Back to the future: Implications of genetic complexity for hybrid breeding strategies

Frank Technow*, Dean Podlich† and Mark Cooper‡

Abstract Commercial hybrid breeding operations can be described as decentralized networks of smaller, more or less isolated breeding programs. There is further a tendency for the disproportionate use of successful inbred lines for generating the next generation of recombinants, which has led to a series of significant bottlenecks, particularly in the history of the North American and European maize germplasm. Both the decentralization and the disproportionate inbred use reduce effective population size and constrain the accessible genetic space. Under these conditions, long term response to selection is not expected to be optimal under the classical infinitesimal model of quantitative genetics. In this study we therefore aim to propose an alternative rational for the success of large breeding operations in the context of genetic complexity arising from the structure and properties of interactive genetic networks. For this we use simulations based on the *NK* model of genetic architecture. We indeed found that constraining genetic space and reducing effective population size, through program decentralization and disproportionate inbred use, is required to expose additive genetic variation and thus facilitate heritable genetic gains. These results introduce new insights into why the historically grown structure of hybrid breeding programs was successful in improving the yield potential of hybrid crops over the last century. We also hope that a renewed appreciation for "why things worked" in the past can guide the adoption of novel technologies and the design of future breeding strategies for navigating biological complexity.

Introduction

Pioneered by Shull (1908), hybrid breeding is credited as one of the most significant factors for the tremendous productivity increases of major field (Duvick, 1999) and horticultural (Silva Dias, 2010) crops that enabled food production to keep pace with population growth. Hybrid breeding programs originally were centred around maximum exploitation of heterosis, a phenomenon that remains largely unexplained even after a century of research (East, 1936; Lippman and Zamir, 2007). This later evolved into the modern concept of hybrid breeding, characterized by its distinctive structuring of germplasm into heterotic groups and patterns (Melchinger and Gumber, 1998). Beyond heterotic groups, the structure of commercial hybrid breeding, particularly in major crops like maize, is characterized by the largely isolated and unique sub-heterotic patterns of the major companies (White et al., 2020) as well as a high degree of decentralization into smaller, more or less disconnected sub-programs within those (Cooper et al., 2014). Plant breeders further have a tendency for relying on only a small set of elite inbred lines for producing the next generation of recombinants (Rasmusson and Phillips, 1997), leading to a series of significant bottleneck events in the history of, for example, the North American maize germplasm (White et al., 2020). These characteristics drastically reduced the effective population size within breeding programs and are not predicted to be promising strategies under the additive, infinitesimal model of quantitative genetics (Gaynor et al., 2017). Nevertheless, consistent long-term genetic gain has been demonstrated (Duvick et al., 2004).

To better describe and quantify the observed genetic variation among hybrids, the concept of general and specific combining ability was developed early on (Sprague and Tatum, 1942). The former, commonly abbreviated as GCA, is a property of the additive effects of contributing genes and describes the average performance of all hybrids derived from an inbred. The latter, commonly abbreviated as SCA, is a non-additive residual term that describes the deviation of the performance of a particular hybrid from the expectation based on the parental GCA values.

Running efficient hybrid breeding programs requires a preponderance of additive genetic variation to maximize response to selection in the next generation of inbred lines (Falconer and Mackay, 1996) as well as the predictability of hybrid performance from the GCA of inbred lines (Reif et al., 2007). A preponderance of GCA variation also allows identification of inbreds that can serve as parents of several high performing hybrids. This greatly simplifies production of commercial seed, which is a major

*Plant Breeding, Corteva agriscience, Tavistock , ON, Canada; frank.technow@corteva.com
†Systems and Innovation for Breeding and Seed Products, Corteva agriscience, Johnston, IA, USA
‡Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD, Australia

challenge for many crops (Technow, 2019). Therefore, hybrid breeding programs have traditionally relied on maximizing and exploiting GCA variation (Falconer and Mackay, 1996; Melchinger, 1999).

The historically grown paradigms around hybrid breeding designs and strategies are now being challenged by innovative concepts (e.g., Gaynor et al., 2017; Wallace et al., 2018; Hickey et al., 2019; Voss-Fels et al., 2019; Seye et al., 2020) devised in the wake of technological advances such as whole genome prediction (Meuwissen et al., 2001), high-throughput phenotyping (Araus and Cairns, 2014) and genotyping (Poland and Rife, 2012), as well as gene editing (Jaganathan et al., 2018). While some of these concepts are highly speculative and might not live up to expectations (Bernardo, 2016), it is clear that the next decades will change plant breeding. However, before implementing drastic changes to breeding programs we require a theoretical and simulation framework to explore and understand the structures and strategies that have contributed to the success of long term genetic gain and germplasm improvement. From this historical basis we can evaluate novel proposals and draw lessons for design of future breeding strategies.

Empirical reports show a preponderance of additive variation in many wild, domesticated and laboratory species (Falconer and Mackay, 1996; Lynch and Walsh, 1998; Hill et al., 2008). This agrees well with published studies showing a preponderance of additive GCA over non-additive SCA variation in hybrid breeding programs (Technow et al., 2014; Larièpe et al., 2017). At the same time, however, advances in plant physiology and molecular and systems biology have stimulated a renewed appreciation of the intricate interactions at the molecular, metabolical and physiological level that underlie complex traits (Hammer et al., 2006; Carlborg and Haley, 2004; Phillips, 2008; Wilkins et al., 2016; Saha et al., 2011; Jiang et al., 2017). Of particular relevance for hybrid breeding are recent studies indicating that heterosis is an emergent property of complex metabolic networks (Fiévet et al., 2010, 2018; Vacher and Small, 2019).

The paradox between the complexity of the underlying biology and the simplicity of the expressed variation can of course be resolved by distinguishing between biological and statistical effects and realizing that the former cannot be inferred from the latter (Wade, 2002; Mackay, 2014; Huang and Mackay, 2016). Statistical effects of genes, as well as their aggregates such as GCA and SCA and their variances depend on the genetic background of the population in which they are evaluated, particularly on allele frequencies and linkage disequilibrium (LD) patterns (Falconer and Mackay, 1996). For example, it was shown that, regardless of the underlying genetic architecture, genetic variances in random mating populations are expected to be predominantly additive when genes are at extreme frequencies and linkage disequilibrium is high (Hill et al., 2008).

Thus, ratios of additive to non-additive variation are not intrinsic properties of biological systems but at least partly a function of allele frequencies and LD patterns and thus dependent on breeding strategies. Because of the importance of additive variation for efficient operation of breeding programs, a framework to evaluate and study breeding strategies should allow for the possibility of additivity arising from high degrees of biological complexity at the genetic level.

In this study we will use simulations based on the *NK* model of genetic complexity (Kauffman, 1993) to explore two *themes* representing key, historically grown, characteristics of hybrid breeding: firstly its decentralization into smaller, more or less independent sub-programs ("*decentrality* theme") and secondly the disproportional use of superior inbred lines for producing the next generation of recombinants ("*inbred usage* theme"). Our goal thereby is not to make specific recommendations for optimal structuring of programs, but rather to gain an appreciation for the properties of these structures in the context of different degrees of genetic complexity.

Material and Methods

Model of genetic complexity

The *NK* model, introduced by Kauffman (1993) will form the basis of the simulations. The *NK* model allows generation of a tunable series of models of trait genetic architecture with increasing dimensionality and complexity by varying the number of genes N (dimensionality) and the degree of interaction among them (K, complexity).

The *genetic landscape* metaphor was introduced and developed by Wright (1932) to aid conceptualizing genetic complexity in high dimensions. As a metaphor it should not be taken literally but can help to gain an intuition for the complexity and ruggedness associated with increasing values of K

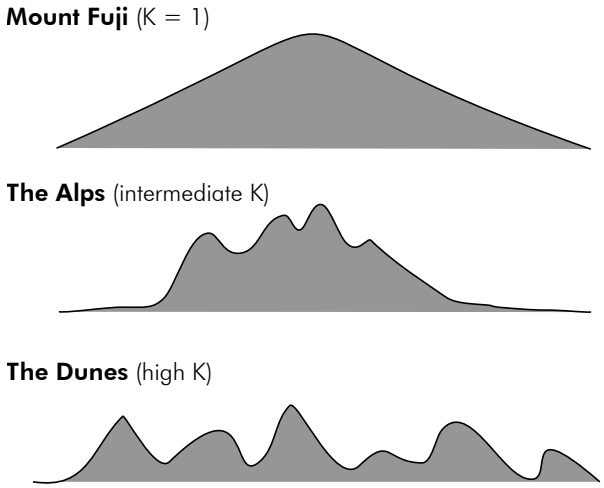

Figure 1 Schematic visualization of genetic landscapes corresponding to different values of complexity parameter K.

(Kauffman, 1993) as well as making the rather abstract concepts discussed henceforth more tangible. At $K = 1$ (special case of additive gene action), the genetic landscape can be imagined as that of Mount Fuji, i.e., a single, clearly distinguished peak with a steady and monotonous incline to the top (Figure 1). At intermediate K levels, the landscape is characterized by multiple peaks clustered together in a certain region of genetic space. This might be visualized as akin to the European Alps, i.e., a mountainous region within an otherwise flat landscape. Finally, at high value of K, the landscape resembles a sea of dunes, i.e., a range of peaks of similar height and shape distributed more or less evenly in space.

We implemented the *NK* model according to the generalized approach described by Altenberg (1994), but adapted the model to accommodate diploid genomes. Here, the complex trait is described as a normalized sum of a set of "fitness components". The value of each fitness component is computed as a function of K interacting genes drawn at random from all N genes. Following Altenberg (1994), the specific fitness values were calculated with *random functions* derived from the *ran4* pseudo-random number generator (Press et al., 1992) and are distributed uniformly between 0 and 1. For this study, the number of fitness components and the number of genes were both set to 500. Genes were biallelic and the simulated organism diploid. The complexity parameter K was varied from 1 to 15 in steps of 1 (i.e. creating genetic landscapes ranging in complexity from Mount Fuji to The Dunes; Figure 1). With the exception of $K = 1$, this parameter was used as the rate parameter in a Poisson distribution from which the number of interacting genes was drawn independently for each fitness component. The sampled values were then truncated to fall within a range of 1 and 15. The identities of the interacting genes were drawn at random from the total set. Thus genes typically influenced multiple fitness components (i.e., act pleiotropically). For $K = 1$, each of the 500 genes was assigned to exactly one fitness component and the values of heterozygous allele configurations constrained to be midway between the homozygous configurations. Thus, $K = 1$ represents the special case of additive gene action. Using order statistics, the expected value of the maximum of two samples from a Uniform distribution between 0 and 1 is $2/3$. Thus, the expected maximum attainable fitness for the $K = 1$ special case is $2/3$.

