# Assessing the potential of genetic resource introduction into elite germplasm: a collaborative multiparental population for flint maize

Dimitri Sanchez, Antoine Allier, Sarah Ben Sadoun, Tristan Mary-Huard, Cyril Bauland, Carine Palaffre, Bernard Lagardère, Delphine Madur, Valérie Combes, Stéphane Melkior, et al.

## HAL Id: hal-04401428
## https://hal.inrae.fr/hal-04401428

Submitted on 4 Apr 2024

## ORIGINAL ARTICLE

# Assessing the potential of genetic resource introduction into elite germplasm: a collaborative multiparental population for flint maize

Dimitri Sanchez[1] · Antoine Allier[1,2] · Sarah Ben Sadoun[1] · Tristan Mary-Huard[1,3] · Cyril Bauland[1] · Carine Palaffre[4] · Bernard Lagardère[4] · Delphine Madur[1] · Valérie Combes[1] · Stéphane Melkior[5] · Laurent Bettinger[6] · Alain Murigneux[7] · Laurence Moreau[1] · Alain Charcosset[1]

## Abstract

***Key message*** **Implementing a collaborative pre-breeding multi-parental population efficiently identifies promising donor x elite pairs to enrich the flint maize elite germplasm.**

**Abstract** Genetic diversity is crucial for maintaining genetic gains and ensuring breeding programs' long-term success. In a closed breeding program, selection inevitably leads to a loss of genetic diversity. While managing diversity can delay this loss, introducing external sources of diversity is necessary to bring back favorable genetic variation. Genetic resources exhibit greater diversity than elite materials, but their lower performance levels hinder their use. This is the case for European flint maize, for which elite germplasm has incorporated only a limited portion of the diversity available in landraces. To enrich the diversity of this elite genetic pool, we established an original cooperative maize bridging population that involves crosses between private elite materials and diversity donors to create improved genotypes that will facilitate the incorporation of original favorable variations. Twenty donor × elite BC1S2 families were created and phenotyped for hybrid value for yield related traits. Crosses showed contrasted means and variances and therefore contrasted potential in terms of selection as measured by their usefulness criterion (UC). Average expected mean performance gain over the initial elite material was 5%. The most promising donor for each elite line was identified. Results also suggest that one more generation, i.e., 3 in total, of crossing to the elite is required to fully exploit the potential of a donor. Altogether, our results support the usefulness of incorporating genetic resources into elite flint maize. They call for further effort to create fixed diversity donors and identify those most suitable for each elite program.

✉ Alain Charcosset
alain.charcosset@inrae.fr

1 Université Paris-Saclay, INRAE, CNRS, AgroParisTech, Génétique Quantitative et Evolution–Le Moulon, 91190 Gif-Sur-Yvette, France

2 Present Address: Syngenta, 12 Chemin de L'Hobit, 31790 Saint-Sauveur, France

3 Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA-Paris Saclay, 91120 Palaiseau, France

4 UE 0394 SMH, INRAE, 2297 Route de l'INRA, 40390 Saint-Martin-de-Hinx, France

5 RAGT2n, 12510 Druelle, France

6 Lidea France, Avenue Gaston Phoebus, 64230 Lescar, France

7 Limagrain Europe, 28 Route d'Ennezat, 63720 Chappes, France

## Introduction

The release of new varieties with constantly improved genetic values has efficiently contributed to the augmentation of the yield production that was needed to meet the increasing demand for agricultural products during the last decades (Lobell et al. 2011; Welcker et al. 2022). This demand will continue to increase with the growth of the human population and the diversification of agricultural product uses (Goddard 2009; Tester and Langridge 2010). This necessitates that plant breeders maintain their programs' genetic gain, while also selecting for adaption to more environment friendly practices such as agroecology (Wezel et al. 2014). Genetic diversity of a breeding population is one of the key drivers of genetic gain as it is a determinant of trait genetic variance, which controls expected response to selection per generation (Lush 1937). Elite germplasm in major cultivated species generally

has a narrow genetic basis because modern breeding has exploited only part of the genetic variability that was available in traditional varieties (Maccaferri et al. 2003; Palmgren et al. 2015). During breeding cycles, new lines and varieties are derived mostly from a limited number of crosses between selected elite lines, which further decreases elite genetic diversity (Reif et al. 2005; Mikel and Dudley 2006). This was illustrated, for example, by Allier et al. (2019) who reported a genetic diversity drop over time in a maize breeding program. Recent selection methods such as genomic selection may worsen diversity depletion in the absence of specific constraints (Jannink 2010; Rutkoski et al. 2015; Lin et al. 2016). These trends toward a decreasing diversity level are expected to restrain future genetic gain and may hamper the ability of breeding programs to address new selection objectives related to climate change and evolutions in agronomical practices (McCouch et al. 2013; Mickelbart et al. 2015). Managing diversity in breeding programs can delay its loss and preserve, to some extent, long-term genetic gain (Allier et al. 2020a). Nevertheless, introduction of external diversity sources is required to bring back genetic variation in breeding programs and counterbalance the negative impact of breeding on genetic diversity in elite germplasm (Wray and Goddard 1994; Meuwissen 1997; Woolliams et al. 2015; Allier et al. 2020a).

Since the beginning of modern selection, an intense effort has been made to collect and store diversity sources such as wild relatives, exotic germplasms, landraces, and first-cycle inbred lines (developed from landraces). Gene banks keep these accessions available to breeders thanks to ex situ and in situ conservation (Wang et al. 2017). Compared with elite materials, these resources present a higher genetic diversity and carry original potentially favorable alleles (Maccaferri et al. 2003; Palmgren et al. 2015). They provide a source of alleles to deal with yield stability and abiotic stress but they suffer from lower performances than elite lines (performance gap) due to unfavorable alleles that have been eliminated in elite materials by recent selection and may not be adapted to local conditions (Strigens et al. 2013; Dwivedi et al. 2016). For monogenic and oligogenic traits, using targeted marker-assisted backcross has been a way to introgress favorable alleles from genetic resources into an elite background genome without being hampered by the performance gap (Visscher et al. 1996; Hospital and Charcosset 1997; Frisch and Melchinger 2005). Experimental studies confirm the efficiency of this approach to recover the elite germplasm after single or multiple introgression events (Peng et al. 2014; Han et al. 2017). This backcross procedure has also helped to improve polygenic traits controlled by a few major genomic regions, such as flowering time and yield components under drought conditions in maize (Ribaut and Ragot 2006). However, the success of the introgression may be compromised by the presence of unfavorable alleles in the residual donor genome or negative interactions with the recipient background (Hospital 2005).

The backcross procedure is based on the incorporation of favorable variation at some genomic regions to improve the value of the targeted trait. Its implementation is difficult for quantitative traits which are determined by many regions with minor effects (e.g., grain yield in maize). In this case, the favorable donor alleles are not easily identifiable and they may be eliminated during the backcrossing process (Cowling 2013). Simmonds (1993) described an alternative genetic resource utilization strategy adapted to the improvement of quantitative traits, called "incorporation" (in opposition to "introgression"). In this case, the objective is not to introgress a few targeted genomic regions but to broaden the genetic basis of the elite programs thanks to the incorporation of extrinsic polygenic favorable variations. Simmonds (1993) proposed to first improve genetic resources by recurrent selection to reduce the performance gap with the elite material, i.e., conduct pre-breeding. A recent pre-breeding program aiming at improving flint landraces illustrated the potential but also the difficulties of this process (Ordás et al. 2023). If the pre-breeding progenies still underperform elite, they can be crossed with elite lines to produce a specific buffer population that complements the elite program, i.e., implement a bridging population. The best bridging individuals become potential parents for introduction in the elite breeding program. This strategy aims to close the gap in performance and limit diversity donor introduction's negative impact on short-term genetic gain. In maize, after a number of preliminary unsuccessful efforts reported by Simmonds et al. (1993), this strategy was implemented to incorporate tropical diversity in temperate North-American germplasm by MM Goodman (Lewis and Goodman 2003) and then in the Germplasm Enhancement of Maize (GEM) project (see below). This illustrates that the efforts and methods dedicated to the pre-breeding and bridging populations must be tuned according to the considered genetic resources (Allier et al. 2020a; Sanchez et al. 2023).

The breeding potential of crosses between donor and elite lines in a bridging population is determined by their abilities to generate transgressive individuals. This question can be addressed considering the usefulness criterion concept (UC), which represents the expected performance of the cross progeny after selection (Schnell and Utz 1976). This expected performance is determined by the progeny genetic mean ($\mu$), the selection intensity ($i$), the selection accuracy ($h$) and the progeny genetic standard deviation ($\sigma$) as: $UC_1 = \mu + ih\sigma$. Wolfe et al. (2021) proposed to take only the additive part of the genetic variance in the progenies ($UC_2 = \mu + ih\sigma_a$ here $\sigma_a$ is the progeny additive genetic standard deviation). $UC_2$ is meaningful to evaluate the interest of the progeny as parents of a new breeding cycle. Adding

back-cross generations with the elite parent is expected to diminish the gap in performance at the cost of a reduction in variance, with an optimum in terms of UC values that needs to be determined.

For a breeding company, allocating time and budget to screen potential donors and deploy pre-breeding and bridging approaches causes a reduction of the efforts devoted to the elite breeding programs, which can reduce short-term competitiveness (Pollak 2003). This has prompted the establishment of pre-competitive collaborations to share costs between public institutes and private companies. For instance, the Latin American Maize Project (LAMP, Pollak 1990; Salhuana et al. 1997; Salhuana and Pollak 2006) was crucial for characterizing and evaluating the maize Latin American tropical germplasm accessions. This international project, involving 12 countries, provided breeders with useful yield evaluation and agronomical characterization of these accessions (Salhuana et al. 1997). Later, the Germplasm Enhancement of Maize project (GEM) was conducted to integrate the LAMP accessions in the US elite germplasm. In this project, bridging individuals were derived from three-way crosses involving each a LAMP accession and two elite lines from different private companies (Pollak 2003). For the first time, maize proprietary inbred materials were used in an extensive collaborative program. This project led to the creation of original lines adapted to the Southern and Midwest USA photoperiod conditions (Wang et al. 2017). Other public–private initiatives inspired by the GEM project have emerged, such as the Soybean Asian Variety Evaluation (SAVE) project to characterize Asian soybean varieties (Manjarrez-Sandoval et al. 1998). Sharma and Carena (2012) recorded the use of GEM materials to create dent inbred lines well adapted to the North Dakota conditions.

Other maize germplasms may benefit from similar initiative, among which the European flint group. This group is characterized by early vigor and cold adaptation. It is commonly crossed with the dent pool to create performing hybrid varieties adapted to the Northern European environmental conditions (Messmer et al. 1993; Cartea et al. 1999; Böhm et al. 2017). Results based on haplotyping illustrate that a significant fraction of the diversity of landraces has not been exploited to form the flint breeding pool (Mayer et al. 2020). This bottleneck effect has been amplified by the extensive use of a few first-cycle lines (F2, F7, EP1, DK105, etc.) to create the elite flint lines used in modern breeding (Messmer et al. 1992; Dubreuil and Charcosset 1999; Stich et al. 2005; Van Inghelandt et al. 2010). The preservation of lines derived from European maize landraces and historical lines has made possible to create several diversity panels (Camus-Kulandaivelu et al. 2006; Rincent et al. 2014) later enriched to form a collection of 1191 inbred flint lines (Gouesnard et al. 2017). This collection appears as a reservoir of potential diversity donors, which can be harnessed to incorporate new favorable variations in the elite flint lines and preserve the efficiency of the dent × flint heterotic pattern.

