SNP Allocation For Estimating Heritability (SAFE-h^2): A tool to explore genomic origins of phenotypes for estimation of SNP heritability using additive-only allelic effects or additive and non-additive allelic effects

Behrooz Darbani*, Mogens Nicolaisen

Department of Agroecology, Faculty of Technical Sciences, Aarhus University, DK-4200 Slagelse, Denmark

Corresponding authors emails: behroozdarbani@gmail.com / bds@rsyd.dk ; mn@agro.au.dk

*Current affiliation: ***Research Unit for Molecular Diagnostic and Clinical Research, Sonderjylland Hospital, DK-6200 Aabenraa, Denmark***

Keywords: GWAS, SNP, heritability, additive effects, non-additive effects, false-positive hits

Abstract

Polygenicity necessitates the inclusion of all SNPs for an accurate estimation of SNP heritability. However, SNPs with large association *p*-values contribute negatively, leading to an underestimation of heritability. This effect is demonstrated using simulated data and further validated with real-world datasets. Biologically, contributions to heritability can only be positive. This implies the exclusion of SNPs with large association *p*-values, as their negative contributions hinder rigorous approximations of heritability. SAFE-h^2 overcomes this dilemma by SNP heritability profiling and establishing an association *p*-value threshold to filter out SNPs from its upper-bound, as they contribute to heritability estimations negatively. By examining 74 phenotypes in nine species, SAFE-h^2 revealed 30 units negative contribution on average imposed by the SNPs from the upper-bounds of *p*-value thresholds. SAFE-h^2 also provides a safe *p*-value interval to minimize false-positive SNP hits and captures intra-locus non-additive effects. The allelic adjustment algorithm of SAFE-h^2 captured additional phenotypic variance on simulated datasets. Likewise, SAFE-h^2 revealed considerable uncaptured phenotypic variance when relying on additive effects and improved the SNP heritability estimations up to 56 units by joint analysis of additive and non-additive allelic effects for 50 phenotypes in seven species. We found EMMAX and GEMMA as the most accurate algorithms for heritability profiling using SAFE-h^2.

Introduction

Single nucleotide polymorphism (SNP) heritability estimations depend on study design and reflect the experimental power for maximum exploitation of phenotypic variance in genomic studies and breeding programs[1,2]. SNP heritability offers valuable insights into the genetic architecture of traits and has significant applications in plant and animal breeding, as well as medicine. Compared to broad sense heritability estimates on individual levels, the SNP heritability may be subject to incomplete representation of genetic effects. However, SNP heritability comes along with useful information on genetic elements that regulate the trait-of-interest. Estimating heritability requires both genetic and phenotypic variance, with heritability values ranging from 0 to 1, often expressed as a percentage (0–100%).

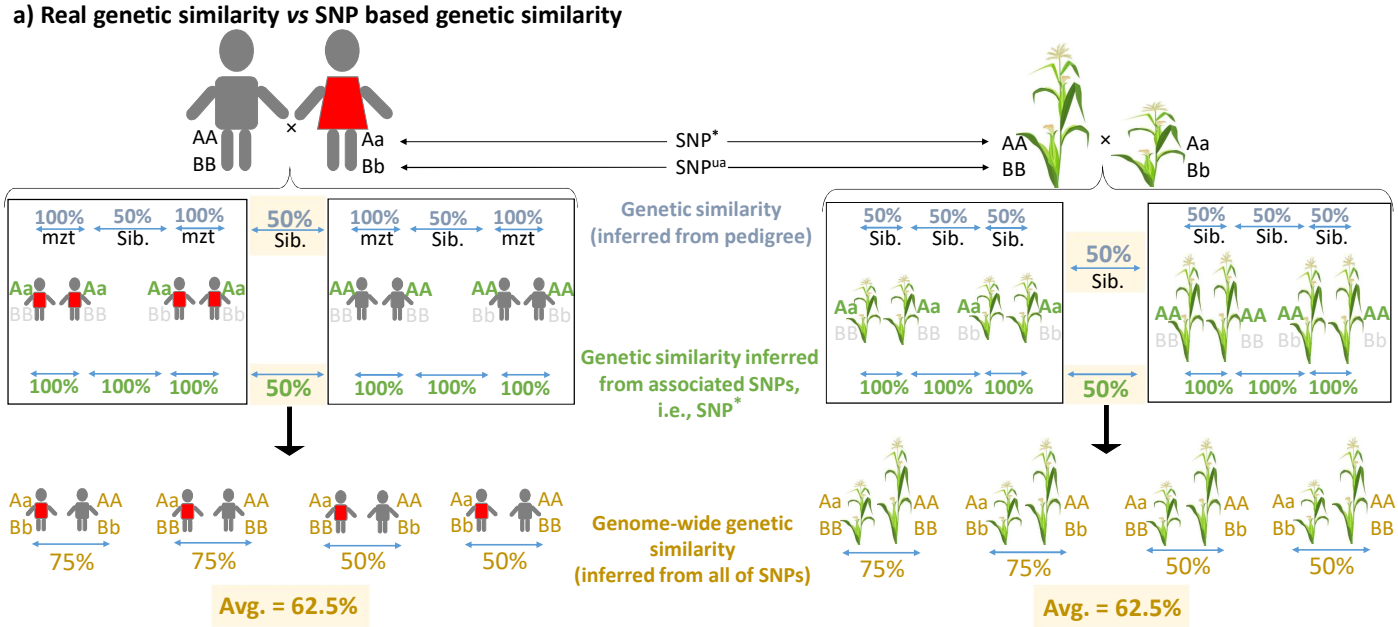
Phenotypic diversity arises through genomic divergence over evolutionary time. These divergent events may lead to intrinsic effects, such as the loss of protein function independent of environmental factors, or to non-constitutive effects in response to specific environmental conditions. As a result, phenotypic variance can exceed genetic variance in multi-environment experiments due to environmental influences, i.e., phenotypic variations for every genotype among environments. In contrast, genetic variance is expected to account for the majority of phenotypic variance, approaching 100% heritability, in single-environment experiments, with any deviations typically attributed to sampling errors. However, SNP heritability estimates in single-environment experiments are typically much lower than 100%, primarily due to the experimental challenges discussed later in this section. Importantly, incorporating additional environments into a multi-environment study can alter the total phenotypic variance, the environmental variance, and consequently, the genetic variance. Likewise, the inclusion of new genotypes, those not previously considered, alters both the total phenotypic and genetic variances in single-environment experiments. Consistent with this, heritability shifts for a given phenotype have been observed across different experimental conditions[3]. An interesting example of heritability shifts is observed with age for various human phenotypes[4]. Therefore, there is no single, definitive heritability value for a given phenotype. Instead, heritability estimates are experiment-specific indicators to forecast the maximum genetic gains under given experimental conditions. It is also important to note that outperformed models for heritability estimation capture the greatest proportion of phenotypic variation when applied under the same statistical power, *i.e.*, using identical phenotypic and genotypic datasets.

In genetic and genomic studies, experimental observations that are under genetic governance can be defined as phenotypes. This is because genetic gains do not occur on observations that have no genetic root and these observations are in fact environmental characteristics and should not be considered as phenotypes. It is also important to recognize that zero or near-zero SNP heritability may result from experimental defects, rather than representing a true absence of genetic influence. Experimental challenges that can affect heritability estimates include inadequate genotypic coverage, partial genomic coverage, imprecise phenotypic scoring scales, assignment of unassociated SNPs to estimation models, exclusion of some associated SNPs from estimation models, and ignoring non-additive allelic effects. For example, approximately $\approx 2^{100,000}$ genotypes (individuals) are required for precise estimation of genetic effects in a genome with 100,000 biallelic SNPs. It is important to note that this calculation is unaffected by genomic linkage, as mutation events are independent. Lethal SNPs, which are conditional on specific suppressor SNPs, are also exceedingly rare. Additionally, imprecise phenotypic scoring scales can be improved, for instance, by defining of possible phenotypic subclasses for qualitative traits. In genomic studies, SNP heritability is typically estimated on additive effects and is thus considered as narrow-sense heritability. Various models have been developed to estimate SNP heritability, for example, using the minor allele frequency (MAF) and the linkage disequilibrium based weighted SNP effects[5–8]. Despite recent technological advancements enabling the detection of millions of variants, SNP heritability still captures 30%–50% less phenotypic variance compared to heritability estimates derived from family and twin studies[9–20]. This underestimation is even more pronounced when compared to the near-complete heritability levels, approaching 100%, typically expected in single-experiment studies that utilize millions of SNPs in large populations (e.g., > 10,000 individuals). To fill this gap, additional variants can be incorporated to cover larger portions of genomes, new genotypes can be added, and more robust phenotype scoring systems can be defined. Controlling population structure and synthetic population re-structuring by crossing for confounded traits have also been introduced to reduce the missing heritability levels[21]. However, these strategies alone may not sufficiently close the gap unless there is a proper allocation of SNPs to estimation models, alongside the consideration of both additive and non-additive effects between reference (*Ref*) and alternative (*Alt*) alleles. Biologically, only associated SNPs (both positive and negative associations) can contribute (positively) to the heritability levels and

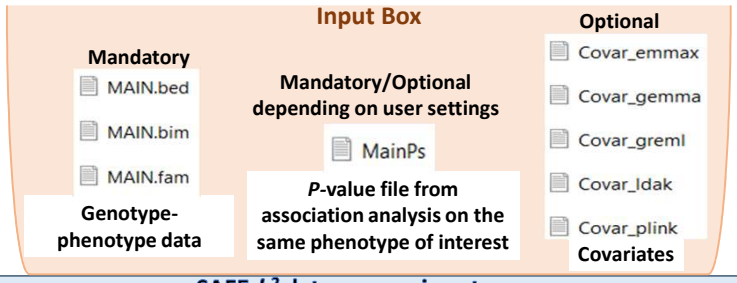
unassociated SNPs have zero contribution. This study however sheds light on the negative contributions of SNPs with large association *p*-values through the commonly used genome-wide (all SNPs-based) relatedness matrices. SNP heritability estimations derived from all SNPs often suffer from genetic variance estimations that are not phenotype-specific. This is because every phenotype-of-interest is governed by a small fraction of the genome, and thus, a small fraction of the SNPs. In contrast, family-based heritability estimations employ the true population statistic of genetic variance, independent of the specific phenotype under study (Fig. 1a). This parameter of genetic variance is constant across all polymorphic genomic positions and can be applied to any phenotype. For example, while monozygotic twins are genetically identical by descent, fraternal twins typically share, on average, 50% of their genomes[22]. In contrast, the SNP-derived genetic similarities and variance estimates are prone to substantial error when applying a genome-wide genetic relatedness matrix to calculate SNP heritability (Fig. 1a). Despite Golan *et al.*'s identification of this common miscalculation over a decade ago[23,24], the majority of studies continue to rely on the old-fashion relatedness matrices obtained genome-wide using all SNPs. As a result, estimates of genetic variance and SNP heritability are heavily depend on the specific set of SNPs chosen, *i.e.*, various pools of SNP combinations can lead to varying levels of genetic similarities (and hence, genetic variance), ultimately resulting in different heritability estimates. In practice, this issue arises because the genetic relationship matrix, which underpins these calculations, is sensitive to the SNPs included, with varying pools of SNPs yielding distinct inter-individual genetic similarity levels. For example, consider a phenotype assessed in 1,000 individuals. If Lab A estimates SNP heritability using 100,000 SNPs, while Lab B employs 200,000 SNPs, the resulting heritability estimates could differ substantially, despite both labs using identical individuals and phenotypic data. It is evident that the inclusion of unassociated SNPs in estimation models leads to overestimation of genetic similarities and underestimation of genetic variance (Fig. 1a). In other words, the rate of random distribution of SNPs among genotypes which have different phenotypes (or different phenotypic levels) becomes more pronounced along the ascending array of association *p*-values. Consequently, SNPs from the upper-bound of a specific *p*-value threshold which is dataset dependent, results in inflated estimates of genetic overlap and, consequently, an underestimation of genetic variance. This genetic variance is not phenotype-specific, as it also reflects genomic regions unrelated to the phenotype-of-interest. In summary, unassociated SNPs (those from the upper-bound of a specific *p*-value threshold) artificially inflate inter-genotype (or inter-individual) genetic similarities by amplifying genetic overlap rates within genetic relationship matrices, ultimately leading to underestimation of SNP heritability. Due to polygenicity, which is characterized by numerous SNPs with small effect sizes, it is essential to include as many variants as possible to capture a larger fraction of SNPs with minimal effects or SNPs having association *p*-values being several orders of magnitude larger than the commonly used significance threshold of 10^{-8}. However, this approach carries the risk of including unassociated SNPs that may have negative contributions to heritability. This technical challenge has led to the widespread practice of using either all SNPs or only those with *p*-value below a specific threshold such as 10^{-8}, depending on which approach yields the highest heritability estimate in every study. One of the remarkable examples for this case is the recent heritability estimates for human height, using two large cohorts[16,17]. Instead of relying on a genome-wide relationship matrix, this dilemma underscores the need for an efficient SNP quota strategy to facilitate the estimation of a phenotype-specific genetic relationship matrix by excluding unassociated SNPs with negative contributions.

This study is accordingly focused on preventing the inclusion of unassociated SNPs (SNPs with negative contributions) in estimation models while ensuring the incorporation of associated SNPs, minimizing false-positive SNP hits' contributions to heritability, and finally on performing a joint analysis that captures both additive and non-additive effects between *Ref* and *Alt* alleles. Here, we introduce SAFE-h^2 (SNP Allocation For Estimating Heritability) that is equipped with a *p*-value based SNP and heritability profiling approach to exclude unassociated genomic regions that have biologically meaningless negative contributions to SNP

a) Real genetic similarity *vs* SNP based genetic similarity



b) SAFE-h^2 pipeline



SAFE-h^2 data processing steps

Step 1: Checking the input for correctness
Step 2: Depending on the user setting: 2-1. Creating genotypic datasets for additive and non-additive scenarios (if selected). This step is a sequence of multiple sub-steps that handle genotype conversion to meet different nonadditive scenarios (see Fig. 4a–f), 2-2. Association *p*-value based profiling and creating related files.
Step 3. Constructing intermediate bfiles specific for SNP clusters
Step 4. Constructing intermediate genetic relatedness matrices specific for SNP clusters
Step 5. Estimating cluster-specific SNP heritability using the algorithms of EMMAX, GCTA, LDAK, and GEMMA (based on the user settings)
Step 6. Identification of the *p*-value threshold to exclude the SNPs with negative contribution to heritability. If pure-effects is selected, a second *p*-value threshold will be identified to report a false-positive free estimations
Step 7. Finalizing the analyses by deleting all the intermediate files and classification of the results as raw data and graphs within specific folders.

Output

- ✓ Heritability profiles & heritability estimations on additive effects
- ✓ Heritability profiles & heritability estimations on additive and non-additive effects
- ✓ False-positive free estimations of heritability
- ✓ SNP information for SNP clusters
- ✓ Related graphs
- ✓ Genotypic files specific for different combinations of additive and non-additive effects that can be used in linear association analyses to efficiently address additive and non-additive effects simultaneously

Fig. 1. Genetic similarities among genotypes (individuals) and pipeline illustration for SAFE-h^2. a) Genetic similarity in family and population based studies. Among genotypes which are differentiated based on a given phenotype-of interest, family-based studies apply real genetic similarities which always match to the genetic similarities that can be inferred from associated SNPs (SNP^*). Genome-wide genetic similarities are overestimated when they are based on all of the SNPs, *i.e.*, associated (SNP^*) and unassociated (SNP^{ua}). The mzt and Sib stand for monozygotic twins and siblings. b) SAFE-h^2 pipeline.

heritability estimates. SAFE-h^2 also defines a safe *p*-value interval to mitigate heritability contributions from false-positive SNP hits. Furthermore, SAFE-h^2 introduces a co-extraction algorithm to capture intra-locus non-additive and additive allelic effects, enhancing the accuracy of SNP heritability estimates. These intra-locus allelic effects involve interactions of *Ref* and *Alt* alleles (*Ref*-allele × *Ref*-allele, *Ref*-allele × *Alt*-allele, and *Alt*-allele × *Alt*-allele). It is important to distinguish these from epistatic genetic effects, which refer to interactions between non-allelic genes or interactions between different SNPs (SNP × SNP).

