RESEARCH

# Controlling Misclassification Rates in Identification of Haploid Seeds from Induction Crosses in Maize with High-Oil Inducers

Albrecht E. Melchinger,* Markus Winter, Xuefei Mi, Hans-Peter Piepho, Wolfgang Schipprack, and Vilson Mirdita


## ABSTRACT

In vivo production of double haploid (DH) lines in maize (*Zea mays* L.) requires reliable identification of haploid (*H*) seeds. A new method for achieving this goal is production of induction crosses with high-oil (HO) inducers and sorting the resulting *H* and diploid crossing (*C*) seeds based on their oil content (OC). Balancing the false discovery rate (FDR) and false negative rate (FNR) by choice of a suitable proportion $\alpha$ of selected seeds represents an unsolved problem with this method. We investigated solutions by applying mixture distribution (MD) analysis to the OC of induction crosses for estimating the means and standard deviation ($\mu_H$, $\mu_C$, and $\sigma$) of *H* and *C* seeds and the haploid induction rate $\kappa$. Moreover, we developed formulas and software for calculating the FDR and FNR from these estimates. Using several induction crosses with HO inducer UH600, parameter estimates from (i) MD analysis in different environments and (ii) gold standard classification (GSC) of plants in the field agreed well for $\mu_H$ and $\mu_C$ , but only moderately for $\sigma$ and $\kappa$. Parameter estimates from the MD provided meaningful guidelines for calculating the expected FDR and FNR. Selecting the $\alpha$ = 7.5% proportion of seeds with lowest OC was optimal for most induction crosses and balanced the FDR and FNR. In conclusion, induction crosses with HO inducers hold great promise for promoting the DH technology in maize, but an automated high-throughput platform for sorting the seeds from the MD into several distinct classes with increasing OC is recommended to take full advantage of this novel approach.


A.E. Melchinger, M. Winter, X. Mi, W. Schipprack, and V. Mirdita, Institute of Plant Breeding, Seed Science and Population Genetics, Univ. of Hohenheim; H.P. Piepho, Bioinformatics Unit, Institute of Crop Science, Univ. of Hohenheim, D-70593 Stuttgart, Germany. Received 19 Sept. 2014. Accepted 11 Nov. 2014. *Corresponding author (melchinger@uni-hohenheim.de).

**Abbreviations:** AIC, Akaike information criterion; *C,* diploid crossing; CDF, computable document format; DH, double haploid; FDR, false discovery rate; FNR, false negative rate; GR, germination rate; GSC, gold standard classification; *H*, haploid; HO, high oil; MD, mixture distribution; OC, oil content; PC, putative diploid crossing; PDF, probability density function; PH, putative haploid..

In vivo haploid induction has become a routine tool in maize breeding for rapid production of homozygous lines. If source germplasm to be used for breeding of lines is pollinated with pollen from so-called inducer genotypes, this sporadically triggers the formation of *H* seeds that inherited their genome exclusively from the maternal parent (Prigge and Melchinger, 2012). Identifying the small fraction (about 10%) of *H* seeds in the vast number of normal *C* seeds produced in induction crosses with modern inducers remains a challenge for maize breeders. All inducers currently available in the public domain are equipped with the *R1-nj* marker gene that enables discrimination of *H* and *C* seeds on the basis of white or purple coloration of the embryo, respectively. This system is very labor intensive and not amenable to automation owing to the varying shape of maize seeds and the variable expression of the *R1-nj* marker. Moreover, the *R1-nj* marker cannot be applied in many germplasm of tropical and temperate origin (V. Chaikam, personal communication, 2014;

Published in Crop Sci. 55:1–11 (2015).
doi: 10.2135/cropsci2014.09.0633

W. Schipprack, unpublished data, 2014) because inhibitor genes suppress the expression of the *R1-nj* marker gene.

Recently, Melchinger et al. (2013) proposed the use of inducers with HO content of seeds for discriminating *H* seeds with low OC from *C* seeds with high OC and provided a proof-of-concept. With this novel approach, automated sorting of seeds could be achieved by high-throughput platforms on the basis of precise weighing of individual seeds and nondestructive measurement of their oil mass by nuclear magnetic resonance. In a study with 10 source germplasm of different type and origin, Melchinger et al. (2014) reported generally lower misclassification rates for the HO seed marker system than the *R1-nj* embryo marker system. Moreover, the user can, in principle, control the rate of false positives, because OC measurements display a quantitative distribution. However, applying thresholds for OC of seeds from induction crosses calculated on the basis of the OC of their parents did not prove successful in discrimination of *H* and *C* seeds (Melchinger et al., 2014).

The seeds in an induction cross represent a mixture of *H* and *C* seeds with an unknown mixture ratio (Melchinger et al., 2013). Hence, it seems promising to apply well-known statistical methods (cf. Frühwirth-Schnatter, 2006) for analyzing the mixture distribution of the OC measurements of all seeds. For example, the available methods could be used to estimate the proportion of *H* seeds, which corresponds to the haploid induction rate as well as the means and standard deviation of OC for the unknown distributions of the *H* and *C* seeds. Estimates of these parameters could be used to calculate integrals needed to determine the theoretically expected misclassification rates for any proportion of putative haploid (PH) seeds selected. This novel approach would hold great promise in practice, but information is needed whether the parameters estimated from the analysis of the mixture distribution can be used directly or must be modified to obtain meaningful guidelines for specifying the proportion PH seeds that comply with a given false discovery rate.

The objectives of our study were to (i) estimate for induction crosses produced in different environments the haploid induction rate $\kappa$ and the means ($\mu_H$, $\mu_C$) and standard deviation ($\sigma_H$, $\sigma_C$) of OC for *H* and *C* seeds on the basis of their MD, (ii) compare these parameter estimates with those obtained from a GSC of plants grown in the field from selected seeds, (iii) detail a mathematical approach and a software for controlling the misclassification rates in sorting of PH seeds with low OC from PC seeds with high OC, and (iv) contrast expected and observed misclassification rates for different proportions of selected PH seeds in a larger number of induction crosses.

# MATERIALS AND METHODS

## Plant Materials, Trait Measurements, and Classification of Seeds

### Experiment 1

This experiment was described in detail in a previous paper (Melchinger et al., 2014) and the seeds from the same induction crosses were reanalyzed here to compare the estimates of parameters $\mu_H$, $\mu_C$, $\sigma_H$, $\sigma_C$, and $\kappa$ obtained from analysis of (i) the MD of OC of all seeds and (ii) GSC of a subset of seeds reported before. Briefly, 10 source germplasm comprising six single crosses, two synthetics, and two landraces of European maize were pollinated in summer 2012 in the nursery with pollen from HO inducer UH600 (https://plant-breeding.uni-hohenheim.de/84531). Individual seeds from the bulk of each induction cross were weighed individually with a ZER-FS Weigh Cell (Wipotec GmbH) high-precision balance and measured for oil mass by nuclear magnetic resonance with a Minispec instrument according to manufacturer's instruction (BRUKER BioSpin GmbH). Subsequently, the seeds from each induction cross were classified according to their OC (in %, calculated from ratio of oil mass and seed weight) in two steps. First, seeds with OC < 2.1% were deemed embryoless with high reliability (Melchinger et al., 2014). Second, a threshold *t* individually chosen for each induction cross, on the basis of prior information about the OC of the source germplasm and inducer UH600, was used to classify the seeds into PH seeds with $2.1 \leq OC \leq t$ and PC seeds with $OC > t$.

From each induction cross, randomly chosen subsets of 300 PH seeds and 300 PC seeds classified in the previous step were grown in a field trial conducted in summer 2013 at the Agricultural Research Station, Eckartsweier, close to Offenburg in the Upper Rhine valley, Germany. In a so-called GSC, we determine the true nature of the seeds by phenotypic evaluation of the respective plants at flowering time as described by Dong et al. (2013). Compared with the *C* phenotype, the *H* phenotype is characterized by a shorter stature, erect and narrow leaves, as well as reduced growth and male fertility. Moreover, *C* plants displayed anthocyanin coloration of the stem in most of the induction crosses.