The objective of this study was to document experimentally the potential of genetic resources to bring suitable variation into the flint elite programs. To do so, we established a cooperative pre-breeding multi-parental population. This population consists of a total of 20 BC1 connected families issued from crosses between (i) diversity donor lines chosen for their originality and initial performance and (ii) elite recipient lines from different private partners. For each family, we compared models to estimate the within-family mean performances and the genetic variances to determine their UCs. Results showed that most crosses have the potential to generate positive transgressive materials. It also revealed contrasted variances and UCs among crosses and made it possible to identify the most promising donor for each partner. Finally, our results suggest that one more generation of crossing to elite should in general be beneficial to fully exploit the potential of the genetic resources which were used.

## Material and methods

### Plant material

The experimental material consisted of flint maize BC1S2:3 individuals derived from donors and seven recipient lines. Recipient lines were flint elite material genotypes, each provided by a different partner (Limagrain, RAGT2n, KWS, Masseed, Euralis, Caussade and INRAE). These lines were named A1, A2, A3, A4, A5, A6 and A7 (each number was randomly assigned to one of the partners). Each partner also supplied a dent line complementary to its flint line (referred to as the tester line in the following). Each pair was previously selected by its owner to provide a performing hybrid adapted to the cultivation in the B-C1 French grain precocity group area (early-mid early).

We pre-selected, as potential donors, 74 lines of the same precocity group from a collection of 1191 lines representative of the flint diversity (Gouesnard et al. 2017). The test-cross yield performances of potential donors were evaluated by each private partner using their own tester line and field network (for further information, see File S1). The estimation of donor General Combining Ability (GCA) and other criteria (no agronomic default, no lodging, etc.) led to the selection of seven donor lines (described in Table S1) after discussion among the different partners. Donor lines were named D1, D2, D3, D4, D5, D6 and D7.

Recipient and donor lines were crossed according to an incomplete factorial design discussed in a concerted way among partners (Fig. 1). This design was chosen to
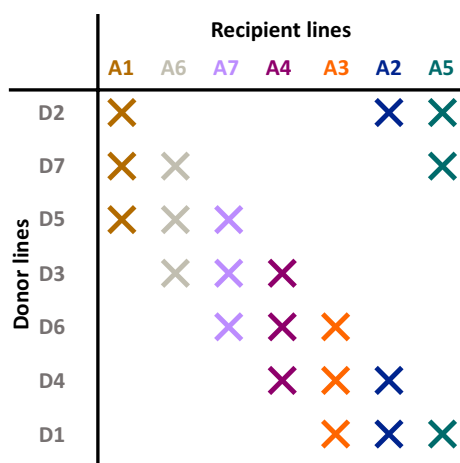
**Fig. 1** Incomplete crossing design between donor lines and recipient lines. Twenty-one F1 single-cross individuals were created by donor × recipient (D×R) crosses (visualized by a cross). Each donor line was crossed with three recipient lines and each recipient line was crossed with three donor lines. Each F1 individual was backcrossed with its recipient parent to produce a BC1 population. For each D×R cross, 60 BC1 plants were self-pollinated during two generations (Single Seed Descent process) to obtain BC1S2 individuals

maximize the connectivity between crosses (i.e., maximize the number of cross failures necessary to disconnect the crossing design into two independent ones). Each F1 single-cross hybrid created by crossing a donor (D) and a recipient line (R) was backcrossed with the recipient line to produce BC1 populations. For each D×R cross, 60 BC1 plants were self-pollinated during two generations (single seed descent process, SSD) to obtain BC1S2 individuals. A total of 1174 BC1S2 plants was obtained due to some failures during the SSD process. The numbers of D×R crosses (21) and of BC1S2 individuals by cross were guided by statistical considerations (e.g., the minimal required number of BC1S2 individuals to estimate the genetic variance of a D×R cross) and total experimental means available. In order to perform testcross phenotypic evaluation (see section *Plant Phenotyping*), each BC1S2 plant was further self-pollinated to generate a BC1S2:3 progeny.

## Plant genotyping

The BC1S2 individuals and the parental lines, with the exception of D3 and D7, were genotyped using a customized Maize Illumina Infinium 25K SNP XT array. D3 and D7 were genotyped with the Maize Illumina Infinium 50K SNP array (Ganal et al. 2011). Thanks to the overlap between both arrays (16587 markers) and the 25K genotyping of their progenies, we imputed the D3 and D7 25K genotyping. The percentage of incoherent markers between parental and progeny genotyping, called

illegitimate rate, was computed. Sixteen BC1S2 individuals were discarded from the dataset because of a high illegitimate rate (> 5%). For A1×D2, A1×D7 and A2×D1 crosses, the average illegitimate rate was initially 8.78%, 7.65% and 8.87%, respectively, which was higher than in other families (0.52% on average). Multiple correspondence analysis (MCA, Fig. S1) showed that D2 and D7 were not the actual donor parents of crosses called A1×D2 and A1×D7, suggesting a labelling error at the beginning of the process. For both crosses, 25K genotyping of a virtual donor line was constructed thanks to the genotyping of A1 and of the BC1S2 individuals. This virtual donor line was compared by identity by state (IBS) with the 74 candidate donor lines. Thus, the QSF3_inra line was identified as the actual A1×D2 donor line; it is named D8 in the following. The A1×D7 donor line was not identified among the 74 lines; the virtual donor line was considered as an additional donor line, named D9 in the following. So, the two corresponding crosses were named as A1×D8 and A1×D9. Lastly, genotypic frequencies of BC1S2 individuals coming from the A2×D1 cross suggested a complex pedigree involving two recipients instead of one (i.e., a three-way cross). This cross was removed from the dataset for variance decomposition analysis (see below). Considering the updated pedigree, genotypic missing data were imputed using AlphaPlantImpute software v1.1 (Gonen et al. 2018). Parameters used for imputation are provided in File S2.

## Plant phenotyping

BC1S2:3 individuals were crossed to the dent tester line associated to the recipient line to produce hybrid progenies. In the following, hybrids derived from a same RxD cross are considered as a family. Hybrids were evaluated in 3 different locations in France (Blois, Loir-et-Cher; Saint-Martin de Hinx, Landes; Villers, Pas-de-Calais) in 2019 (further referred to as Blo19, Smh19 and Vil19). Each trial was composed of 1512 elementary two-row plots with an area of 9.6 m$^2$, 9.3 m$^2$ or 10.9 m$^2$, for Blo19, Smh19 and Vil19, respectively. Plant density was 9.4, 8.6 and 9.2 seed.m$^{-2}$, respectively. Hybrids derived from the same recipient line were gathered in a same sub-trial to minimize experimental errors in the evaluation of hybrids derived from the same recipient line. To balance the number of plots allocated to each recipient line, each sub-trial had the same dimension (216 plots). Each sub-trial was divided into 12 blocks. In each block, a commercial hybrid (ADEVEY) and the reference hybrid, corresponding to the single-cross hybrid between the recipient line and the tester line, were sown and considered as checks. In each sub-trial, the average number of experimental hybrid repetitions varied from 1.07 to 1.56

depending of the family. Hybrid randomization in each sub-trial was performed using the nursery function of the Digger R package to ensure a homogeneous repartition of checks and repeated hybrids in each sub-trial (Coombes 2009).

Hybrids were phenotyped for male and female flowering times (FLOM and FLOF in days after planting), anthesis-silking interval (ASI in days), grain yield at 15% moisture (GY in Mg/ha), grain humidity at harvest ($H_2O$ in %) and plant height (PH in cm). FLOM and FLOF are the day at which 50% of the plants exhibited anthers or silks, respectively. ASI is the difference in days between FLOM and FLOF. PH was not measured in Vil19. In Blo19 and Smh19, plots with a number of plants lower than the median number minus 15 plants were eliminated. In Vil19, final plant density was not recorded and fresh grain weight was used as a proxy to filter aberrant plots (plots with a fresh grain weight bellow 7 kg were eliminated). A lower germination rate was observed in the A7 sub-trial which led to the elimination of 33%, 28% and 14% of plots in this sub-trial in Blo19, Smh19 and Vil19, respectively. On average, 1.3% of the data were eliminated in the other sub-trials.

The raw phenotypic data were corrected for spatial effects predicted using AR1×AR1 models (see File S3 for details). These models were fitted using ASREML-R v4 (Butler et al. 2017). A commercial grain yield index (YI) was calculated afterwards according to the following index:

$$YI_m = GY_m + 0.25 * \left(H_2O_{ADEVEY} - H_2O_m\right)$$

where $GY_m$ and $H_2O_m$ are the corrected grain yield and humidity values of plot $m$ and $H_2O_{ADEVEY}$ is the trial mean grain humidity value of the check hybrid ADEVEY. This index corresponds to the one usually considered for variety registration in France. It penalizes the experimental hybrids that mature later than the reference (i.e., ADEVEY).

## Variance decomposition

Different variance decomposition models were used for multi-trial analysis (Table 1). They allowed us to estimate the total genetic variance and the additive genetic variance, in the whole population and within each family. The variance–covariance matrix of parameter estimates was computed and used to test the effects of recipients and donors on these parameters. Parameters of all models were estimated with the package MM4LMM in R (Laporte et al. 2022).

## Estimation of global genetic variance

The following model was used to estimate global genetic variance:

$$Y_{ilhm} = \mu + \alpha_i + \tau_l + (\alpha\tau)_{il} + G_h + GE_{ih} + e_{ilhm}(M\_G)$$

$$G \sim N\left(0,\ I\sigma_G^2\right),\ GE_i \sim N\left(0,\ I\sigma_{GE_i}^2\right) \text{ ind,}$$

$$e_i \sim N\left(0,\ I\sigma_{e_i}^2\right) \text{ ind,}\quad G \perp GE_i \perp e_i$$

where $Y_{ilhm}$ is the phenotypic value of the repetition $m$ of check $l$ or experimental hybrid $h$ in trial $i$. $\alpha_i$ is the fixed effect of the trial, $\tau_l$ is the fixed genetic effect accounting for the difference between checks (7 reference hybrids and ADEVEY) and experimental hybrids (a factor with 9 levels: one level for each check and a supplemental level for the experimental hybrids),$(\alpha\tau)_{il}$ is an interaction term between the trials and the checks, $G_h$ is the random genetic effect of experimental hybrids, $GE_{ih}$ is an interaction term between the trial $i$ and the experimental hybrid $h$ and $e_{ilhm}$ is the error term (the errors are considered independent between trials). Symbol ($\perp$) indicates random effects which are considered independent of each other. In the following, this model is referred to as M_G (Table 1).

**Table 1** Summary of fixed and random effects declared to model genetic effects in the presented linear mixed models of variance decomposition

| Model | Pedigree structure in fixed part | Random genetic effects | |
|---|---|---|---|
| | | Genetic (or permanent) effect | Additive effect |
| M_G | | Global | – |
| M_FG$_S$ | ✓ | Family specific | – |
| M_FG | ✓ | Global | – |
| M_FA$_S$P$_S$ | ✓ | Family specific | Family specific |
| M_FA$_S$ | ✓ | – | Family specific |
| M_FAP | ✓ | Global | Global |
| M_FA | ✓ | – | Global |

The second column indicates the presence of a fixed recipient effect and a fixed family effect in the model. The third column sums up how the random genetic effects are declared in each model

## Estimation of within-family genetic variance

Each hybrid family was characterized by a recipient line (with its associated tester) and a donor line. In the following models, the pedigree structure was modeled by two fixed effects: a recipient effect and a family (recipient × donor combination) effect. As each tester was associated to a specific recipient line, the recipient effect captured the effect of a recipient-tester combination.