Results

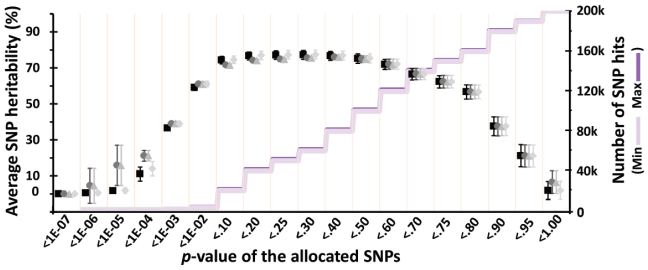
Polymorphic genomic positions such as SNPs should either be neutral or positive when included in models estimating SNP heritability. However, our findings revealed that weakly associated SNPs, characterized by large association *p*-values, contribute to heritability estimates negatively. Here, we propose a method for excluding SNPs with negative contributions from heritability calculations. This approach involves defining an experiment and dataset-specific threshold for association *p*-values. SNPs with *p*-values above this threshold exhibit random distributions with respect to phenotypic variations across genotypes (individuals), leading to biologically meaningless negative contributions to heritability estimates when combined with SNPs having smaller *p*-values than the threshold. We will henceforth call the unassociated SNPs as those with *p*-values from the upper-bound of data-specific *p*-value thresholds. We introduce SAFE-h^2 application tool (Fig. 1b) which performs SNP profiling and calculates SNP heritability for *p*-value based SNP-clusters. SAFE-h^2 can leverage heritability estimation models used in EMMAX[25], LDAK-GCTA, LDAK-Thin[5], GCTA-GREML[26], and GEMMA[27]. SAFE-h^2 determines the *p*-value threshold used for omitting SNPs with negative contributions to heritability (Fig. 1b), but it also allows users to define lower *p*-value thresholds. Additionally, SAFE-h^2 enables the identification of a second *p*-value threshold at which contributions of false-positive SNP hits to heritability can be minimized efficiently (Fig. 1b). False-positive SNP hits, which are not genuinely associated with phenotypes, still inflate heritability estimates due to low statistical power levels. Finally, SAFE-h^2 enhances the detection of non-additive effects by incorporating a joint analysis approach for intra-locus allelic effects of additive and non-additive. This is achieved through an allelic adjustment strategy based on the observed genetic effects for every SNP, further improving the precision of heritability estimates.

Within an array of association p-values, SNPs with large p-values neutralize the contribution of SNPs with small p-values to heritability estimates

Simulated genomic and phenotypic datasets were used to demonstrate the negative contributions by SNPs with large *p*-values from the upper-bound of an ascending array of association *p*-values. It is worth to note that the critical factor in assessing negative contributions to heritability is the levels of association *p*-values of SNPs in relation to each other and not the level of association *p*-value for each SNP alone. In reality, within any given range of association *p*-values, there will inevitably be SNPs with smaller and larger *p*-values relative to one another. Accordingly, the purpose of simulating random phenotypes and genotypes was to generate arrays of association *p*-values, rather than to represent true associations. First, two analyses were performed on 1,000 and 10,000 simulated genotypes each with 200k SNPs (MAF = 1.6%–10%). The simulated phenotypes consisted of random integers within certain ranges. Therefore, there were random chances of

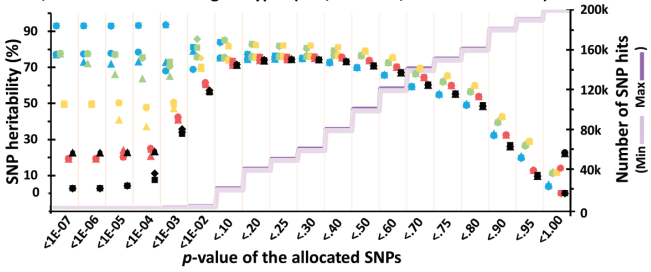
a) Analysis of 10 artificial random phenotypes on a panel of 1,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Random phenotypes	Range (N)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-100 (1,000)	■	●	▲	◆
2nd	40-100 (1,000)	■	●	▲	◆
3rd	40-100 (1,000)	■	●	▲	◆
4th	40-100 (1,000)	■	●	▲	◆
5th	40-100 (1,000)	■	●	▲	◆
6th	40-100 (1,000)	■	●	▲	◆
7th	40-100 (1,000)	■	●	▲	◆
8th	40-100 (1,000)	■	●	▲	◆
9th	40-100 (1,000)	■	●	▲	◆
10th	40-100 (1,000)	■	●	▲	◆



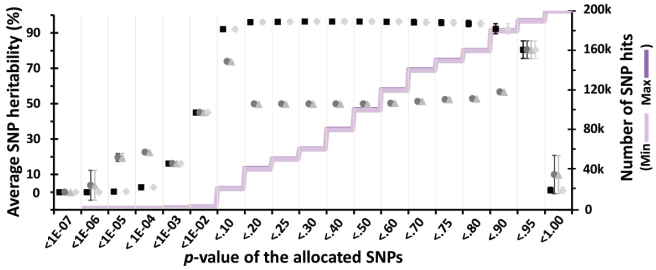
b) Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Phenotypes	The introduced non-random SNP position: Ref/Ref Genotype range (N)	Alt/Alt Genotype range (N)	σ_e^2 (%)	σ_g^2 (%)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-50 (900)	80-100 (100)	6.995	93.005	■	●	▲	◆
2nd	40-60 (900)	75-100 (100)	22.389	77.611	■	●	▲	◆
3rd	40-70 (900)	70-100 (100)	50.267	49.733	■	●	▲	◆
4th	40-80 (900)	60-100 (100)	80.690	19.310	■	●	▲	◆
5th	40-90 (900)	50-100 (100)	97.170	2.830	■	●	▲	◆



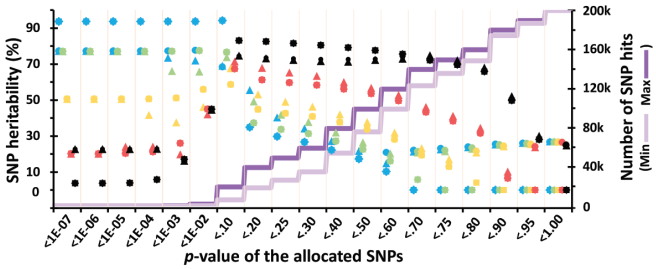
c) Analysis of 10 artificial random phenotypes on a panel of 10,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Random phenotypes	Range (N)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-100 (10,000)	■	●	▲	◆
2nd	40-100 (10,000)	■	●	▲	◆
3rd	40-100 (10,000)	■	●	▲	◆
4th	40-100 (10,000)	■	●	▲	◆
5th	40-100 (10,000)	■	●	▲	◆
6th	40-100 (10,000)	■	●	▲	◆
7th	40-100 (10,000)	■	●	▲	◆
8th	40-100 (10,000)	■	●	▲	◆
9th	40-100 (10,000)	■	●	▲	◆
10th	40-100 (10,000)	■	●	▲	◆



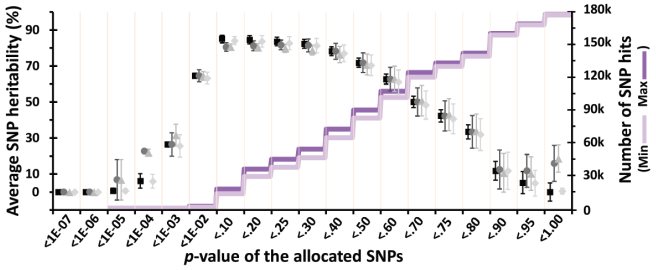
d) Analysis of 5 artificial non-random phenotypes on a panel of 10,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Phenotypes	The introduced non-random SNP position: Ref/Ref Genotype range (N)	Alt/Alt Genotype range (N)	σ_e^2 (%)	σ_g^2 (%)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-50 (9,000)	80-100 (1,000)	6.605	93.395	■	●	▲	◆
2nd	40-60 (9,000)	75-100 (1,000)	22.738	77.262	■	●	▲	◆
3rd	40-70 (9,000)	70-100 (1,000)	48.937	51.063	■	●	▲	◆
4th	40-80 (9,000)	60-100 (1,000)	79.675	20.325	■	●	▲	◆
5th	40-90 (9,000)	50-100 (1,000)	96.309	3.691	■	●	▲	◆



e) Analysis of 10 artificial random phenotypes on a panel of 1,092 HapMap phase 1 genotypes (177,675 SNPs + 1 simulated SNP; MAF ≥1%)

Random phenotypes	Range (N)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-150 (1,092)	■	●	▲	◆
2nd	40-150 (1,092)	■	●	▲	◆
3rd	40-150 (1,092)	■	●	▲	◆
4th	40-150 (1,092)	■	●	▲	◆
5th	40-150 (1,092)	■	●	▲	◆
6th	40-150 (1,092)	■	●	▲	◆
7th	40-150 (1,092)	■	●	▲	◆
8th	40-150 (1,092)	■	●	▲	◆
9th	40-150 (1,092)	■	●	▲	◆
10th	40-150 (1,092)	■	●	▲	◆



f) Analysis of 5 artificial non-random phenotypes on a panel of 1,092 HapMap phase 1 genotypes (177,675 SNPs + 1 simulated SNP; MAF ≥1%)

Phenotypes	The introduced non-random SNP position: Ref/Ref Genotype range (N)	Ref/Alt Genotype range (N)	Alt/Alt Genotype range (N)	σ_e^2 (%)	σ_g^2 (%)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-50 (982)	80-100 (103)	120-150 (7)	5.704	94.296	■	●	▲	◆
2nd	40-60 (982)	75-100 (103)	112-150 (7)	21.037	78.963	■	●	▲	◆
3rd	40-70 (982)	70-100 (103)	105-150 (7)	41.501	58.499	■	●	▲	◆
4th	40-80 (982)	60-100 (103)	90-150 (7)	72.099	27.901	■	●	▲	◆
5th	40-90 (982)	50-100 (103)	75-150 (7)	91.416	8.584	■	●	▲	◆

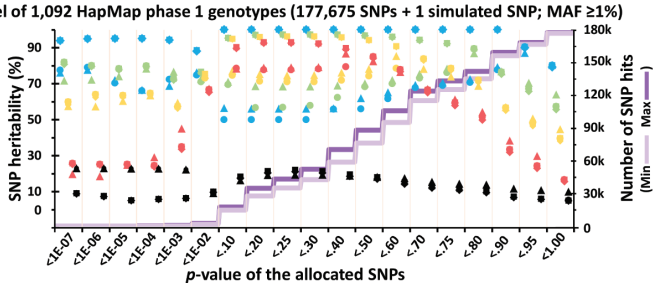


Fig. 2. The SNP heritability profiles for artificial random and non-random quantitative phenotypes using simulated SNPs and variants of chr. 22 from the human HapMap phase 1 genotypes. a–f) Initial GWAS was performed using EMMAX to realize association levels for the SNPs. The four different approaches of EMMAX, LDAK-GCTA, LDAK-Thin, and GCTA-GREML were then applied to estimate SNP heritability for the clusters of SNP hits captured using the association *p*-values. N represents the sample sizes, *i.e.*, number of genotypes. MAF stands for the minor allele frequency. a,c,e) The average SNP heritability estimates were obtained for 10 independent random phenotypes within identical ranges. The error bars represent 99% confidence intervals. b,d,f) The SNP heritability estimates were obtained for five artificial non-random phenotypes which were associated with a simulated non-random SNP position introduced within both the simulated and HapMap genomes. By ensuring additive-only effects, the non-random phenotypes were defined at different quantitative ranges among the genotypic groups of the introduced non-random SNP position, *i.e.*, *Ref*/*Ref* and *Alt*/*Alt* for the simulated inbred genotypes and *Ref*/*Ref*, *Ref*/*Alt*, and *Alt*/*Alt* for HapMap data. The intra-genotype assignment of the phenotypes, while being within their specific ranges, was still random. The inter-genotype variances of the phenotypes are shown as σ_g^2. σ_e^2 represents the intra-genotype variations. a–d) The analyses of inbred genotypes using GCTA-GREML were performed based on the "inbred" algorithm of GCTA.

varying association levels between phenotypes and SNPs. Analysis of these datasets revealed a consistent pattern of negative contribution to SNP heritability imposed by the SNPs having large association *p*-values (Fig. 2a,c). Increasing statistical power, as anticipated, resulted in greater phenotypic contributions from SNPs with smaller *p*-values (Fig. 2a,c). This also shifted the *p*-value threshold to higher values, at which SNPs with larger *p*-values exhibited negative contributions (Fig. 2a,c). However, the negative contribution by SNPs with large association *p*-values persisted, and including all SNPs in the estimation models led to a sharp underestimation of heritability (Fig. 2c). A similar pattern was observed by including more SNPs (two millions SNPs, Supplementary Fig. 1a). In contrast to the increased number of genotypes (higher genotypic coverage), mimicking lower genomic coverage by analyzing limited number of SNP, as expected, reduced the likelihood of observing random associations (Supplementary Fig. 1b,c). Further validation obtained through a similar analysis conducted on the human HapMap phase 1 genomic data[28], which includes 1,092 genotypes (individuals) and 177,675 SNPs on chromosome 22 (MAF ≥1%). Artificial phenotypes were generated as random integers for the HapMap genotypes, and heritability profiling revealed similar patterns of negative contributions (Fig. 2e). In all these genomic datasets that were analyzed *via* the four different heritability estimation models, significant declines in the SNP heritability estimates were observed after allocating SNPs with large association *p*-values (Fig. 2a,c,e, Supplementary Fig. 1). The key finding across all these analyses was the consistent pattern in which SNPs with large association *p*-values exhibited negative genetic contributions. As previously mentioned, this is biologically implausible for contributions to heritability to be negative, and any genetic contribution, if present, are expected to be positive.

Simulating non-random phenotypes also revealed similar patterns within the same genomic datasets, *i.e.*, the simulated and HapMap chromosome 22 SNP datasets, each with an introduced associated SNP position. The predesignated SNP position was introduced to be associated, exclusively through additive effects, with the simulated non-random phenotypes. The phenotypic variance across the genotype groups for the associated SNP position was set from 2.8% to 94.3% (Fig. 2b,d,f). As a result, SNP heritability estimations were predictable at or above these predefined levels, given that other SNPs could also, by chance, account for additional phenotypic variance. However, as anticipated, the most weakly associated SNPs exhibited negative contributions, leading to a substantial underestimation of SNP heritability across all scenarios and heritability estimation models, including EMMAX[25], LDAK-GCTA, LDAK-Thin[5], and GCTA-GREML[26] (Fig. 2b,d,f). When all SNPs were included in the models, SNP heritability estimates were consistently lower than the predefined levels (Fig. 2b,d,f). As discussed previously, SNP heritability estimation involves a trade-off between associated SNPs (with small *p*-values) and unassociated SNPs (with larger *p*-values), with latter resulting in heritability underestimations through negative contributions. So, the negative genetic effects of SNPs with large association *p*-values are anticipated to be offset when multiple highly associated SNPs are

included in the analysis. It is exceedingly rare, and virtually impossible, to identify multiple SNPs in linkage equilibrium and with high contributions to phenotypic variance (*i.e.*, highly associated with *p* values < 10^{-20}) when analyzing a limited number of SNPs, such as 200k across 1000 genotypes (individuals) in real-world experiments. To simulate such rare scenarios, we designed and performed identical analyses as in Figure 2b but this time in the presence of 12, 24, and 48 strongly associated SNPs that were incorporated into the genomic datasets (Supplementary Fig. 2). In addition to these predesignated associations (1, 12, 24, and 48), randomly associated SNPs also contributed to the heritability estimates, though with smaller effects. Specifically, there were 400 to 39000 associated SNPs among the 200k SNPs that displayed positive contributions to the heritability estimates. Consistent with the previous analyses, negative contribution patterns were observed for SNPs with large association *p*-values (Supplementary Fig. 2). It is important to note that we observed the same pattern of negative contributions to heritability in simulations involving SNPs with no distance between adjacent SNPs, *i.e.*, 200k SNPs on a 200kb chromosome (Supplementary Fig. 3). In these analyses, the simulated genotypic datasets had minor allele frequencies from 1.6%–10%. Although, the real genotypic dataset of chromosome 22 had minor allele frequencies of up to 50%, we additionally performed similar analysis on simulated genotypes with minor allele frequencies ranging from 1.6%–50% and found the same pattern of negative contributions (Supplementary Fig. 4). Taken together, the analyses using both simulated random and non-random phenotypes highlight the importance of excluding SNPs with larger association *p*-values, as they impose negative contributions to heritability. We also noticed instability in SNP heritability estimations using LDAK-GCTA and LDAK-Thin, as these models failed to accurately estimate the predefined heritability levels of 93.4%, 93%, 3.7%, and 2.8% when the estimations were solely based on the predesignated associated SNPs, with all other SNPs excluded (Fig. 2b,d, Supplementary Fig. 2). In contrast, EMMAX and GCTA provided more stable estimates and performed better, yielding the closest approximations.