The complexity of the generated *NK* models was quantified following the "one-mutant neighbour" hill-climbing algorithm described by Kauffman (1993), but adapted to diploid organisms. A randomly generated genotype was used as the starting value. From there, all possible genotypes were generated that differ from the initial genotype by one allele at one of the 500 loci. Thus, a homozygous locus was changed to the heterozygous state while a heterozygous locus was changed to both alternate homozygotes. Then, the fitness values of all one-allele neighbours were evaluated according to the defined *NK* model and an improved genotype chosen at random from all fitter one-allele neighbours. This process was repeated until no fitter one-mutant neighbour could be found, meaning that the search reached a local or global optimum. For each level of K, 100 *NK* models were generated independently and a minimum of 65 searches, each starting at a random initial genotype, were conducted for each. The statistics recorded were the average number of steps until a local optimum was reached, the average Hamming genotypic distance, i.e., the normalized number of differing genome positions (Pinheiro et al., 2005), among optima and the correlation between the fitness values of the optima and the Hamming distance to the highest optimum identified (Kauffman, 1993).

The average Hamming distance between the local peaks increased from just below 0.5 at $K = 2$ to $2/3$ at around K of 6 or 7 and remained constant at this value from there on (Figure 2A). Note that with three different genotypes at each locus, $2/3$ is the expected value of the Hamming distance between

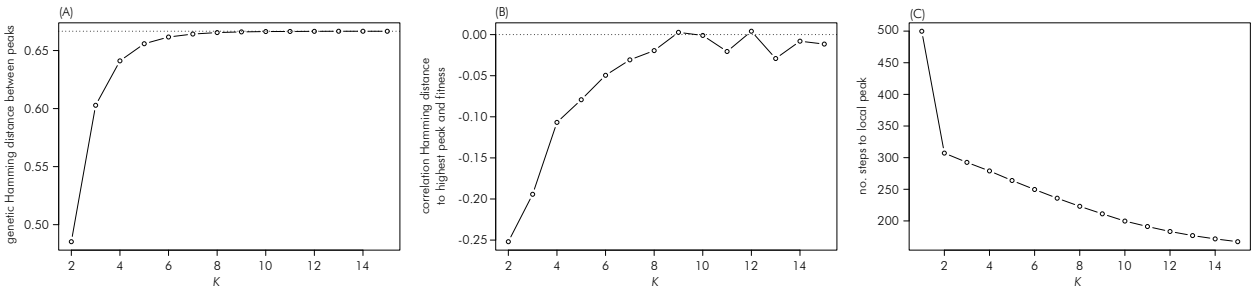

Figure 2 Relationship between the *NK* model complexity parameter K and (A) the average genetic Hamming distance between local peaks, (B) the average correlation between the fitness values of local peaks and their genetic Hamming distance to the highest peak and (C) the average number of steps to a local peak. Panels (A) and (B) omit results for $K = 1$, for which only a single peak exists. The dotted lines in these panels indicate values of $2/3$ and 0.0, respectively.

randomly generated genotypes. Similarly, the correlation between the fitness values of the local peaks and their Hamming distance to the highest identified peak increased from -0.25 at $K = 2$ to zero at $K = 9$ (Figure 2B). Here, a negative correlation means that local peaks with higher fitness tend to be found near each other and clustered around the highest peak. Further, a zero correlation indicates that there is no clustering of the peaks and proximity to the highest peak. Therefore, local peaks are randomly distributed throughout the genetic landscape. Thus, somewhere between $K = 6$ and $K = 9$, the landscape shifts from one in which local peaks tend to cluster together, to one where local peaks of arbitrary height can exist anywhere in genetic space. The average number of steps until a local peak was reached decreased with K from 500 at $K = 1$ to just 167 at $K = 15$ (Figure 2C). Note that 500 is the expectation at $K = 1$, the special case of additive gene action, when starting from randomly generated genotypes, because $1/3$ of the 500 loci are already at their highest possible value, $1/3$ are one step removed (the heterozygous genotypes) and $1/3$ are two steps removed (the lower homozygotes). Thus, the complexity and ruggedness of the genetic landscapes increase further after they become uncorrelated around K of 6 to 9.

The simulated genome comprised 10 diploid chromosomes of 1 Morgan length each. Each of the chromosomes received a random subset of 50 of the 500 genes, which were distributed evenly across the chromosome. Recombination was simulated according to the Haldane mapping function with the R package "hypred" (Technow, 2013), in the version available from the supplement of Technow and Gerke (2017).

Simulation of hybrid breeding process

The simulation process is visualized in Figure 3. The starting point of the simulation was a base population of inbred lines of size 1,000. This population was simulated stochastically as described by Montana (2005) to result in an expected LD between two loci t Morgan apart equal to $r^2 = 0.5 \cdot 2^{-t/0.1}$ and with minor allele frequencies distributed uniformly between 0.35 and 0.50. The lines from the base population were then separated at random into two heterotic groups (arbitrarily labelled '1' and '2') and further into sub-populations within those. The size of those sub-populations depended on the scenario. One sub-population from one heterotic group was then paired with one sub-population from the other group to form sub-heterotic patterns. These population pairs will henceforth be referred to as "breeding programs". Hybrids were produced strictly across heterotic groups, by crossing lines from one sub-population of a program with lines from the other. Breeding crosses, i.e., crosses to generate a new generation of recombinant lines were done within and among sub-populations, depending on the scenario but strictly within heterotic groups. The simulation of the breeding process described above is an approximation of the structure and evolution of long term hybrid breeding akin to what we have observed in practice; i.e. starting from an initial germplasm base, separation into distinct heterotic groups and future separation into sub-populations.

The GCA of the lines was evaluated with an incomplete mating design (Melchinger et al., 1987; Seye et al., 2020) by performing 10 crosses per line with random partners from the opposite sub-population of the same program. The performance of the resulting hybrids, as determined according to

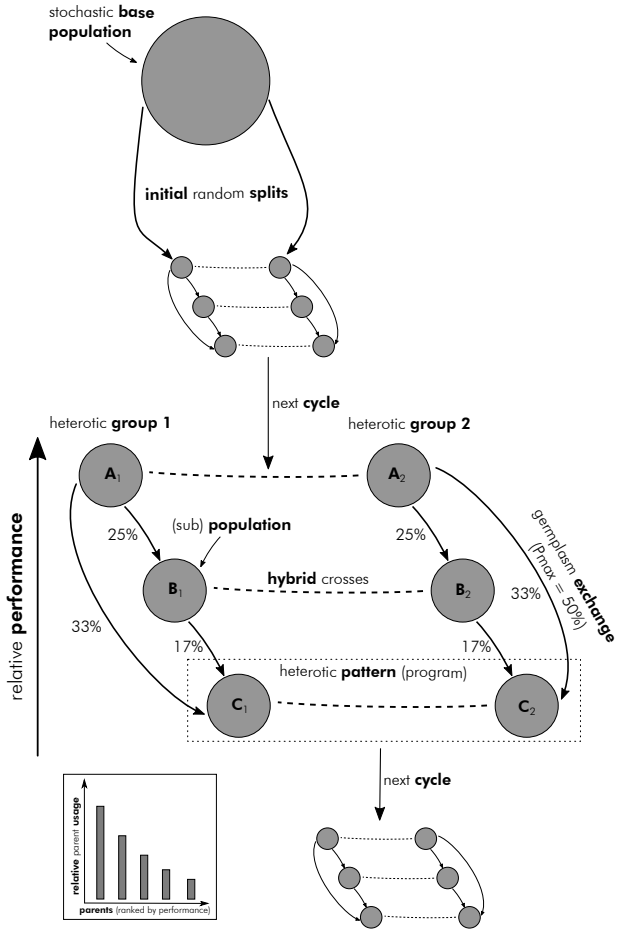

Figure 3 Schematic visualization of simulated hybrid breeding process.

the defined *NK* model, was then averaged. Finally, a normally distributed noise variable with zero mean and variance equal to one third of the variance of the GCA values of that sub-population was added to those averages to represent experimental and environmental noise. The so obtained values were used as observed GCA values. Those GCA estimates were then used to predict the performance of all possible inter-group hybrids of that program. The top hybrids, how many exactly depended on the scenario, were then selected and their true performance determined according to the *NK* model. The average of this select group of hybrids, which represents a set of advanced experimental hybrids, was used to quantify the overall performance of the program in the current cycle. The maximum true performance of the selected hybrids from all programs was defined as the peak performance of the whole breeding operation in the current cycle and used as a metric of genetic gain. This metric reflects that commercial breeding programs release only a handful of hybrid products each cycle.

Breeding crosses among inbred lines for initiating the next recombination cycle were chosen by assigning each inbred line a usage probability, which was a product between an individual and population level relative contribution value. To determine the former, the lines within each sub-population were ranked according to their observed GCA values. Only the top lines, how many depended on the scenario, were selected as potential parents, the remainder given an individual contribution value of zero. The relative contributions of the selected lines from a given sub-population were drawn from a Dirichlet distribution. The concentration parameters of this distribution were used to modulate the relationship between selection rank and relative contribution. Further details about this will be given later when describing the setting for the "inbred usage theme".

The population level contribution values describe the overall contribution of lines from one sub-population to the breeding crosses of another. They are thus defined anew for each target population and hence the contribution value of population 'A' to the crosses for population 'B' might be different than that to the crosses of population 'C'. The process will be explained using the example visualized in Figure 3. Here there are three programs (labelled 'A', 'B' and 'C', with subscript 1 or 2 indicating the heterotic group).

The programs are ranked from highest to lowest performing according to the average performance of the selected set of experimental hybrids, as described above. Germplasm, in the form of lines used as crossing partners, is exchanged only from higher performing to lower performing programs. Specifically, the amount of crosses with lines from other programs increased from zero for the best performing program (A) to a proportion of P_{max} for the lowest performing program (C), with intermediate programs staggered equidistantly between. In the example, $P_{max} = 50\%$. Thus, program A will perform no crosses with lines from other programs, program B will use lines from other programs in 25% of its new crosses and program C in 50% of its crosses. How much of that overall proportion was derived from each of the other programs was proportional to the relative performance differences. In the example, the difference between program C and A is twice as large as that between C and B, thus, lines from program A were used in twice as many crosses

than lines from program B (33% from A and 17% from B for at total of 50%). This process thus reflects that highly successful programs tend to exploit their own genetics while less successful programs have more of an incentive to explore superior genetics from other programs.