The following model was used to estimate the within-family genetic variances:

$$Y_{ijkhm} = \mu + \alpha_i + \rho_j + \theta_k + (\alpha\rho)_{ij} + (\alpha\theta)_{ik} + G_{kh} + GE_{ikh} + e_{ijkhm} \ (M\_FG_S)$$

$$G_k \sim N\left(0, \ I\sigma^2_{G_k}\right) \text{ ind}, \ GE_{ik} \sim N\left(0, \ I\sigma^2_{GE_{ik}}\right) \text{ ind},$$

$$e_i \sim N\left(0, \ I\sigma^2_{e_i}\right) \text{ ind}, \ G_k \perp GE_{ik} \perp e_i$$

where $Y_{ijkhm}$ is the phenotype of the repetition $m$ of hybrid $h$ in family $k$ derived from the recipient line $j$ (or the recipient line if the hybrid is a reference hybrid) in trial $i$. $\alpha_i$ is the fixed effect of the trial, $\rho_j$ is the fixed effect of the recipient line (a factor with 8 levels: one level for each recipient line and an additional level for ADEVEY) and $\theta_k$ is the fixed effect of the family (a factor with 21 levels: one level for each hybrid family and an additional level for ADEVEY and the reference hybrids). $(\alpha\rho)_{ij}$ is a fixed interaction term between the trial $i$ and the recipient $j$ and $(\alpha\theta)_{ik}$ is a fixed interaction term between the trial $i$ and the family $k$. $G_{kh}$ is a random genetic effect (with a specific variance per family) and $GE_{ikh}$ is a random interaction term between the trial $i$ and the genotype $h$. $e_{ijkhm}$ is the error term. In the following, this model is referred to as $M\_FG_S$ (Table 1).

The same model was also fitted considering that the genetic effect and the associated interaction term had variances common to all families. It is referred to as M_FG in the following (Table 1). Multi-trial adjusted means were estimated for each hybrid using an alternative M_FG model considering the hybrid genetic effect as fixed. For each trait, the difference between the multi-trial mean values of each family and the value of its reference hybrid has been tested following the procedure described in the File S4.

## Estimation of within-family additive genetic variance

The within-family additive variances were estimated thanks to the following model:

$$Y_{ijkhm} = \mu + \alpha_i + \rho_j + \theta_k + (\alpha\rho)_{ij} + (\alpha\theta)_{ik} + A_{kh} + P_{kh} + AE_{ikh} + PE_{ikh} + e_{ijkhm} \ (M\_FA_SP_S)$$

$$A_k \sim N\left(0, \ K_k\sigma^2_{A_k}\right) \text{ ind}, \ P_k \sim N\left(0, \ I\sigma^2_{P_k}\right) \text{ ind},$$

$$AE_{ik} \sim N\left(0, \ K_k\sigma^2_{AE_{ik}}\right), PE_{ik} \sim N\left(0, \ I\sigma^2_{PE_{ik}}\right) \text{ ind},$$

$$e_i \sim N\left(0, \ I\sigma^2_{e_i}\right) \text{ ind, and } A_k \perp P_k \perp AE_{ik} \perp PE_{ik} \perp e_i$$

The fixed terms of this model were similar to the fixed terms of the model ($M\_FG_S$). The random part of the model was composed of an additive genetic effect $A_{kh}$, a genetic permanent effect $P_{kh}$, an interaction term between the trials and the additive genetic effect $AE_{ikh}$ and an interaction term between the trials and the genetic permanent effect $PE_{ikh}$. $e_{ijkhm}$ is the error term. The permanent effect modeled non-additive genetic effects within each family. It can be estimated because the experimental hybrids have been partially repeated. A specific variance was estimated per family for each random term. This model is referred to as $M\_FA_SP_S$ (Table 1).

In this model, the variance–covariance matrix of additive terms depends on marker-based kinship matrices ($K_k$). Each $K_k$ matrix was specific to a family. Its coefficients were computed using the genotyping of the BC1S2 individuals derived from the cross, according to the Natural and Orthogonal Interaction Approach (NOIA, Álvarez-Castro and Carlborg 2007), as recommended by Vitezica et al. (2017). For a given individual $i$ (of the family $k$), at a given marker $j$, coefficients were calculated using genotypic frequencies as follows:

$$h_{A_{ijk}} = \begin{cases} -\left(-p_{Bb_{jk}} - 2p_{bb_{jk}}\right) \\ -\left(1 - p_{Bb_{jk}} - 2p_{bb_{jk}}\right) \\ -\left(2 - p_{Bb_{jk}} - 2p_{bb_{jk}}\right) \end{cases} \text{ for genotypes} \begin{cases} BB \\ Bb \\ bb \end{cases}$$

where $p_{Bb_{jk}}$ and $p_{bb_{jk}}$ are the frequencies of genotypes Bb and bb in the family $k$, respectively.

Additive kinship matrices were obtained with the formula:

$$K_k = \frac{H_{A_k}H'_{A_k}}{tr\left(H_{A_k}H'_{A_k}\right)/n_k}$$

where $n_k$ is the number of individuals in family $k$ and

$$H_{A_k} = \begin{pmatrix} h_{A_{11k}} & \cdots & h_{A_{1m_kk}} \\ \vdots & \cdots & \vdots \\ h_{A_{n1k}} & \cdots & h_{A_{nm_k}} \end{pmatrix}$$

where $m_k$ is the number of polymorphic markers in the family $k$.

To test the pertinence of the permanent effect in this context, we also fitted a sub model $M\_FA_S$ including only the additive genetic effects ($A_k$ and $AE_{ik}$). We considered also two

other models (M_FA and M_FAP) where the additive and permanent effect variances were declared homogeneous between families (Table 1).

## Heritability estimation

A multi-trial heritability was computed for each trait using the variance parameter estimates of the M_G model:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\overline{\sigma_{GE}^2}}{n_{\text{trial}}} + \frac{\sum_{i=1}^{n_{\text{trial}}} \sigma_{e_i}^2 / n_{\text{rep}_i}}{n_{\text{trial}}}}$$

where $\sigma_G^2$ is the genetic variance, $\overline{\sigma_{GE}^2}$ is the average of the genotype × trial variances. $n_{\text{trial}}$ is the number of trials ($n_{\text{trial}} = 3$), $\sigma_{e_i}^2$ is the error variance in the trial $i$ and $n_{\text{rep}_i}$ is the mean number of repetition in the trial $i$ (Piepho and Möhring 2007).

## Usefulness criterion calculation

For performance traits (GY and YI), we computed two usefulness criteria (UC$_1$ and UC$_2$) to provide information about the expected response to selection in each family $k$:

$$UC_{1k} = \hat{\mu}_k + ih\hat{\sigma}_{G_k}$$

$$UC_{2k} = \hat{\mu}_k + ih\hat{\sigma}_{A_k}$$

where $\hat{\mu}_k$ is the adjusted mean for the family effect in the model M_FA$_S$P$_S$, $\hat{\sigma}_{A_k}^2$ is the additive variance and $\hat{\sigma}_{G_k}^2 = \hat{\sigma}_{A_k}^2 + \hat{\sigma}_{P_k}^2$ is the total genetic variance where $\hat{\sigma}_{P_k}^2$ is the permanent effect variance. $\hat{\sigma}_{A_k}^2$ and $\hat{\sigma}_{P_k}^2$ were estimated with the model M_FA$_S$P$_S$. $h$ is the selection accuracy and $i$ is the selection intensity. The selection accuracy was assumed to be one, as would be the case when selecting directly on genetic effects (Zhong and Jannink 2007). Usefulness criteria were calculated with $i = 2.07$ (selection rate of 5%).

UC$_{1k}$ is an estimation of the expected performance of individuals selected in the family $k$. UC$_{2k}$ is an estimation of the expected additive value transmitted by selected individuals to the next generation.

## Results

### Genetic variance and multi-trial heritability

Global variance components were estimated thanks to the model M_G (Table 2). The broad-sense heritability was high for all traits. Flowering traits (FLOM and FLOF) had higher heritabilities (0.82 and 0.86) than GY (0.68). FLOM had a lower genetic variance than FLOF. For flowering traits, we noticed a lower G×E variance and a stronger error term variance for trial Vil19. For PH, variance decomposition revealed a clear difference between the two trials where this trait was measured, error variance in Blo19 being more

**Table 2** Variance decomposition with the models M_G (genetic effects modelled as a global genetic random effect) and M_FG (addition of recipient and family fixed effects) for each trait

| Trait | Model | AIC | BIC | H$^2$ | Variance components | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | | $\sigma_G^2$ | $\sigma_{GE}^2$ | | | $\sigma_e^2$ | | |
| | | | | | | Blo19 | Smh19 | Vil19 | Blo19 | Smh19 | Vil19 |
| FLOM | M_G | 5684 | 5738 | 0.82 | **1.40** (0.07) | **0.62** (0.09) | **0.63** (0.06) | **0.02** (0.08) | **0.67** (0.06) | **0.26** (0.02) | **0.91** (0.07) |
| | M_FG | 4963 | 5109 | – | **0.94** (0.05) | **0.20** (0.07) | **0.37** (0.04) | **0** (−) | **0.66** (0.05) | **0.27** (0.02) | **0.89** (0.04) |
| FLOF | M_G | 6468 | 6523 | 0.86 | **2.30** (0.11) | **1.15** (0.1) | **0.40** (0.07) | **0.22** (0.08) | **0.57** (0.05) | **0.39** (0.03) | **0.85** (0.07) |
| | M_FG | 5640 | 5785 | – | **1.28** (0.07) | **0.45** (0.07) | **0.33** (0.06) | **0.15** (0.08) | **0.57** (0.05) | **0.39** (0.03) | **0.87** (0.07) |
| ASI | M_G | 2934 | 2989 | 0.63 | **0.51** (0.02) | **0.55** (0.07) | **0** (−) | **0.72** (0.1) | **0.62** (0.05) | **0.01** (0) | **1.05** (0.08) |
| | M_FG | 640 | 786 | – | **0.04** (0.01) | **0.29** (0.06) | **0.02** (0.01) | **0.31** (0.09) | **0.63** (0.05) | **0.01** (0) | **1.09** (0.08) |
| PH | M_G | 16273 | 16306 | 0.69 | **112.97** (7.64) | **70.52** (14.02) | **18.75** (6.5) | – | **128.80** (10.66) | **27.98** (2.29) | – |
| | M_FG | 15481 | 15568 | – | **47.07** (4.53) | **18.27** (12.47) | **24.51** (5.04) | – | **134.76** (11.13) | **28.07** (2.29) | – |
| GY | M_G | 22552 | 22606 | 0.68 | **0.48** (0.03) | **0.26** (0.05) | **0.17** (0.06) | **0.17** (0.05) | **0.50** (0.04) | **0.61** (0.05) | **0.47** (0.04) |
| | M_FG | 21498 | 21644 | – | **0.20** (0.02) | **0.11** (0.04) | **0.11** (0.05) | **0.06** (0.04) | **0.49** (0.04) | **0.61** (0.05) | **0.46** (0.03) |
| H$_2$O | M_G | 5870 | 5925 | 0.75 | **0.96** (0.06) | **0.86** (0.09) | **0.20** (0.06) | **1.96** (0.12) | **0.56** (0.05) | **0.43** (0.03) | **0.36** (0.03) |
| | M_FG | 4200 | 4346 | – | **0.45** (0.03) | **0.53** (0.07) | **0** (−) | **0.64** (0.06) | **0.57** (0.05) | **0.39** (0.02) | **0.36** (0.03) |
| YI | M_G | 22782 | 22836 | 0.68 | **0.48** (0.03) | **0.26** (0.06) | **0.18** (0.06) | **0.22** (0.05) | **0.53** (0.04) | **0.66** (0.05) | **0.47** (0.04) |
| | M_FG | 21765 | 21910 | – | **0.19** (0.02) | **0.10** (0.05) | **0.11** (0.06) | **0.10** (0.04) | **0.53** (0.04) | **0.67** (0.05) | **0.47** (0.04) |

AIC and BIC criterion are indicated for each model. H$^2$ is the broad-sense multi-trial heritability estimated thanks to the model M_G. The estimations of variance components are indicated in bold and their standard errors are given in parenthesis

than four times larger than in Smh19. We observed a similar trend for G×E terms. YI and GY had similar heritabilities and genetic variances. However, YI had stronger G×E variances than GY. For both YI and GY, the Smh19 trial showed higher error variances than the two other trials.