SAFE-h^2 effectively prevents the illogical negative contributions of SNPs to heritability estimations

To address polygenicity and prevent negative contributions of unassociated SNPs to heritability estimates, we introduce SAFE-h^2, a method that employs association *p*-values and estimates SNP heritability through *p*-value-guided clustering of SNPs. The heritability profiling by SAFE-h^2 facilitates the most accurate approximation of SNP heritability levels by determining an optimal *p*-value threshold. This threshold is used to exclude unassociated SNPs, which, if included, lead to an overestimation of genetic similarities within the genetic relationship matrix and thereby, introduce negative contributions to heritability. In essence, this *p*-value threshold is the best border that we can assume between associated and unassociated SNPs for accurate estimation of SNP heritability. Further validation was accomplished on public data from *Arabidopsis* and rice genotypes. SNP heritability profiling was carried out for 14 different phenotypes from 328 to 1058 *Arabidopsis* genotypes[29]. The heritability profiling was performed on 2,621k to 2,840k *Arabidopsis* SNPs[30,31], specific to the phenotypes. We also estimated SNP heritability for 10 phenotypes observed in 1,422 to 2,102 rice genotypes[32], using 6,131k to 6,414k SNPs[32–34], which varied across the phenotypes. A minimum MAF of 0.01 was applied in all analyses. Consistent with our findings in in the *Arabidopsis* and rice datasets, the genomic contributions of strongly associated SNPs were effectively neutralized by weakly associated SNPs, *i.e.*, SNPs with large *p*-values (Table 1). Notably, weakly associated SNPs with negative contributions masked the heritability estimates for all the examined *Arabidopsis* and rice phenotypes with an average reduction of 32.3 units, ranging from 5.6 to 64.7 units across individual phenotypes and the four heritability estimation models (Table 1). The SNP heritability was also underestimated when it was solely based on SNPs with association *p*-values of <10^{-8} or even <10^{-5}. By calculating SNP heritability at the identified *p*-value thresholds, SAFE-h^2 improved heritability estimates by an average of 28.7 units, ranging from 13.6 to 63.7 units for individual phenotypes, when compared to the heritability estimates based on SNPs with association *p*-values less than 10^{-5} (Table 1). We used SAFE-h^2 to analyze 50 additional phenotypes from three plant and four

Table 1. SAFE-h^2 improves the heritability estimations by preventing the negative contributions of weakly-associated genomic regions.

		Number of Genotypes/Variants	SAFE-h^2_{SNP} ([SAFE-h^2_{SNP}] - [h^2_{SNP} based on all the SNPs]; [SAFE-h^2_{SNP}] - [h^2_{SNP} based on the SNP hits with p-value < 10^{-5}])				
			EMMAX	LDAK GCTA-model	LDAK Thin-model	GCTA-GREML[a]	Average
Arabidopsis thaliana phenotypes	CLN	904 / 2,807,143	**83.4%** (44.5%; 30.8%)	**75.4%** (36.4%; 22.8%)	**73.5%** (32.2%; 24.5%)	**83.4%** (44.5%; 30.8%)	**78.9%** (39.4%; 27.2%)
	FBE	936 / 2,824,927	**93.1%** (42.1%; 19.4%)	**77.0%** (26.0%; 3.3%)	**80.1%** (17.0%; 12.1%)	**93.1%** (42.1%; 19.4%)	**85.9%** (31.8%; 13.6%)
	DTF	1,058 / 2,830,731	**62.9%** (43.4%; 46.2%)	**62.8%** (43.4%; 46.1%)	**61.5%** (40.5%; 38.1%)	**62.9%** (43.4%; 46.2%)	**62.5%** (42.7%; 44.1%)
	FN	409 / 2,628,158	**62.2%** (62.2%; 53.0%)	**62.1%** (39.8%; 39.0%)	**65.9%** (44.2%; 42.8%)	**62.1%** (62.1%; 52.9%)	**63.1%** (52.1%; 46.9%)
	GR	420 / 2,631,629	**73.4%** (66.4%; 40.4%)	**73.4%** (66.4%; 40.4%)	**66.0%** (59.6%; 32.4%)	**73.4%** (66.4%; 40.4%)	**71.5%** (64.7%; 38.4%)
	RLN	850 / 2,820,131	**89.4%** (23.2%; 13.1%)	**80.9%** (14.7%; 4.6%)	**81.1%** (11.7%; 24.5%)	**89.4%** (23.2%; 13.1%)	**85.2%** (18.2%; 13.8%)
	RL	328 / 2,621,362	**71.5%** (71.5%; 49.6%)	**71.5%** (49.6%; 49.7%)	**72.6%** (59.8%; 44.6%)	**71.4%** (71.4%; 46.9%)	**71.8%** (63.1%; 48.4%)
	SL	680 / 2,789,485	**54.6%** (54.6%; 30.9%)	**54.6%** (54.6%; 31.0%)	**56.2%** (53.7%; 40.4%)	**54.6%** (54.6%; 30.9%)	**55.0%** (54.4%; 33.3%)
	As	344 / 2,839,190	**100%** (30.8%; 41.3%)	**78.8%** (9.8%; 20.2%)	**93.5%** (30.1%; 34.6%)	**99.5%** (30.3%; 40.9%)	**93.0%** (25.3%; 34.3%)
	Cd	345 / 2,840,072	**99.3%** (44.0%; 24.9%)	**79.4%** (24.1%; 5.0%)	**83.2%** (18.3%; 10.8%)	**99.3%** (44.0%; 24.9%)	**90.3%** (32.7%; 16.4%)
	Li	345 / 2,840,072	**97.6%** (56.2%; 69.8%)	**87.6%** (46.3%; 59.8%)	**84.8%** (49.9%; 55.2%)	**97.6%** (56.2%; 69.8%)	**91.9%** (52.2%; 63.7%)
	Mn	345 / 2,840,072	**100%** (45.5%; 65.3%)	**79.4%** (24.9%; 44.8%)	**84.5%** (31.4%; 50.6%)	**100%** (45.5%; 65.4%)	**91.0%** (36.8%; 56.5%)
	Zn	345 / 2,840,072	**100%** (67.4%; 35.2%)	**88.0%** (55.4%; 23.2%)	**81.5%** (56.4%; 35.4%)	**100%** (67.4%; 35.2%)	**92.4%** (61.7%; 32.3%)
	P	345 / 2,840,072	**96.8%** (44.1%; 46.5%)	**85.8%** (33.4%; 35.6%)	**80.0%** (32.3%; 44.0%)	**96.8%** (44.2%; 46.6%)	**89.8%** (38.5%; 43.2%)
Rice (*O. sativa*) phenotypes	DTF	2,102 / 6,414,800	**99.9%** (10.0%; 16.1%)	**90.0%** (0.0%; 18.1%)	**92.2%** (1.5%; 19.2%)	**99.9%** (11.7%; 17.8%)	**95.6%** (5.6%; 17.8%)
	GL	2,099 / 6,411,145	**99.9%** (14.0%; 16.0%)	**88.3%** (2.4%; 16.3%)	**91.3%** (3.1%; 17.3%)	**99.9%** (15.2%; 17.2%)	**94.9%** (8.6%;16.7%)
	GWE	2,098 / 6,410,047	**99.9%** (13.5%; 17.4%)	**88.8%** (2.4%; 16.6%)	**91.4%** (3.3%; 17.3)	**99.9%** (14.6%; 19.1%)	**95.0%** (8.4%; 17.6%)
	PL	1,762 / 6,269,207	**99.9%** (17.3%; 14.7%)	**86.9%** (4.2%; 14.4%)	**88.9%** (3.3%; 14.3%)	**99.9%** (17.8%; 14.7%)	**93.9%** (10.7%; 14.5%)
	GWI	2,099 / 6,411,145	**99.9%** (14.8%; 15.7%)	**87.6%** (2.4%; 15.4%)	**90.6%** (3.1%; 16.5%)	**99.9%** (15.9%; 16.8%)	**94.6%** (9.1%; 16.1%)
	SD (EC18)[b]	1,998 / 6393347	**99.9%** (54.3%; 4.6%)	**85.6%** (39.9%; 16.6%)	**88.3%** (29.7%; 20.1%)	**99.9%** (53.1%; 17.3%)	**93.5%** (44.2%; 14.6%)
	SD (EC12)[b]	1,998 / 6,393,347	**99.9%** (23.5%; 19.4%)	**88.4%** (11.9%; 15.11%)	**90.5%** (9.8%; 13.5%)	**99.9%** (23.8%; 19.9%)	**94.7%** (17.3%; 17.0%)
	SH	1,743 / 6,254,739	**99.9%** (16.4%; 13.7%)	**87.6%** (4.0%; 14.5%)	**89.5%** (3.1%; 15.0%)	**99.9%** (17.0%; 14.7%)	**94.3%** (10.1%; 14.4%)
	SPB	1,422 / 6,131,325	**99.9%** (35.3%; 2.2%)	**94.6%** (29.9%; 27.2%)	**98.3%** (26.9%; 30.0%)	**99.9%** (33.9%; 39.6%)	**98.2%** (31.5%;24.8%)
	SF	2,100 / 6,411,911	**99.9%** (21.2%; 27.7%)	**90.8%** (12.0%; 18.5%)	**92.5%** (10.4%; 16.2%)	**99.9%** (21.8%; 28.2%)	**95.8%** (16.3%; 22.6%)
Average			..... (38.2%; 29.8%)	..... (26.4%; 24.9%)	..... (26.3%; 27.9%)	..... (38.3%; 32.0%)	..... **(32.3%; 28.7%)**
Minimum			..... (10.0%; 2.2%)	..... (0.0%; 3.3%)	..... (1.5%; 10.8%)	..... (11.7%; 13.1%)	..... **(5.6%; 13.6%)**
Maximum			..... (71.5%; 69.8%)	..... (66.4%; 59.8%)	..... (59.8%; 55.2%)	..... (71.4%; 69.8%)	..... **(64.7%; 63.7%)**

CLN: cauline leaf number; FBE: days until the emergence of flowering buds; DTF: days until first open flower; FN: number of fruits per plant; GR: average growth rate; RLN: rossette leaf number; RL: root length; SL: stem length; As: arsenic content in leaf; Cd: cadmium content in leaf; Li: lithium content in leaf; Mn: manganese content in leaf; Zn: zinc content in leaf; P: phosphorus content in leaf; GL: grain length; GWE: grain weigth; PL: panicle length; GWI: grain width; SD (EC18): salt injury at EC of 18; SD (EC12): salt injury at EC of 12; SH: seedling height; SPB: secondary panicle branching; SF: spikelet fertility. [a]Analyzing inbred genotypes of rice and Arabidopsis was performed under the "inbred" algorithm of GCTA. [b]The electrical conductivity.

animal species (Supplementary Table 1). In agreement with our previous results, these analyses also revealed that SNPs with large association *p*-values imposed up to 30 units of negative contributions to the heritability estimates, on average (Supplementary Table 1).

The performance of our method was also examined in comparison with BayesR[35,36] , a Bayesian model that simultaneously facilitates variant discovery and heritability estimation. We first examined the same simulated phenotypes with predefined genetic variances (see Fig. 2b,d,f) among the genotypes of an associated SNP position. This SNP was introduced into both the simulated and HapMap (Chr22) genomes. In contrast to SAFE-h^2, which accurately extracted the assigned genetic variances (Fig. 2b,d,f), BayesR either failed to detect any genetic variance or underestimated the genetic variances (Supplementary Tables 2 and 3). These analyses were conducted under both linkage equilibrium and linkage disequilibrium scenarios. The processing time by BayesR was 8–10 times longer than that of SAFE-h^2 (Supplementary Tables 2 and 3). We extended the comparison by analyzing data for eight real phenotypes from four organisms. We observed a similar pattern, with SAFE-h^2 consistently extracting more genetic variance than BayesR (Supplementary Table 4).

SAFE-h2 reports a safe p-value interval to minimize positive contributions of false-positive SNP hits to heritability estimates

False-positive rates are dependent on statistical power of experiments, which can be particularly challenging in genome-wide association studies. This difficulty arises not only from the resource-intensive nature of GWAS but also from the impracticality of achieving near-complete genotypic coverage. In theory, we approximately expect $2^{(\text{number-of-SNPs})}$ possible genotypes. Additionally the availability of unrelated genotypes within populations is often limited. As a result, varying levels of false-positive rates are to be expected, depending on the experimental design. To address this challenge, SAFE-h^2 employs an innovative approach to exclude false-positive SNP hits. SAFE-h^2 defines a safe *p*-value interval, where upper-end acts as an alternative *p*-value threshold for heritability estimation with minimized contributions of false-positive SNP hits. To achieve this, the range of phenotype-of-interest is used to simulate six independent random phenotypes for all the genotypes. Afterwards, SAFE-h^2 creates SNP heritability profiles for all the generated random phenotypes as well as the original phenotype-of-interest using the same genomic dataset. Analyzing of the random phenotypes then reveals a safe *p*-value interval with minimum SNP heritability level at the lower-tail of heritability profile. The SNP heritability for the original phenotype can subsequently be estimated at the upper-end of this interval that acts as a *p*-value threshold to omit the false-positive SNP hits from the estimation model. Figure 3 illustrates the same analysis on four different genomic datasets from tomato, rice, foxtail millet, and *Arabidopsis*. We found the *p*-value threshold of 10^{-4} to estimate false-positive free SNP heritability for betaine and hesperetin contents in tomato (Fig 3a). This threshold was 10^{-6} for flowering-bud emergence and zinc content in *Arabidopsis* (Fig. 3b). The SNP heritability estimations with minimized false-positive SNP contribution were also possible at the *p*-value thresholds of 10^{-6} and 10^{-5} for salt-injury and secondary-panicle branching in rice, respectively (Fig. 3c). Finally, the analysis determined the threshold of 10^{-5} for two independent operational taxonomic units of root-microbiome in foxtail millet (Fig. 3d). In all these examples, SNP heritability estimates based on all SNPs were consistently lower and underestimated compared to those obtained after minimizing the false-positive contribution risk (Fig. 3). It was evident that the degree of negative contribution reflect a trade-off between the number of associated SNPs and the number of unassociated SNPs with negative contributions, resulting in variability across different phenotypes (Table 1, Supplementary Table 1, Fig. 3).