The relative individual contributions were then multiplied with the relative population contributions to arrive at a final relative contribution value for each line to the crosses of a given population. The actual breeding crosses were then determined by sampling the lines with probabilities proportional to their contribution values. This was done with replacement, meaning that the same cross could have been made multiple times, but excluding crosses that would result in selfings. One recombinant line was derived from each crossing through seven generations of single seed descent selfing, followed by a final doubled haploid step (Dwivedi et al., 2015) to result in fully homozygous inbred lines. This new generation of recombinants fully replaced the previous generations, i.e., a line was considered as a crossing partner in only one generation. The so obtained new recombinants then form the next breeding cycle. The simulations were conducted for 30 cycles in total and repeated independently at least 500 times for each scenario studied. All computations were conducted in the R environment for statistical computing (R Core Team, 2018).

Recorded metrics

In addition to the already described true performance of the best identified hybrid, which was used as a measure of *peak performance* in a given cycle, several other measures were recorded to describe and understand the dynamics of the system.

The *proportion of GCA to total genetic variance* (%GCA) describes the amount of exploitable additive genetic variation currently available. It was estimated using the hybrids generated for evaluating the GCA of the inbred lines. For this, the following mixed model was fitted: $h_{ij} = \mu + g_i + g_j + e_{ij}$, where h_{ij} was the true performance of the hybrids, μ the overal mean, g_i and g_j were the GCA effects of the parents from the two heterotic groups and e_{ij} a residual term. Because the true genetic performances of the hybrids were used, e_{ij} corresponds to the SCA component. The model was fitted using the R package "lme4" (Bates et al., 2015) and %GCA then calculated as $\left(V_{g_i} + V_{g_j}\right) / \left(V_{g_i} + V_{g_j} + V_{e_{ij}}\right)$, where V_{g_i} etc. were the estimated variance components. In scenarios with multiple programs, %GCA was estimated separately for each and then averaged to arrive at a single estimate for each cycle.

The *modified Rogers' distances* (Reif et al., 2005) between the heterotic groups within each program were used as measures of heterotic group divergence. The distances were calculated for all programs and averaged to arrive at representative value for that cycle.

To describe the distribution of allele frequencies within each sub-population and hence the amount of available allelic diversity we calculated the proportion of loci with a minor allele frequency of less than 5%. This probability measures the thickness of the extreme tail of the allele frequency distribution and thus reflects the degree with which it follows a '*U-shape*' (Hill et al., 2008). This metric was evaluated for all sub-populations in each cycle and then averaged.

As a more high-level diversity metric we considered the *effective population size* (N_e) of each sub-population. N_e was calculated according to the method described by Corbin et al. (2012) for estimating constant effective population size. The so obtained values were averaged across sub-populations.

Hybrid breeding 'themes'

All previously described parameters, such as parameters related to the *NK* model and genetic architecture, parameters related to testcross evaluation, etc, were kept constant across the themes investigated.

In the *decentrality* theme we explored consequences of separating hybrid breeding programs into smaller, more or less isolated, units. We defined three distinct strategies for 'searching' (Podlich and Cooper, 1999) genetic space (Figure 4): a single large program (*centralized search*) to multiple smaller, fully isolated programs (*isolated search*). Between these two extremes we considered a strategy with multiple smaller programs that exchange germplasm in the form of breeding crosses (*distributed search*).

The centralized search was characterized by a single program consisting of one sub-population per heterotic group. The size of each was 500, for a total of 1,000 lines generated in each cycle. The number of lines selected to contribute to the next generation was 125 per sub-population. The relative individual contributions of these lines decreased proportionally with their performance ranks. The number

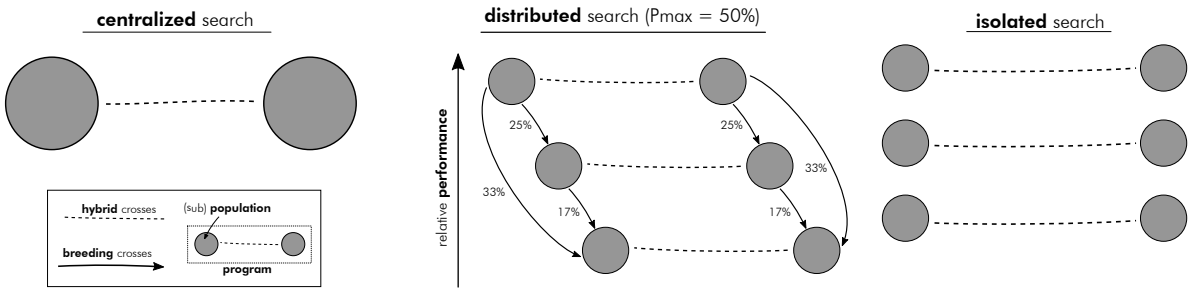

Figure 4 Schematic visualization of the three general search strategies explored in the *decentrality* theme.

of selected experimental hybrids was 125. The isolated search strategy comprised five programs, each with one sub-population per heterotic group. The sub-population size was 100 of which 25 were selected. Also here, the relative individual contributions of the lines were proportional to their performance ranks. The number of experimental hybrids selected per program was 25.

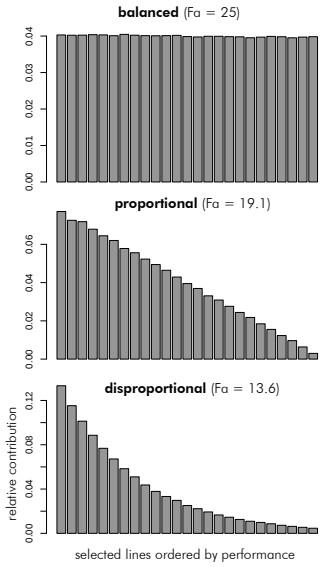

Figure 5 Distributions of relative individual contributions of selected inbred lines considered in the *inbred usage* theme.

In the distributed search we considered three levels of Pmax: 25%, 50% and 75%. The number of programs as well as lines and hybrids created and selected for each followed those of the isolated search. Note, thus, that the total number of lines and hybrids were the same across all scenarios as was the selection intensity.

In the *inbred usage* theme we explored the consequences of different degrees of imbalance in the relative contributions of the selected inbred lines to the next generation. The different scenarios explored correspond to the distributed search strategy with Pmax = 50%. Only the relative usage of inbred lines was varied. As described above, the observed relative contributions were drawn from a Dirichlet distribution with concentration parameter chosen in a way to result in a certain average relationship between relative contribution and performance rank. Three scenarios were considered (Figure 5). In the *balanced* usage scenario, all selected inbreds contributed equally on average, in the *proportional* scenario, the relative contribution declined proportional with the performance rank of the lines. In the *disproportional* scenario, contributions halved with every 5 ranks, meaning that the highest performing line will contribute twice as much to the next generation as the 5th ranked line. The increasing imbalance in contributions can be quantified as $1/\boldsymbol{b}'\boldsymbol{b}$ (with $\boldsymbol{b}$ being the vector of relative contributions), which is an estimate of the effective number of contributing lines (Boichard et al., 1997). For the balanced scenario, this was 25 and thus equal to the actual number of selected lines within each sub-population. It decreased to 19.1 and 13.6 for the proportional and disproportional scenarios, respectively.

Results

Decentrality theme

Which strategy achieved the highest peak performance depended on the value of the complexity parameter K, with the centralized strategy being superior at low $K < 5$, the distributed strategy at intermediate K and the isolated strategy at high values of K above eight (Figure 6A). The differences between the strategies tended to increase with increasing K. The centralized and distributed search

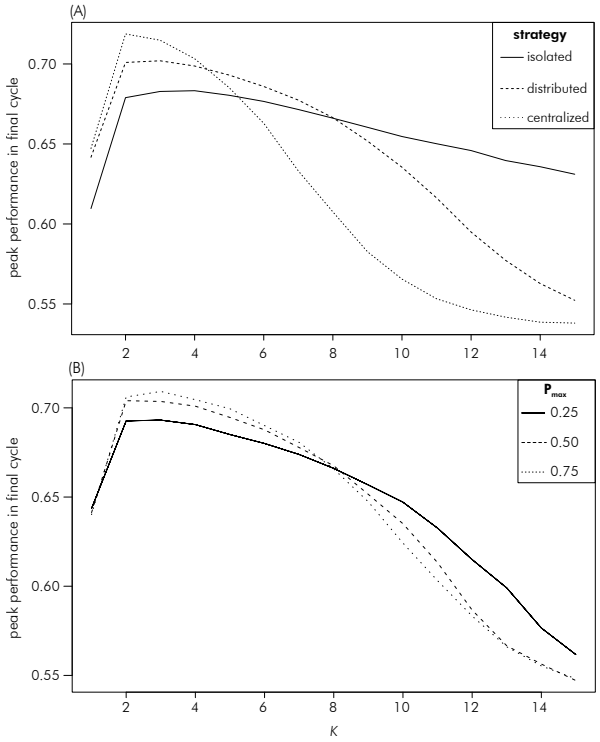

Figure 6 Relationship between the *NK* model complexity parameter K (average number of interacting genes) and peak genetic performance in the last cycle for the strategies explored in the decentrality theme: (A) comparing the isolated, distributed and centralized strategies and (B) the different values of Pmax within the distributed strategy. The curve of the distributed strategy in (A) is an average across the three Pmax scenarios within it.

strategies came very close to reaching the theoretical maximum peak performance at the special case of additivity ($K = 1$), but the isolated strategy remained considerably below that. Within the distributed strategy, the highest Pmax value of 75% was superior at K values below eight and the lowest Pmax of 25% at high K (Figure 6B). The case of Pmax = 50% had peak performance in between the two extremes, but more similar to Pmax = 75%. All Pmax scenarios achieved virtually identical peak performance at $K = 1$.

For brevity, trajectories across cycles are shown only for K values of 1, 6 and 15, representing the additive, multi-peaked but clustered and fully uncorrelated landscapes, respectively (Figure 1). Results for all values of K are available as supplemental information (File S1). At $K = 1$, the centralized search strategy had the highest peak performance in all cycles, closely followed by the three versions of the distributed search (Figures 8A, B, C). The peak performance of the isolated search strategy was considerably lower than that of the other strategies as it increased at a lower rate and seemed to reach a plateau at around cycle 20. At $K = 6$, the isolated search strategy achieved the highest peak performances in the earlier cycles but was overtaken by the distributed search strategies later. Those had very similar peak performances until the last few cycles when the version with Pmax of 25% fell behind. The centralized search had the lowest peak performances throughout, with the differences to the other strategies being particularly large between the intermediate cycles 15–20. Finally, at $K = 15$, only the isolated search had a sizable increase in peak performance cycle over cycle. The distributed search strategies showed an increase only in the last few cycles and the centralized strategy did not increase peak performance at all.