For all traits, AIC and BIC values (the smaller the better) showed that the inclusion of a recipient and family fixed effects were beneficial (model M_FG in Table 2). The error variances were similar between M_G and M_FG. The estimated genetic variances were lower with model M_FG than with model M_G, highlighting that including pedigree structure as fixed effects in model M_FG absorbed part of the genetic variation. The remaining within-family genetic variance was particularly low for ASI. For other traits, it represented between 40 and 67% of the global genetic variance. We also noticed a diminution of the G×E interaction variances when including family pedigree and corresponding interaction terms as fixed effects.

## Mean performance of hybrid families

Adjusted means of the reference hybrids and hybrid families were computed by environment and over environments

with the model M_FG (Tables S2 and S3). We observed a strong variability of the phenotypic values of the reference hybrids across environments. The Smh19 trial, located in a warmer and drier climate than the two other trials, showed a faster flowering (55 DAP on average against 92 DAP in other trials) and an earlier maturity at harvest (mean $H_2O$ value of 23% in Smh19 against 28% in other trials). Hybrids were also taller in Smh19 (306 cm on average against 270 cm in other trials). For GY, the hybrids were less productive in Vil19, with a yield reduction of 2.20 Mg/ha on average compared to Smh19 and Blo19. We noticed systematic lower productivity of the reference hybrid corresponding to the recipient line A7 and its associated hybrid families (A7D3, A7D5, and A7D6).

For the flowering time (FLOF and FLOM), the average values of hybrid families were close to the reference hybrid value with a slight trend toward later flowering (one or two days, Fig. 2 and Table S4). This difference was significant for 12 families for FLOM and 18 families for FLOF. For all families, except A4D3 and A5D2, more than half of individuals had later male and female flowering time than the corresponding reference hybrid (Table S5). Half of the hybrid families did not show significant different H2O mean values
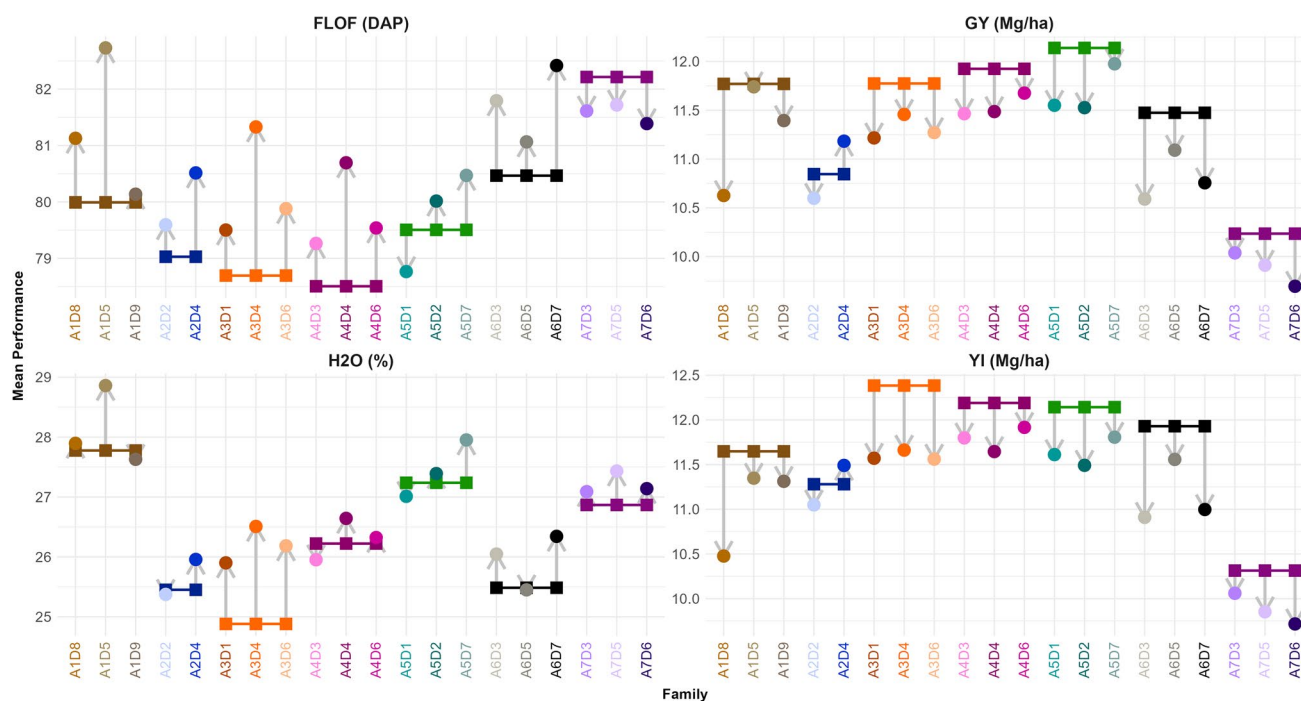


**Fig. 2** Mean performances of families and reference hybrids for FLOF, GY, $H_2O$ and YI. The dots indicate the family mean performances (colored by family) and the squares indicate the reference hybrid performances (family derived from the same recipient line are compared to the same reference). Mean performance of a family or a reference hybrid corresponds to its adjusted mean across the trials computed thanks to the estimated parameters of the model M_FG. Grey arrows highlight the difference between each family and its reference (color figure online)
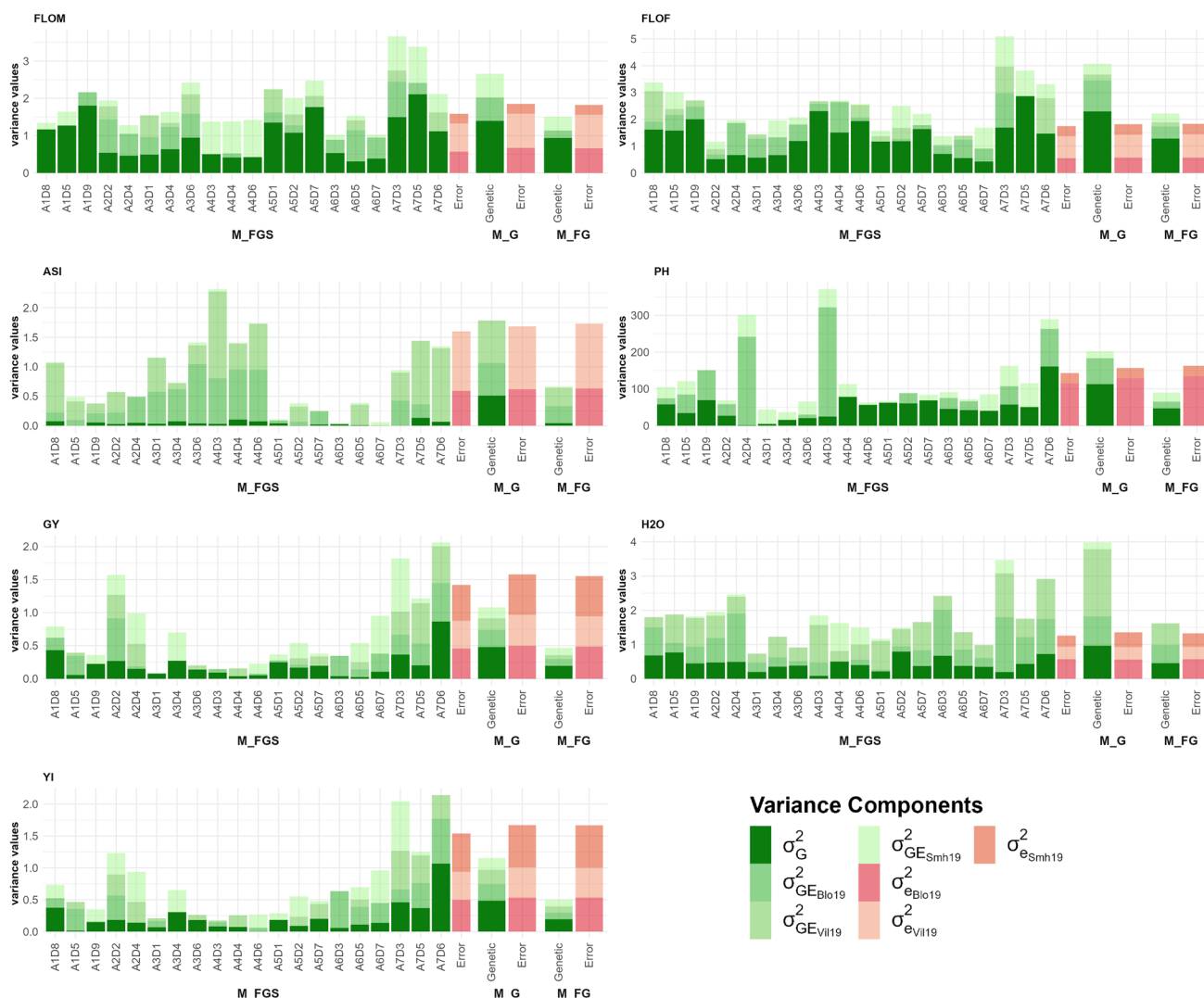
**Fig. 3** Variance decomposition with the model M_FG$_S$ and within-family genetic variance estimation for each trait. For each family, genetic (in dark green) and G×E (in shades of light green) variances are showed. Error term variances are also displayed (in shades of pink). For comparison, variances estimated with the models M_G and M_FG are presented on the right part of each graph with the same color code (color figure online)

compared to their reference hybrid (Table S4). For the other half, the difference ranged from 0.4 to 1.6%.

The donor introgressions came with a reduction of GY and YI for the different families (−0.42 and −0.53 Mg/ha on average, respectively). Only the A2D4 family had significantly higher average GY than its corresponding reference hybrid. Others displayed a large variation in performance loss: between 0.03 and 1.14 Mg/ha for GY and 0.3 and 1.17 Mg/ha for YI. We noticed that the loss of GY and YI compared to the reference hybrid was significant for the majority of families (13 families for GY and 16 families for YI). We distinguished different patterns of loss according to the recipient line. For example, the cross of A3 with three distinct donor lines led to hybrid families with similar average GY (A3D1, A3D4, and A3D6). In contrast, the hybrid

families derived from A1 had contrasted mean productivities, with a difference of 1.10 Mg/ha between A1D8 and A1D9 for GY. Only one family (A6D7) had no individual with better YI performance than the corresponding reference hybrid. Other families displayed between 5 and 63% of individuals with YI adjusted mean values superior to that of the corresponding reference hybrid (Table S5).