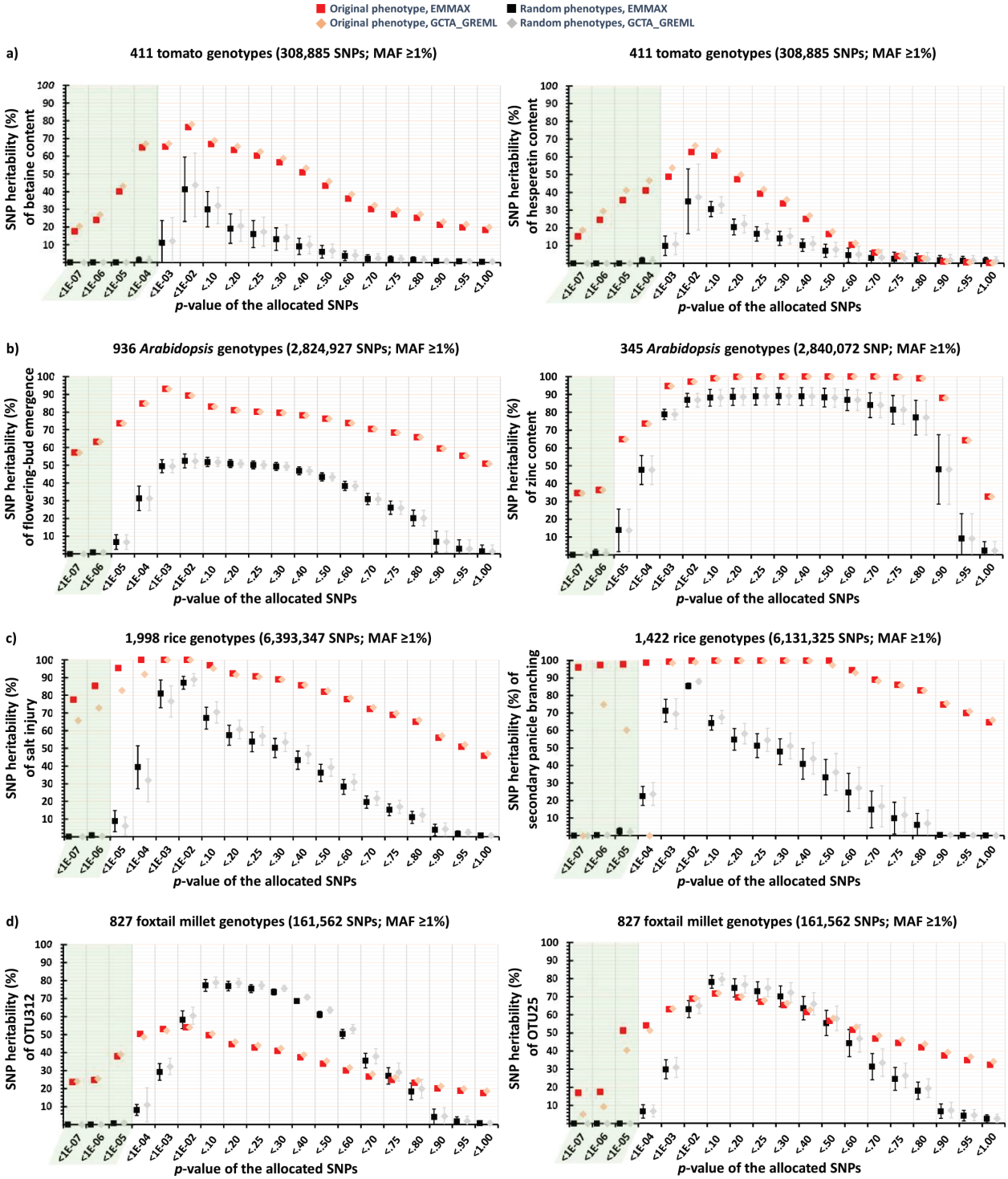

Original phenotype, EMMAX
Original phenotype, GCTA_GREML
Random phenotypes, EMMAX
Random phenotypes, GCTA_GREML
a)
411 tomato genotypes (308,885 SNPs; MAF ≥1%)
SNP heritability (%) of betaine content
411 tomato genotypes (308,885 SNPs; MAF ≥1%)
SNP heritability (%) of hesperetin content
b)
936 Arabidopsis genotypes (2,824,927 SNPs; MAF ≥1%)
SNP heritability (%) of flowering-bud emergence
345 Arabidopsis genotypes (2,840,072 SNP; MAF ≥1%)
SNP heritability (%) of zinc content
c)
1,998 rice genotypes (6,393,347 SNPs; MAF ≥1%)
SNP heritability (%) of salt injury
1,422 rice genotypes (6,131,325 SNPs; MAF ≥1%)
SNP heritability (%) of secondary panicle branching
d)
827 foxtail millet genotypes (161,562 SNPs; MAF ≥1%)
SNP heritability (%) of OTU312
827 foxtail millet genotypes (161,562 SNPs; MAF ≥1%)
SNP heritability (%) of OTU25
p-value of the allocated SNPs

Fig. 3. SAFE-h^2 estimates SNP heritability by minimizing the contributions from false-positive SNP hits. a–d) The safe-intervals for *p*-value are highlighted in light green representing the area with minimum contributions to heritability by false-positive SNP hits. The two different approaches of EMMAX and GCTA-GREML were applied to estimate the SNP heritability for the clusters of SNP hits captured using the association *p*-values. MAF stands for the minor allele frequency. The average estimations are illustrated for six independent random phenotypes (simulated within the range of original phenotype) and the error bars represent standard deviations among the six estimations.

SNP heritability estimation through joint-analysis of intra-locus allelic effects of additive and non-additive

The genetic scoring system for SNP heritability estimations is designed to predict intra-locus additive effects between *Ref* and *Alt* alleles, encompassing SNP-specific allelic interactions of *Ref*-allele × *Ref*-allele, *Ref*-allele × *Alt*-allele, and *Alt*-allele × *Alt*-allele. Linear models can partially capture intra-locus non-additive allelic effects, but some non-additive effects remain uncaptured[37–39]. The significance of non-additive genetic effects has been reviewed[40] and methods like GCTA-GREMLd has also been introduced to address dominance effects in heritability estimations[41]. Simultaneous consideration of both additive and non-additive genetic effects is valuable in various applications, including parental selection in breeding programs, predictive performance for quantitative traits, and polygenic risk scores for disease-related categorical outcomes[40]. However, the relatively small genetic contributions estimated for non-additive effects have led to the widespread assumption that these effects are negligible[40]. In theory, there is no inherent genetic or biological rationale to support the view that intra-locus allelic effects of non-additive should be trivial. Therefore, the estimated small non-additive allelic effects may be attributable to methodological limitations and artifacts, aligning with findings of our study.

SAFE-h^2 employs an allelic adjustment strategy to extract non-additive effects, including dominance, overdominance, and heterosis-like effects, within the framework of linear models. Figure 3a and 3b provide examples of different intra-locus allelic interactions, demonstrating that the exponential model outperforms the linear approach in the presence of non-additive effects. However, when linear models are augmented with allelic adjustments tailored to specific non-additive effects, they can capture the same levels of genetic variance as the exponential model (Fig. 4c–f). SAFE-h^2 utilizes the same algorithm to identify the best fitted linear model among adjusted allele-sets for every SNP position and the phenotype-of-interest. The method includes allelic adjustments, *i.e.*, allele conversions or replacements, specific for either of the dominance, overdominance, or heterosis-like effects (see Fig. 4a–f). For every SNP position, SAFE-h^2 examines different scenarios, *i.e.*, additive (original genotypes), adjusted genotypes for dominance effects, adjusted genotypes for overdominance effects, and adjusted genotypes for heterosis-like effects, to find the smallest association *p*-value to the phenotype of interest. Before heritability profiling, combinatory genotype datasets can then be built based on one selected adjustment or the original genotype, *i.e.*, the one with smallest *p*-value, for every SNP position. Using the allelic adjustments, SAFE-h^2 improved the SNP heritability estimations for the 10 artificial random phenotypes which was assigned to 1,092 HapMap 1 genotypes (chr.22 SNPs, see Fig. 2e). This resulted in up to 9.7 unit improvement when compared to the additive-only strategy (Fig. 4g). On the same datasets, the GCTA-GREMLd showed no improvements over its additive based model of GCTA-GREML (Fig. 4g). A similar pattern, although at a lower rate, was also observed on 10,000 simulated genotypes which were derived from 1,092 HapMap 1 genotypes (Fig. 4h). On this dataset, GREMLd led to 3.46 unit underestimation on average when compared to the additive approach (Fig. 4h). The allele adjustment strategy employed by SAFE-h^2 proved robust enough that no further improvements in heritability estimates were observed even after increasing the population size from 1,092 to 10,000 (Fig. 4i). On the contrary, the additive and GREMLd based estimates were improved by population expansion (Fig. 4i). The SAFE-h^2 performance on SNP heritability estimation through joint analysis of additive and non-additive allelic effects

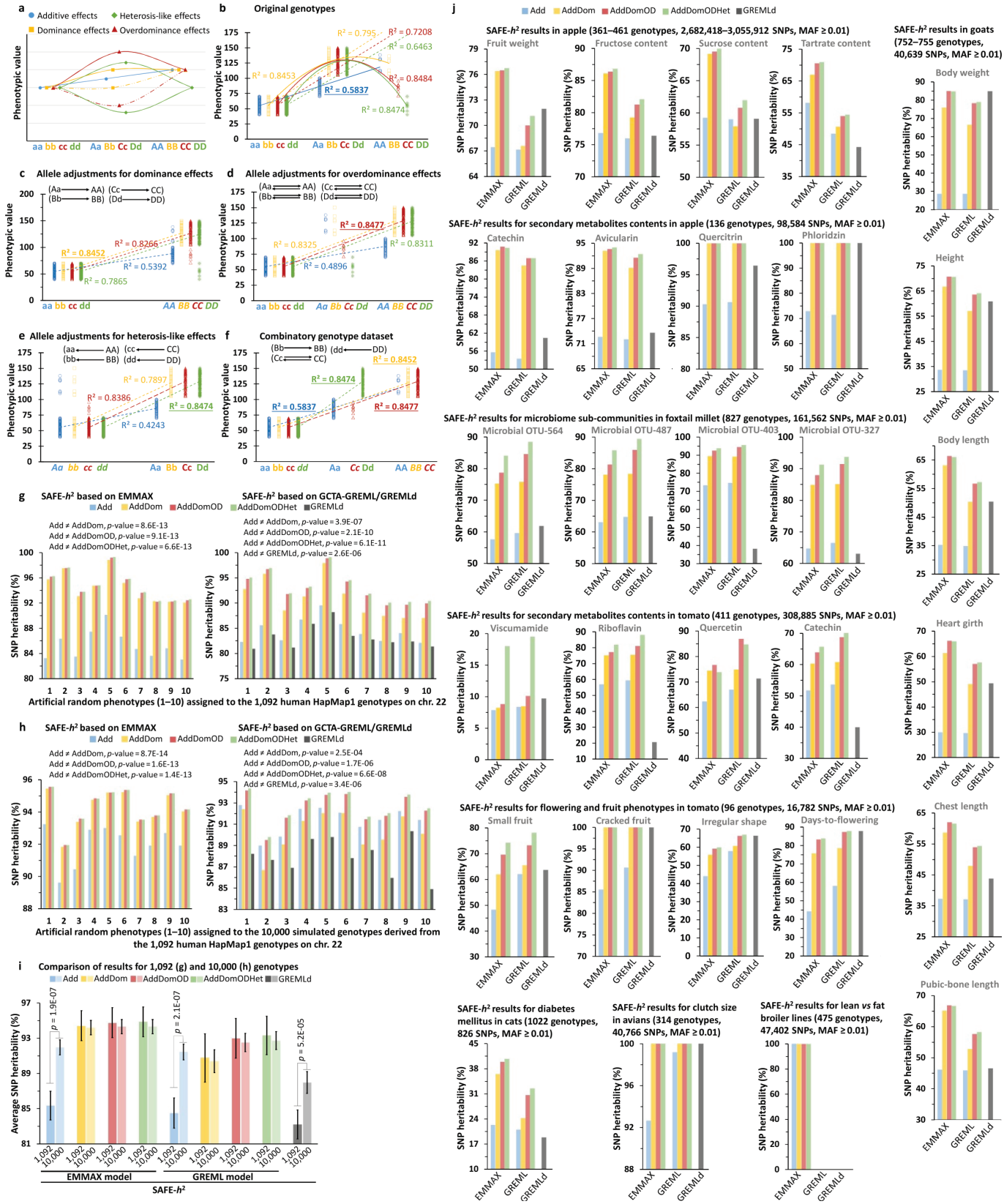

a
Additive effects
Heterosis-like effects
Dominance effects
Overdominance effects
Phenotypic value
aa bb cc dd
Aa Bb Cc Dd
AA BB CC DD
b
Original genotypes
R² = 0.795
R² = 0.7208
R² = 0.6463
R² = 0.8453
R² = 0.5837
R² = 0.8484
R² = 0.8474
c
Allele adjustments for dominance effects
(Aa AA) (Cc CC)
(Bb BB) (Dd DD)
R² = 0.8452
R² = 0.8266
R² = 0.5392
R² = 0.7865
d
Allele adjustments for overdominance effects
R² = 0.8477
R² = 0.8325
R² = 0.8311
R² = 0.4896
e
Allele adjustments for heterosis-like effects
(aa AA) (cc CC)
(bb BB) (dd DD)
R² = 0.7897
R² = 0.8386
R² = 0.8474
R² = 0.4243
f
Combinatory genotype dataset
(Bb BB) (dd DD)
(Cc CC)
R² = 0.8452
R² = 0.8474
R² = 0.5837
R² = 0.8477
g
SAFE-h² based on EMMAX
SAFE-h² based on GCTA-GREML/GREMLd
Add
AddDom
AddDomOD
AddDomODHet
GREMLd
Add ≠ AddDom, p-value = 8.6E-13
Add ≠ AddDomOD, p-value = 9.1E-13
Add ≠ AddDomODHet, p-value = 6.6E-13
Add ≠ AddDom, p-value = 3.9E-07
Add ≠ AddDomOD, p-value = 2.1E-10
Add ≠ AddDomODHet, p-value = 6.1E-11
Add ≠ GREMLd, p-value = 2.6E-06
SNP heritability (%)
Artificial random phenotypes (1–10) assigned to the 1,092 human HapMap1 genotypes on chr. 22
h
SAFE-h² based on EMMAX
SAFE-h² based on GCTA-GREML/GREMLd
Add ≠ AddDom, p-value = 8.7E-14
Add ≠ AddDomOD, p-value = 1.6E-13
Add ≠ AddDomODHet, p-value = 1.4E-13
Add ≠ AddDom, p-value = 2.5E-04
Add ≠ AddDomOD, p-value = 1.7E-06
Add ≠ AddDomODHet, p-value = 6.6E-08
Add ≠ GREMLd, p-value = 3.4E-06
Artificial random phenotypes (1–10) assigned to the 10,000 simulated genotypes derived from the 1,092 human HapMap1 genotypes on chr. 22
i
Comparison of results for 1,092 (g) and 10,000 (h) genotypes
Average SNP heritability (%)
p = 1.9E-07
p = 2.1E-07
p = 5.2E-05
1,092
10,000
EMMAX model
GREML model
SAFE-h²
j
SAFE-h² results in apple (361–461 genotypes, 2,682,418–3,055,912 SNPs, MAF ≥ 0.01)
Fruit weight
Fructose content
Sucrose content
Tartrate content
SAFE-h² results in goats (752–755 genotypes, 40,639 SNPs, MAF ≥ 0.01)
Body weight
Height
Body length
Heart girth
Chest length
Pubic-bone length
SAFE-h² results for secondary metabolites contents in apple (136 genotypes, 98,584 SNPs, MAF ≥ 0.01)
Catechin
Avicularin
Quercitrin
Phloridzin
SAFE-h² results for microbiome sub-communities in foxtail millet (827 genotypes, 161,562 SNPs, MAF ≥ 0.01)
Microbial OTU-564
Microbial OTU-487
Microbial OTU-403
Microbial OTU-327
SAFE-h² results for secondary metabolites contents in tomato (411 genotypes, 308,885 SNPs, MAF ≥ 0.01)
Viscumamide
Riboflavin
Quercetin
Catechin
SAFE-h² results for flowering and fruit phenotypes in tomato (96 genotypes, 16,782 SNPs, MAF ≥ 0.01)
Small fruit
Cracked fruit
Irregular shape
Days-to-flowering
SAFE-h² results for diabetes mellitus in cats (1022 genotypes, 826 SNPs, MAF ≥ 0.01)
SAFE-h² results for clutch size in avians (314 genotypes, 40,766 SNPs, MAF ≥ 0.01)
SAFE-h² results for lean vs fat broiler lines (475 genotypes, 47,402 SNPs, MAF ≥ 0.01)
SNP heritability (%)
EMMAX
GREML
GREMLd