### Experiment 2A

This experiment served for comparing the OC of seeds from the same induction crosses produced in different environments. Fourteen elite inbred lines, comprising six flint and eight dent lines from the maize breeding program of the University of Hohenheim, were grown in summer 2013 in single-row plots with 20 plants in two environments: a breeding nursery in Eckartsweier and a plastic greenhouse in Hohenheim, Stuttgart, Germany. At flowering time, these plants were emasculated and pollinated with HO inducer UH600 grown in adjacent rows. Hand-pollination and open-pollination were practiced in the nursery and greenhouse, respectively. After harvest and shelling of ears, individual seeds from the bulk of each induction cross were weighed and measured for oil mass as described above to determine the OC. Seeds with OC < 2.1% were again classified as embryoless and culled. The remaining seeds from each induction cross in each environment were subject to a mixture distribution analysis of their OC.

### Experiment 2B

The goal of this experiment was to determine by GSC the haploid induction rate κ and the misclassification rates of the 14 induction crosses produced in the greenhouse in Experiment 2A. Starting with $N$ seeds with OC $\geq$ 2.1% from an induction cross, the 18% lower quantile ($Q_{<18\%}$) of the mixture distribution for OC was sorted in $n_{18\%}$ bags, each containing 20 seeds, where $n_{18\%}$ is the largest integer $\leq 0.18\ N$. The first bag contained the PH seeds with lowest OC and the $n\text{th}_{18\%}$ bag contained the PH seeds from $Q_{<18\%}$ with highest OC. A random sample of 200 seeds from the remaining upper 82% quantile ($Q_{>18\%}$) was also sorted in ascending order of OC in 10 bags of 20 seeds so that the total number of bags sown from each induction cross was $n_{18\%} + 10$.

These seeds were planted for GSC in a field trial conducted at the Agricultural Research Station, Eckartsweier, in summer 2014. The 20 seeds in each bag were planted in single-row plots. Plots were arranged according to a split-plot design, with the source material comprising the main plots and individual bags corresponding to the subplots. Besides recording the number of plants that developed from germinated seeds in each plot ($S_i$), all plants were visually scored at flowering time according to the GSC procedure described above for determining the number of genuine $H$ plants ($H_i$) in each plot $i \in \{1,2,\ldots n_{18} + 10\}$.

## Statistical Analyses

### Parameter Estimation

For estimating the five parameters $\mu_H$, $\mu_C$, $\sigma_H$, $\sigma_C$, and κ from the OC measurements of seeds of an induction cross, we assumed that the MD follows a probability density function (PDF) $f(x)$ with the following:

$$f(x) = \kappa\, \phi_H(x) + (1 - \kappa)\, \phi_C(x) \qquad [1]$$

where, $\phi_H(x)$ and $\phi_C(x)$ are Gaussian normal PDFs with $N\left(\mu_H, \sigma_H^2\right)$ and $N\left(\mu_C, \sigma_C^2\right)$ for OC of $H$ and $C$ seeds, respectively. Maximum likelihood estimates of the parameters, subsequently denoted by a hat, were obtained on the basis of the measured OC of all seeds with OC $\geq$ 2.1%. The models were fitted using the NLMIXED procedure of the SAS System (SAS Institute, 2008). Standard errors of all parameter estimates were determined by the delta method (Cox, 1998). Heterogeneity of $\sigma_H^2$ and $\sigma_C^2$ were assessed by the Akaike information criterion (AIC) in a comparison of the full and reduced model (Burnham and Anderson, 2002). Since the AIC was mostly lower and parameter estimates more robust for the model with homogenous variances, all further analyses and results reported were restricted to parameter estimates from the reduced model assuming $\sigma^2 = \sigma_H^2 = \sigma_C^2$.

Estimates of $\mu_H$, $\mu_C$, $\sigma_H$, $\sigma_C$, and κ from the GSC, subsequently denoted by a tilde, were for Experiment 1 taken from the publication Melchinger et al. (2014). These authors estimated the parameters from the observed OC of the genuine $H$ and $C$ seeds in the subsets of 300 PH and 300 PC seeds with adjustments accounting for the fact that the numbers of PH and PC seeds in the subsamples planted in the field trial were not proportional to their ratio in the original sample.

In Experiment 2B, we estimated the haploid induction rate $\tilde{\kappa}$ from the GSC as follows:

$$\tilde{\kappa} = \frac{\left(\sum_{i=1}^{n_{18}} H_i + \frac{0.82}{200} N \sum_{i=n_{18}+1}^{n_{18}+10} H_i\right)}{\left(\sum_{i=1}^{n_{18}} S_i + \frac{0.82}{200} N \sum_{i=n_{18}+1}^{n_{18}+10} S_i\right)} \qquad [2]$$

In addition, we calculated the germination rate (GR) $GR_{>18\%}$ and $GR_{<\alpha\%}$ for the $Q_{>18\%}$ and $Q_{<\alpha\%}$ quantiles of seeds as follows:

$$GR_{>18\%} = \frac{1}{200} \sum_{i=n_{18}+1}^{n_{18}+10} S_i \qquad [3]$$

$$GR_{<\alpha\%} = \frac{1}{\alpha N}\left(\sum_{i=1}^{n_\alpha} S_i + \frac{r_\alpha}{20} S_{n_\alpha+1}\right), \qquad [4]$$

where $r_\alpha = \text{mod}(\alpha N, 20)$, that is, $r_\alpha = \alpha N - n_\alpha \times 20$, with $0 \leq r_\alpha < 20$ and $n_\alpha$ being the largest integer closest to $0.05\alpha N$.

Parameters estimated for the same set of source germplasm in different environments or with different approaches were compared by calculating Pearson correlations ($r$) and examining two-dimensional plots.

## Determination of Misclassification Rates

In Experiment 2A, the expected FDR and FNR were calculated for different values of α by inserting the maximum likelihood estimates of parameters $\mu_H$, $\mu_C$, and $\sigma = \sigma_H = \sigma_C$ obtained from the MD as well different estimates of κ in the following formulas:

$$\text{FDR}(\alpha) = \frac{\int_{-\infty}^{t(\alpha)} \varphi_C(x)\,dx}{\int_{-\infty}^{t(\alpha)} f(x)\,dx} \qquad [5]$$

$$\text{and, } \text{FNR}(\alpha) = 1 - \int_{-\infty}^{t(\alpha)} \varphi_H(x)\,dx, \qquad [6]$$

where $t(\alpha)$ corresponds to the abscissa of $Q_{<\alpha\%}$ of the PDF $f(x)$ defined in Eq. [1].

In Experiment 2B, the observed FDR and FNR were calculated for α = 5.0, 7.5, and 10.0% as follows:

$$\text{FDR}(\alpha) = 1 - \frac{\sum_{i=1}^{n_\alpha} H_i + \frac{H_{n_\alpha+1}}{20} r_\alpha}{\sum_{i=1}^{n_\alpha} S_i + \frac{r_\alpha}{20} S_{n_\alpha+1}} \qquad [7]$$