## Within-family genetic variance comparison

The model M_FG$_S$ gave us access to the genetic and G×E variances associated with each family (Fig. 3 and Table S6). Likelihood ratio tests between M_FG and M_FG$_S$ showed an advantage of considering specific genetic variances between families for all traits (Table S7). For all traits, the mean of
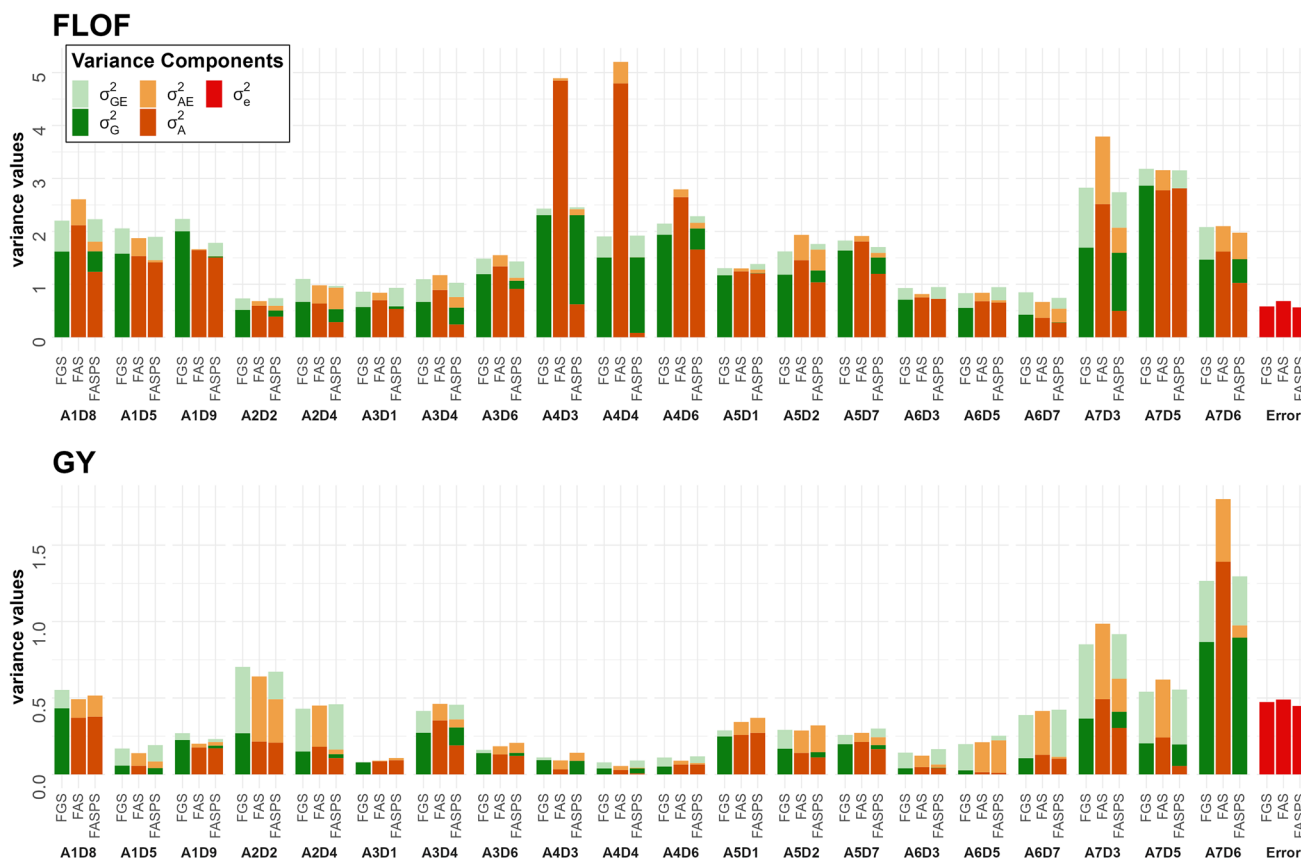
within-family genetic variances estimated with M_FG$_S$ was similar to the common genetic variance estimated with the model M_FG. The error variances were alike between both models (Tables 2 and S6). AIC values were lower with the model M_FG$_S$. However, we noticed the BIC values were higher for this model.

For FLOM, the within-family genetic variances varied from 0.31 to 2.11. These values depended to a large extent on the recipient line used to generate the families (see contrast test procedure explained in File S4 and associated results of pairwise tests in Fig. S2). The families derived from the recipient lines A1, A5, and A7 had significantly higher genetic variances than other families. We also observed the influence of some donor lines on FLOM genetic variances. For example, the families derived from D4 presented significantly lower genetic variances than those derived from D1, D5, D7, D8 or D9. Compared to FLOM, genetic variance values were slightly higher for FLOF (from 0.43 to 2.86). The families derived from A1, A4, A5, A7 presented a higher FLOF variance than others. For ASI, all families had low genetic variances and high G×E variances, in agreement

with the global variance decomposition provided by the model M_FG. We did not detect any systematic effect of parental lines on the genetic variance for this trait (Fig. S2).

The within family variances estimated with M_FG$_S$ model were rather similar for GY and YI, with a Pearson coefficient correlation of 0.95. The ranges of genetic variances were 0.02–0.87 and 0.03–1.07 for GY and YI, respectively. For GY, we noticed significantly higher genetic variances for families derived from A1 and A7 than from A4 and A6 (Fig. S2). For YI, families whose recipient parent was A7 had significantly larger genetic variance values than all other families. We observed no systematic effect of donor lines on the within-family genetic variances. However, we observed variation between families sharing the same recipient line. For instance, for GY, A1D9 had a genetic variance four times higher than A1D5.

The range of the genetic variances was reduced for H$_2$O (0.09–0.79). PH genetic variances were also relatively homogenous between families. We identified three families with extreme variance values: low for A2D4 (1.37) and A3D1 (4.85), and high A7D6 (160.99). A2D4 and A4D3



**Fig. 4** Within-family genetic and additive variance estimation with the models M_FG$_S$, M_FA$_S$ and M_FA$_S$P$_S$ for FLOF and GY. For M_FA$_S$ and M_FA$_S$P$_S$, within-family additive variances are indicated in orange. The within-family genetic (model M_FG$_S$) and permanent effect (model M_FA$_S$P$_S$) variances are in dark green. The interaction (A×E: light orange and G×E: light green) and error (red) terms are trial specific and represented by their respective average values (color figure online)
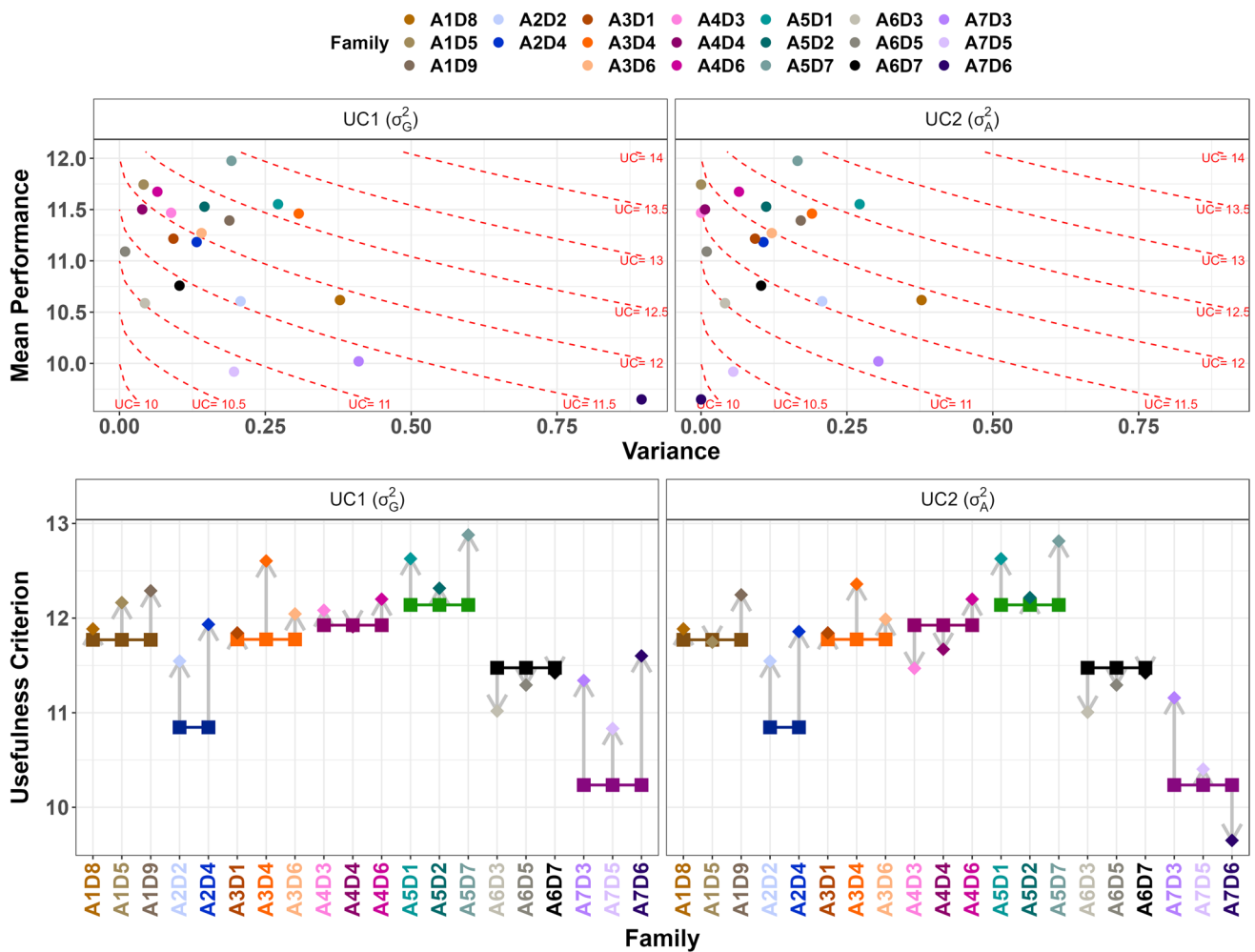
**Fig. 5** Usefulness Criterion (UC) prediction for each family and comparison with the reference hybrid performance for GY. The top charts present the relation between family mean performance and genetic variance components (left: within-family genetic variances, right: within-family additive variances). Dots are colored by family and dotstood out from other families with a high G × E interaction. We observed no systematic effect of donor or recipient lines on variances for these traits.

ted red lines are isoclines of UC. The bottom charts display the UC values for the different families (diamonds). For comparison, the performances of reference hybrids were added on the graphics (squares). Grey arrows visualize the potential gain for each family (color figure online)

## Within-family additive variance comparison

We obtained estimations of the within-family additive variances using the models M_FA$_S$ and M_FA$_S$P$_S$, which included a family-specific permanent effect (Tables S8 and S9). For all traits, M_FA$_S$ yielded lower AIC and BIC values than M_FA$_S$P$_S$. Both values were close for FLOM and FLOF. M_FA$_S$P$_S$, presented AIC values lower than M_FG$_S$ for all traits but YI. Similarly to the genetic variances (see above), likelihood ratio tests between M_FA and M_FA$_S$, and between M_FAP and M_FA$_S$P$_S$, showed that declaring family specific variances led to a better fit than considering

a homogenous variance for the additive and permanent effect terms.

We noticed large variations in the ratio between the additive variances estimated with M_FA$_S$ and the corresponding genetic variances estimated with M_FG$_S$ across traits and families (Figs. 4, S3 and Table S10). On average over families, this ratio ranged from 0.67 (ASI) to 1.28 (FLOF) excluding A7D3 for ASI which had close-to-zero genetic and additive variances. For FLOF and FLOM, some families such as A4D4 had additive variances which were up to three times their genetic variance values. Adding a family-specific permanent effect (model M_FA$_S$P$_S$) reduced the mean ratio between the additive variances and the M_FG$_S$ genetic variances. For example, it dropped from 1.23 and 1.28 to 0.71 for FLOF and FLOM. This diminution appeared stronger for GY and YI,

notably because of an estimation close to zero of the additive variance in some families.

The sum per family of the additive and permanent effect variances (model M_FA$_S$P$_S$) was close to the M_FG$_S$ genetic variances for all traits except ASI (Table S10). For this trait, the permanent effect variances were estimated close to zero, and the average ratio of additive variances over genetic variances was 0.58 (excluding A7D3). The additive part of the genetic variances was higher for FLOM, FLOF, and H$_2$O (from 0.73 to 0.83) than for PH, GY, and YI (from 0.6 to 0.7). For these six traits, this ratio varied strongly across families. For example, A1D8 and A7D3 had similar GY genetic variance values, but the additive part was higher for the first one (Fig. 4, Table S10). This variation was also observed between families derived from the same recipient.