Fig. 4. Allele adjustments capture intra-locus non-additive allelic effects. a–f) Demonstration of four biallelic loci; locus with A/a alleles having additive effect, locus with B/b alleles having dominance effect, locus with C/c alleles having overdominance effect, and locus with D/d alleles having heterosis-like effect. The best linear-fits are highlighted by bold and underlined R^2. The adjusted alleles are in italic. Single arrows indicate the allele adjustment by conversion and double arrows indicate allele adjustment by replacement. a) Illustration of allelic effects in a biallelic scenario. Dominance and overdominance effects can be on either alternative or reference alleles (*Alt* or *Ref*) which are differentiated by continues and dashed extrapolations. b) The artificial phenotype no. 3 that was assigned to the 1,092 human HapMap phase 1 genotypes (see Fig. 2f), and its derivatives are used to represent the additive and non-additive effects, respectively. Continues lines represent the most fitted models; for the locus with A/a alleles this was linear. Low-fit linear approximates (dashed lines) are also shown for dominance, overdominance, and heterosis-like effects in the other loci with alleles B/b, C/c, and D/d, respectively. c) Allelic adjustment for the dominance effects of Ref alleles (A, B, C, & D) leads to an improved linear-fit only on the locus having the B allele with dominance effect. d) Allelic adjustment for the overdominance effects of Ref alleles (A, B, C, & D) leads to an improved linear-fit only on the locus having the C allele with overdominance effect. e) Allelic adjustment for the heterosis-like effects leads to an improved linear-fit on the locus with D/d alleles with heterosis-like effect. f) A combinatory genotype dataset contains every SNP position in either of original or one of the adjusted states determined by having the best linear-fit, *i.e.*, bigger coefficient of determination R^2. g) SAFE-h^2 estimates of SNP heritability for the 10 artificial random phenotypes assigned to the 1,092 HapMap 1 genotypes on chr. 22 (see Fig. 2e) while considering additive-only effects, additive plus non-additive effects by allele adjustment, and additive plus dominance effects by GREMLd model. The *p*-values are from two-tailed paired t-tests comparing additive-only heritability estimates with either of additive plus non-additive combinations. h) SAFE-h^2 estimates of SNP heritability for the 10 artificial random phenotypes (at the same range as in G and Fig. 2e) assigned to the simulated 10,000 genotypes derived from the chr. 22 of 1,092 HapMap genotypes. Again, SAFE-h^2 profiling was performed by applying the three approaches of additive-only effects, additive plus non-additive effects by allele adjustment, and additive plus dominance effects by GREMLd model. The *p*-values are from two-tailed paired t-tests comparing the SNP heritability estimates on additive-only effect with the estimates from either of additive plus non-additive effects combinations. i) Comparing the results for 1,092 and 10,000 genotypes. The error bars represent 95% confidence intervals. j) SAFE-h^2 estimates of SNP heritability under additive-only model, combined additive and non-additive models, and GREMLd model for different phenotypes in plant and animal species. The analyses of tomato and foxtail millet genotypes using GCTA-GREML model were performed based on the "inbred" algorithm of GCTA. The average inbreeding coefficients (Fhat) for tomato and foxtail millet genotypes were > 0.8. g–j) Add: heritability based on additive effects, AddDom: heritability based on additive and dominance effects, AddDomOD: heritability based on additive, dominance and overdominance effects, AddDomODHet: heritability based on additive, dominance, overdominance, and heterosis-like effects, GREMLd: heritability estimated by GREMLd model, MAF: minor allele frequency, OUT: operational taxonomic unit.

was also evaluated using artificial non-random phenotypes (Table 2). These six sets of artificial phenotypes (simulated phenotypes within specific ranges, also see Fig. 2f) were associated with a biallelic locus at different rates and through different genetic effects including additive, dominance, overdominance, and heterosis-like effects (Table 2). In this example (*i.e.*, Table 2), the biallelic locus represented the only SNP position within the analyzed genomic dataset. When the underlying genetic effects were additive on the phenotypes, the estimates of SNP heritability based on additive and non-additive allelic effects were identical to those derived solely from additive effects (Table 2). When the underlying genetic effects were non-additive, the allelic adjustments for non-additive effects by SAFE-h^2 improved the heritability estimations using all the four models, *i.e.*, EMMAX, GCTA-GREML, LDAK-Thin, and LDAK-GCTA (Table 2). GCTA-GREML and LDAK were however less effective at capturing additive-only or additive plus non-additive allelic effects at different association levels (Table 2). SAFE-h^2 obtained the best approximates of the genetic variance, *i.e.*, here equivalent to the broad-sense heritability, when applying the EMMAX model (Table 2). Similar results were observed when SAFE-h^2 estimates were based on GEMMA model (Supplementary Table 5). This performance was noticeable by considering the type of allelic effects being analyzed for estimating SNP heritability as well as the type of real allelic effects being responsible for the phenotypes (Table 2). SAFE-h^2 improved estimates of SNP heritability by capturing 6.22% extra broad-sense heritability on the allelic effects of dominance, 16.22% extra broad-sense heritability on the allelic effects of overdominance, and 24.31% extra broad-sense heritability on the heterosis-like effects on average. The efficiency of SAFE-h^2 was further

Table 2: SNP heritability estimations on additive or additive + non-additive effects of a biallelic locus.

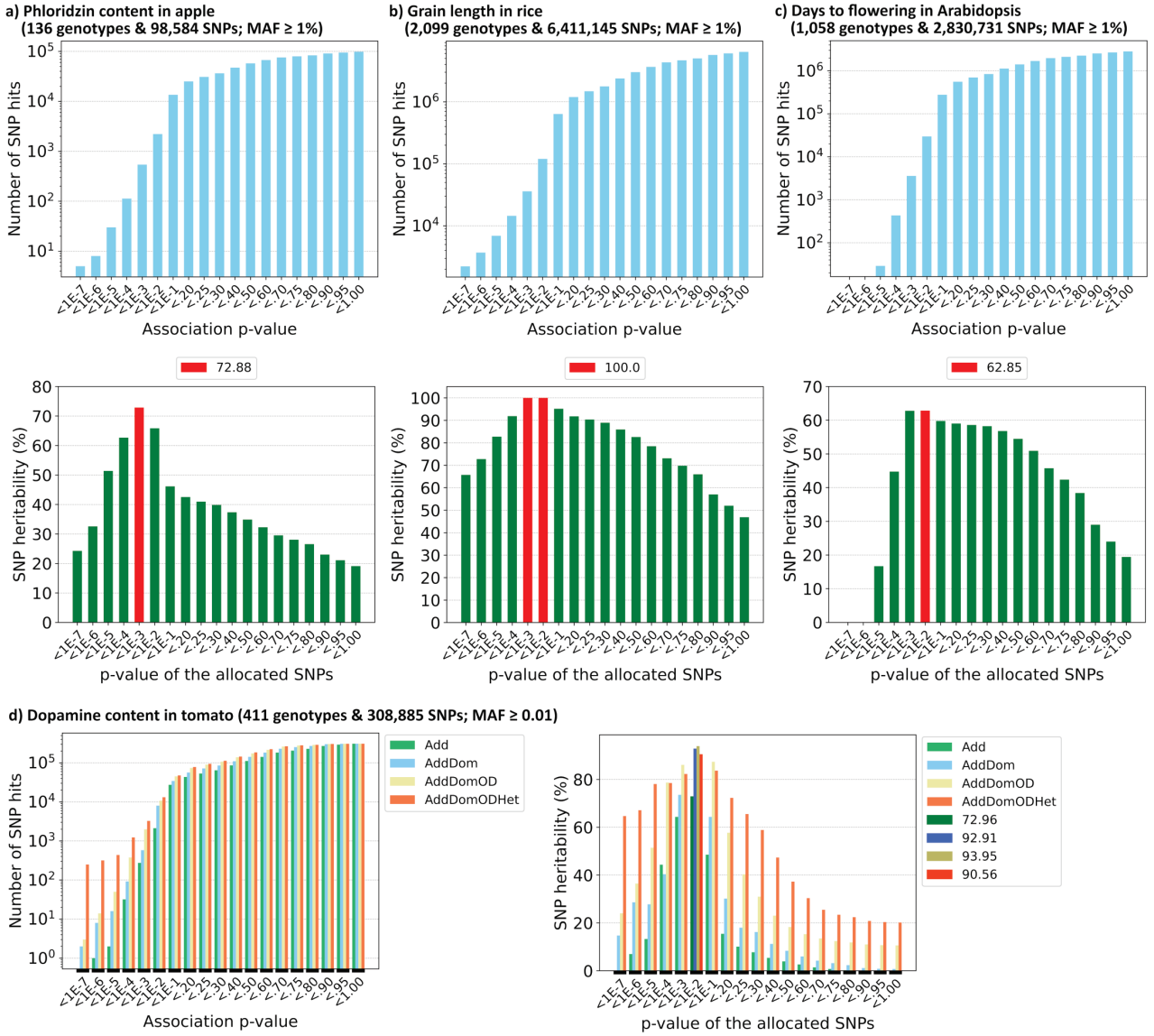
Phenotypes						SAFE-h^2 estimates of SNP-heritability																GCTA GREMLd -model
	Ref/Ref Genotype	Ref/Alt Genotype	Alt/Alt Genotype			using EMMAX model				using GCTA GREML-model				using LDAK GCTA-model				using LDAK Thin-model				
Data_Genetic effect	range (N)	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	A[a]	AD[b]	ADO[c]	ADOH[d]	A[a]	AD[b]	ADO[c]	ADOH[d]	A[a]	AD[b]	ADO[c]	ADOH[d]	A[a]	AD[b]	ADO[c]	ADOH[d]	
1st_Additice	40-50 (982)	80-100 (103)	120-150 (7)	5.7	94.3	94.29	94.29	94.29	94.29	93.91	93.91	93.91	93.91	77.22	77.22	77.22	77.22	77.22	77.22	77.22	77.22	93.92
1st_Dominance	40-50 (982)	120-150 (103)	120-150 (7)	2.16	97.84	92.8	▲ 97.83	▲ 97.83	▲ 97.83	92.77	▲ 95.74	▲ 95.74	▲ 95.74	77.16	▲ 77.53	▲ 77.53	▲ 77.53	77.16	▲ 77.53	▲ 77.53	▲ 77.53	97.89[e]
1st_Overdominance	40-50 (982)	120-150 (103)	80-100 (7)	2.35	97.65	84.77	▲ 95.9	▲ 97.65	▲ 97.65	84.29	▲ 92.12	▲ 95.48	▲ 95.48	85.21	▼ 77.35	▼ 77.51	▼ 77.51	85.21	▼ 77.35	▼ 77.51	▼ 77.51	97.77[e]
1st_Heterosis-like	40-50 (982)	120-150 (103)	40-50 (7)	6.11	93.89	70.51	▲ 86.47	▲ 91.72	▲ 93.86	69.34	▲ 76.16	▲ 84.93	94.14[e]	70.77	▲ 76.61	▲ 77.02	▲ 77.19	70.77	▲ 76.61	▲ 77.02	▲ 77.19	94.24[e]
2nd_Additive	40-60 (982)	75-100 (103)	112-150 (7)	21.04	78.96	78.86	78.86	78.86	78.86	77.7	77.7	77.7	77.7	78.84	78.84	78.84	78.84	78.84	78.84	78.84	78.84	77.53
2nd_Dominance	40-60 (982)	112-150 (103)	112-150 (7)	12.75	87.25	82.1	▲ 87.23	▲ 87.23	▲ 87.23	81.2	▼ 77.34	▼ 77.34	▼ 77.34	82.22	▼ 76.67	▼ 76.67	▼ 76.67	82.22	▼ 76.67	▼ 76.67	▼ 76.67	87.57[e]
2nd_Overdominance	40-60 (982)	112-150 (103)	75-100 (7)	14.7	85.3	71.58	▲ 82.7	▲ 85.13	▲ 85.13	70.33	▲ 70.49	▲ 74.44	▲ 74.44	71.74	▲ 82.69	▲ 76.5	▲ 76.5	71.74	▲ 82.69	▲ 76.5	▲ 76.5	86.02[e]
2nd_Heterosis-like	40-60 (982)	112-150 (103)	40-60 (7)	8.1	91.9	70.56	▲ 85.58	▲ 90.28	▲ 91.89	69.36	▲ 74.78	▲ 82.54	92.26[e]	70.79	▲ 76.54	▲ 76.91	▲ 77.04	70.79	▲ 76.54	▲ 76.91	▲ 77.04	92.35[e]
3rd_Additive	40-70 (982)	70-100 (103)	105-150 (7)	41.5	58.5	58.34	58.34	58.34	58.34	56.67	56.67	56.67	56.67	58.32	58.32	58.32	58.32	58.32	58.32	58.32	58.32	56.31
3rd_Dominance	40-70 (982)	105-150 (103)	105-150 (7)	15.47	84.53	79.48	▲ 84.51	▲ 84.51	▲ 84.51	78.45	▼ 73.15	▼ 73.15	▼ 73.15	79.57	▲ 84.49	▲ 84.49	▲ 84.49	79.57	▲ 84.49	▲ 84.49	▲ 84.49	84.91[e]
3rd_Overdominance	40-70 (982)	105-150 (103)	70-100 (7)	15.16	84.84	72.05	▲ 82.65	▲ 84.76	▲ 84.76	70.82	▼ 70.41	▲ 73.88	▲ 73.88	72.21	▲ 82.63	▲ 84.74	▲ 84.74	72.21	▲ 82.63	▲ 84.74	▲ 84.74	85.57[e]
3rd_Heterosis-like	40-70 (982)	105-150 (103)	40-70 (7)	16.44	83.56	62.59	▲ 76.85	▲ 81.56	▲ 83.52	61.12	▲ 62.39	▲ 69.24	84.19[e]	62.73	▲ 76.83	▲ 81.54	▲ 83.5	62.73	▲ 76.83	▲ 81.54	▲ 83.5	84.38[e]
4th_Additive	40-80 (982)	60-100 (103)	90-150 (7)	72.1	27.9	27.34	27.34	27.34	27.34	26.01	26.01	26.01	26.01	27.33	27.33	27.33	27.33	27.33	27.33	27.33	27.33	25.07
4th_Dominance	40-80 (982)	90-150 (103)	90-150 (7)	33.36	66.64	62.6	▲ 66.6	▲ 66.6	▲ 66.6	61.01	▼ 49.91	▼ 49.91	▼ 49.91	62.61	▲ 66.59	▲ 66.59	▲ 66.59	62.61	▲ 66.59	▲ 66.59	▲ 66.59	67.29[e]
4th_Overdominance	40-80 (982)	90-150 (103)	60-100 (7)	32.11	67.89	56.78	▲ 65.73	▲ 67.72	▲ 67.72	55.15	▼ 48.93	▼ 51.63	▼ 51.63	56.82	▲ 65.71	▲ 67.7	▲ 67.7	56.82	▲ 65.71	▲ 67.7	▲ 67.7	69.16[e]
4th_Heterosis	40-80 (982)	90-150 (103)	40-80 (7)	35.4	64.6	51.1	▲ 60.96	▲ 63.84	▲ 64.5	49.34	▼ 43.82	▼ 47.32	65.64[e]	51.03	▲ 60.94	▲ 63.82	▲ 64.49	51.03	▲ 60.94	▲ 63.82	▲ 64.49	65.98[e]
5th_Additive	40-90 (982)	50-100 (103)	75-150 (7)	91.42	8.58	7.7	7.7	7.7	7.7	7.23	7.23	7.23	7.23	22.95[e]	27.95[e]	22.95[e]	22.95[e]	22.95[e]	27.95[e]	22.95[e]	22.95[e]	na
5th_Dominance	40-90 (982)	75-150 (103)	75-150 (7)	61.67	38.33	36.61	▲ 38.24	▲ 38.24	▲ 38.24	35.04	▼ 23.62	▼ 23.62	▼ 23.62	36.59	▲ 38.22	▲ 38.22	▲ 38.22	36.59	▲ 38.22	▲ 38.22	▲ 38.22	38.65[e]
5th_Overdominance	40-90 (982)	75-150 (103)	50-100 (7)	55.12	44.88	34.61	▲ 41.87	▲ 44.1	▲ 44.82	33.09	▼ 26.46	▼ 28.64	46.06[e]	34.61	▲ 41.84	▲ 44.08	▲ 44.8	34.61	▲ 41.84	▲ 44.08	▲ 44.8	46.34[e]
5th_Heterosis-like	40-90 (982)	75-150 (103)	40-90 (7)	56	44	34.25	▲ 41.23	▲ 43.33	▲ 43.93	32.73	▼ 25.95	▼ 28	45.16[e]	34.24	▲ 41.21	▲ 43.3	▲ 43.91	34.24	▲ 41.21	▲ 43.3	▲ 43.91	45.46[e]
6th_Additive	40-95 (982)	45-100 (103)	50-150 (7)	97.56	2.44	2.3	2.3	2.3	2.3	2.15	2.15	2.15	2.15	22.59[e]	22.59[e]	22.59[e]	22.59[e]	22.59[e]	22.59[e]	22.59[e]	22.59[e]	na
6th_Dominance	40-95 (982)	50-150 (103)	50-150 (7)	76.6	23.4	20.67	▲ 23.03	▲ 23.31	▲ 23.31	19.57	▼ 13.02	▼ 13.39	▼ 13.39	20.65	▲ 23.05	▲ 23.29	▲ 23.29	20.65	▲ 23.05	▲ 23.29	▲ 23.29	24.16[e]
6th_Overdominance	40-95 (982)	50-150 (103)	45-100 (7)	76.64	23.36	19.5	▲ 22.56	▲ 23.24	▲ 23.24	18.45	▼ 12.71	▼ 13.35	▼ 13.35	19.44	▲ 22.55	▲ 23.22	▲ 23.22	19.49	▲ 22.55	▲ 23.22	▲ 23.22	24.29[e]
6th_Heterosis-like	40-95 (982)	50-150 (103)	40-95 (7)	76.57	23.43	16.19	▲ 20.62	▲ 22.57	▲ 23.21	15.28	▼ 11.49	60.84[e]	24.11[e]	16.18	▲ 20.61	▲ 22.56	▲ 23.19	16.18	▲ 20.61	▲ 22.56	▲ 23.19	24.31[e]