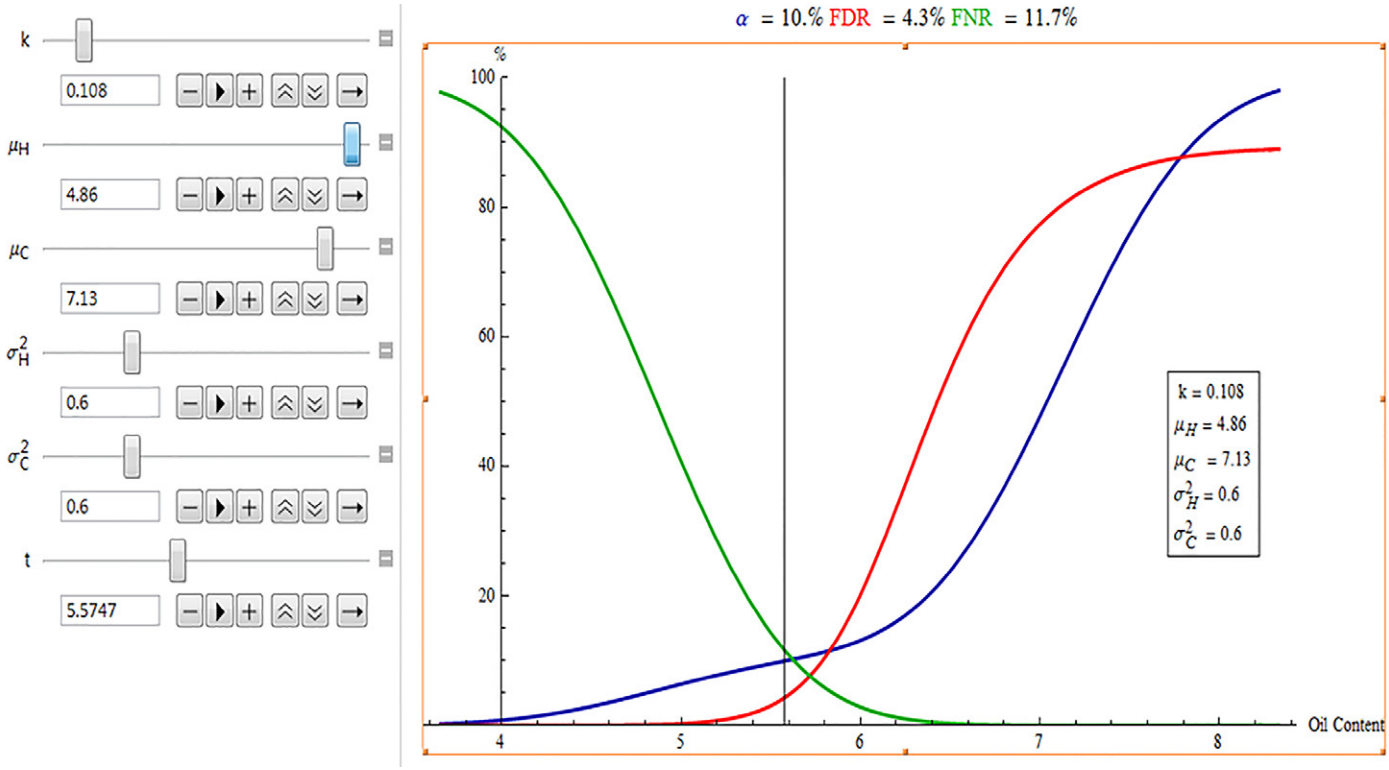


Figure 1. Interactive graphical user interface of the software for calculating the selected proportion (α), the expected false discovery rate (FDR), and false negative rate (FNR). The software (Supplementary File S1) can be used with the freely available computable document format (CDF) player (Wolfram research Inc., www.wolfram.com/cdf-player/). Parameters κ (haploid induction rate [HIR]), $\mu_H$ , $\mu_C$, $\sigma_H$, and $\sigma_H$ (means and standard deviations for haploid [*H*] and diploid crossing [*C*] seeds) can be adjusted using sliders at the left panel. The graphical output can be deactivated to speed up computation.

$$\mathrm{FNR}(\alpha)=\frac{\left(\sum_{i=n_\alpha+1}^{n_{18}} H_i-\frac{r_\alpha}{20}H_{n_\alpha+1}\right)+\frac{0.82}{200}N\sum_{i=n_{18}+1}^{n_{18}+10}H_i}{\sum_{i=1}^{n_{18}}H_i+\frac{0.82}{200}N\sum_{i=n_{18}+1}^{n_{18}+10}H_i}\quad [8]$$

## Software

For computing the expected FDR and FNR according to Eq. [5–6], we provide a software in an interactive computable document format (CDF) usable with the freely available CDF player (www.wolfram.com/cdf-player) (Fig. 1). Users are able to adjust values of all parameters ($\mu_H$, $\mu_C$, $\sigma_H$, $\sigma_C$, and κ) and the threshold $t(\alpha)$ on the left panel of the interface. The curves for the FDR(α) and FNR(α) as a function of the truncation point $t(\alpha)$, as well as the corresponding percentage values of α, FDR and FNR are given as output on the right panel of the interface.

## RESULTS

In Experiment 1, the number of seeds measured for OC from each induction cross ranged between 995 and 5680 (Table 1). Estimates $\hat{\mu}_H$ from the MD had a mean of 3.76% OC and ranged between 3.29 and 5.22% OC. They were almost in perfect agreement with the $\tilde{\mu}_H$ estimates from the GSC ($r = 0.97$, $P < 0.01$). Estimates $\hat{\mu}_C$ and $\tilde{\mu}_C$ were also tightly correlated ($r = 0.95$, $P < 0.01$) between both approaches and ranged between 5.15 and 6.52% OC with slightly higher values for $\hat{\mu}_C$ than $\tilde{\mu}_C$. The standard deviation σ of *H* and *C* seeds averaged 0.55% OC for $\hat{\sigma}$ and 0.71% OC for $\tilde{\sigma}$, with a lower range for the $\hat{\sigma}$ values. Estimates of κ varied between 9.0 and 19.4% and were on average 2.3% higher for $\hat{\kappa}$ than $\tilde{\kappa}$. Correlations of $\hat{\sigma}$ with $\tilde{\sigma}$ and $\hat{\kappa}$ with $\tilde{\kappa}$ were moderate but increased considerably when induction crosses involving source germplasm (F047 × F151) and UHF were eliminated as outliers (σ: $r = 0.77$, $P < 0.05$; κ: $r = 0.92$, $P < 0.01$). The germination rate of the 300 PH seeds and 300 PC seeds averaged 72 and 98%, respectively.

In Experiment 2A, the number of seeds measured for OC ranged between 763 and 4258 in Hohenheim and between 580 and 3399 in Eckartsweier (Table 2). Estimates of $\hat{\mu}_H$ and $\hat{\mu}_C$ were both on average 0.85% OC higher for the induction crosses produced in the greenhouse in Hohenheim than those produced in the nursery in Eckartsweier and varied from 2.72 to 4.94% OC for $\hat{\mu}_H$ and from 4.37 to 7.55% OC for $\hat{\mu}_C$. For both parameters, the correlation between environments was very tight ($r = 0.84$, $P < 0.01$). The mean of $\hat{\sigma}$ was 0.51% OC in both environments and estimates ranged from 0.32 to 0.75%. The mean $\hat{\kappa}$ was 13.4% in Hohenheim and 13.9% in Eckartsweier and the estimates varied between 7.9 and 18.1% and between 5.0 and 29.3%, respectively. Estimates from the two environments were moderately correlated

**Table 1. Number of seeds (*N*) with oil content (OC) ≥2.1% in Experiment 1, means ($\mu_H$ , $\mu_C$), and standard deviation ($\sigma$) of haploid (*H*) and diploid crossing (*C*) seeds and haploid induction rate $\kappa$ estimated by (i) mixture distribution (MD) analysis of OC of seeds and (ii) gold standard classification (GSC) of plants grown from 300 putative haploid (PH) and 300 putative diploid crossing (PC) seeds from induction crosses of 10 source germplasm pollinated with high-oil inducer UH600. The germination rates (GR) of the PH and PC seeds from each induction cross are also given.**