As expected %GCA was equal to one for all scenarios at $K = 1$ (Figures 8D, E, F). At $K = 6$, %GCA started at just below 10% and increased from there with each cycle. The rate of increase was greatest for the isolated strategy which reached almost 100% in the final cycles. The centralized search strategy had the slowest increase and was still below 50% in the final cycle. The distributed search strategies were intermediate between these two extremes. The increase was steepest for Pmax = 25% case, which translated to it having a markedly higher %GCA than the Pmax = 50% case and Pmax = 75% case during intermediate cycles 15–20. However, all three converged to a similar value of around 80% in the final cycle. At $K = 15$, %GCA started at zero and only the isolated strategy saw a marked increase in early cycles. The distributed search strategies saw an increase in %GCA noticeably above zero only in the final cycles and the centralized strategy remained at zero throughout.

The percent of loci with MAF < 0.05 increased over cycles for all strategies and complexity levels (Figures 8G, H, I). In all cases, the increase over cycles was strongest for the isolated search strategy, where it reached close to 100% in the final cycles and weakest in the centralized search strategy. The curves for the three Pmax levels of the distributed search strategy were similar to each other and intermediate compared to the two other strategies. The differences between the strategies increased with K because the increase in the proportion of loci at extreme frequencies slowed for the distributed and centralized strategies with increasing K. At the highest levels of complexity, only between 30% and

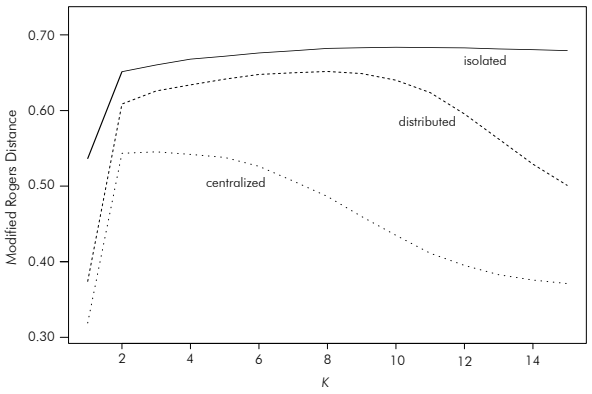

Figure 7 Modified Rogers Distance between heterotic groups as a function K, averaged across programs, in the connectivity theme. The curve of the distributed strategy represents the average across the three Pmax levels.

40% of loci showed a MAF < 0.05 in the different distributed strategies and less than 20% in the centralized strategy.

The modified Rogers Distance between heterotic groups increased over cycles for all strategies (Figures 8J, K, L). For all levels of complexity, this distance was highest for the isolated strategy and lowest for the centralized strategy, with the three versions of the distributed search having similar values that were intermediate to the two extremes (Figure 7).

The N_e differences among the strategies remained largely constant across cycles and levels of K. For the sake of brevity results will only be reported for cycle 15 and $K = 7$. The estimated N_e for each-sub-population for the isolated search strategy was 20.0, for the three versions of the distributed search it was 23.7 (Pmax = 25%), 31.3 (Pmax = 50%) and 35.4 (Pmax = 75%), respectively, and for the centralized search strategy 98.3.

Inbred usage theme

For brevity sake, only results for K of 1, 6 and 15 are shown (results for all values of K are provided in supplemental file S2). Again, which inbred usage scenario achieved the highest peak performance depended on the complexity level K (Figure 9). At the additive case of $K = 1$, all strategies achieved very similar peak performances close to the theoretical maximum of 2/3. Until $K = 8$, the highest peak performances were reached with proportional usage of selected inbred lines. For $K > 8$, disproportional use of inbred lines resulted in the highest peak performances. Balanced use generally resulted in the lowest peak performance, except for $K < 4$, where this strategy was slightly ahead of the disproportional usage strategy. The differences between the strategies tended to increase with K.

The cycle over cycle increase in peak performance was initially higher the more disproportional the use of the inbreds (Figures 10A, B, C). However, except for the highest level of complexity, this did not result in the highest maximum performance for this scenario, because the increase started to level off in the last five to ten cycles. Scenarios with proportional and balanced use of inbreds therefore had the highest peak performance at $K = 1$, though the differences were small. At the intermediate level of $K = 6$, the disproportional use scenario was overtaken by the proportional use scenario in the last cycles. The differences between these two were small, however. Finally, at $K = 15$, only the disproportional use scenario achieved a sizable increase in peak performance.

At $K = 1$, %GCA stayed constant at one for all scenarios, as expected. At $K = 6$, %GCA increased most strongly for disproportional use, followed by proportional and balanced use (Figures 10D, E, F). Reaching above 90% for the former, and above 80% and 50% for the latter two, respectively. At $K = 15$, %GCA remained near zero for the balanced and proportional use scenarios throughout. For the disproportional use scenario, it remained at zero as well until cycle ten and increased from there to almost 60%.

For all strategies and values of K, the percent of loci with a MAF < 0.05 increased from its initial value of zero (Figures 10G, H, I). The increase over cycles was strongest for disproportional inbred use, for which it reached close to 100% at K of 1 and 6. The proportional usage strategy had the second strongest increase and the balanced strategy the weakest. As was the case in the decentrality theme, the differences between the strategies tended to increase with K. At the highest level of $K = 15$, the proportional and balanced strategies stayed below 40%, whereas the disproportional usage strategy reached close to 80%.

The modified Rogers distance between heterotic groups increased over cycles in all scenarios (Figures 10J, K, L). Throughout it was highest for disproportional use, followed by proportional and balanced use. Overall, the distance was greatest for the intermediate complexity level of $K = 6$.

N_e, again reported only for cycle 15 at $K = 7$, was

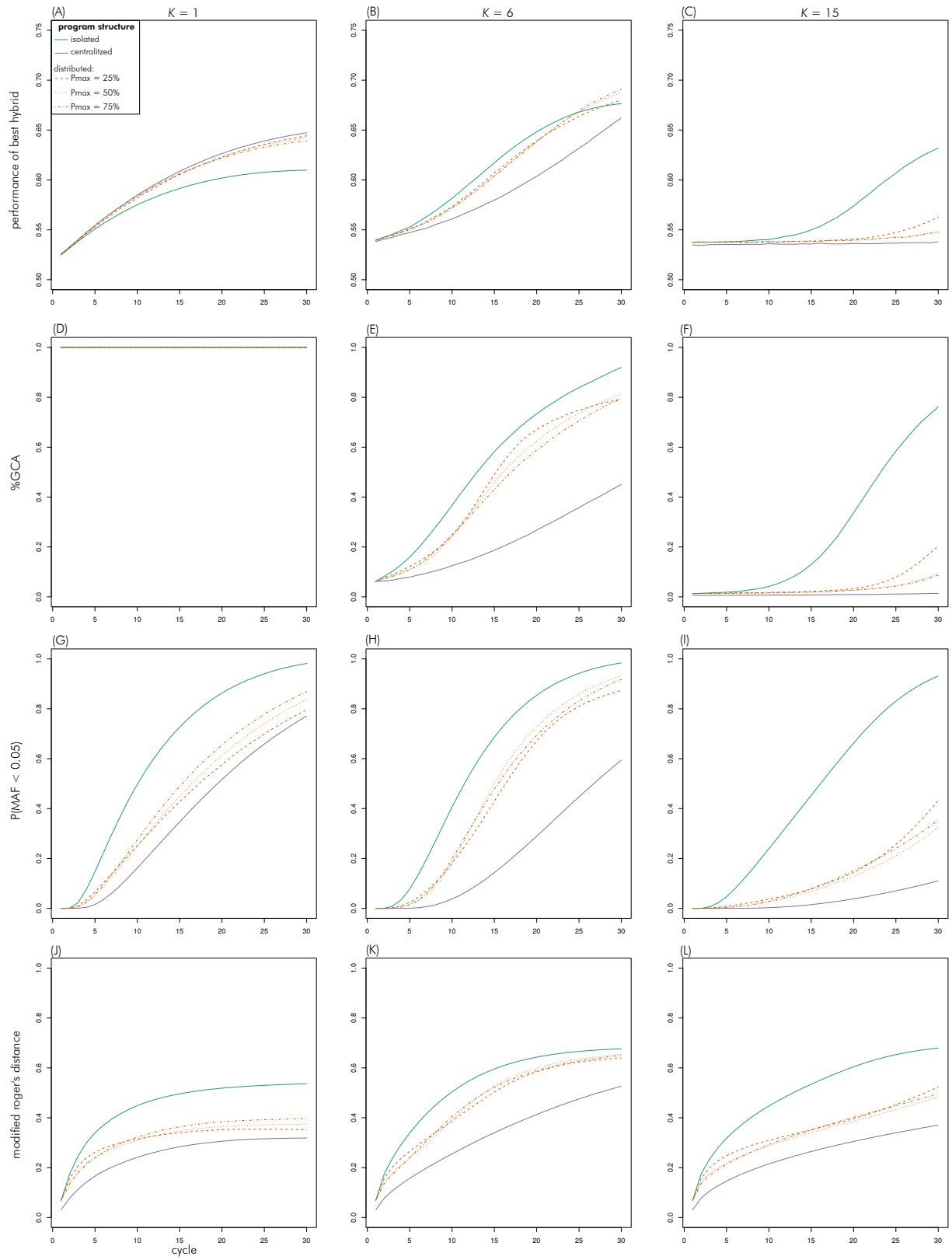

Figure 8 Evolution of metrics over cycles in the decentrality theme for scenarios with K of 1 (left column), 6 (middle) and 15 (right).

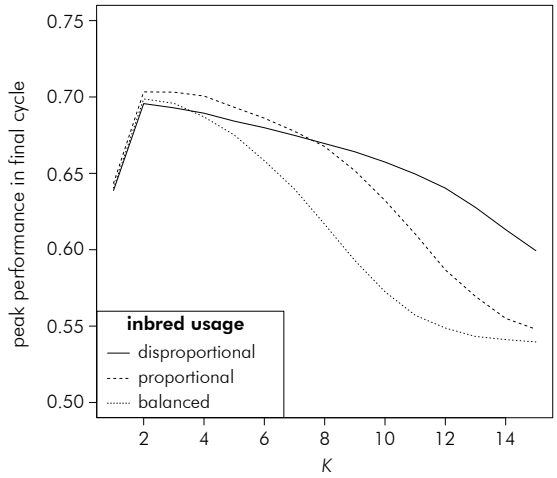

Figure 9 Relationship between the *NK* model complexity parameter K and peak genetic performance in the last cycle for the strategies explored in the the inbred usage theme.