Green triangles indicate improved estimation when compared to the analysis on the original genotypes (a). Red triangles indicate underestimations when compared to the analysis on the original genotypes (a). σ_e^2 & σ_g^2 represent the the enviromental and genetic variances, respectively. σ_g^2 is equivalent to the broad-sense heritability. [a]On original genotypes to capture additive effects; [b]On adjusted genotypes for dominance effects to capture additive and dominance effects; [c]On adjusted genotypes for dominance and overdominance effects to capture additive, dominance, and overdominace effects; [d]On adjusted genotypes for dominance,overdominance, and heterosis-like effects to capture additive, dominance, overdominace, and heterosis-like effects; [e]Overestimated.

validated using 10 published datasets[42–51], including 50 phenotypes observed on three plants and four animal species (Fig. 4j, Supplementary Fig. 5). Again, SAFE-h^2 was able to extract extra intra-locus non-additive allelic effects at different rates depending on the dataset and the phenotype of interest (Fig. 4j, Supplementary Fig. 5). These results also confirmed the superior performance of SAFE-h^2 allelic adjustment model versus GREMLd (Fig. 4j, Supplementary Fig. 5). As observed in the simulated datasets, SAFE-h^2 led to the most stable performance and substantial improvements when analyzing the real datasets (Fig. 4j, Supplementary Fig. 5). It is important to emphasize that the overestimation of SNP heritability can only occur through positive contributions from false-positive SNP hits, which depends on statistical power levels. Under an identical statistical power level, *i.e.*, identical genomic and genotypic coverage, the most effective method for heritability estimation is the one that captures the greatest proportion of phenotypic variation.

Discussion

SNP heritability estimations are typically obtained using one of two following approaches: either by including all SNPs or by selecting SNPs having a *p*-value less than 10^{-8}. It appears that researchers often choose the method that yields the highest estimate for each experiment. A remarkable example of this is the recent heritability estimates for human height in two large cohorts[16,17]; two different strategies for heritability estimation of an identical phenotype. Functional, regional, and MAF stratified SNPs have also been employed in SNP heritability calculations[8–10,12,14,16]. Nevertheless, these approaches provide estimates of genetic variance at sub-genome levels rather than phenotype-specific estimates, *i.e.*, associated SNP specific estimates. Importantly, these approaches do not exclude SNPs having upper-tail association *p*-values. This issue also applies to SNP heritability estimations that are adjusted for linkage disequilibrium. It is evident that all of these SNP heritability estimation strategies are vulnerable to the inclusion of unassociated SNPs with upper-tail association *p*-values, potentially introducing illogical negative contributions to heritability estimates. Consequently, these methods will remain of limited utility until strategies are developed that can accurately estimate heritability while avoiding the negative contribution of unassociated SNPs. Interestingly, it has been suggested that excluding a given SNP from genetic relatedness matrix (GRM) calculations may be an effective strategy to mitigate covariate effects on association test of that given SNP[52]. Although the genetic relatedness matrices ideally require phenotype-associated SNPs, the trade-off between the detrimental effects of excluding associated SNPs and including irrelevant variants has made the SNP selection challenging, often resulting in inaccurate genomic predictions[53]. However, the SAFE-h^2 method bypasses this trade-off by identifying a *p*-value threshold below which no SNP can be found with deflationary effects on genetic variance. In summary, the findings of this study suggest that for accurate heritability estimation, it is crucial not only to include SNPs beyond the user-defined significant hits, but also to perform SNP-pruning for large *p*-values which is essential to prevent biologically meaningless negative contributions. To meet these challenges, SAFE-h^2 performs a dynamic "**S**NP **A**llocation **F**or **E**stimating the narrow-sense **H**eritability". SAFE-h^2 leverages the SNP heritability estimating tools EMMAX[25], LDAK[5], GCTA-GREML[26], and GEMMA[27] to estimate SNP heritability for *p*-value-guided SNP clusters. SAFE-h^2 generates a list of *p*-value-guided SNP clusters and outputs a profiling graph of heritability estimates (Fig. 5a–c). This approach facilitates the inclusion of SNP hits as their association *p*-values are ascending, up to a threshold at which the SNP heritability reaches its maximum for the given experiment or dataset. Heritability profiling by SAFE-h^2 identified the EMMAX algorithm as the most stable model for accurate estimation of SNP heritability (Fig. 2, Supplementary Fig. 2–4, Table 2). Although the *p*-values used for SNP clustering in the analyses were obtained exclusively by EMMAX to ensure proper comparisons, users are free to perform any association test of their choice to supply data for SAFE-h^2. Furthermore, we conducted analyses on eight phenotypes across four species to align the heritability estimates obtained based on the calculated *p*-values by EMMAX and GEMMA (Supplementary Table 6). In general, we found larger negative contributions to heritability


a) Phloridzin content in apple
(136 genotypes & 98,584 SNPs; MAF ≥ 1%)
b) Grain length in rice
(2,099 genotypes & 6,411,145 SNPs; MAF ≥ 1%)
c) Days to flowering in Arabidopsis
(1,058 genotypes & 2,830,731 SNPs; MAF ≥ 1%)
Number of SNP hits
Association p-value
72.88
100.0
62.85
SNP heritability (%)
p-value of the allocated SNPs

d) Dopamine content in tomato (411 genotypes & 308,885 SNPs; MAF ≥ 0.01)
Number of SNP hits
Association p-value
Add
AddDom
AddDomOD
AddDomODHet
SNP heritability (%)
p-value of the allocated SNPs
72.96
92.91
93.95
90.56

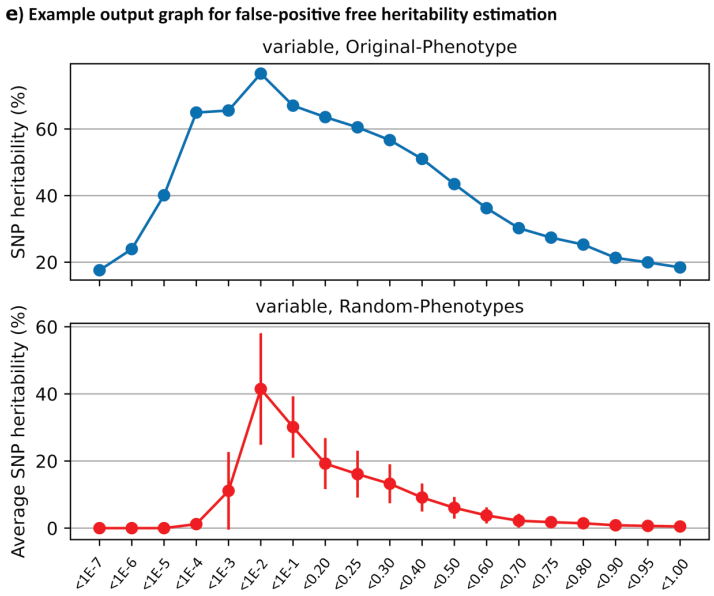

e) Example output graph for false-positive free heritability estimation
variable, Original-Phenotype
SNP heritability (%)
variable, Random-Phenotypes
Average SNP heritability (%)

Fig. 5. SAFE-h^2 generates bar graphs for both the *p*-value-guided SNP clusters and the SNP heritability profiles. The SNP heritability estimates are highlighted by red color in (a), (b), (c), and by darker colors in (d). Add: heritability based on additive effects, AddDom: heritability based on additive and dominance effects, AddDomOD: heritability based on additive, dominance and overdominance effects, AddDomODHet: heritability based on additive, dominance, overdominance, and heterosis-like effects, MAF: minor allele frequency. e) Bars represent the standard deviations among six random phenotypes.

estimations when the analyses were based on *p*-values calculated by GEMMA, which employs an exact mixed model[27]. Notably, SAFE-h^2 utilizes the Balding-Nichols (BN) kinship matrix for EMMAX-based estimation of heritability. The BN kinship matrix has been reported to be highly correlated with FaST-LMM and GEMMA kinship matrices[54].

A key challenge in GWAS studies is the low statistical power due to insufficient genotypic coverage, which can lead to false-positive SNP hits. To address this, SAFE-h^2 restricts the contributions of false-positive SNP hits to SNP heritability estimations. This is achieved by comparing heritability profiles of random phenotypes to that of the phenotype-of-interest (Fig. 3, 5e). The hidden heritability due to false-negative SNP hits has also been discussed previously[55]. In its process of false-positive free heritability estimation, SAFE-h^2 also ensures that weakly associated SNPs, which might otherwise be excluded as false negatives using a predefined *p*-value cut-off (*e.g.*, 10^{-8}), are included in the analysis. Replication of experiments, such as selecting associated SNPs in one cohort and calculating polygenic risk scores for each individual in a second cohort, is a common strategy to mitigate overfitting, which is a common challenge in GWAS[56]. It is important to note that overfitting typically occurs for the genotype-level, not at the cohort-level. Provided that both cohorts originate from the same population (as opposed to distinct populations), the risk of overfitting is minimized. In contrast, performing replicated analyses using cohorts from different populations does not mitigate overfitting, as the differences between populations may introduce confounding variables. Nevertheless, overfitting was not a concern worth considering in our analyses. This is because heritability is a population or cohort-level statistic (*i.e.*, not a genotype-level measure) and, an estimate that is specific to the dataset and experiment, as outlined in the introduction.

SAFE-h^2 is capable of extracting intra-locus non-additive allelic effects, leading to more accurate estimates of SNP heritability. An example of SAFE-h^2 output graphs for joint analysis of additive and non-additive allelic effects is shown in Fig. 5d. Unlike the GCTA-GREMLd model, SAFE-h^2 employs an allele adjustment algorithm, which produces identical results to the original genotypic data when only additive allelic effects are present (Table 2). SAFE-h^2 achieved the best performance using EMMAX and GEMMA models to capture additive and non-additive allelic effects of the predefined single biallelic-locus, without introducing overestimation errors (Table 2, Supplementary Table 5). In contrast, GCTA-GREMLd model, while still providing useful estimates, resulted in small overestimation errors across all measurements for the predefined single biallelic-locus analyses (Table 2). GCTA-GREMLd was also prone to SNP heritability underestimations for the real phenotypes, as its estimates for joint additive and non-additive effects were smaller than those for additive-only effect (Fig. 4j, Supplementary Fig. 5). It is important to distinguish SNP × SNP genetic effects or interactions between non-allelic genes from intra-locus (SNP-specific *Ref* × *Ref, Ref* × *Alt, and Alt* × *Alt*) allelic effects. The mapping efficiency of epistatic gene × gene interactions has been addressed previously[57] and the recently developed GWAIS tool offers a major improvement by speeding up the processing of SNP × SNP effects[58]. In contrast, SAFE-h^2 deals with intra-locus genetic effects of additive and non-additive. Available methods including GCTA-GREMLd only consider intra-locus effects of dominance and do not include other intra-locus non-additive effects[41,59]. Our analyses also confirms that GCTA-GREMLd is inadequate for properly and efficiently extracting dominance effects. In contrast, SAFE-h^2 offers a significant advantage through its assumption-free pure mathematics to extract all types of non-additive effect. SAFE-h^2 adjusts genotypes

based on observed intra-locus effects. In summary, SAFE-h^2 facilitates phenotype-specific rather than genome or sub-genome specific estimates of SNP heritability with the ability to minimize the likelihood of false-positive contributions. SAFE-h^2 further refines the estimations by joint analysis of intra-locus non-additive and additive allelic effects. By adjusting alleles, SAFE-h^2 generates genotypic datasets that facilitate the linear association tests to address intra-locus additive and non-additive effects jointly.

Materials and Methods

Datasets

The simulated homozygote genotypes were arranged across 29 chromosomes, with 300kb distance between SNPs to meet linkage equilibrium. For each genotype, reference and alternative alleles were randomly assigned at different chromosomal positions. VCF files for 1,000 and 10,000 genotypes were assembled with 200k SNPs with MAF ranging from 1% to 10%. Within these datasets, one of the 200k SNP positions was designated as non-random to be associated with the simulated non-random phenotypes. For having multiple predefined associations, 11, 23, or 47 extra predefined SNPs were integrated into the same genomic datasets. Further proof-of-concept analyses were accomplished using human chromosome 22 from the HapMap phase 1 genomic data [28]. This included 1,092 genotypes and 177,675 SNPs after dataset filtering (MAF ≥1%). As for the simulated genotypes, one simulated non-random SNP position was introduced among the HapMap variants to be associated with the simulated non-random phenotypes. The human chr.22 SNP-set was also extended by simulation to 10,000 individuals. For this, the original SNP dataset was shuffled three times among the chromosomal positions to get a new genotype set. This was repeated to have 9 new datasets. The original and new datasets were then horizontally assembled to build a dataset of 10,000 genotypes represented by 177,675 SNPs. The final simulated dataset had 177,445 variants after filtering for MAF of ≥1%. The artificial phenotypes were produced as random or non-random integers assigned randomly among the genotypes or among the genotypic groups of the predefined associated SNPs. Comprehensive analyses were also conducted using *Arabidopsis* and rice public datasets. For *Arabidopsis*, 14 different phenotypes were obtained from the AraPheno database [29]. These phenotypes were from a panel of 328 to 1,058 genotypes. The imputed genomic variants for these panels [30,31] contained 2,621,362 to 2,840,072 post-filtering SNPs. The public rice data was obtained from the rice SNP-Seek database [32]. The analyses were performed using 10 phenotypes that were previously recorded in 1,422 to 2,102 rice genotypes. The rice genomic data was also publicly available in rice SNP-Seek database [32–34]. The genotypes contained from 6,131,325 to 6,414,800 post-filtering SNPs. Additional datasets from both animal and plant species were analyzed to examine the negative contributions to heritability as well as the inclusion of intra-locus allelic effects of non-additive in the SNP heritability estimates. These included clutch size in 314 wild avians [42] genotyped by 40,766 SNPs (MAF ≥0.01), 203 lean broiler lines versus 272 fat broiler lines [43] genotyped by 47,402 SNPs (MAF ≥0.01), four fruit and flowering related phenotypes in 96 tomato genotypes [44] represented by 16,782 SNPs (MAF ≥0.01), the contents of 14 secondary metabolites in 411 tomato genotypes [50] represented by 308,885 SNPs (MAF ≥0.01), the contents of 7 secondary metabolites in 136 apple genotypes [45] represented by 98,584 SNPs (MAF ≥0.01), apple fruit weight, sugar and dicarboxylic acid contents in 368 genotypes [46] represented by 3,055,912 SNPs (MAF ≥0.01), six growth related phenotypes in 752–755 goat genotypes [47] represented by 40,639 SNPs (MAF ≥0.01), the abundance of 12 microbiome sub-communities in 827 genotypes of foxtail millet represented by 161,562 SNPs [51], and a merged dataset for susceptibility to diabetes mellitus in 1,022 genotypes represented by 826 SNPs (MAF ≥0.01) that derived from two independent studies in cats [48,49].