| Source germplasm | MD | | | | | GSC | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | $\hat{\mu}_H$ | $\hat{\mu}_C$ | $\hat{\sigma}$ | $\hat{\kappa}$ | $\tilde{\mu}_H$ | $\tilde{\mu}_C$ | $\tilde{\sigma}$† | $\tilde{\kappa}$ | $GR_{PH}$ | $GR_{PC}$ |
| | | % | | | | % | | | | | |
| F047 × F151 | 995 | 4.13 | 6.08 | 0.48 | 19.4 | 4.39 | 5.69 | 0.84 | 10.2 | 84 | 100 |
| F047 × LH127 | 5271 | 3.29 | 5.34 | 0.50 | 9.3 | 3.34 | 5.15 | 0.60 | 9.0 | 68 | 100 |
| F104 × PHN37 | 4220 | 3.31 | 5.55 | 0.44 | 9.5 | 3.41 | 5.39 | 0.60 | 9.2 | 57 | 99 |
| P135 × GEMS-0115 | 1644 | 3.67 | 5.61 | 0.62 | 16.8 | 3.80 | 5.36 | 0.71 | 15.3 | 83 | 98 |
| P159 × P167 | 5680 | 3.31 | 5.39 | 0.46 | 17.3 | 3.32 | 5.22 | 0.59 | 16.5 | 75 | 100 |
| P072 × W608S | 5457 | 3.33 | 5.56 | 0.51 | 18.6 | 3.36 | 5.37 | 0.70 | 16.4 | 59 | 98 |
| Flint-synthetic UHF | 4027 | 5.22 | 6.52 | 0.60 | 13.3 | 4.89 | 6.26 | 0.74 | 12.0 | 85 | 99 |
| Dent-synthetic UHD1 | 4601 | 3.75 | 5.82 | 0.60 | 16.2 | 3.67 | 5.80 | 0.64 | 16.0 | 61 | 97 |
| Schindelmeiser | 1983 | 3.79 | 6.17 | 0.65 | 13.3 | 3.80 | 5.75 | 0.85 | 9.0 | 72 | 96 |
| Gelber Badischer | 1663 | 3.79 | 6.05 | 0.66 | 12.2 | 3.86 | 5.67 | 0.82 | 9.2 | 76 | 99 |
| Mean | 3554 | 3.76 | 5.81 | 0.55 | 14.6 | 3.78 | 5.57 | 0.71 | 12.3 | 72 | 98 |
| SE‡ | | 0.038 | 0.013 | 0.009 | 1.1 | 0.073 | 0.044 | 0.035 | 0.6 | 2.5 | 0.7 |

† Calculated as $\tilde{\sigma} = \sqrt{\left(\tilde{\sigma}_H^2 + \tilde{\sigma}_C^2\right)/2}$.

‡ SE, standard error averaged over all 10 induction crosses.

**Table 2. Number of seeds (*N*) with oil content (OC) ≥2.1% in Experiment 2A. Means ($\mu_H$ , $\mu_C$) and standard deviation ($\sigma$) of haploid (*H*) and diploid crossing (*C*) seeds and haploid induction rate $\kappa$ estimated by mixture distribution (MD) analysis of OC of seeds from induction crosses of 14 maize inbred lines pollinated with high-oil (HO) inducer UH600 in two environments (Hohenheim and Eckartsweier). The mean OC of 100 seeds of the source germplasm ($\mu_S$) produced by self-pollination in Eckartsweier is also given.**

| Source germplasm | $\mu_S$ | Hohenheim | | | | | Eckartsweier | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *N* | $\hat{\mu}_H$ | $\hat{\mu}_C$ | $\hat{\sigma}$ | $\hat{\kappa}$ | *N* | $\hat{\mu}_H$ | $\hat{\mu}_C$ | $\hat{\sigma}$ | $\hat{\kappa}$ |
| | | | % | | | | | % | | | |
| F047 | 3.47 | 1653 | 4.59 | 6.66 | 0.53 | 9.6 | 1619 | 3.35 | 5.78 | 0.51 | 5.0 |
| F087 | 3.82 | 1681 | 4.77 | 6.58 | 0.50 | 9.7 | 1833 | 3.75 | 5.78 | 0.52 | 5.8 |
| F137 | 4.25 | 1291 | 4.93 | 7.12 | 0.50 | 9.3 | 1531 | 3.50 | 6.55 | 0.54 | 7.6 |
| F154 | 4.58 | 4258 | 4.94 | 7.55 | 0.47 | 12.9 | 3399 | 4.53 | 6.82 | 0.62 | 11.8 |
| F157 | 3.95 | 1666 | 4.94 | 7.10 | 0.59 | 8.0 | 580 | 4.00 | 6.06 | 0.51 | 13.2 |
| F159 | 3.31 | 1545 | 4.30 | 6.09 | 0.49 | 7.9 | 2137 | 3.35 | 5.44 | 0.40 | 6.3 |
| P140 | 4.45 | 1052 | 4.59 | 6.61 | 0.75 | 15.9 | 743 | 3.64 | 5.57 | 0.65 | 14.6 |
| P159 | 3.73 | 763 | 4.40 | 6.38 | 0.49 | 16.3 | 2111 | 3.46 | 5.32 | 0.51 | 21.0 |
| P167 | 3.17 | 3404 | 3.85 | 5.46 | 0.50 | 14.7 | 2633 | 3.35 | 4.85 | 0.50 | 11.6 |
| P213 | 3.33 | 1142 | 3.71 | 5.79 | 0.50 | 17.5 | 672 | 3.21 | 5.23 | 0.53 | 29.3 |
| P245 | 2.61 | 1855 | 3.23 | 5.48 | 0.49 | 15.5 | 1194 | 2.79 | 4.51 | 0.39 | 27.5 |
| P290 | 4.25 | 3207 | 4.79 | 7.06 | 0.59 | 15.5 | 2577 | 4.04 | 5.92 | 0.57 | 15.5 |
| P312 | 4.41 | 1070 | 4.69 | 6.42 | 0.45 | 18.1 | 1618 | 3.46 | 5.38 | 0.57 | 10.5 |
| S072 | 2.59 | 2810 | 3.26 | 5.06 | 0.37 | 16.5 | 1120 | 2.72 | 4.37 | 0.32 | 15.7 |
| Mean | 3.71 | 1957 | 4.36 | 6.38 | 0.51 | 13.4 | 1698 | 3.51 | 5.54 | 0.51 | 13.9 |
| SE† | 0.45 | | 0.047 | 0.016 | 0.012 | 1.0 | | 0.060 | 0.017 | 0.012 | 1.1 |

† SE, standard error averaged across all 14 induction crosses.

for both $\hat{\sigma}$ ($r = 0.59$, $P < 0.05$) and $\hat{\kappa}$ ($r = 0.66$, $P < 0.01$). The ratio $\left(\hat{\mu}_C - \hat{\mu}_H\right)/\hat{\sigma}$ ranged between 2.7 and 5.6 and averaged 4.0 in both environments (data not shown).

Estimates of $\tilde{\kappa}$ in Experiment 2B averaged 7.1% and ranged between 5.2 and 12.4% (Table 3). The mean germination rate amounted to 53.2% for $GR_{<7.5\%}$ and 88.7% for $GR_{>18\%}$; $GR_{<7.5\%}$ was even below 32% for three of the 14 induction crosses. Estimates $\tilde{\kappa}$ and $GR_{<7.5\%}$ were weakly associated ($r = 0.42$).

Averaged across all 14 induction crosses, the FDR($\alpha$) determined from the GSC amounted to 9.9, 11.3, and 16.9% for $\alpha$ = 5.0, 7.5, and 10.0%, respectively (Table 3). In 12 of the induction crosses, FDR(7.5%) was below 20% and in nine even below 7.5%. The two induction crosses

Table 3. Haploid induction rate (HIR) $\tilde{\kappa}$, false discovery rate FDR(α), and false negative rate FNR(α) determined by gold standard classification (GSC) of plants in Experiment 2B for different proportions α of putative haploid (PH) seeds selected on the basis of the oil content (OC) from induction crosses of 14 maize inbred lines in Experiment 2A. The germination rate (GR) of the 7.5% lower quantile ($GR_{<7.5\%}$) and the 82.0% upper quantile ($GR_{>18\%}$) of seeds are also given.