## Family ranking based on the usefulness criterion

The GY UC$_1$, which estimates the expected mean of the top 5% selected individuals, was superior to the adjusted mean of the reference line for all families except for the families derived from A6 and A4D4 (Fig. 5). For YI, eleven families had UC$_1$ values higher than the reference value. The average expected gain was 0.41 Mg/ha for GY and 0.29 Mg/ha for YI. These values rose to 0.54 and 0.55 Mg/ha if we considered only the families with higher UC$_1$ values than the reference value. They corresponded to an expected mean gain of 5% for both traits. The maximal potential gain was observed for A7D6, which reached 13% for GY and 14% for YI. The family expected gains decreased as the reference line's initial performance increased, with a correlation coefficient of $-0.72$ for GY and $-0.89$ for YI. We also noticed that the UC$_1$ values across families were inferior to the corresponding best individual adjusted means, with a mean difference of 0.6 Mg/ha for GY and YI. Nevertheless, both quantities correlated well (0.78 for GY and 0.80 for YI).

For GY, the family rankings based on UC$_1$ were identical to those based on mean performance for recipients A2, A3, and A5, whereas they differed for other recipients (Figs. 2 and 5, Table S11). For example, for A1, the family A1D9 had a lower mean performance than A1D5 but a higher genetic variance, leading to a superior UC$_1$ value. We also noticed that A1D3 had a poor mean performance compared to A1D5 and A1D9 (10.6 Mg/ha against 11.7 and 11.4 Mg/ha) but a rather similar UC$_1$ value. Using UC$_1$ allowed us to highlight the difference between A5D1 and A5D2 too, which had similar mean performances (1.15 Mg/ha) but contrasted genetic variance values (A5D1: 0.27 and A5D2: 0.15). For YI, the rankings established with UC$_1$ and mean performance were similar for families derived from A2, A3, A5, and A6. Similarly to GY, using UC$_1$ allowed us to underline families

with high genetic variances, such as A1D9, which appeared superior to A1D8 and A1D5 (Table S11). Although the A2 reference hybrid had lower performance than the A1, A3, A4 and A5 reference hybrids (Fig. 1), A2D4 had similar GY and YI UC$_1$ values than families issued from other recipients.

Considering only the additive part of the genetic variance to compute the usefulness criterion (UC$_2$) led, by definition, to lower values. It also affected the within-recipient family ranking, although UC$_1$ and UC$_2$ correlated well (0.80 for GY and 0.73 for YI, Fig. 5; Table S11). The ranking differed for four recipients for GY and three for YI. For example, the families from A4 had similar GY UC$_1$ values, but only A4D6 had a substantial additive variance, leading to a higher UC$_2$. The higher the genetic variance and the lower the additive part was, the more the family was penalized with UC$_2$ compared to UC$_1$. The most extreme loss was observed on YI for A7D6 (UC$_1$ = 11.8 Mg/ha and UC$_2$ = 9.6 Mg/ha).

# Discussion

## Impact of the introduction of non-elite material on flowering time and yield performance

The selection of the donor lines was driven by a collective choice made by the seven partners of the project, which led to consider (i) originality relative to main founders of the flint genetic groups, (ii) hybrid performance, (iii) limited agronomic defaults, in particular lodging and (iv) phenology compatible with the targeted environment (File S1). Regarding this fourth factor, choice was restrained among those with the same precocity as the recipient line (B-C1 maturity groups according to French nomenclature) with the aim to produce progeny with phenology adapted to targeted environmental conditions. This pre-selection of the donor lines for the precocity was made possible by the assembly and evaluation of an European flint panel encompassing more than 1000 lines (Gouesnard et al. 2017). The incorporation of donors had a limited impact on flowering time average in progenies (a mean delay of one or two days) which will facilitate their use in the breeding programs of the private partners. This illustrates that characterization of genetic resources for adaptive traits such as flowering time to preselect donor lines is one of the keys to their efficient use in breeding programs.

Even though they passed a first selection step, diversity donors used in our study showed an average YI performance gap with elite lines of 2 Mg/ha (File S1). It was expected that their incorporation into elite material would result in a loss of performance (Pollak 2003; Allier et al. 2020a) but we also expected to identify transgressive individuals in the

different families. Experimental evaluation of progenies in our study showed an average loss 0.57 Mg/ha excluding A2D4. This is consistent with the initial performance gap between donor and recipient lines and the performed cross type (BC1), which leads to anticipate 25% of the initial gap. For the A2D4 family, donor introduction led to a gain of 0.21 Mg/ha. This average gain was mainly due to a lower performance of the A2 reference hybrid in the Vil19 environment (Table S2). For the other families, the average performance loss was variable, ranging from 0.21 to 1.2 Mg/ha. Beyond the mean value, the progeny performances depend on the genetic variance generated by the cross. We noticed substantial YI within-family genetic variances. This is in accordance with several simulation studies which showed that crosses between parents with large performance differences may lead to high progeny variance (Mohammadi et al. 2015; Lado et al. 2017; Neyhart and Smith 2019). However, a large performance difference is not sufficient as illustrated by the presence of families with genetic variances close to zero. Wide differences in genetic variance were observed both between families derived from different recipients or the same one. Similar to our study, the experimental evaluation of a nested association mapping population of European flint maize identified a large range of genetic variances between half-sibling families sharing all a common parent (Bauer et al. 2013; Lehermeier et al. 2014). For most of the families, the level of variance in progenies was high enough to counterbalance the average performance loss and led to transgressive progenies with better performances than their recipient parent. The presence of transgressive individuals confirmed the interest of incorporating diversity donors in the flint maize elite material. Consistently with the observation of such transgressive individuals, the computation of the usefulness criterion showed an expected mean gain of 5% over the recipient parents after selecting the top 5% individuals within each family.

## Causes of genetic variance variations between families

We noticed a disparity in genetic variance values across families and traits. On average, the additive part of the genetic variance was higher for FLOF, FLOM and $H_2O$ than for other traits. The level of genetic variance in the progeny appeared to be more impacted by the recipient parent than by the donor parent. This trend was clear for FLOF and FLOM, for which the families from A1, A5 and A7 had a higher genetic variance. For yield-related traits, crossing donor lines with A1 and A7 led to higher genetic variances than other recipient lines. This preponderant influence of the recipient parent may first be linked to the backcrossing procedure. For some recipient lines, an involuntary selection in

BC1 individuals during material creation might have led to the loss of some donor alleles (Ødegård et al. 2009; Neyhart and Smith 2019), causing a reduction of variance in progenies. However, the observed within-family marker-based diversity values are only slightly inferior to the theoretical ones which suggests a limited selection has been made during the material creation (Table S12).

The genetic variance differences between families may be linked to a difference in parental genetic distances: simulation work by Beckett et al. (2019) showed that the range of possible genetic variances increases as the genetic parent distance between parents becomes larger. The genetic distance between the pairs of donor and recipient used as parents in our design was slightly correlated to the FLOM within-family genetic variance (Pearson coefficient: 0.40, Fig. S4). For other traits, the genetic distance was poorly correlated to the genetic variance, which confirmed the results of other studies (Mohammadi et al. 2015; Beckett et al. 2019). This lack of correlation may be due to the genetic distance which is computed as a whole-genome relatedness based on neutral markers and not on the QTLs that contribute to the variance (Hung et al. 2012). One can expect that this relationship would be improved by accounting for QTL effects, as supported by results regarding heterosis prediction in wheat (Boeven et al. 2020).

The difference between within-family genetic variances may also be attributable to a characteristic of our experimental design: a specific tester line was used to evaluate the progenies of a given recipient line. This specificity allowed each private company to compare its new materials to its reference hybrid. In presence of dominance, using a tester that accumulates a high number of dominant favorable alleles leads to a reduction of the genetic variance in the hybrid population (Rawlings and Thompson 1962; Hallauer et al. 1988). This may explain why some families sharing the same recipient parent have comparable variances (e.g., low for A4 for GY). Differences in genetic variances may also be caused by donor specific epistatic interactions with the recipient alleles which may hide part of the new variations due to the incorporation of donor alleles. Complementary test crosses involving several testers may be necessary to test these hypotheses.

## Interest of the usefulness criterion to rank donor × recipient crosses

Zhong and Jannink (2007) highlighted that the interest of UC to compare bi-parental cross performances is restrained because the variation of mean performances of crosses is much higher than the variance of crosses genetic variances. In addition, the simulation work carried out by Beckett et al (2019) highlighted a strong correlation between the

mean of parental performances and that of best progenies in a bi-parental cross. Our experimental results showed that crossing elite lines with diversity donors leads to families with variable genetic variances. In this context, using UC is necessary to consider this variation. In our study, the ranking of the crosses involving the same recipient line was largely reshaped when it was based on the usefulness criterion instead of the GY or YI mean performances. Using the $UC_1$ allowed us to identify the most interesting crosses to generate transgressive progenies for each private company (e.g., A2D4 for A2, A3D4 for A3, A5D7 for A5). Note that UC has also been recommended for the selection of crosses within a selection program and was proved useful to increase the genetic gain (Lehermeier et al. 2017; Yao et al. 2018; Allier et al. 2020a).

We also proposed a ranking based on $UC_2$. This indicator, which considers the additive variance rather than the genetic variance, enables one to project the future response to selection that could be achieved in the progeny of the best individuals of each family. The computation of $UC_2$ required to estimate accurately the GY and YI within-family additive variances. We estimated these variances without ($M\_FA_S$) or with ($M\_FA_SP_S$) a permanent effect in the model, which represented the non-additive genetic effects such as epistasis effect (Kruuk 2004; Vitezica et al. 2018). Adding this effect to the model increased the AIC and BIC. This could be due to the low proportion of repeated hybrids in our experiment, preventing accurate estimations of variances linked to the permanent effect. Nevertheless, the comparison of both models revealed the importance to consider such permanent genetic effect, to avoid the overestimation of the additive variances (as also observed by González-Diéguez et al. 2021). The $UC_2$ appeared as a good tool to distinguish between families having similar $UC_1$ values and privilege those with higher additive variances (e.g., A1D9 rather than A1D5 for the A1 recipient).

More individuals derived from the most promising crosses could be created to maximize the opportunity to find progenies with high performance. The choice of the number of new individuals may be specific to each company and may be guided using the expected maximum breeding value (EMBV, Müller et al. 2018). This indicator gives the expected performance of the best individual for a DH population of a given size estimating their breeding values.

## Implementation in breeding programs and future work

The creation and evaluation of a large multi-parental population confirmed experimentally the interest of introducing genetic resources into elite material, therefore supporting recent simulation results (Allier et al. 2020a; Vanavermaete

et al. 2021; Sanchez et al. 2023). Our mating design can be viewed as a cooperative bridging population fulfilling a dual task: identifying promising D×R crosses and delivering performing new lines directly usable as flint parents in breeding programs. To this aim it was required to deal with two constraints: (i) incorporating a large enough proportion of donor genome in progeny to explore new variations and (ii) minimizing the loss in the global performance due to the lower donor performance. Backcrossing the D×R crosses with the recipient parents turned out to be a good compromise as it generated, for most populations, transgressive individuals superior to the recipient elite line. One may however wonder whether a lower or higher proportion of donor genome would have been preferable. We addressed this question under a simple genetic model aiming at extrapolating our results to other possible pedigrees for the same donor x recipient combinations, as described in appendix File S5. This comparison of F1, BC1 and BC2 cross types indicates that the maximal UC values for the observed crosses in this study should be reached in general with two backcrosses with the recipient parent (BC2), as the decrease of the variance in the progeny is counterbalanced by the mean performance gain (Fig. 6, File S5). Note however, that if such populations are considered as a bridging step before introduction into an elite pool, the next breeding generations should also be considered. Simulation work showed that progenies of D×R crosses selected for introduction into the elite program can be preserved and improved in the elite breeding program provided selection is performed under a diversity constraint (Allier et al. 2020a; Sanchez et al. 2023). Such genetic resources generally contribute to varietal release after three crosses with the elite material in total (Sanchez et al. 2023). F1 and BC1 populations that may be sub-optimal compared to BC2 in terms of UC may nevertheless be a good option for the bridging step as individuals will carry more introduced segments.