Data analysis

Plink[60] and EMMAX[25] were used for the initial processing of genomic data and association analyses. MAF was ≥ 1% in all the analyses. A complete record on the latitudes of geographical origins was available for Arabidopsis genotypes when analyzing the stem length and leaf mineral contents. This geographical information was applied as quantitative covariates in the analyses. SAFE-h^2 was developed to perform a dynamic SNP allocation for estimation of narrow-sense heritability. A comprehensive manual is provided together with the software. In brief, SAFE-h^2 begins by scanning the input files and executing the user-defined combination of heritability estimation models on *p*-value guided SNP clusters. It determines a dataset- and experiment-specific *p*-value threshold at which SNPs with larger *p*-values are distributed randomly across genotypes in relation to the phenotypic variation, imposing negative contributions to heritability estimates. We define these SNPs as "unassociated". They lead to higher genetic similarities and thereby, genetic variance and SNP heritability underestimations. In fact, the p-value threshold serves as the most effective criterion to differentiate the associated and unassociated SNPs under standard experimental conditions, *i.e.*, the genomic and genotypic coverage. SAFE-h^2 requires a *p*-value file for SNPs, the Plink b-files ("bim", "bed", and "fam"), and optional covariate files specific for every estimation model. SAFE-h^2 leverages five different models of EMMAX[25], LDAK-GCTA, LDAK-Thin[5], GCTA-GREML[26], and GEMMA[27] to estimate SNP heritability. GCTA-GREML inbred algorithm can also be applied to inbred genotypes or populations with considerable levels of inbreeding. To operate, SAFE-h^2 requires the binary files of the mentioned tools (see the SAFE-h^2 manual with full description). In return, SAFE-h^2 generates the output data files and related graphs for clustered SNP hits and SNP heritability profiles. It also features the *p*-value threshold and the estimated SNP heritability at that threshold on the graphs. Given that different levels of false-positive rates are typically expected in GWAS, SAFE-h^2 employs an innovative approach to exclude these false-positive SNP hits. SAFE-h^2 determines a "safe" *p*-value interval, with its upper-end to be used as an alternative p-value threshold for heritability estimations, ensuring minimal contributions from false-positive SNP hits. To perform the analysis, the user should simulate six independent random phenotypes within the range of the original phenotype-of-interest for all genotypes in the dataset. SAFE-h^2 requires several input files: a "bim" file, a "bed" file, a "fam" file containing the original phenotypic records, six "fam" files for simulated random phenotypes, SNP p-value files for the real (one file) and random phenotypes (six files), and optional covariate files specific for estimation models. Using these input files, SAFE-h^2 creates the SNP heritability profiles for all the random phenotypes as well as the real phenotype-of-interest using an identical genomic dataset. Users can select the estimation models of interest. Analyzing the random phenotypes subsequently determines a safe *p*-value interval having minimum heritability level at the lower-tail of heritability profile. The upper-end of this interval can function as a *p*-value threshold to omit the false-positive SNP hits from the estimation models and keep their contribution to heritability at minimum. Finally, SAFE-h^2 generates an aligned graph of heritability profiles for the random and original phenotypes. The SNP heritability of the original phenotype can accordingly be defined at upper-end of the determined *p*-value interval to keep the false-positive contributions at its minimum.

SAFE-h^2 also offers allelic adjustments for linear approximation on intra-locus allelic effects including dominance, overdominance, and heterosis-like effects. This facilitates capturing of non-additive allelic effects. SAFE-h^2-preADOH builds two VCF files for the dominance adjustments by converting the heterozygote nucleotide positions into one of the homozygotes. It also creates two VCF files for the overdominance adjustments, where heterozygote nucleotide positions are converted into either of homozygotes and, the same homozygotes are converted into their heterozygote counterparts. For the heterosis-like effects, SAFE-h^2-preADOH builds an additional VCF file by converting *Ref* × *Ref* homozygotes into the corresponding homozygote of *Alt* × *Alt*, and vice versa, at each nucleotide position. Here, the user can perform an independent association test or alternatively can command SAFE-h^2-preADOH to execute the Plink's glm function on the five adjusted and one original genotype datasets. To construct the combinatory

genotype datasets, SAFE-h^2-preADOH applies the association *p*-values to select the best linear-fit (with smallest association *p*-value) among the original and the five adjusted versions for every SNP position. At the end, SAFE-h^2-preADOH provides three genotypic datasets that can be used to estimate the additive-dominance effects, additive-dominance-overdominance effects, and additive-dominance-overdominance-heterosis. Here, every SNP-ID gets either of the following specific tags; [Add] for having additive effects on the phenotype of interest, [Dom] or [dom] for having dominance effects on the phenotype of interest, [OD] or [od] for having overdominance effects on the phenotype of interest, or [Het] for having heterosis-like effects on the phenotype of interest. Finally, SAFE-h^2-ADOH performs the heritability profiling as explained for SAFE-h^2 and on the three adjusted genotype datasets as well as the original genotype dataset. It is important to emphasize that SAFE-h^2 considers very SNP through either of additive, dominance, overdominance, or heterosis-like effects for inclusion in the heritability model. Therefore, the file for additive-dominance effects is used to examine the SNPs tagged with [Add] only for additive effect and the SNPs tagged with [Dom] or [dom] only for dominance effect. The file for additive-dominance-overdominance effects is used to examine the SNPs tagged with [Add] only for additive effect, the SNPs tagged with [Dom] or [dom] only for dominance effect, and the SNPs tagged with [OD] or [od] only for overdominance effect. Finally, the file for additive-dominance-overdominance-heterosis effects is used to examine the SNPs tagged with [Add] only for additive effect, the SNPs tagged with [Dom] or [dom] only for dominance effect, the SNPs tagged with [OD] or [od] only for overdominance effect, and the SNPs tagged with [Het] only for heterosis-like effect.

Author contributions

B.D conceived the study. B.D performed the analyses and compiled the SAFE-h^2 script. B.D and M.N wrote the manuscript. B.D and M.N supervised the study.

Competing interests

The authors declare no competing interests.

Supplementary figures and tables

Supplementary Fig. 1: The SNP heritability profile for simulated random (quantitative) phenotypes and SNPs.

Supplementary Fig. 2: The SNP heritability profile for simulated quantitative phenotypes and SNPs, with predesignated associated SNPs at varying rates.

Supplementary Fig. 3: The SNP heritability profiles for simulated random and non-random quantitative phenotypes and SNPs, with a comparison of linkage disequilibrium and linkage equilibrium.

Supplementary Fig. 4: The SNP heritability profiles for simulated random and non-random quantitative phenotypes and SNPs with minor allele frequencies ranging from 1.6%–50%.

Supplementary Fig. 5: SAFE-h^2 estimates of SNP heritability using additive-only model and the combined additive and non-additive models, compared to GREMLd estimates.

Supplementary Table 1: Heritability profiling by SAFE-h^2 reveals negative contributions of SNPs with large association *p*-values.

Supplementary Table 2: The SNP heritability estimates for simulated quantitative phenotypes and 200K SNPs in LE and LD.

Supplementary Table 3: The SNP heritability estimates for simulated quantitative phenotypes and the chr. 22 SNPs from human HapMap phase 1 genotypes.

Supplementary Table 4: The SNP heritability estimates for real phenotypes using BayesR and SAFE-h^2.

Supplementary Table 5: SNP heritability estimations on additive or additive + non-additive effects of a biallelic locus.

Supplementary Table 6: Heritability estimates based on p-values, calculated using EMMAX and GEMMA.

Code availability

SAFE-h^2 is available for download at: https://github.com/SAFE-h2/SAFE-h2-v2024-1. The example dataset can also be obtained from figshare (DOI: 10.6084/m9.figshare.24025728, Link: https://figshare.com/s/7193d332894f11deb1a5).

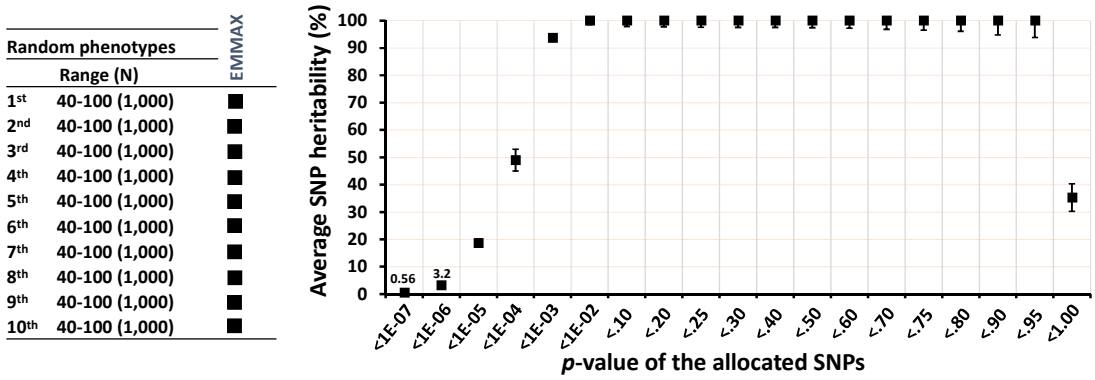
References

1. Tenesa, A. & Haley, C. S. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet* **14**, 139–149 (2013).
2. Zuk, O. *et al.* Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences* **111**, E455–E464 (2014).
3. Dahl, A. *et al.* A Robust Method Uncovers Significant Context-Specific Heritability in Diverse Complex Traits. *Am J Hum Genet* **106**, 71–91 (2020).
4. Cohen, N. M. *et al.* Personalized lab test models to quantify disease potentials in healthy individuals. *Nat Med* **27**, 1582–1591 (2021).
5. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved Heritability Estimation from Genome-wide SNPs. *Am J Hum Genet* **91**, 1011–1021 (2012).
6. Tang, M., Wang, T. & Zhang, X. A review of SNP heritability estimation methods. *Briefings in Bioinformatics* **23**, bbac067 (2022).
7. Yang, J., Zeng, J., Goddard, M. E., Wray, N. R. & Visscher, P. M. Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet* **49**, 1304–1310 (2017).
8. Evans, L. M. *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet* **50**, 737–745 (2018).
9. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565–569 (2010).
10. Lee, S. H. *et al.* Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* **44**, 247–250 (2012).
11. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* **51**, 793–803 (2019).
12. Speed, D., Cai, N., Johnson, M. R., Nejentsev, S. & Balding, D. J. Reevaluation of SNP heritability in complex human traits. *Nat Genet* **49**, 986–992 (2017).
13. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
14. Speed, D., Kaphle, A. & Balding, D. J. SNP-based heritability and selection analyses: Improved models and new results. *BioEssays* **44**, 2100170 (2022).
15. Baselmans, B. M. L., Yengo, L., van Rheenen, W. & Wray, N. R. Risk in Relatives, Heritability, SNP-Based Heritability, and Genetic Correlations in Psychiatric Disorders: A Review. *Biological Psychiatry* **89**, 11–19 (2021).
16. Yengo, L. *et al.* A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
17. Wainschtein, P. *et al.* Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nat Genet* **54**, 263–273 (2022).
18. Elks, C. E. *et al.* Variability in the Heritability of Body Mass Index: A Systematic Review and Meta-Regression. *Front Endocrinol (Lausanne)* **3**, 29 (2012).
19. Carmichael, C. M. & McGue, M. A cross-sectional examination of height, weight, and body mass index in adult twins. *J Gerontol A Biol Sci Med Sci* **50**, B237-244 (1995).
20. Young, A. I. *et al.* Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat Genet* **50**, 1304–1310 (2018).
21. Brachi, B., Morris, G. P. & Borevitz, J. O. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology* **12**, 232 (2011).
22. Barry, C.-J. S. *et al.* How to estimate heritability: a guide for genetic epidemiologists. *International Journal of Epidemiology* **52**, 624–632 (2023).
23. Golan, D. & Rosset, S. Accurate estimation of heritability in genome wide studies using random effects models. *Bioinformatics* **27**, i317–i323 (2011).