| Source germplasm | HIR | GR | | FDR(α) | | | FNR(α) | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\tilde{\kappa}$ | $GR_{<7.5\%}$ | $GR_{>18\%}$ | α = 5.0% | α = 7.5% | α = 10.0% | α = 5.0% | α = 7.5% | α = 10.0% |
| | % | | | | | | | | |
| F047 | 7.1 | 61 | 91 | 2.0 | 2.6 | 9.0 | 51.6 | 27.7 | 6.3 |
| F087 | 5.2 | 58 | 95 | 4.4 | 7.1 | 26.0 | 38.7 | 13.9 | 0.0 |
| F137 | 5.7 | 69 | 88 | 15.4 | 19.7 | 38.8 | 44.6 | 16.1 | 11.0 |
| F154 | 6.7 | 62 | 91 | 2.4 | 2.5 | 7.5 | 50.0 | 21.9 | 0.0 |
| F157 | 5.6 | 66 | 91 | 4.9 | 16.7 | 37.3 | 41.3 | 17.6 | 10.7 |
| F159 | 5.3 | 63 | 86 | 14.6 | 14.0 | 38.6 | 40.9 | 8.0 | 4.8 |
| P140 | 5.3 | 22 | 85 | 32.7 | 29.5 | 23.2 | 79.1 | 71.6 | 48.7 |
| P159 | 7.5 | 41 | 77 | 6.9 | 4.2 | 2.8 | 67.0 | 44.8 | 15.6 |
| P167 | 7.4 | 59 | 88 | 9.2 | 6.6 | 8.7 | 58.3 | 34.3 | 11.1 |
| P213 | 6.2 | 32 | 91 | 0.0 | 1.0 | 5.1 | 71.1 | 54.2 | 45.7 |
| P245 | 12.4 | 76 | 77 | 2.4 | 1.9 | 2.4 | 61.5 | 41.3 | 19.0 |
| P290 | 7.6 | 55 | 95 | 7.0 | 5.3 | 3.8 | 63.2 | 42.5 | 18.6 |
| P312 | 6.0 | 18 | 96 | 32.1 | 42.6 | 30.1 | 89.4 | 85.0 | 68.4 |
| S072 | 10.9 | 65 | 94 | 5.1 | 4.4 | 3.2 | 66.6 | 52.2 | 34.5 |
| Mean | 7.1 | 53.2 | 88.7 | 9.9 | 11.3 | 16.9 | 58.8 | 37.9 | 21.0 |

with FDR(α) ≥ 25% for α = 7.5% had also the lowest $GR_{<7.5\%}$. The average of FDR(α) was 58.8, 37.9, and 21.0% for α = 5.0, 7.5, and 10.0%, respectively. Estimates of FDR(α) and FNR(α) varied considerably among induction crosses for given values of α.

## DISCUSSION

Mixture distributions occur frequently in biology and many technical disciplines and their analyses with modern statistical tools have received considerable attention in recent years (Früwirth-Schnatter, 2006). One of the most prominent examples represents the analyses of molecular assays, where the genotype of individuals must be assigned to two or three genotype classes on the basis of quantitative measurements of polymerase chain reaction products or single-nucleotide polymorphism signal intensities (Piepho and Koch, 2000). Seeds from induction crosses provide a further example in as much as *H* and *C* seeds differ in their expression of quantitative traits such as OC, protein content, or other traits. This motivated us to apply established statistical methods to the analysis of mixture distributions of OC in induction crosses produced with HO inducers. Our goal was to examine whether the acquired information could be used for balancing the misclassification rates in the discrimination of *H* and *C* seeds as a means to optimize the allocation of resources in the production of DH lines with this novel method following a similar approach as Riedelsheimer and Melchinger (2013) for genomic selection.

### Misclassification Rates and their Importance in Sorting of Haploid and Diploid Crossing Seeds

According to theoretical results (Melchinger et al., 2013), the FDR and FNR in sorting seeds from induction crosses with HO inducers depend on the following items: (i) the difference $\mu_C - \mu_H$ between the mean OC of the *H* and *C* seeds relative to the size of the standard deviation σ of OC within each fraction, (ii) the haploid induction rate κ, and (iii) the proportion α of seeds with lowest OC selected from the MD, which is determined by the threshold $t_\alpha$ chosen to discriminate *H* from *C* seeds. In all induction crosses in Experiment 2A, the ratio $(\hat{\mu}_C - \hat{\mu}_H)/\hat{\sigma}$ exceeded 1.80, which was considered as minimum for reliable discrimination of PH and PC seeds (Melchinger et al., 2013). The haploid induction rate κ is mainly a property of the inducer but may also vary among source germplasm (Penghao et al., 2014).

Different from the *R1-nj* embryo marker employed so far in practice for sorting of PH and PC seeds, use of HO inducers and production of induction crosses and discrimination of PH and PC seeds on the basis of their OC enables controlling the misclassification rates by suitable choice of the proportion α of selected PH seeds (Melchinger et al., 2014). Since FDR(α) and FNR(α) depend inversely on α (Melchinger et al., 2013), the breeder must balance both by a suitable choice of α. If the FDR is high, this means that many PH seeds treated with colchicine and transplanted to the $D_0$ nursery later turn out to be false positives (i.e., *C* plants). By contrast, a high FNR implies that many genuine *H* seeds are discarded as

PC seeds, which increases the required number of seeds in induction crosses. If the success rates of the various steps in the DH procedure are known, choice of $\alpha$ conditions FNR($\alpha$) and, as outlined in the Appendix, both together determine (i) the total number of seeds $N(\alpha)$ that must be produced in induction crosses and (ii) the total number of PH seeds $n(\alpha)$ that must be treated for chromosome doubling to reach a targeted number of $D_1$ lines. Commonly, $\alpha$ is chosen such that a low FDR is warranted, usually at the expense of a high FNR, but a rigorous economic optimization has to take into account the costs for all steps in the DH procedure (Melchinger et al., 2013). In the following, we discuss whether information from the parents or the induction crosses themselves can be of use for choosing $\alpha$ to balance the FDR and FNR.

## Can Parents Provide Information on Properties of Induction Crosses?

From Eq. [5, 6] it follows that FDR($\alpha$) and FNR($\alpha$) depend only on the ratio $(\mu_C - \mu_H)/\sigma$ but not on the absolute size of $\mu_H$, $\mu_C$, and $\sigma$. Melchinger et al. (2013) proposed estimating $\mu_C - \mu_H$ by $0.5(\mu_I - \mu_S)$, where $\mu_I$ and $\mu_S$ are the mean OC of the inducer and source germplasm, respectively. The rationale underlying this proposal is that $\mu_C = 0.5(\mu_I + \mu_S)$ and $\mu_H = \mu_S$ under the assumption of additive gene action and absence of gene dosage effects for OC, as suggested by results from the literature (Curtis et al., 1956; Wassom et al., 2008).

In Experiment 2A, we could test these assumptions because in the nursery at Eckartsweier, the parent lines used as source germplasm were grown and self-pollinated in plots adjacent to the production of the induction crosses. The OC of the parents $\tilde{\mu}_S$ agreed closely with $\hat{\mu}_H$ ($r = 0.83$, $P < 0.01$) (Table 2). Likewise, $0.5(\tilde{\mu}_I + \tilde{\mu}_S)$ was tightly correlated ($r = 0.87$, $P < 0.01$) with $\hat{\mu}_C$, but the predictions were on average about 0.88% OC higher than $\hat{\mu}_C$, considering $\tilde{\mu}_I = 10.8\%$ OC for inducer UH600. As a consequence, $0.5(\tilde{\mu}_I + \tilde{\mu}_S)$ overestimated $\hat{\mu}_C - \hat{\mu}_H$ by 1.52% OC, which is in harmony with the results of Experiment 1 reported by Melchinger et al. (2014). Interestingly, $\tilde{\mu}_S$ was moderately associated ($r = 0.53$) with the observed FDR(7.5%) in Experiment 2B. In particular, parents P140 and P312 with FDR(7.5%) $\geq$ 25% in their induction crosses had also high OC in addition to low values for $GR_{<7.5\%}$. These results strongly support recording the OC of the parents before production of induction crosses. If $\mu_S$ is high, the breeder can react in advance by increasing the number $N(\alpha)$ of seeds produced in induction crosses so that later a more stringent selection, corresponding to a low value of $\alpha$, can be applied to control the FDR.

## Precision of Parameter Estimates from Mixture Distributions

According to statistical theory, the precision of the estimates obtained from a MD differs among the parameters and depends in addition to the sample size on the magnitude of parameters themselves because the second derivative of the log-likelihood and its expectation are not free of the parameters (McLachlan and Peel, 2000). Generally, first-degree statistics ($\mu_H$ and $\mu_C$) are estimated with higher precision than second degree statistics ($\sigma$). Precision of the estimated mixture weight $\kappa$ is usually lowest and decreases as $\kappa$ deviates from 0.5 towards the borders of the interval [0, 1]. Obviously, undetected outliers in the mixture distribution can cause deviations from these general trends.