Finally, our results highlight large differences in UC across donor recipient pairs. This supports experimentally the conclusion of Sanchez et al. (2023) that bridging is a key step to select donors prone to improving a given elite pool. Despite its large size, our design permitted to characterize only a restrained number of D×R crosses and its deployment to a larger scale could be expensive. Pre-selection of promising D×R crosses therefore would be of key interest before producing and evaluating them. Genomic selection approaches have been considered to predict usefulness criterion for such a pre-selection (Civan et al. 2021). The prediction of genetic variance for a given cross remains challenging as methods developed so far demonstrated inconsistent accuracies (Tiede et al. 2015; Adeyemo and Bernardo 2019; Neyhart and Smith 2019). Nevertheless, the use of our interconnected multi-parental populations to calibrate such prediction models would deserve evaluation.
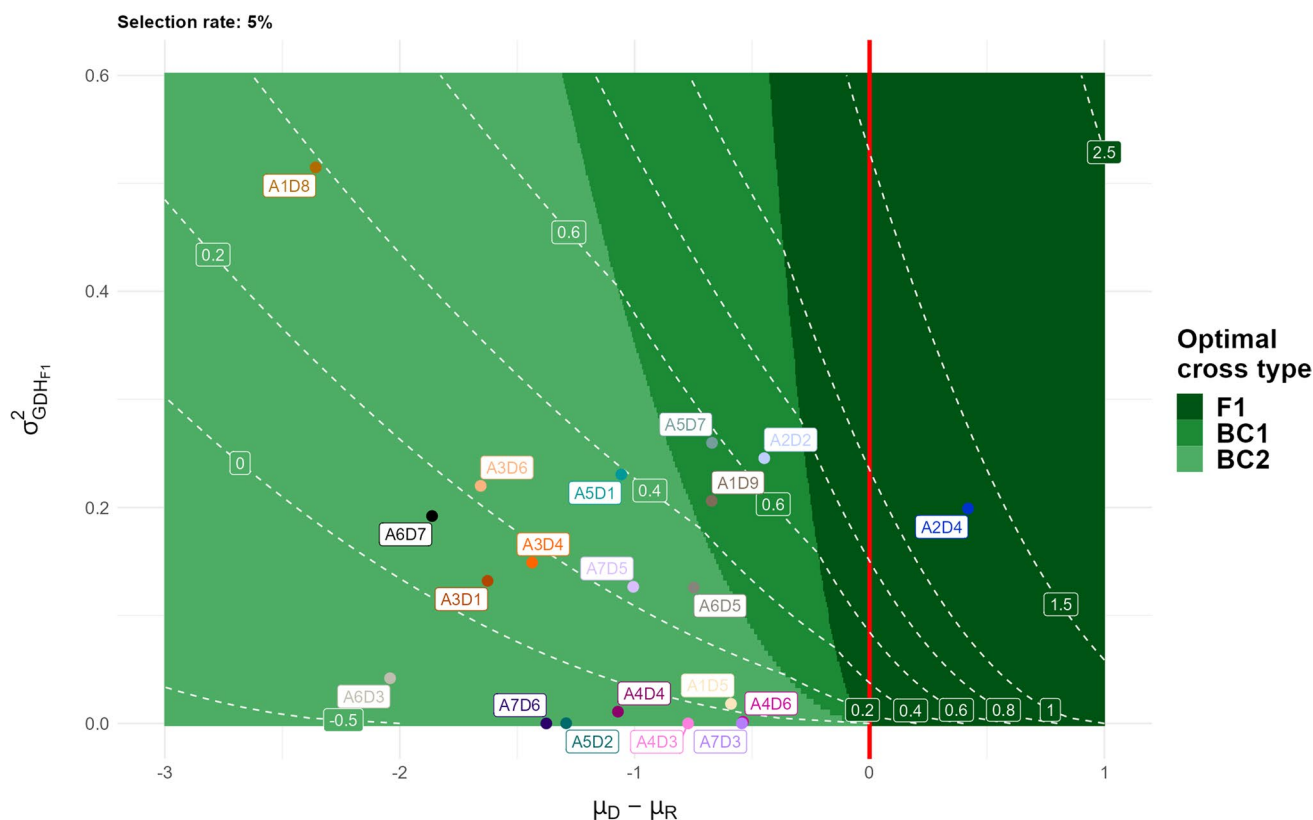
**Fig. 6** D×R cross type choice to maximizing the expected gain after selection. $\mu_D$ and $\mu_R$ are the performance values of the donor and recipient parents. $\sigma^2_{GDH_{F1}}$ is the theoritical genetic variance in a DH population derived from F1 cross between both parents. For pairs of $\sigma^2_{GDH_{F1}}$ and $\mu_D - \mu_R$ values, we compared three cross types (F1, BC1 and BC2) computing the usefulness criteria in DH populations derived from each. For the backrosses, the recurrent parent was the recipient line. The usefulness criteria were computed for a selection rate of 5%. For each parameter pair, the cross type which maximized the performance gain in comparison with the recipient line was considered as optimal. This is visualised through the green areas. The white dashed lines give the expected gain after selection in the DH population derived from the optimal cross. The red line indicates the case were $\mu_D$ and $\mu_R$ are equal. The experimental evaluated crosses were positionned in the graph using their mean performance and genetic variance values (color figure online)

## Conclusion

Altogether, our results support the usefulness of incorporating genetic resources into elite flint maize. Given the genetic gap between genetic resources and elite material observed at the beginning of this project (approx. 2 Mg/ha), we estimated that donors should be crossed three times to elite materials to maximize the expected value of selected progenies. We observed contrasted genetic variances and components (additive vs. permanent) across crosses, which can lead to large differences in both short-term and longer-term selection potential. This confirms the role of bridging programs as ours to detect the most suitable donor × elite pairs. In our experiment, only 20 such crosses could be evaluated despite a large size design of 1174 hybrids. An interesting prospect would be to evaluate the potential of genomic prediction based methods (Allier et al. 2020b) to predict the variance of additional crosses and identify

the most promising ones. Finally, our results support further effort to create fixed diversity donors in European flint maize (Böhm et al. 2017; Mayer et al. 2020) and evaluate them to conduct an efficient preselection step, which appears key in simulation work (Sanchez et al. 2023).

## Author contributions statement

CB, LM, AM and AC initiated this project. CB and AC coordinated it with the help of SM, LB and AM. CB, CP and BL contributed to the development of the plant material. DM and VC provided the genotyping data. AA analyzed the results to determine the crossing plan. DS analyzed the results and prepared the manuscript. AC, LM, TMH and SBS supervised this work. All authors revised and approved the manuscript.

## Declarations

## References

Adeyemo E, Bernardo R (2019) Predicting genetic variance from genomewide marker effects estimated from a diverse panel of maize inbreds. Crop Sci 59:583–590. https://doi.org/10.2135/cropsci2018.08.0525

Allier A, Teyssèdre S, Lehermeier C et al (2019) Assessment of breeding programs sustainability: application of phenotypic and genomic indicators to a North European grain maize program. Theor Appl Genet 132:1321–1334

Allier A, Teyssèdre S, Lehermeier C et al (2020a) Optimized breeding strategies to harness genetic resources with different performance levels. BMC Genomics 21:349. https://doi.org/10.1186/s12864-020-6756-0

Allier A, Teyssèdre S, Lehermeier C et al (2020b) Genomic prediction with a maize collaborative panel: identification of genetic resources to enrich elite breeding programs. Theor Appl Genet 133:201–215. https://doi.org/10.1007/s00122-019-03451-9

Álvarez-Castro JM, Carlborg Ö (2007) A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. Genetics. https://doi.org/10.1534/genetics.106.067348

Bauer E, Falque M, Walter H et al (2013) Intraspecific variation of recombination rate in maize. Genome Biol 14:R103. https://doi.org/10.1186/gb-2013-14-9-r103

Beckett TJ, Rocheford TR, Mohammadi M (2019) Reimagining maize inbred potential: identifying breeding crosses using genetic variance of simulated progeny. Crop Sci 59:1457–1468. https://doi.org/10.2135/cropsci2018.08.0508

Boeven PH, Zhao Y, Thorwarth P et al (2020) Negative dominance and dominance-by-dominance epistatic effects reduce grain-yield heterosis in wide crosses in wheat. Sci Adv 6:eaay4897

Böhm J, Schipprack W, Utz HF, Melchinger AE (2017) Tapping the genetic diversity of landraces in allogamous crops with doubled haploid lines: a case study from European flint maize. Theor Appl Genet 130:861–873. https://doi.org/10.1007/s00122-017-2856-x

Butler DG, Cullis BR, Gilmour AR et al (2017) ASReml-R reference manual version 4. VSN International Ltd, Hemel Hempstead

Camus-Kulandaivelu L, Veyrieras J-B, Madur D et al (2006) Maize adaptation to temperate climate: relationship between population structure and polymorphism in the *Dwarf8* gene. Genetics 172:2449–2463. https://doi.org/10.1534/genetics.105.048603

Cartea ME, Revilla P, Butrón A et al (1999) Do second cycle maize inbreds preserve the European flint heterotic group? Crop Sci 39:1060–1064. https://doi.org/10.2135/cropsci1999.0011183X003900040017x

Civan P, Rincent R, Danguy-Des-Deserts A et al (2021) Population genomics along with quantitative genetics provides a more efficient valorization of crop plant genetic diversity in breeding and pre-breeding programs. Springer International Publishing, Cham

Coombes NE (2009) DiGGer, a spatial design program, Biometric bulletin NSW department of primary industries, Orange, NSW

Cowling WA (2013) Sustainable plant breeding. Plant Breed 132:1–9. https://doi.org/10.1111/pbr.12026

Dubreuil P, Charcosset A (1999) Relationships among maize inbred lines and populations from European and North-American origins as estimated using RFLP markers. Theor Appl Genet 99:473–480. https://doi.org/10.1007/s001220051259

Dwivedi SL, Ceccarelli S, Blair MW et al (2016) Landrace germplasm for improving yield and abiotic stress adaptation. Trends Plant Sci 21:31–42. https://doi.org/10.1016/j.tplants.2015.10.012

Frisch M, Melchinger AE (2005) Selection theory for marker-assisted backcrossing. Genetics 170:909–917. https://doi.org/10.1534/genetics.104.035451

Ganal MW, Durstewitz G, Polley A et al (2011) A large maize (Zea mays L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6:e28334. https://doi.org/10.1371/journal.pone.0028334

Goddard M (2009) Genomic selection: prediction of accuracy and maximisation of long term response. Genetica 136:245–257. https://doi.org/10.1007/s10709-008-9308-0

Gonen S, Wimmer V, Gaynor RC et al (2018) A heuristic method for fast and accurate phasing and imputation of single-nucleotide polymorphism data in bi-parental plant populations. Theor Appl Genet 131:2345–2357. https://doi.org/10.1007/s00122-018-3156-9

González-Diéguez D, Legarra A, Charcosset A et al (2021) Genomic prediction of hybrid crops allows disentangling dominance and epistasis. Genetics 218:iyab026. https://doi.org/10.1093/genetics/iyab026

Gouesnard B, Negro S, Laffray A et al (2017) Genotyping-by-sequencing highlights original diversity patterns within a European collection of 1191 maize flint lines, as compared to the maize USDA genebank. Theor Appl Genet 130:2165–2189. https://doi.org/10.1007/s00122-017-2949-6

Hallauer AR, Russell WA, Lamkey KR (1988) Corn breeding. In: Sprague GF, Dudley JW (eds) Agronomy monographs. American

society of agronomy, crop science society of America, soil science society of America, Madison, WI, USA, pp. 463–564