23.2, 31.3 and 44.6, for the disproportional, proportional and balanced usage scenarios, respectively.

Discussion

The objective of this study was to explore properties of the historically grown structure of commercial hybrid breeding programs, particularly in maize, and aid the understanding of why these structures successfully generated significant amounts of genetic gain in the past and thereby impact global food security. The infinitesimal framework (Barton et al., 2017), in which traits are described as the sum of a large number of genes all having additive, context independent effects of similar magnitude, is our starting point. However, we seek extensions to account for the empirical observations that a) there are results observed from operating a long-term breeding effort that are not consistent with or not easily explained within the infinitesimal model framework (Rasmusson and Phillips, 1997) and b) reflect the reality of a highly complex trait biology (Hammer et al., 2006). Therefore, we are motivated to consider the influence of complexity of trait genetic architecture on breeding strategies from the perspective of a long-term commercial breeding program (Duvick et al., 2004; Cooper et al., 2014).

Emergence of additivity

As a representation of genetic complexity we chose the *NK* model framework developed by Kauffman (1993), which allows exploration of the full continuum from complete additivity to deep and almost intractable genetic complexity. For reference, the *NK* models used in this study, corresponded to the Mount Fuji landscape at $K = 1$ (Figure 1) and to the 'Alps' landscape from $K = 2$ to K of 7 or 8. After this the genetic models transitioned from the multi-peaked but correlated 'Alps' landscape to the uncorrelated landscape represented by the 'Dunes' metaphor (Figures 1 and 2).

In complex genetic landscapes, additive genetic variance, the sine qua non of genetic gain, is not a constant factor of trait biology (i.e., deducible from the molecular properties of genes) but rather emerging from the interplay of biology and natural or artificial properties of population structure (Wade, 2002; Cooper et al., 2005). In particular, additivity emerges in response to a constraining of the dimensionality of genetic space, or, in other words, by limiting genetic diversity. In practice, such constraints in dimensionality are achieved through fixation or near fixation of genes (Wade, 2002; Hill et al., 2008). This process is illustrated in Figure 11 for a simple epistatic network consisting of two genes. Thus, as genetic complexity increases, the breeder needs practical ways to reduce this complexity to a manageable level that allows genetic progress. This study explored two particular practical approaches that have been adopted within commercial hybrid breeding, particularly in maize. With the availability of genomics and novel thinking about genetic complexity, we can now study the genetic implications of these practical approaches, many of which were devised and adopted prior to the availability of a theoretical and empirical framework to study their effects.

Two processes in particular accelerate such constrainment, namely the creation of population bottlenecks and the subdivision of larger populations into more or less independent 'demes' (Katz and Young, 1975; Goodnight, 1995). Equivalent processes in the context of plant breeding programs are the degree of connectivity between breeding programs and the relative use of superior inbred lines in breeding crosses for producing the next generation, both 'themes' were explored in this study.

Decentrality theme

Classical quantitative genetics infinitesimal theory was used to design and optimize commercial hybrid breeding strategies, in combination with em-

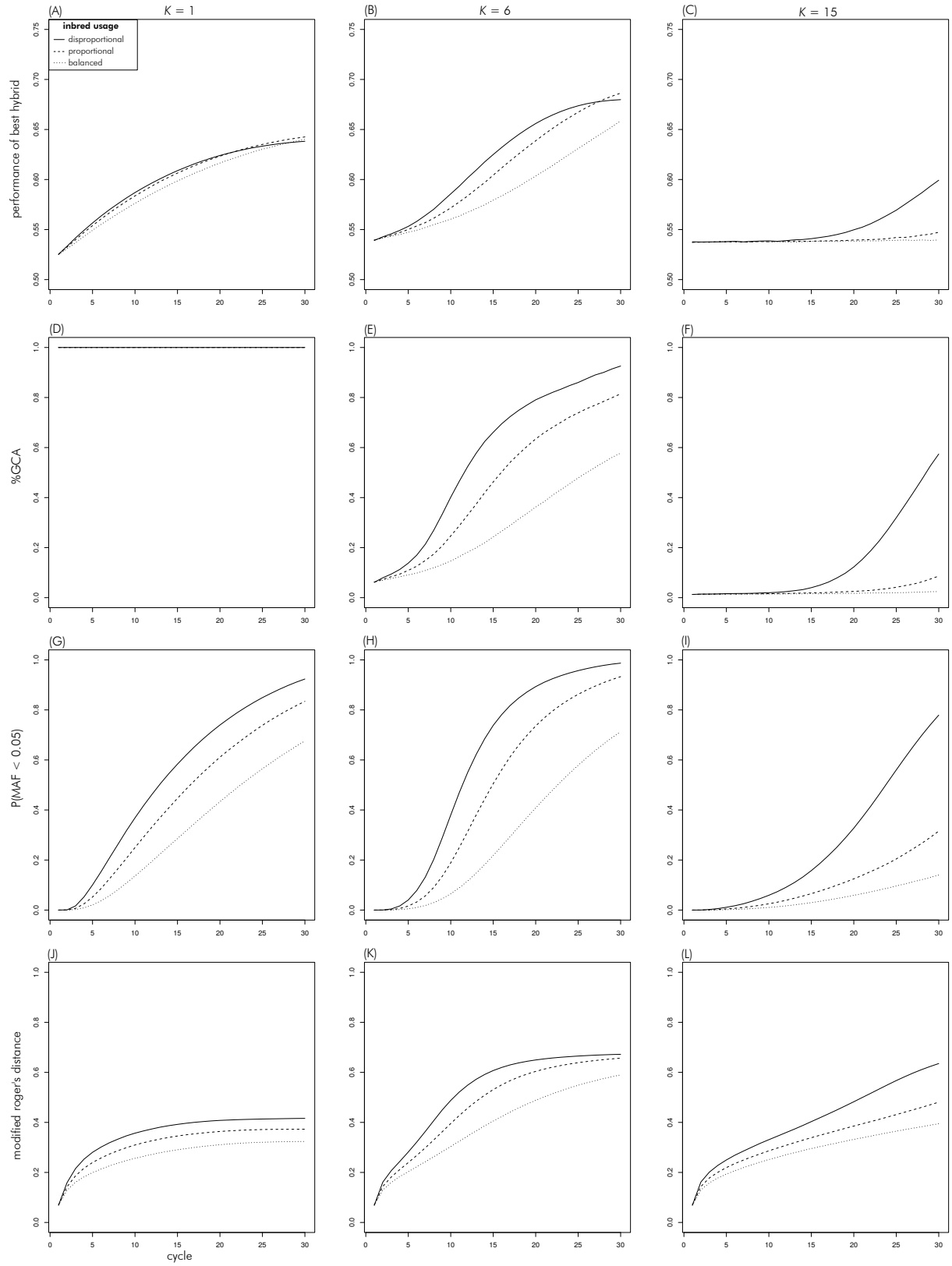

Figure 10 Evolution of metrics over cycles in the inbred usage theme for scenarios with K of 1 (left column), 6 (middle) and 15 (right).

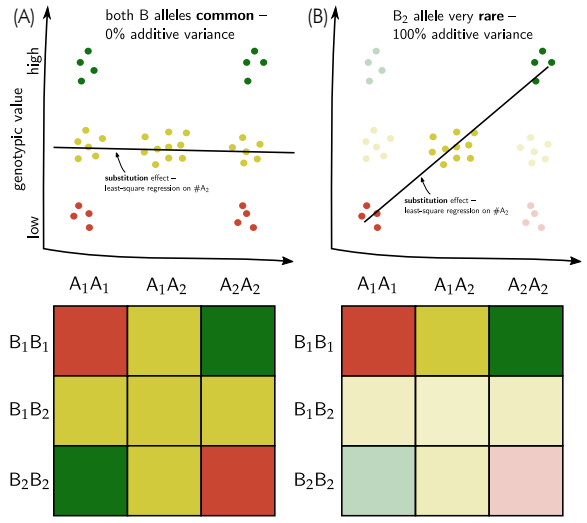

Figure 11 Illustration of a two-locus epistatic network in which neither the A nor the B locus exhibit any additive variation when all alleles are common (A) but collapses to a perfectly additive system in which the substitution effect of the A locus explains 100% of the variation when the B_2 allele becomes rare (B). Note that the substitution effect of the A locus would be reversed in sign when the B_1 allele became rare instead. Colours green, yellow and red represent high, intermediate and low phenotypic values, respectively.

pirical experience of what worked and what did not (Hallauer et al., 2010). Yet, even though the infinitesimal model implies optimality of a single, homogenous population, there were discussions about the relative merits of large centralized vs. decentralized breeding programs early on (Baker and Curnow, 1969). Later on, Podlich and Cooper (1999) explored this problem on the basis of Sewall Wright's *shifting balance theory*, (Wright, 1931, 1977; Wade and Goodnight, 1998). The shifting balance theory describes an evolutionary process in which genetic drift resulting from population subdivision enables random movements across genetic space (i.e., against selection gradients) and also converts epistatic to additive genetic variation through constraining genetic space as described above. This then enables local adaptation in complex genetic landscapes which is followed by differential migration from higher to lower performing sub-populations and thus 'spreading' of superior gene complexes across the whole *meta population*.

While this theory remains controversial as a model of natural evolution (Coyne et al., 1997), there are remarkable similarities between meta populations in the context of the shifting balance theory and the population structure of large commercial breeding operations. The latter also do not operate as a centralized unit but rather as a decentralized network of smaller programs with the most successful germplasm being shared across (Cooper et al., 2014). The same seems to be the case at the industry level, with the major commercial breeding operations being based on unique and distinct germplasm backgrounds, with only occasional exchange of elite material, e.g., through ex-PVP lines (Mikel and Dudley, 2006; White et al., 2020). As a result of this decentralization, plant breeding programs are also characterized by having low effective population sizes (Cowling, 2013), which makes them more susceptible to genetic drift.