The standard errors of the parameters estimated from the MD in Experiment 1 and 2A agreed well with these theoretical expectations and were very small in comparison with the magnitude of the parameters, suggesting that the sample size was sufficient in our induction crosses. In practical breeding, the number $N(\alpha)$ of seeds produced per induction cross is most likely even much higher than in our study to reach the number of targeted $D_1$ lines (Prigge and Melchinger, 2012). Thus, the precision in the estimates of the means and standard deviation should not be a limiting factor in application of the MD approach.

In both Experiment 1 and 2A, we found that the robustness of parameter estimates from the MD improved considerably for the reduced model with $\sigma = \sigma_H = \sigma_C$. In addition to the Akaike criterion, experimental estimates of $\sigma_H^2$ and $\sigma_C^2$ from the GSC in Experiment 1 are in harmony with this assumption (Melchinger et al., 2014). Using the model with constant variance of the two mixture components is also preferable because of better convergence properties. With heterogeneous variances, multiple local maxima, and unboundedness of the likelihood can cause severe numerical problems (cf. Piepho and Koch, 2000). As reported by Melchinger et al. (2014), the environmental and measurement errors are by far the major sources of variation of OC among individual *H* and *C* seeds in comparison to the influence of the genetic component attributable to segregation. In Experiment 2A, the latter source of variation was even absent, because *H* and *C* seeds are both expected to be genetically uniform as the source germplasm and inducer were homozygous lines.

## Comparison of Mixture Distribution Results from Different Environments

In Experiment 2A, the production conditions were much better in the greenhouse than in the nursery, and this resulted in higher values of $\hat{\mu}_H$ and $\hat{\mu}_C$ for OC (Table 2). Since the seed weight, averaged over all induction crosses, amounted to 278 mg for both environments, the higher OC in the greenhouse was mainly attributable to the 20% higher oil mass observed under the more favorable

conditions. The values for $\hat{\sigma}$ and $\hat{\kappa}$ were of similar size for both environments but showed lower correlations between environments in comparison with $\hat{\mu}_H$ and $\hat{\mu}_C$ as a consequence of their lower precision. Moreover, the moderate correlation between environments for $\hat{\kappa}$ could reflect genotype-by-environment interactions for haploid induction rate, as reported for tropical germplasm by Kebede et al. (2011). Altogether, the means of $(\hat{\mu}_C - \hat{\mu}_H)/\hat{\sigma}$ and $\hat{\kappa}$ were almost identical across environments. Thus, even if the level in the OC of seeds varies among environments, the difference in the mean OC of both fractions and, consequently, the power to discriminate *H* from *C* seeds is at best marginally affected by the environment, in which the induction crosses are produced, but the choice of the threshold must be determined on the basis of the mean OC of each fraction. Since seed set and quality are better under favorable conditions, we recommend producing the induction crosses in environments most appropriate for maize production, because a higher OC in *H* and *C* seeds seems not to entail a smaller difference between both fractions.

## Comparison of Mixture Distribution and Gold Standard Classification Results

In Experiment 1, the values for $(\hat{\mu}_C - \hat{\mu}_H)$ and $\hat{\sigma}$ agreed well in magnitude with those of $(\tilde{\mu}_C - \tilde{\mu}_H)$ and $\tilde{\sigma}$, respectively, and were also moderately correlated with each other (Table 1). Thus, it seems safe to conclude that these parameters and the ratio $(\mu_C - \mu_H)/\sigma$ can be estimated with sufficient precision by MD analysis of the OC of seeds from the induction crosses.

Parameter κ was an exception in that values of $\hat{\kappa}$ were generally much higher than those of $\tilde{\kappa}$ (Table 1, 2, 3). The former estimates rely on data of seeds, whereas the latter estimates rely on data of plants surviving until flowering time. Hence, differences in germination of PH and PC seeds could explain the discrepancies in estimates of κ from the MD and GSC. This hypothesis is in harmony with the substantial differences between the germination rates $GR_{<7.5\%}$ and $GR_{>18\%}$ in Experiment 2B (Table 3). In Experiment 1 too, the GR of the 300 PH and the 300 PC seeds averaged over all 10 induction crosses amounted to 72 and 98%, respectively (Table 1). Seeds with low OC tend to have lower seed weight, as reflected by the positive correlation between both traits (Melchinger et al., 2015). Consequently, PH seeds with low OC have less energy reserves than PC seeds with high OC and may encounter greater problems in germination or survival of the juvenile stage under low temperatures and other stress conditions as prevailed in Experiment 2B and to a lower extent in Experiment 1. Since *C* seeds have a heterozygous diploid embryo, whereas *H* seeds have a haploid embryo, they can also benefit from a higher fitness due to heterosis for germination and early vigor in stress environments. Both these causes are expected to be less important for germination in the greenhouse, as is routine in the DH method, and explain the higher value of κ = 10.2% reported for UH600, assessed in testcrosses with a tester carrying the liguleless (*lg2*) marker gene grown in the greenhouse (https://plant-breeding.uni-hohenheim.de/84531), in comparison with the mean $\tilde{\kappa}$ = 7.1% observed in Experiment 2B.

The main reason for the higher values of $\hat{\kappa}$ in comparison with $\tilde{\kappa}$ and the low GR of PH seeds in both experiments is presumably a substantial proportion of seeds with aborted embryos in the PH fraction. Results of Xu et al. (2013) showed that the *sed1* locus triggers not only haploid induction but also causes defective kernels. Thus, we speculate if the defect becomes effective at an early developmental stage, this results in embryoless seeds with OC < 2.1%, which we discarded from the MD. However, if embryo abortion occurs at a later stage of embryo development, the seed may have already reached an OC ≥ 2.1% and, therefore, would be included in the PH fraction, but later fail to germinate. We were not able to test this hypothesis by discriminating embryo-aborted seeds from intact *H* seeds in our experiments. However, in accordance with this hypothesis, when determining κ of inducers with testers carrying the ligueless (*lg2*) marker gene, we observed under optimum germination conditions in the greenhouse generally a lower GR for the PH than for the PC fraction (V. Mirdita, unpublished data, 2014).

## Options for Controlling the False Discovery Rate

Theoretical results given in the Appendix demonstrate that α and FDR(α) determine inter alia the number $N(\alpha)$ of seeds to be produced in induction crosses and the number $n(\alpha)$ of seeds to be treated with colchicine. A first clue about FDR(α) for different values of α can be obtained by inserting estimates of $\mu_H$, $\mu_C$, and σ from previous experiments and prior information about κ of the inducer into Eq. [5] and calculate FDR(α) by our software (Fig. 1). Using $\hat{\mu}_H$, $\hat{\mu}_C$, and $\hat{\sigma}$ from Experiment 1 together with κ = 10.2% for inducer UH600 gave a fairly good idea about the magnitude of FDR(α) and FNR(α) observed in Experiment 2B (data not shown). Induction crosses, for which α should be reduced and $N(\alpha)$ be increased to warrant an acceptable FDR(α), could be identified in most cases by examining the OC $\mu_S$ of the source germplasm, as exemplified by lines P140 and P312 in Experiment 2.

