotyping-by-sequencing approaches to characterize crop genomes: Choosing the right tool for the right application. Plant Biotechnol. J. 2017, 15, 149–161.

17. Spindel, J.E.; Begum, H.; Akdemir, D.; Collard, B.; Redoña, E.; Jannink, J.L.; McCouch, S. Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. Heredity 2016, 116, 395–408.

18. Röber, F.K.; Gordillo, G.A.; Geiger, H.H. In vivo haploid induction in maize-Performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 2005, 50, 275–283.

19. Kebede, A.Z.; Dhillon, B.S.; Schipprack, W.; Araus, J.L.; Bänziger, M.; Semagn, K.; Alvarado, G.; Melchinger, A.E. Effect of source germplasm and season on the in vivo haploid induction rate in tropical maize. Euphytica 2011, 180, 219–226.

20. Prigge, V.; Schipprack, W.; Mahuku, G.; Atlin, G.N.; Melchinger, A.E. Development of in vivo haploid inducers for tropical maize breeding programs. Euphytica 2012, 185, 481–490.

21. Prigge, V.; Sánchez, C.; Dhillon, B.S.; Schipprack, W.; Araus, J.L.; Bänziger, M.; Melchinger, A.E. Doubled haploids in tropical maize: I. effects of inducers and source germplasm on in vivo haploid induction rates. Crop Sci. 2011, 51, 1498–1506.

22. Burr, F.A.; Burr, B.; Scheffler, B.E.; Blewitt, M.; Wienand, U.; Matz, E.C. The maize repressor-like gene intensifier1 shares homology with the r1/b1 multigene family of transcription factors and exhibits missplicing. Plant Cell 1996, 8, 1249–1259.

23. Della Vedova, C.B.; Lorbiecke, R.; Kirsch, H.; Schulte, M.B.; Scheets, K.; Borchert, L.M.; Scheffler, B.E.; Wienand, U.; Cone, K.C.; Birchler, J.A. The dominant inhibitory chalcone synthase allele C2-Idf (Inhibitor diffuse) from Zea mays (L.) acts via an endogenous RNA silencing mechanism. Genetics 2005, 170, 1989–2002.

24. Paz-Ares, J.; Ghosal, D.; Saedler, H. Molecular analysis of the C1-I allele from Zea mays: A dominant mutant of the regulatory C1 locus. EMBO J. 1990, 9, 315–321.