Here, we expanded on the work of Baker and Curnow (1969) and Podlich and Cooper (1999) by exploring breeding population structures with differing degrees of decentrality (Figure 4), ranging from a large centralized program with high N_e to a isolated set of smaller programs with low N_e, with a series of scenarios with decentralized but connected programs with N_e values in between these two extremes. We indeed found that strategies resulting in low within program N_e through increased decentralization and isolation became increasingly superior in terms of peak hybrid performance, as genetic complexity K increased, while a centralized strategy with high N_e was superior in less complex landscapes. These results thus confirm the findings of Podlich and Cooper (1999) that a decentralized search strategy is superior in complex genetic landscapes. Increasing isolation and decentralization and the associated N_e reduction led to quicker increases over cycles and higher overall values of of %GCA (Figure 8). At the highest levels of complexity, only complete isolation generated amounts of GCA variation sufficient for making genetic improvements cycle over cycle. The better ability to expose additivity in the form of GCA variation of the more isolated and decentralized strategies was expected as per the discussion at the beginning of this section outlining the relationship between amounts of additive variation and constrainment of genetic space. This explains the clear advantages in terms of genetic peak performance of the isolated strategy at the highest levels of K.

The corollary of the constrainment of genetic space of course is a more rapid decline in genetic diversity and susceptibility to genetic drift, which ultimately limits the selection potential of the programs. Indeed, for values of K below eight, which marked the switch from an uncorrelated to multi-peaked but

correlated genetic landscape (Figure 2), decentralized programs with increasing rates of germplasm exchange became superior. Accordingly, having a large centralized program became the optimal strategy at lower values of K. Here, the genetic landscape was simple enough to not require severe constrainment of genetic space to expose sufficient amounts of GCA variation. The genetic drift experienced by small, isolated programs then unnecessarily led to the fixation of unfavourable alleles. This was most apparent at $K = 1$ where all variation is additive by definition and a decentralized strategy is not expected to have any advantage (Rathie and Nicholas, 1980). Here the isolated strategy led to fixation of almost all loci from cycle twenty onward and to a stalling of genetic gain significantly below the theoretically achievable maximum (Figure 8).

The establishment of genetically divergent heterotic groups has always been a central tenant of hybrid breeding (Melchinger and Gumber, 1998). Originally, optimal exploitation of heterosis was the main driver of their establishment (East, 1936). Later, however, maximization of GCA vs. SCA variation was identified as an import secondary feature of heterotic groups (Melchinger and Gumber, 1998). While this is well established for dominant gene action (Reif et al., 2007; Fischer et al., 2009), there are also indications for the conservation of favourable epistatic patterns that are disrupted when lines from different heterotic groups are recombined (Bernardo, 2001). Often, heterotic groups are established from populations that evolved in isolation for a long time. One of the best examples for this is the Dent by Flint heterotic pattern in maize which is prevalent in Central Europe and is comprised of populations that evolved in separation for centuries (Rebourg et al., 2003). Heterotic groups are thus a different and additional form of constrainment of genetic space through historically grown genetic isolation. In our simulations, the different heterotic groups were originated from the same base population, yet we still observed a significant degree of genetic differentiation evolve over cycles (Figure 8), as expected in recurrent, reciprocal selection regimes (Labate et al., 1999; Longin et al., 2013). A portion of this differentiation can be attributed to genetic drift (Gerke et al., 2015), as evidenced by the non-zero genetic distance at $K = 1$, where all effects are additive and increasing genetic differentiation between heterotic groups would have no effect on the proportion of GCA variance. However, the genetic differentiation was considerably higher for $K > 1$ (Figure 7), indicating that there indeed was a selection advantage to increased heterotic group divergence in complex genetic landscapes. This was particularly clear for the isolated scenario, where heterotic patterns could form uninterrupted within programs. For the centralized and decentralized strategies, the differentiation was maximal at lower values of K, because %GCA, and hence the effectiveness of recurrent, reciprocal selection, declined afterwards.

Inbred usage theme

The history of North American and European maize germplasm can be described as a succession of key inbreds that were heavily used in breeding crosses and had a distinct and lasting impact on germplasm (Mikel and Dudley, 2006; Technow et al., 2014; White et al., 2020). Those inbreds owe their success either to the outstanding general combining ability relative to their peers at the time, such as was case for the important North American line B73 (Mikel and Dudley, 2006) or their unique adaptation to specific climatic conditions, such as the European Flint lines *F2* and *F7* (Messmer et al., 1992; Böhm et al., 2014). The highly disproportionate importance of successful inbreds led to a significant reduction in genetic diversity (Rasmusson and Phillips, 1997; White et al., 2020), particularly relative to the source populations from which they were derived (Böhm et al., 2017). However, this constrainment also might be responsible for the emergence of additive genetic variation from complex gene action through the so called founder or bottleneck effect (Goodnight, 1988; Cheverud and Routman, 1996; Naciri-Graven and Goudet, 2003; van Heerwaarden et al., 2008). We indeed observed that %GCA increased faster over cycles and reached higher values overall the more uneven the use of selected parents in breeding crosses (Figure 10), with the exception of $K = 1$, where all variance is additive by definition. At the highest degrees of landscape complexity only disproportionate use of inbreds, resulting in very low N_e, succeeded in generating amounts of %GCA sufficient for genetic improvements. Like in the decentrality theme, the higher values of %GCA of the disproportional use strategy translated into superior peak performances only at the values $K > 8$, i.e., after the landscape transitioned from multi-peaked but correlated to uncorrelated. Before that, balanced and particularly proportional use, both having higher N_e, achieved superior peak performances.

Maintenance of diversity

In our simulations, the constrainment of genetic space and reduction of N_e through decentralization and isolation or disproportionate use of inbred lines, while necessary for exposing additive genetic variation, led to a rapid fixation of alleles and a slowing of genetic gain in later cycles. This was partly a consequence of genetic drift caused by the low N_e (Cowling, 2013). However, the reduction of N_e was also caused in part by the effects of selection, particularly once the majority of the genetic variation was additive. This has not generally happened in commercial breeding programs, where genetic gain continues apace (Rasmusson and Phillips, 1997; Fischer et al., 2008; Duvick et al., 2004; Pfeiffer et al., 2019). Several factors that maintain diversity in practical programs were not included in the simulation model. For example, the simulation implicitly assumed that the environment and management conditions remained constant across all cycles, whereas both change more or less rapidly in reality. Changing selection environments imply changing selection targets and trajectories (Messina et al., 2011; Hammer et al., 2009), which reduce the pressure on particular alleles or allele complexes and thus slow or prevent fixation. Long-term selection experiments have shown that selection response can be maintained even in isolated and narrow populations (Dudley and Lambert, 2010; Durand et al., 2010, 2015). Several hypotheses were proposed for these surprising results, including epistasis (Carlborg et al., 2006), de novo genetic mutations, particularly when magnified through effects on epistatic complexes (Rasmusson and Phillips, 1997; Durand et al., 2010), creation of heritable epigenetic variation (Hauben et al., 2009), activity of transposable elements (Dubin et al., 2018), as well as the presence of 'cryptic genetic variation' through phenomena such as canalization (Gibson and Dworkin, 2004). Of these, only epistasis was present in our simulations. While highly speculative, these biological phenomena might explain the presence of abundant genetic variation and continued genetic gain in largely isolated and genetically narrow commercial plant breeding programs.

Applications of the *NK* for plant breeding

The *NK* model, originally developed by the theoretical biologist Stuart Kauffman to study evolution in complex genetic landscapes, has found applications for modelling complex systems in disparate fields such as business administration (Csaszar, 2018), organizational learning theory (Lazer and Friedman, 2007), infrastructure design (Grove and Baumann, 2012) and physics (Qu et al., 2002). Following the example of (Podlich and Cooper, 1999), we here used the *NK* model to represent genetic complexity navigated by commercial hybrid breeding operations to study the effect of the degree of isolation between programs as well as the degree of imbalance in inbred usage, both key aspects of breeding strategies. We propose that this model can serve as an ideal starting point to study other aspects of hybrid breeding strategies. For example, Cooper and Podlich (2002) proposed an extension to the *NK* model that adds an environmental dimension and thus allows modelling concepts related to genotype by environment interaction (Cooper and DeLacy, 1994), yield stability (Piepho, 1998; Tollenaar and Lee, 2002), product placement (Messina et al., 2018) and the target population of environments (Comstock, 1977). These so called *E(NK)* models represent different environments or management practices through a series of more or less similar genetic landscapes. This of course adds considerable complexity to the already complex static landscapes studied here and poses interesting dilemmas. For example, rapidly exposing additive variation, e.g., through isolation, might be even more important than in static landscapes because local optima have to be exploited quickly before they disappear once the environment shifts, for example through changes in management such as the historical increases in plant population for commercial maize production (Hammer et al., 2009). However, retaining allelic diversity, which hampers the exposing of additivity, is required to enable renewed adaptation to the changed environmental conditions.

A high degree of genetic complexity also implies a high degree of context dependency of genetic effects. Observe, for example, that the additive substitution effect of the 'A' locus in Figure 11 changes sign when the B_1 allele instead of the B_2 allele becomes rare. This has consequences on the persistence of accuracy of estimated QTL effects and genomic prediction models and can be addressed through iteratively updating training populations for genetic model parameterization (Podlich et al., 2004; Wolc et al., 2016; Forneris et al., 2017). The *NK* model framework can help address questions about the frequency with which this has to happen and whether data from previous generations

can be used. Recently, approaches were proposed that attempt to capture those context dependencies through biological models representing the interdependencies underlying the traits of interest (Technow et al., 2015). Such models are only approximations of the true biological complexity. However, Cooper et al. (2005), using the *NK* framework, have shown that even incomplete knowledge of biological networks can improve predictability of genetic effects and genetic gain. The context dependency of genetic effects, i.e., the effects being neither universally positive or negative (Wade, 2002), also has implications on innovative proposals for using CRISPR-Cas9 gene editing (Jaganathan et al., 2018; Gao et al., 2020) to either target recombination to create superior hypothetical linkage groups (Brandariz and Bernardo, 2019) or even the large scale "editing away" of deleterious mutations (Wallace et al., 2018). Finally, this framework might help devise strategies for the efficient introduction of exotic or ancient germplasm (Yu et al., 2016; Böhm et al., 2017), which evolved not just in a very different environmental, but also a different genetic context from the current elite breeding germplasm.