After seeds of an induction cross have been produced and measured for OC, prediction of FDR(α) could be fine-tuned by inserting the values $\hat{\mu}_H$, $\hat{\mu}_C$, and $\hat{\sigma}$ obtained from their MD into Eq. [5]. However, in view of the upward bias in $\hat{\kappa}$ found in both experiments, we recommend using either prior information about κ of the inducer or multiplying the $\hat{\kappa}$ with a shrinkage factor to correct for the expected bias. Using prior information (κ = 10.2%) reported for inducer UH600 yielded poorer results for prediction of the FDR

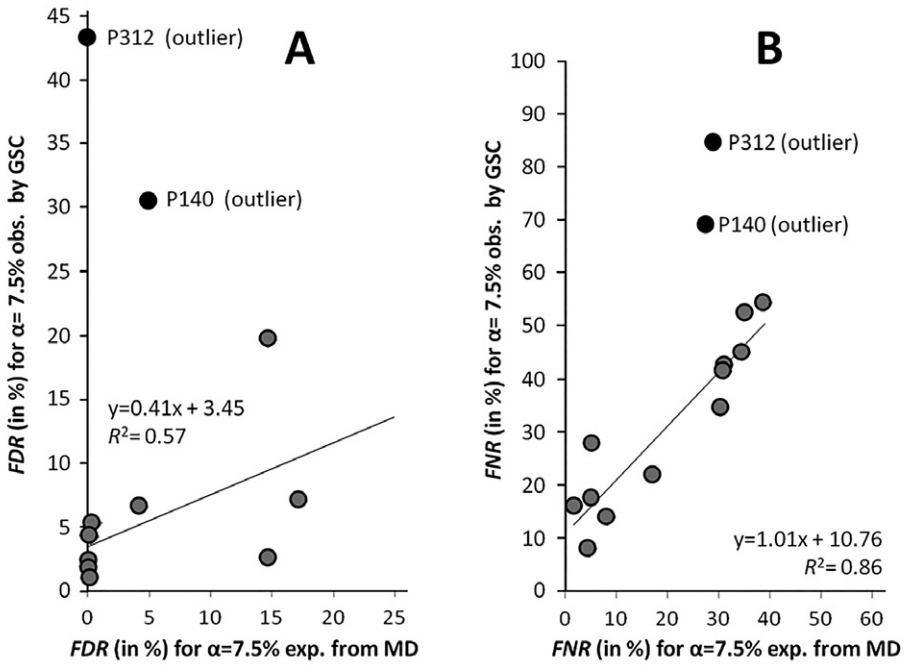


Figure 2. Plots of (A) the expected false discovery rate $FDR_{7.5\%}$ vs. the observed $FDR_{7.5\%}$ for gold standard classification (GSC) and (B) the expected false negative rate $FNR_{7.5\%}$ vs. the observed $FNR_{7.5\%}$ for GSC in Experiment 2B. The expected values were obtained by inserting in Eq. [5] the parameter estimates $\breve{\mu}_H = 4.46\%$ OC, $\breve{\mu}_C = 6.38\%$ OC, and $\breve{\sigma} = 0.51\%$ OC taken from Table 2, but replacing $\hat{\kappa} = 13.4\%$ with $\hat{\kappa} = 0.7 \times 13.4\% = 9.4\%$ to correct for the bias. The $R^2$ values and regression equations were calculated after deleting P140 and P312 as outliers.

and FNR observed in Experiment 2B than multiplying the $\hat{\kappa}$ of each induction cross with common shrinkage factor $\beta$. For example, choosing $\beta = 0.7$ on the basis of the ratio $\overline{\kappa} : \hat{\kappa}$ averaged over both experiments yielded fairly good agreement between the observed and predicted FDR(7.5%) and FNR(7.5%) in Experiment 2 (Fig. 2). After elimination of the induction crosses of P140 and P313 as outliers, which could be identified a priori owing to their low $GR_{<7.5\%}$ and high OC of the source germplasm, the coefficient of determination amounted to $R^2 = 0.57$ for FDR(7.5%) and $R^2 = 0.86$ for FNR(7.5%). Choosing $\alpha = 7.5\%$ turned out to be the best choice, because the FDR and FNR were both at acceptable levels, whereas $\alpha = 5.0\%$ yielded a much higher FNR and $\alpha = 10.0\%$ a considerably higher FDR (Table 3). This is also supported by the results in the Appendix showing that for $\alpha = 7.5\%$, the curves for FDR($\alpha$) and FNR($\alpha$) intersect and N($\alpha$) and $n(\alpha)$ are in a balanced ratio (Fig. 3).

## CONCLUSIONS

Our findings from Experiment 2 corroborate the proof-of-concept (Melchinger et al., 2014), founded on Experiment 1 that in induction crosses with HO inducers, *H* seeds can be discriminated from *C* seeds on the basis of their OC. In all cases, the 7.5% lower quantile of the MD was highly enriched for *H* seeds and in the vast majority of induction crosses <7.5% of the seeds germinating from this fraction were false positives. Thus, unless the source germplasm has extremely high OC and low haploid induction or germination rates, both misclassification rates were at an acceptable level for practical purposes.

Nevertheless, we found ample variation in Experiment 2 regarding FDR($\alpha$) and FNR($\alpha$) for values of $\alpha$ that come into consideration. Equation [5, 6] together with our software (Fig. 1) can provide suitable guidelines for predicting FDR($\alpha$) and FNR($\alpha$) on the basis of prior information about the parent or parameter estimates obtained from the MD of the induction cross and for choosing N($\alpha$), the total number of seeds produced in the induction cross, and $n(\alpha)$, the total number of PH seeds subject to chromosome doubling (see Appendix). However, this information becomes available only a posteriori, that is, after all seeds have been measured for OC. Moreover, $\hat{\kappa}$ values estimated from the MD can be inflated owing to an unknown fraction of embryo-aborted seeds in the fraction of PH seeds. Hence, in parallel to the measurements of OC we recommend sorting the seeds into several classes corresponding to different segments of the MD such as $Q_{<4\%}$, ($Q_{<6\%} - Q_{<4\%}$), ..., and $Q_{>12\%}$. The latter will be discarded immediately because it contains only few *H* seeds, if at all. Depending on the results of the MD and the number of seeds available in each class, the breeder could decide to employ only the $Q_{<4\%}$ class for colchicine treatment and add, if needed, further classes with increasingly higher OC at the risk of a higher proportion of false positives. We are currently developing an automated platform for high-throughput sorting of seeds on the basis on their OC that meets exactly this requirement.