Han Y, Cameron JN, Wang L, Beavis WD (2017) The predicted cross value for genetic introgression of multiple alleles. Genetics 205:1409–1423. https://doi.org/10.1534/genetics.116.197095

Hospital F (2005) Selection in backcross programmes. Phil Trans R Soc B 360:1503–1511. https://doi.org/10.1098/rstb.2005.1670

Hospital F, Charcosset A (1997) Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485

Hung H-Y, Browne C, Guill K et al (2012) The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. Heredity 108:490–499. https://doi.org/10.1038/hdy.2011.103

Jannink J-L (2010) Dynamics of long-term genomic selection. Genet Sel Evol 42:35. https://doi.org/10.1186/1297-9686-42-35

Kruuk LEB (2004) Estimating genetic parameters in natural populations using the 'animal model.' Phil Trans R Soc Lond B 359:873–890. https://doi.org/10.1098/rstb.2003.1437

Lado B, Battenfield S, Guzmán C et al (2017) Strategies for selecting crosses using genomic prediction in two wheat breeding programs. Plant Genome. https://doi.org/10.3835/plantgenome2016.12.0128

Laporte F, Charcosset A, Mary-Huard T (2022) Efficient ReML inference in variance component mixed models using a min-max algorithm. PLoS Comput Biol 18:e1009659. https://doi.org/10.1371/journal.pcbi.1009659

Lehermeier C, Krämer N, Bauer E et al (2014) Usefulness of multi-parental populations of maize (*Zea mays* L.) for genome-based prediction. Genetics 198:3–16. https://doi.org/10.1534/genetics.114.161943

Lehermeier C, Teyssèdre S, Schön C-C (2017) Genetic gain increases by applying the usefulness criterion with improved variance prediction in selection of crosses. Genetics 207:1651–1661. https://doi.org/10.1534/genetics.117.300403

Lewis R, Goodman M (2003) Incorporation of tropical maize germplasm into inbred lines derived from temperate× temperate-adapted tropical line crosses: agronomic and molecular assessment. Theor Appl Genet 107:798–805

Lin Z, Cogan NOI, Pembleton LW et al (2016) Genetic gain and inbreeding from genomic selection in a simulated commercial breeding program for perennial ryegrass. Plant Genome. https://doi.org/10.3835/plantgenome2015.06.0046

Lobell DB, Schlenker W, Costa-Roberts J (2011) Climate trends and global crop production since 1980. Science 333:616–620. https://doi.org/10.1126/science.1204531

Lush JL (1937) Animal breeding plans. ames. Collegiate Press Inc

Maccaferri M, Sanguineti MC, Donini P, Tuberosa R (2003) Microsatellite analysis reveals a progressive widening of the genetic basis in the elite durum wheat germplasm. Theor Appl Genet 107:783–797. https://doi.org/10.1007/s00122-003-1319-8

Manjarrez-Sandoval P, Carter Jr TE, Nelson RL, et al (1998) Soybean Asian variety evaluation (SAVE): agronomic performance of modern Asian cultivars in the US 1997, USDA-ARS, Raleigh, NC

Mayer M, Hölker AC, González-Segovia E et al (2020) Discovery of beneficial haplotypes for complex traits in maize landraces. Nat Commun 11:4954. https://doi.org/10.1038/s41467-020-18683-3

McCouch S, Baute GJ, Bradeen J et al (2013) Agriculture: feeding the future. Nature 499:23–24. https://doi.org/10.1038/499023a

Messmer MM, Melchinger AE, Boppenmaier J et al (1992) Relationships among early European maize inbreds: I. genetic diversity among flint and dent lines revealed by RFLPs. Crop Sci 32:1301–1309. https://doi.org/10.2135/cropsci1992.0011183X003200060001x

Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J (1993) Relationships among early European maize inbreds: II. comparison of pedigree and RFLP data. Crop Sci 33:944–950

Meuwissen TH (1997) Maximizing the response of selection with a predefined rate of inbreeding. J Anim Sci 75:934. https://doi.org/10.2527/1997.754934x

Mickelbart MV, Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nat Rev Genet 16:237–251. https://doi.org/10.1038/nrg3901

Mikel MA, Dudley JW (2006) Evolution of North American dent corn from public to proprietary germplasm. Crop Sci 46:1193–1205. https://doi.org/10.2135/cropsci2005.10-0371

Mohammadi M, Tiede T, Smith KP (2015) PopVar: a genome-wide procedure for predicting genetic variance and correlated response in biparental breeding populations. Crop Sci 55:2068–2077. https://doi.org/10.2135/cropsci2015.01.0030

Müller D, Schopp P, Melchinger AE (2018) Selection on expected maximum haploid breeding values can increase genetic gain in recurrent genomic selection. G3 Genes Genomes Genet 8:1173–1181. https://doi.org/10.1534/g3.118.200091

Neyhart JL, Smith KP (2019) Validating genomewide predictions of genetic variance in a contemporary breeding program. Crop Sci 59:1062–1072. https://doi.org/10.2135/cropsci2018.11.0716

Ødegård J, Yazdi MH, Sonesson AK, Meuwissen THE (2009) Incorporating desirable genetic characteristics from an inferior into a superior population using genomic selection. Genetics 181:737–745. https://doi.org/10.1534/genetics.108.098160

Ordás B, Malvar RA, Revilla P, Ordás A (2023) Effect of three cycles of recurrent selection for yield in four Spanish landraces of maize. Euphytica 219:77

Palmgren MG, Edenbrandt AK, Vedel SE et al (2015) Are we ready for back-to-nature crop breeding? Trends Plant Sci 20:155–164. https://doi.org/10.1016/j.tplants.2014.11.003

Peng T, Sun X, Mumm RH (2014) Optimized breeding strategies for multiple trait integration: I. minimizing linkage drag in single event introgression. Mol Breed 33:89–104. https://doi.org/10.1007/s11032-013-9936-7

Piepho HP, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. https://doi.org/10.1534/genetics.107.074229

Pollak LM (2003) The history and success of the public–private project on germplasm enhancement of maize (GEM). Advances in agronomy. Elsevier, pp 45–87

Pollak LM (1990) Evaluation of Caribbean maize accessions in Puerto Rico

Rawlings JO, Thompson DL (1962) Performance level as criterion for the choice of maize testers[1]. Crop Sci 2:217–220. https://doi.org/10.2135/cropsci1962.0011183X000200030012x

Reif JC, Zhang P, Dreisigacker S et al (2005) Wheat genetic diversity trends during domestication and breeding. Theor Appl Genet 110:859–864. https://doi.org/10.1007/s00122-004-1881-8

Ribaut JM, Ragot M (2006) Marker-assisted selection to improve drought adaptation in maize: the backcross approach, perspectives, limitations, and alternatives. J Exp Bot 58:351–360. https://doi.org/10.1093/jxb/erl214

Rincent R, Nicolas S, Bouchet S et al (2014) Dent and Flint maize diversity panels reveal important genetic potential for increasing biomass production. Theor Appl Genet 127:2313–2331. https://doi.org/10.1007/s00122-014-2379-7

Rutkoski J, Singh RP, Huerta-Espino J et al (2015) Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. Plant Genome. https://doi.org/10.3835/plantgenome2014.10.0074

Salhuana W, Sevilla R, Eberhart SA (1997) Latin American maize project (LAMP): final report

Salhuana W, Pollak L (2006) Latin American maize project (LAMP) and germplasm enhancement of maize (GEM) project: generating useful breeding germplasm [Zea mays L.; USA]. Maydica (Italy) 51(2):339–355

Sanchez D, Sadoun SB, Mary-Huard T et al (2023) Improving the use of plant genetic resources to sustain breeding programs' efficiency. Proc Natl Acad Sci USA 120:e2205780119. https://doi.org/10.1073/pnas.2205780119

Schnell F, Utz H (1976) F1 Leistung und Elternwahl in der Zuchtung von Selbstbefruchtern. Ber Arbeitstag Arbeitsgem Saatzuchtleiter

Sharma S, Carena MJ (2012) NDSU EarlyGEM: incorporating tropical and temperate elite exotic germplasm to increase the genetic diversity of short-season maize. Maydica 57(1):24–42

Simmonds NW (1993) Introgression and incorporation. Strategies for the use of crop genetic resources. Biol Rev 68:539–562. https://doi.org/10.1111/j.1469-185X.1993.tb01243.x

Stich B, Melchinger AE, Frisch M et al (2005) Linkage disequilibrium in European elite maize germplasm investigated with SSRs. Theor Appl Genet 111:723–730. https://doi.org/10.1007/s00122-005-2057-x

Strigens A, Schipprack W, Reif JC, Melchinger AE (2013) Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PLoS One 8:e57234. https://doi.org/10.1371/journal.pone.0057234

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science 327:818–822. https://doi.org/10.1126/science.1183700

Tiede T, Kumar L, Mohammadi M, Smith KP (2015) Predicting genetic variance in bi-parental breeding populations is more accurate when explicitly modeling the segregation of informative genomewide markers. Mol Breed 35:199. https://doi.org/10.1007/s11032-015-0390-6

Van Inghelandt D, Melchinger AE, Lebreton C, Stich B (2010) Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor Appl Genet 120:1289–1299. https://doi.org/10.1007/s00122-009-1256-2

Vanavermaete D, Fostier J, Maenhout S, De Baets B (2021) Deep scoping: a breeding strategy to preserve, reintroduce and exploit genetic variation. Theor Appl Genet. https://doi.org/10.1007/s00122-021-03932-w

Visscher PM, Haley CS, Thompson R (1996) Marker-assisted introgression in backcross breeding programs. Genetics 144:1923–1932. https://doi.org/10.1093/genetics/144.4.1923

Vitezica ZG, Legarra A, Toro MA, Varona L (2017) Orthogonal estimates of variances for additive, dominance, and epistatic effects in populations. Genetics 206:1297–1307. https://doi.org/10.1534/genetics.116.199406

Vitezica ZG, Reverter A, Herring W, Legarra A (2018) Dominance and epistatic genetic variances for litter size in pigs using genomic models. Genet Sel Evol 50:71. https://doi.org/10.1186/s12711-018-0437-3

Wang C, Hu S, Gardner C, Lübberstedt T (2017) Emerging avenues for utilization of exotic germplasm. Trends Plant Sci 22:624–637. https://doi.org/10.1016/j.tplants.2017.04.002

Welcker C, Spencer NA, Turc O et al (2022) Physiological adaptive traits are a potential allele reservoir for maize genetic progress under challenging conditions. Nat Commun 13:3225. https://doi.org/10.1038/s41467-022-30872-w

Wezel A, Casagrande M, Celette F et al (2014) Agroecological practices for sustainable agriculture. A review. Agron Sustain Dev 34:1–20. https://doi.org/10.1007/s13593-013-0180-7

Wolfe MD, Chan AW, Kulakow P et al (2021) Genomic mating in outbred species: predicting cross usefulness with additive and total genetic covariance matrices. Genetics 219(3):iyab122

Woolliams JA, Berg P, Dagnachew BS, Meuwissen THE (2015) Genetic contributions and their optimization. J Anim Breed Genet 132:89–99. https://doi.org/10.1111/jbg.12148

Wray N, Goddard M (1994) Increasing long-term response to selection. Genet Sel Evol 26:431. https://doi.org/10.1186/1297-9686-26-5-431

Yao J, Zhao D, Chen X et al (2018) Use of genomic selection and breeding simulation in cross prediction for improvement of yield and quality in wheat (Triticum aestivum L.). Crop J 6:353–365. https://doi.org/10.1016/j.cj.2018.05.003

Zhong S, Jannink JL (2007) Using quantitative trait loci results to discriminate among crosses on the basis of their progeny mean and variance. Genetics 177:567–576. https://doi.org/10.1534/genetics.107.075358