Back to the future

The structure of commercial plant breeding programs, particularly in major crops like maize, is characterized by a large degree of decentralization with exchange of successful germplasm within companies (Cooper et al., 2014), while isolation is the norm among companies (Mikel and Dudley, 2006). Plant breeders further have a tendency for relying on only a small set of elite inbred lines for producing the next generation (Rasmusson and Phillips, 1997), leading to a series of significant bottleneck events (White et al., 2020). All of these features lead to a drastically reduced effective population size and are not predicted to be promising strategies under the additive, infinitesimal model of quantitative genetics. Yet commercial hybrid breeding has delivered incredible amounts of genetic gain over the last century, and has thus contributed to food security and resource conservation (Duvick, 1999). Here we postulated that the described structure of plant breeding programs, with its constrainment of genetic space, is in fact necessary for enabling the exploration and exploitation of genetic variation in complex genetic landscapes and that the success of a breeding program is not only determined by its germplasm per se, but by the structures that allow it to evolve. The breeding program structures described here grew historically and we do not claim that it was set up with this intention. However, by doing "what worked", breeders in preceding generations might have nonetheless been able to take advantage of the process described and postulated in this study. Understanding why these historic structures "worked" will be critical for designing breeding programs that can tackle the challenges of the century ahead.

References

Altenberg, L. (1994). Evolving better representations through selective genome growth. In *Proceedings of the First Ieee Conference on Evolutionary Computation IEEE World Congress on Computational Intelligence ICEC-94*, pp. 182–187. IEEE.

Araus, J. L. and J. E. Cairns (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci 19*, 52–61.

Baker, L. H. and R. N. Curnow (1969). Choice of Population Size and Use of Variation Between Replicate Populations in Plant Breeding Selection Programs 1. *Crop Sci 9*, 555–560.

Barton, N. H., A. M. Etheridge, and A. Véber (2017). The infinitesimal model: Definition, derivation, and implications. *Theor Pop Biol 118*, 50–73.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. *J Stat Soft 67*(1), 1–48.

Bernardo, R. (2001). breeding potential of intra- and interheterotic group crosses in maize. *Crop Sci 41*, 68–71.

Bernardo, R. (2016). Bandwagons I, too, have known. *Theor Appl Genet 129*, 2323âĂŞ2332.

Böhm, J., W. Schipprack, V. Mirdita, H. F. Utz, and A. E. Melchinger (2014). Breeding potential of european flint maize landraces evaluated by their testcross performance. *Crop Sci 54*, 1665–1672.

Böhm, J., W. Schipprack, H. F. Utz, and A. E. Melchinger (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.

Boichard, D., L. Maignel, and É. Verrier (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genet Sel Evol 29*, 5–23.

Brandariz, S. P. and R. Bernardo (2019). Predicted genetic gains from targeted recombination in elite biparental maize populations. *Plant Genom 12*, 10.3835/plantgenome2018.08.0062.

Carlborg, O. and C. S. Haley (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet 5*, 618–625.

Carlborg, O., L. Jacobsson, P. Åhgren, P. Siegel, and L. Andersson (2006). Epistasis and the release of genetic variation during long-term selection. *Nat Genet 38*, 418–420.

Cheverud, J. M. and E. J. Routman (1996). Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution 50*, 1042–1051.

Comstock, R. E. (1977). Quantitative genetics and the design of breeding programs. In *Proceedings of the International Conference on Quantitative Genetics*, Ames, IA, pp. 16–21. Iowa State University Press.

Cooper, M. and I. H. DeLacy (1994). Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theor Appl Genet 88*, 561–572.

Cooper, M., C. D. Messina, D. Podlich, L. R. Totir, A. Baumgarten, N. J. Hausmann, D. Wright, and G. Graham (2014). Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop Pasture Sci 65*, 311.

Cooper, M. and D. W. Podlich (2002). The E(NK) model: Extending the NK model to incorporate gene-by-environment interactions and epistasis for diploid genomes. *Complexity 7*, 31–47.

Cooper, M., D. W. Podlich, and O. S. Smith (2005). Gene-to-phenotype models and complex trait genetics. *Aust J Agric Res 56*, 895–918.

Corbin, L. J., A. Y. H. Liu, S. C. Bishop, and J. A. Woolliams (2012). Estimation of historical effective population size using linkage disequilibria with marker data. *J Anim Breed Genet 129*, 257–270.

Cowling, W. A. (2013). Sustainable plant breeding. *Plant Breed 132*, 1–9.

Coyne, J. A., N. H. Barton, and M. Turelli (1997). Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution 51*, 643–671.

Csaszar, F. A. (2018). A note on how NK landscapes work. *J Org Design 7*, 10.1186/s41469–018–0039–0.

Dubin, M. J., O. Mittelsten Scheid, and C. Becker (2018). Transposons: a blessing curse. *Curr Opin Plant Biol 42*, 23–29.

Dudley, J. W. and R. J. Lambert (2010). 100 Generations of Selection for Oil and Protein in Corn. In *Plant Breeding Reviews*, pp. 79–110. John Wiley & Sons, Ltd.

Durand, E., M. I. Tenaillon, X. Raffoux, S. ThÃľpot, M. Falque, P. Jamin, A. Bourgais, A. Ressayre, and C. Dillmann (2015). Dearth of polymorphism associated with a sustained response to selection for flowering time in maize. *BMC Evol Biol 15*, 10.1186/s12862–015–0382–5.

Durand, E., M. I. Tenaillon, C. Ridel, D. Coubriche, P. Jamin, S. Jouanne, A. Ressayre, A. Charcosset, and C. Dillmann (2010). Standing variation and new mutations both contribute to a fast response to selection for flowering time in maize inbreds. *BMC Evol Biol 10*, 10.1186/1471–2148–10–2.

Duvick, D. (1999). Heterosis: feeding people and protecting natural resources. In J. Coors and S. Pandey (Eds.), *The genetics and exploitation of heterosis in crops*, pp. 19–29. Madison, WI: CSSA.

Duvick, D., J. Smith, and M. Cooper (2004). Long-term selection in a commercial hybrid maize breeding program. In J. Janick (Ed.), *Plant Breeding Reviews*, pp. 109–152. Hoboken, NJ: John Wiley & Sons, Inc.

Dwivedi, S. L., A. B. Britt, L. Tripathi, S. Sharma, H. D. Upadhyaya, and R. Ortiz (2015). Haploids: constraints and opportunities in plant breeding. *Biotechnol Adv 33*, 812–829.

East, E. M. (1936). Heterosis. *Genetics 21*, 375–397.

Falconer, D. S. and T. F. C. Mackay (1996). *Introduction to Quantitative Genetics* (4 ed.). London: Pearson.

Fiévet, J. B., C. Dillmann, and D. de Vienne (2010). Systemic properties of metabolic networks lead to an epistasis-based model for heterosis. *Theor Appl Genet 120*, 463–473.

Fiévet, J. B., T. Nidelet, C. Dillmann, and D. de Vienne (2018). Heterosis is a systemic property emerging from non-linear genotype-phenotype relationships: evidence from in vitro genetics and computer simulations. *Front Genet 9*, 10.3389/fgene.2018.00159.

Fischer, S., J. Möhring, H. P. Maurer, H.-P. Piepho, E.-M. Thiemt, C. C. Schön, A. E. Melchinger, and J. C. Reif (2009). Impact of genetic divergence on the ratio of variance due to specific vs. general combining ability in winter triticale. *Crop Sci 49*, 2119–2122.

Fischer, S., J. Möhring, C. C. Schön, H.-P. Piepho, D. Klein, W. Schipprack, H. F. Utz, A. E. Melchinger, and J. C. Reif (2008). Trends in genetic variance components during 30 years of hybrid maize breeding at the University of Hohenheim. *Plant Breed 127*, 446–451.

Forneris, N. S., Z. G. Vitezica, A. Legarra, and M. Pérez-Enciso (2017). Influence of epistasis on response to genomic selection using complete sequence data. *Genet Sel Evol 49*, 10.1186/s12711–017–0340–3.

Gao, H., M. J. Gadlage, H. R. Lafitte, B. Lenderts, M. Yang, M. Schroder, J. Farrell, K. Snopek, D. Peterson, L. Feigenbutz, S. Jones, G. S. Clair, M. Rahe, N. Sanyour-Doyel, C. Peng, L. Wang, J. K. Young, M. Beatty, B. Dahlke, J. Hazebroek, T. W. Greene, A. M. Cigan, N. D. Chilcoat, and R. B. Meeley (2020). Superior field performance of waxy corn engineered using CRISPR-Cas9. *Nat Biotechnol 38*, 10.1038/s41587–020–0444–0.

Gaynor, R. C., G. Gorjanc, A. R. Bentley, E. S. Ober, P. Howell, R. Jackson, I. J. Mackay, and J. M. Hickey (2017). A two-part strategy for using genomic selection to develop inbred lines. *Crop Sci 57*, 2372–2386.

Gerke, J. P., J. W. Edwards, K. E. Guill, J. Ross-Ibarra, and M. D. McMullen (2015). The genomic impacts of drift and selection for hybrid performance in maize. *Genetics 201*, 1201–1211.

Gibson, G. and I. Dworkin (2004). Uncovering cryptic genetic variation. *Nat Rev Genet 5*, 681–690.

Goodnight, C. J. (1988). Epistasis and the effect of founder events on the additive genetic variance. *Evolution 42*, 441–454.

Goodnight, C. J. (1995). Epistasis and the increase in additive genetic variance: implications for phase 1 of Wright's shifting-balance process. *Evolution 49*, 502–511.

Grove, N. and O. Baumann (2012). Complexity in the telecommunications industry: When integrating infrastructure and services backfires. *Telecomm Policy 36*, 40–50.

Hallauer, A. R., M. J. Carena, and J. B. M. Filho (2010). *Quantitative Genetics in Maize Breeding* (3 ed.). Handbook of Plant Breeding. New York: Springer-Verlag.

Hammer, G., M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. van Eeuwijk, S. Chapman, and D. Podlich (2006). Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci 11*, 587–593.

Hammer, G. L., Z. Dong, G. McLean, A. Doherty, C. Messina, J. Schussler, C. Zinselmeier, S. Paszkiewicz, and M. Cooper (2009). Can changes in canopy and/or root system architecture explain historical maize yield trends in the U.S. corn belt? *Crop Sci 49*, 299–312.

Hauben, M., B. Haesendonckx, E. Standaert, K. Van Der Kelen, A. Azmi, H. Akpo, F. Van Breusegem, Y. Guisez, M. Bots, B. Lambert, B. Laga, and M. De Block (2009). Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proc Natl Acad Sci 106*, 20109–20114.

Hickey, L. T., A. N. Hafeez, H. Robinson, S. A. Jackson, S. C. M. Leal-Bertioli, M. Tester, C. Gao, I. D. Godwin, B. J. Hayes, and B. B. H. Wulff (2019). Breeding crops to feed 10 billion. *Nat Biotechnol 37*, 744–754.

Hill, W. G., M. E. Goddard, and P. M. Visscher (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet 4*, e1000008.