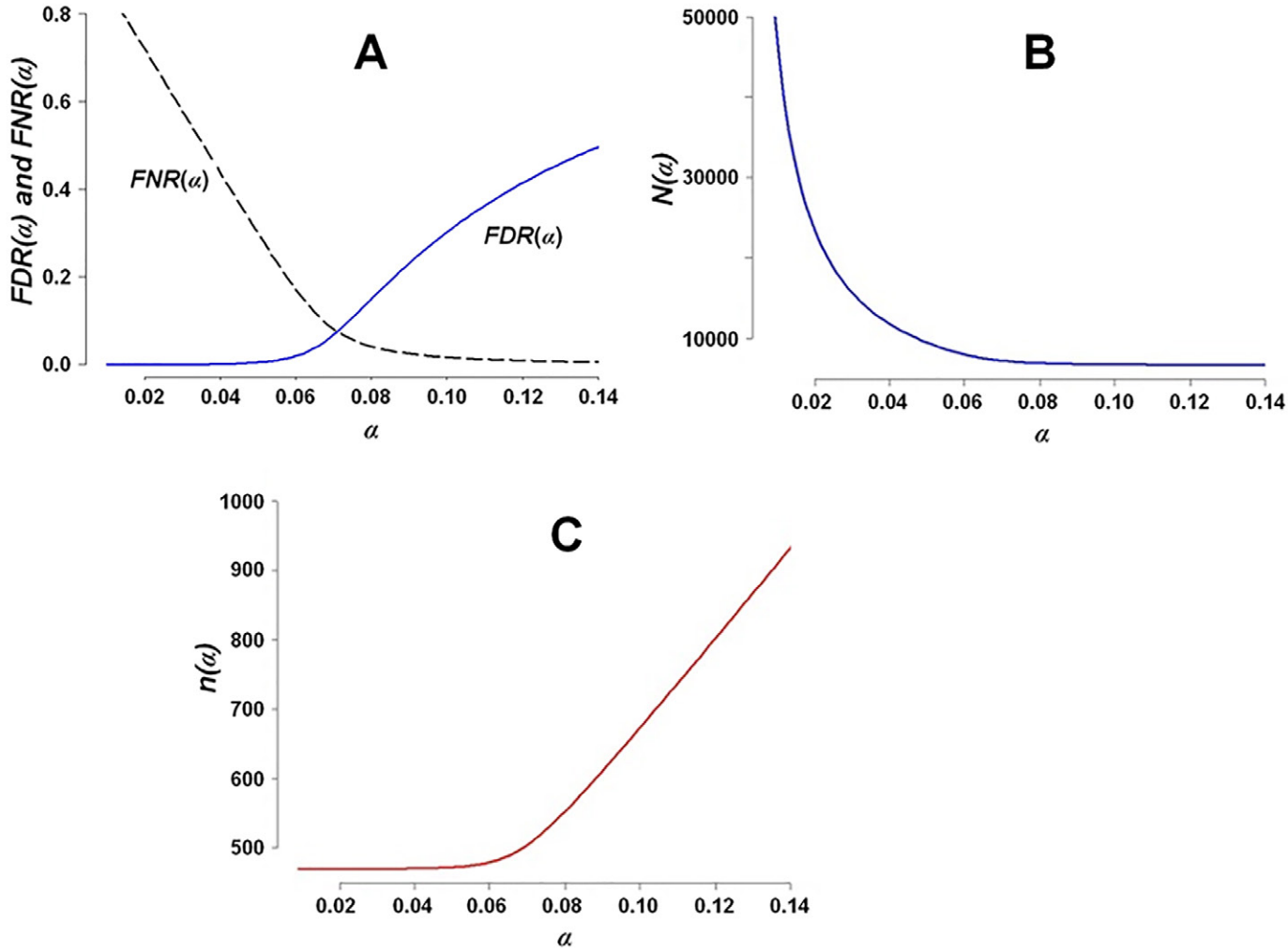


Figure 3. Curves for (A) false discovery rate (FDR[α]) and false negative rate (FNR[α]); (B) number of seeds produced for the induction cross to reach finally number of targeted $D_1$ lines from the induction cross (*F*) $D_1$ lines (*N*[α]); and (C) number of seeds used for germination to produce $D_1$ lines (*n*[α]) as a function of the proportion α of selected putative haploid (PH) seeds calculated by using Eq. [5, 9, 10]. Assumptions were *F* = 50, θ = 20 (W. Schipprack, unpublished data, 2014), and parameters $\hat{\mu}_H = 4.36\%$ OC, $\hat{\mu}_C = 6.38\%$ OC, and $\hat{\sigma} = 0.51\%$ OC, taken from Table 2, and $\tilde{\kappa}$ = 7.1% and $GR_{<7.5\%}$ = 53.2%, taken from Table 3.

## APPENDIX
### Derivations for Calculating N(α) and *n*(α)

Definitions:

- *F* number of targeted $D_1$ lines from the induction cross
- α proportion of seeds with lowest oil content selected from the mixture distribution
- N(α) number of seeds produced for the induction cross to reach finally *F* $D_1$ lines
- *n*(α) number of seeds used for germination to produce $D_1$ lines
- FDR(α) false discovery rate = rate of false positives in the germinated fraction of the α% selected seeds
- FNR(α) false negative rate = rate of *H* seeds in the (100 − α)% of culled seeds
- GR(α) germination rate of haploid (*H*) seed in the α% fraction of selected seeds
- θ success rate of obtaining a $D_1$ line from a germinated haploid (*H*) seed, depending on the method of chromosome doubling applied

With these definitions, we get:

$$n(\alpha) = \frac{F}{\theta\ GR(\alpha)\ (1 - FDR(\alpha))} \quad [9]$$

and $$N(\alpha) = \frac{n(\alpha)}{\alpha} = \frac{F}{\alpha\ \theta\ GR(\alpha)(1 - FDR(\alpha))}. \quad [10]$$

A numerical example showing the graphs of (A) FDR(α) and FNR(α), (B) number of seeds N(α), and (C) *n*(α) based on parameters from Experiment 2A is shown in Fig. 3.

## Acknowledgments

We are grateful to the technical staff of the maize breeding program at the University of Hohenheim; and N. Friedl, R. Volkhausen, and F. Mauch for their dedicated work in the field trials. This research was funded by grants from the Universität Hohenheim. V. Mirdita was financed by the Bill & Melinda Gates Foundation under the grant "A Double Haploid Facility for Strengthening Maize Breeding Programs in Africa".

## References

Burnham, K.P., and D.R. Anderson. 2002. Model selection and multimodel inference: A practical–theoretic approach, 2nd ed. Springer-Verlag, Berlin. ISBN 0-387-95364-7

Cox, C. 1998. Delta method. In: P. Armitage and T. Colton, editors, Encyclopedia of biostatistics. John Wiley Press, New York. p. 1125–1127.

Curtis, J.J., A.M. Brunson, J.E. Hubbard, and F.R. Earle. 1956. Effect of the pollen parent on oil content of the corn kernel. Agron. J. 12:551–555. doi:10.2134/agronj1956.00021962004800120005x

Dong, X., X. Xu, J. Miao, L. Li, D. Zhang, X. Mi, C. Liu, X. Tian, A.E. Melchinger, and S. Chen. 2013. Fine mapping of *qhir1* influencing in vivo haploid induction in maize. Theor. Appl. Genet. 126:1713–1720. doi:10.1007/s00122-013-2086-9

Frühwirth-Schnatter, S. 2006. Finite mixture and markov switching models. Springer Press, Berlin. ISBN 978-1-4419-2194-9.

Kebede, A.Z., B.S. Dhillon, W. Schipprack, J.L. Araus, M. Bänziger, K. Semagn, G. Alvarado, and A.E. Melchinger. 2011. Effect of source germplasm and season on the in vivo haploid induction rate in tropical maize. Euphytica 180:219–226. doi:10.1007/s10681-011-0376-3

McLachlan, G., and D. Peel. 2000. Finite mixture models. John Wiley & Sons, New York.

Melchinger, A.E., W. Schipprack, X. Mi, and V. Mirdita. 2015. Oil content is superior to oil mass for identification of haploid seeds in maize produced with high-oil inducers. Crop Sci. 55: 188–195. doi:10.2135/cropsci2014.06.0432

Melchinger, A.E., W. Schipprack, H.F. Utz, and V. Mirdita. 2014. In vivo haploid induction in maize: Identification of haploid seeds by their oil content. Crop Sci. 54:1497–1504. doi:10.2135/cropsci2013.12.0851

Melchinger, A.E., W. Schipprack, T. Würschum, S. Chen, and F. Technow. 2013. Rapid and accurate identification of in vivo-induced haploid seeds based on oil content in maize. Scientific Reports 3:2129. doi:10.1038/srep02129

Penghao, W., L. Haochuan, and R. Jiaojiao. 2014. Mapping of maternal QTLs for in vivo haploid induction rate in maize (*Zea mays* L.). Euphytica 196:413–421. doi:10.1007/s10681-013-1043-7

Piepho, H.P., and G. Koch. 2000. Codominant analysis of banding data from a dominant marker system by normal mixtures. Genetics 155:1459–1468.

Prigge, V., and A.E. Melchinger. 2012. Production of haploids and doubled haploids in maize. In: V.M. Loyola-Vargas and N. Ochoa-Alejo, editors, Plant cell culture protocols. 3rd ed. Humana Press, Totowa, NJ. p. 161–172.

Riedelsheimer, C., and A.E. Melchinger. 2013. Optimizing the allocation of resources for genomic selection in one breeding cycle. Theor. Appl. Genet. 126:2835–2848. doi:10.1007/s00122-013-2175-9

SAS Institute. 2008. SAS/STAT 9.2 user's guide. SAS Inst., Cary, NC.

Wassom, J.J., V. Mikkelineni, M.O. Bohn, and T.R. Rocheford. 2008. QTL for fatty acid composition of maize kernel oil in Illinois high oil × B73 backcross-derived lines. Crop Sci. 48:69–78. doi:10.2135/cropsci2007.04.0208

Xu, X., L. Li, X. Dong, W. Jin, A.E. Melchinger, and S. Chen. 2013. Gametophytic and zygotic selection leads to segregation distortion associated with in vivo maternal haploids in maize. J. Exp. Bot. 64:1083–1096. doi:10.1093/jxb/ers393