Huang, W. and T. F. C. Mackay (2016). The genetic architecture of quantitative traits cannot be inferred from variance component analysis. *PLOS Genet 12*, 10.1371/journal.pgen.1006421.

Jaganathan, D., K. Ramasamy, G. Sellamuthu, S. Jayabalan, and G. Venkataraman (2018). CRISPR for crop improvement: an update review. *Front Plant Sci 9*, 10.3389/fpls.2018.00985.

Jiang, Y., R. H. Schmidt, Y. Zhao, and J. C. Reif (2017). A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nature Genet 49*, 1741–1746.

Katz, A. and S. Young (1975). Selection for high adult body weight in Drosophila populations with different structures. *Genetics 81*, 163–175.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Labate, J., K. R. Lamkey, M. Lee, and W. L. Woodman (1999). Temporal changes in allele frequencies in two reciprocally selected maize populations. *Theor Appl Genet 99*, 1166–1178.

Larièpe, A., L. Moreau, J. Laborde, C. Bauland, S. Mezmouk, L. Décousset, T. Mary-Huard, J. B. Fiévet, A. Gallais, P. Dubreuil, and A. Charcosset (2017). General and specific combining abilities in a maize (Zea mays L.) test-cross hybrid panel: relative importance of population structure and genetic divergence between parents. *Theor Appl Genet 130*, 403–417.

Lazer, D. and A. Friedman (2007). The network structure of exploration and exploitation. *Adm Sci Q 52*, 667–694.

Lippman, Z. B. and D. Zamir (2007). Heterosis: revisiting the magic. *Trends Genet 23*, 60–66.

Longin, C. F. H., M. Gowda, J. Mühleisen, E. Ebmeyer, E. Kazman, R. Schachschneider, J. Schacht, M. Kirchhoff, Y. Zhao, and J. C. Reif (2013). Hybrid wheat: quantitative genetic parameters and consequences for the design of breeding programs. *Theor Appl Genet 126*, 2791–2801.

Lynch, M. and B. Walsh (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates.

Mackay, T. F. C. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet 15*, 22–33.

Melchinger, A. E. (1999). Genetic diversity and heterosis. In J. Coors and S. Pandey (Eds.), *The genetics and exploitation of heterosis in crops*, pp. 99–118. Madison, WI: ASA, CSSA, and SSSA.

Melchinger, A. E., H. H. Geiger, G. Seitz, and G. A. Schmidt (1987). Optimum prediction of three-way crosses from single crosses in forage maize (Zea mays L.). *Theor Appl Genet 74*, 339–345.

Melchinger, A. E. and R. K. Gumber (1998). Overview of heterosis and heterotic groups in agronomic crops. In *Concepts and Breeding of Heterosis in Crop Plant*, pp. 29–44. CSSA.

Messina, C., F. Technow, T. Tang, R. Totir, C. Gho, and M. Cooper (2018). Leveraging biological insight and environmental variation to improve phenotypic prediction: Integrating crop growth models (CGM) with whole genome prediction (WGP). *Eur J Agron 100*, 151–162.

Messina, C. D., D. Podlich, Z. Dong, M. Samples, and M. Cooper (2011). Yield-trait performance landscapes: from theory to application in breeding maize for drought tolerance. *J Exp Bot 62*, 855–868.

Messmer, M. M., A. E. Melchinger, J. Boppenmaier, E. Brunklaus-Jung, and R. G. Herrmann (1992). Relationships among early european maize inbreds: I. genetic diversity among flint and dent lines revealed by RFLPs. *Crop Sci 32*, 1301–1309.

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

Mikel, M. A. and J. W. Dudley (2006). Evolution of North American dent corn from public to proprietary germplasm. *Crop Sci 46*, 1193–1205.

Montana, G. (2005). Hapsim: a simulation tool for generating haplotype data with pre-specified allele frequencies and ld coefficients. *Bioinformatics 21*, 4309–4311.

Naciri-Graven, Y. and J. R. M. Goudet (2003). The additive genetic variance after bottlenecks is affected by the number of loci involved in epistatic interactions. *Evolution 57*, 706–716.

Pfeiffer, B. K., D. Pietsch, R. W. Schnell, and W. L. Rooney (2019). Long-term selection in hybrid sorghum breeding programs. *Crop Sci 59*, 150–164.

Phillips, P. C. (2008). Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet 9*, 855–867.

Piepho, H. P. (1998). Methods for comparing the yield stability of cropping systems. *J Agron Crop Sci 180*, 193–213.

Pinheiro, H. P., A. de Souza Pinheiro, and P. K. Sen (2005). Comparison of genomic sequences using the Hamming distance. *J Stat Plan Inference 130*, 325–339.

Podlich, D. and M. Cooper (1999). Modelling plant breeding programs as search strategies on a complex response surface. *Lecture Notes in Computer Science 1585*, 171–178.

Podlich, D. W., C. R. Winkler, and M. Cooper (2004). Mapping as you go. *Crop Sci 44*, 1560–1571.

Poland, J. A. and T. W. Rife (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome 5*, 92–102.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge: Cambridge University Press.

Qu, X., M. Aldana, and L. P. Kadanoff (2002). Numerical and theoretical studies of noise effects in the Kauffman model. *J Stat Phys 109*, 967–986.

R Core Team (2018). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rasmusson, D. C. and R. L. Phillips (1997). Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. *Crop Sci 37*, 303–310.

Rathie, K. A. and F. W. Nicholas (1980). Artificial selection with differing population structures. *Genet Res 36*, 117–131.

Rebourg, C., M. Chastanet, B. Gouesnard, C. Welcker, P. Dubreuil, and A. Charcosset (2003). Maize introduction into Europe: the history reviewed in the light of molecular data. *Theor Appl Genet 106*, 895–903.

Reif, J. C., F.-M. Gumpert, S. Fischer, and A. E. Melchinger (2007). Impact of interpopulation divergence on additive and dominance variance in hybrid populations. *Genetics 176*, 1931–1934.

Reif, J. C., A. E. Melchinger, and M. Frisch (2005). Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. *Crop Sci 45*, 1–7.

Saha, R., P. F. Suthers, and C. D. Maranas (2011). Zea mays iRS1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS ONE 6*, 10.1371/journal.pone.0021784.

Seye, A. I., C. Bauland, A. Charcosset, and L. Moreau (2020). Revisiting hybrid breeding designs using genomic predictions: simulations highlight the superiority of incomplete factorials between segregating families over topcross designs. *Theor Appl Genet 133*, 1995–2010.

Shull, G. H. (1908). The composition of a field of maize. *J Hered os-4*, 296–301.

Silva Dias, J. C. (2010). Impact of improved vegetable cultivars in overcoming food insecurity. *Euphytica 176*, 125–136.

Sprague, G. F. and L. A. Tatum (1942). General vs. specific combining ability in single crosses of corn. *Agron J 34*, 923–932.

Technow, F. (2013). *hypred: Simulation of genomic data in applied genetics*. R package. version 0.4.

Technow, F. (2019). Use of f2 bulks in training sets for genomic prediction of combining ability and hybrid performance. *G3 9*, 1557–1569.

Technow, F. and J. P. Gerke (2017). Parent-progeny imputation from pooled samples for cost-efficient genotyping in plant breeding. *PLOS ONE 12*, e0190271.

Technow, F., C. D. Messina, L. R. Totir, and M. Cooper (2015). Integrating crop growth models with whole genome prediction through approximate bayesian computation. *PLOS ONE 10*, e0130855.

Technow, F., T. A. Schrag, W. Schipprack, E. Bauer, H. Simianer, and A. E. Melchinger (2014). Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics 197*, 1343–1355.

Technow, F., T. A. Schrag, W. Schipprack, and A. E. Melchinger (2014). Identification of key ancestors of modern germplasm in a breeding program of maize. *Theor Appl Genet 127*, 2545–2553.

Tollenaar, M. and E. A. Lee (2002). Yield potential, yield stability and stress tolerance in maize. *Field Crop Res 75*, 161–169.

Vacher, M. and I. Small (2019). Simulation of heterosis in a genome-scale metabolic network provides mechanistic explanations for increased biomass production rates in hybrid plants. *Syst Biol Appl 5*(1), 10.1038/s41540–019–0101–8.

van Heerwaarden, B., Y. Willi, T. N. Kristensen, and A. A. Hoffmann (2008). Population bottlenecks increase additive genetic variance but do not break a selection limit in rain forest drosophila. *Genetics 179*, 2135–2146.

Voss-Fels, K. P., M. Cooper, and B. J. Hayes (2019). Accelerating crop genetic gains with genomic selection. *Theor Appl Genet 132*, 669–686.

Wade, M. J. (2002). A gene's eye view of epistasis, selection and speciation. *J Evol Biol 15*, 337–346.

Wade, M. J. and C. J. Goodnight (1998). The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. *Evolution 52*, 1537–1553.

Wallace, J. G., E. Rodgers-Melnick, and E. S. Buckler (2018). On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. *Annu Rev Genet 52*, 421–444.

White, M. R., M. A. Mikel, N. d. Leon, and S. M. Kaeppler (2020). Diversity and heterotic patterns in North American proprietary dent maize germplasm. *Crop Sci 60*, 100–114.

Wilkins, O., C. Hafemeister, A. Plessis, M.-M. Holloway-Phillips, G. M. Pham, A. B. Nicotra, G. B. Gregorio, S. K. Jagadish, E. M. Septiningsih, R. Bonneau, and M. Purugganan (2016). EGRINs (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. *Plant Cell 28*, 2365–2384.

Wolc, A., A. Kranis, J. Arango, P. Settar, J. E. Fulton, N. P. O'Sullivan, A. Avendano, K. A. Watson, J. M. Hickey, G. de los Campos, R. L. Fernando, D. J. Garrick, and J. C. M. Dekkers (2016). Implementation of genomic selection in the poultry industry. *Anim Fron 6*, 23–31.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics 16*, 97–159.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the Sixth International Congress of Genetics*, Ithaca, NY, pp. 356–366. Brooklyn Botanic Garden.

Wright, S. (1977). *Evolution and the genetics of populations*. Chicago: University of Chicago Press.

Yu, X., X. Li, T. Guo, C. Zhu, Y. Wu, S. E. Mitchell, K. L. Roozeboom, D. Wang, M. L. Wang, G. A. Pederson, T. T. Tesso, P. S. Schnable, R. Bernardo, and J. Yu (2016). Genomic prediction contributing to a promising global strategy to turbocharge gene banks. *Nat Plant 2*, 10.1038/NPLANTS.2016.150.