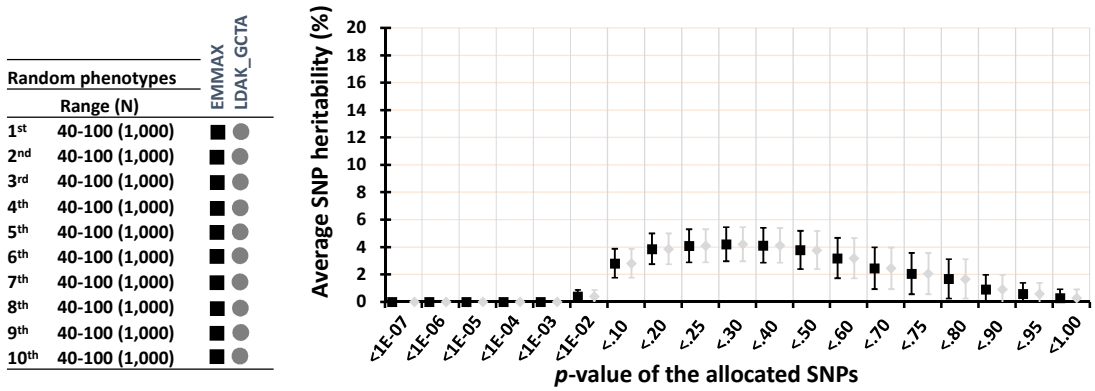
24. Golan, D., Lander, E. S. & Rosset, S. Measuring missing heritability: Inferring the contribution of common variants. *Proceedings of the National Academy of Sciences* **111**, E5272–E5281 (2014).
25. Kang, H. M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* **42**, 348–354 (2010).
26. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *The American Journal of Human Genetics* **88**, 76–82 (2011).
27. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**, 821–824 (2012).
28. McVean, G. A. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
29. Seren, Ü. *et al.* AraPheno: a public database for Arabidopsis thaliana phenotypes. *Nucleic Acids Research* **45**, D1054–D1059 (2017).
30. Alonso-Blanco, C. *et al.* 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166**, 481–491 (2016).
31. Arouisse, B., Korte, A., van Eeuwijk, F. & Kruijer, W. Imputation of 3 million SNPs in the Arabidopsis regional mapping population. *The Plant Journal* **102**, 872–882 (2020).
32. Mansueto, L. *et al.* Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Research* **45**, D1075–D1081 (2017).
33. McCouch, S. R. *et al.* Open access resources for genome-wide association mapping in rice. *Nat Commun* **7**, 10532 (2016).
34. The 3, 000 rice genomes project. The 3,000 rice genomes project. *GigaScience* **3**, 2047-217X-3–7 (2014).
35. Moser, G. *et al.* Simultaneous Discovery, Estimation and Prediction Analysis of Complex Traits Using a Bayesian Mixture Model. *PLOS Genetics* **11**, e1004969 (2015).
36. MacLeod, I. M. *et al.* Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics* **17**, 144 (2016).
37. Piaskowski, J. *et al.* Genomic heritability estimates in sweet cherry reveal non-additive genetic variance is relevant for industry-prioritized traits. *BMC Genetics* **19**, 23 (2018).
38. Guindo-Martínez, M. *et al.* The impact of non-additive genetic associations on age-related complex diseases. *Nat Commun* **12**, 2436 (2021).
39. Yadav, S. *et al.* Improved genomic prediction of clonal performance in sugarcane by exploiting non-additive genetic effects. *Theor Appl Genet* **134**, 2235–2252 (2021).
40. Varona, L., Legarra, A., Toro, M. A. & Vitezica, Z. G. Non-additive Effects in Genomic Selection. *Frontiers in Genetics* **9**, (2018).
41. Zhu, Z. *et al.* Dominance Genetic Variation Contributes Little to the Missing Heritability for Human Complex Traits. *The American Journal of Human Genetics* **96**, 377–385 (2015).
42. Husby, A. *et al.* Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. *Proc Biol Sci* **282**, 20150156 (2015).
43. Dou, D. *et al.* Genome-wide association studies for growth traits in broilers. *BMC Genomic Data* **23**, 1 (2022).
44. Yamamoto, E. *et al.* A simulation-based breeding design that uses whole-genome prediction in tomato. *Sci Rep* **6**, 19454 (2016).
45. McClure, K. A. *et al.* Genome-wide association studies in apple reveal loci of large effect controlling apple polyphenols. *Hortic Res* **6**, 1–12 (2019).
46. Liao, L. *et al.* Unraveling a genetic roadmap for improved taste in the domesticated apple. *Molecular Plant* **14**, 1454–1471 (2021).
47. Moaeen-ud-Din, M., Danish Muner, R. & Khan, M. S. Genome wide association study identifies novel candidate genes for growth and body conformation traits in goats. *Sci Rep* **12**, 9891 (2022).
48. Forcada, Y., Boursnell, M., Catchpole, B. & Church, D. B. A genome-wide association study identifies novel candidate genes for susceptibility to diabetes mellitus in non-obese cats. *PLOS ONE* **16**, e0259939 (2021).
49. Hernandez, I. *et al.* Complex Feline Disease Mapping Using a Dense Genotyping Array. *Frontiers in Veterinary Science* **9**, (2022).
50. Zhu, G. *et al.* Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell* **172**, 249-261.e12 (2018).
51. Wang, Y. *et al.* GWAS, MWAS and mGWAS provide insights into precision agriculture based on genotype-dependent microbial effects in foxtail millet. *Nat Commun* **13**, 5913 (2022).
52. Listgarten, J. *et al.* Improved linear mixed models for genome-wide association studies. *Nat Methods* **9**, 525–526 (2012).
53. Lippert, C. *et al.* The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. *Sci Rep* **3**, 1815 (2013).
54. Eu-ahsunthornwattana, J. *et al.* Comparison of Methods to Account for Relatedness in Genome-Wide Association Studies with Family-Based Data. *PLoS Genet* **10**, e1004445 (2014).
55. Williams, S. M. & Haines, J. L. Correcting away the hidden heritability. *Ann Hum Genet* **75**, 348–350 (2011).

56. Subramanian, J. & Simon, R. Overfitting in prediction models - is it a problem only in high dimensions? *Contemp Clin Trials* **36**, 636–641 (2013).
57. Viana, J. M. S. & Souza, C. A. S. Efficiency of mapping epistatic quantitative trait loci. *Heredity* **131**, 25–32 (2023).
58. Wienbrandt, L., Prieß, C. & Ellinghaus, D. GWAIS-Web: A fast and secure web service for epistasis detection in genome-wide association interaction studies. bioRxiv (2023). Doi: https://doi.org/10.1101/2023.05.16.540964.
59. Vitezica, Z. G., Varona, L. & Legarra, A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* **195**, 1223–1230 (2013).
60. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047–8 (2015).

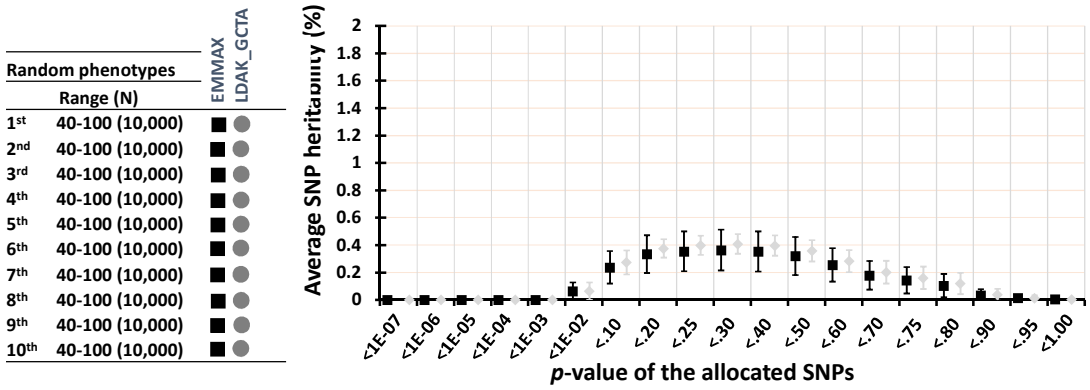
a) Analysis of 10 artificial random phenotypes on a panel of 1,000 simulated inbred genotypes (2000k SNPs; MAF = 1.6%–10%)



b) Analysis of 10 artificial random phenotypes on a panel of 1,000 simulated inbred genotypes (100 SNPs; MAF = 1.6%–10%)



c) Analysis of 10 artificial random phenotypes on a panel of 10,000 simulated inbred genotypes (100 SNPs; MAF = 1.6%–10%)

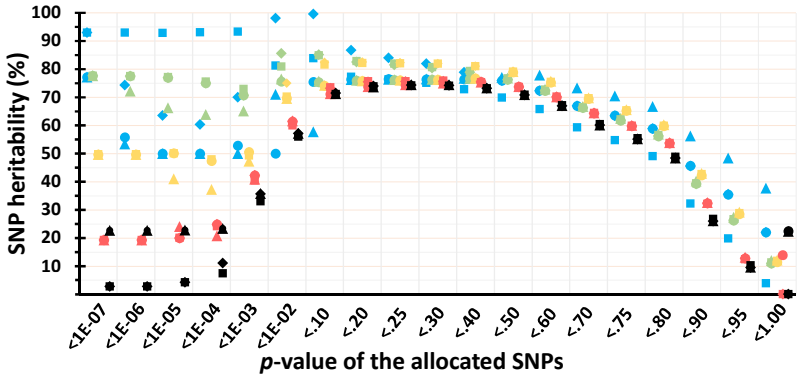


Supplementary Fig. 1: The SNP heritability profile for simulated random (quantitative) phenotypes and SNPs. Initial GWAS was performed using EMMAX to determine the association levels for the SNPs. The four different approaches of EMMAX, LDAK-GCTA, LDAK-Thin, and GCTA-GREML (by "inbred" algorithm) were then applied to estimate the SNP heritability for the clusters of SNP hits captured using the association *p*-values. N represents the sample sizes, *i.e.*, number of genotypes. MAF stands for the minor allele frequency. The average SNP heritability estimates were obtained for 10 independent random phenotypes within identical ranges. The error bars represent 99% confidence intervals.

a) Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,011 SNPs; MAF = 1.6%—10%)

Phenotypes	The introduced non-random SNP positions				
	Ref*i*/Ref*i* Genotype	Alt*i*/Alt*i* Genotype			EMMAX / LDAK_GCTA / LDAK_Thin / GCTA_GREML
	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	
1st	40-50 (900)	80-100 (100)	6.995	93.005	
2nd	40-60 (900)	75-100 (100)	22.389	77.611	
3rd	40-70 (900)	70-100 (100)	50.267	49.733	
4th	40-80 (900)	60-100 (100)	80.690	19.310	
5th	40-90 (900)	50-100 (100)	97.170	2.830	

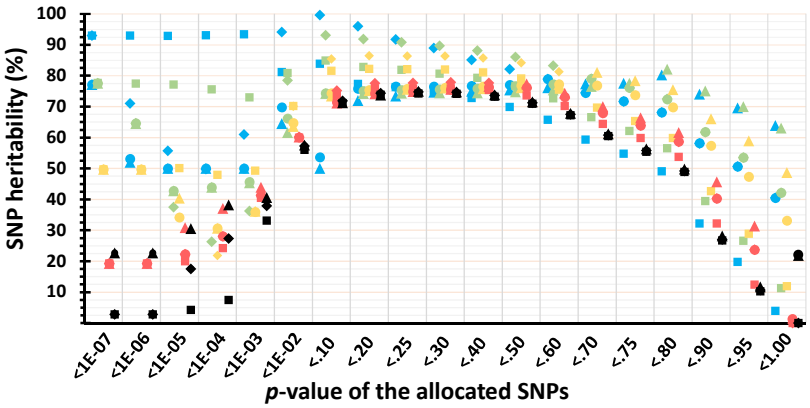
i: 1st-12th associated SNP.



b) Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,023 SNPs; MAF = 1.6%—10%)

Phenotypes	The introduced non-random SNP positions				
	Ref*i*/Ref*i* Genotype	Alt*i*/Alt*i* Genotype			EMMAX / LDAK_GCTA / LDAK_Thin / GCTA_GREML
	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	
1st	40-50 (900)	80-100 (100)	6.995	93.005	
2nd	40-60 (900)	75-100 (100)	22.389	77.611	
3rd	40-70 (900)	70-100 (100)	50.267	49.733	
4th	40-80 (900)	60-100 (100)	80.690	19.310	
5th	40-90 (900)	50-100 (100)	97.170	2.830	

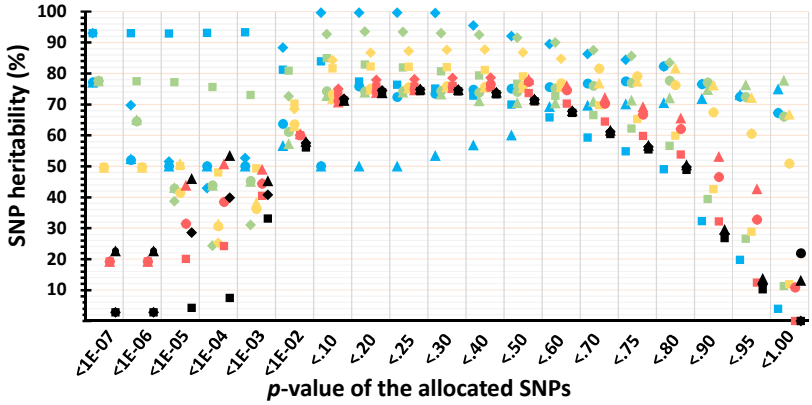
i: 1st-24th associated SNP.



c) Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,047 SNPs; MAF = 1.6%—10%)

Phenotypes	The introduced non-random SNP positions				
	Ref*i*/Ref*i* Genotype	Alt*i*/Alt*i* Genotype			EMMAX / LDAK_GCTA / LDAK_Thin / GCTA_GREML
	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	
1st	40-50 (900)	80-100 (100)	6.995	93.005	
2nd	40-60 (900)	75-100 (100)	22.389	77.611	
3rd	40-70 (900)	70-100 (100)	50.267	49.733	
4th	40-80 (900)	60-100 (100)	80.690	19.310	
5th	40-90 (900)	50-100 (100)	97.170	2.830	

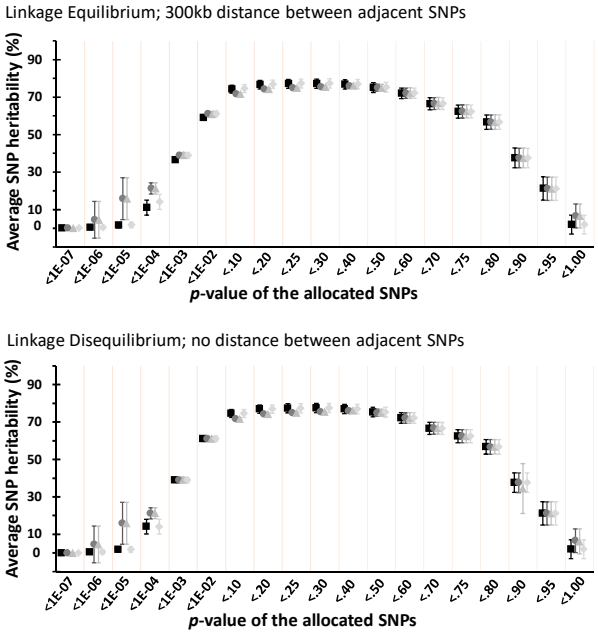
i: 1st-48th associated SNP.



Supplementary Fig. 2: The SNP heritability profile for simulated quantitative phenotypes and SNPs, with predesignated associated SNPs at varying rates. Initial GWAS was performed using EMMAX to determine the association levels for the SNPs. The four different approaches of EMMAX, LDAK-GCTA, LDAK-Thin, and GCTA-GREML (by "inbred" algorithm) were then applied to estimate the SNP heritability for the clusters of SNP hits captured using the association *p*-values. N represents the sample sizes, *i.e.*, number of genotypes. MAF stands for the minor allele frequency. Ensuring additive-only effects, the non-random phenotypes were defined among the genotypic groups of the introduced non-random SNP positions, *i.e.*, Ref*i*/Ref*i* and Alt*i*/Alt*i* for the simulated inbred genotypes. The intra-genotype assignment of the phenotypes, while being within their specific ranges, was random. The inter-genotype variances of the phenotypes are shown as σ_g^2. σ_e^2 represents the intra-genotype variations.

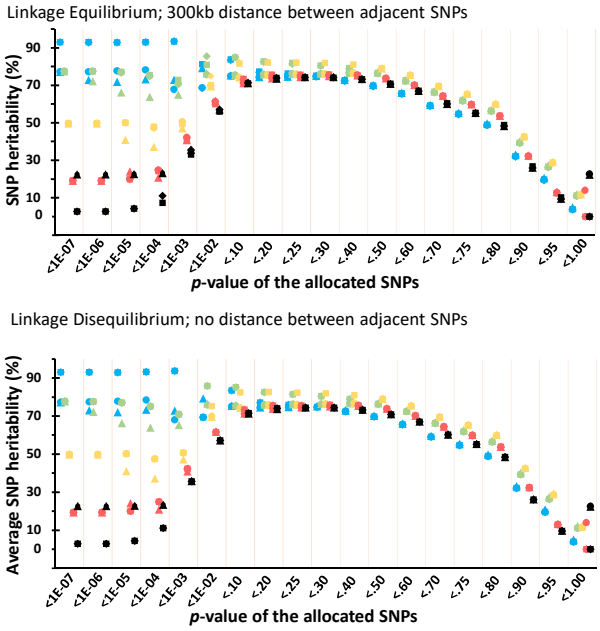
Analysis of 10 artificial random phenotypes on a panel of 1,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Random phenotypes	Range (N)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-100 (1,000)	■	●	▲	◆
2nd	40-100 (1,000)	■	●	▲	◆
3rd	40-100 (1,000)	■	●	▲	◆
4th	40-100 (1,000)	■	●	▲	◆
5th	40-100 (1,000)	■	●	▲	◆
6th	40-100 (1,000)	■	●	▲	◆
7th	40-100 (1,000)	■	●	▲	◆
8th	40-100 (1,000)	■	●	▲	◆
9th	40-100 (1,000)	■	●	▲	◆
10th	40-100 (1,000)	■	●	▲	◆



Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Phenotypes	The introduced non-random SNP position: Ref/Ref Genotype range (N)	Alt/Alt Genotype range (N)	σ_e^2 (%)	σ_g^2 (%)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-50 (900)	80-100 (100)	6.995	93.005	■	●	▲	◆
2nd	40-60 (900)	75-100 (100)	22.389	77.611	■	●	▲	◆
3rd	40-70 (900)	70-100 (100)	50.267	49.733	■	●	▲	◆
4th	40-80 (900)	60-100 (100)	80.690	19.310	■	●	▲	◆
5th	40-90 (900)	50-100 (100)	97.170	2.830	■	●	▲	◆

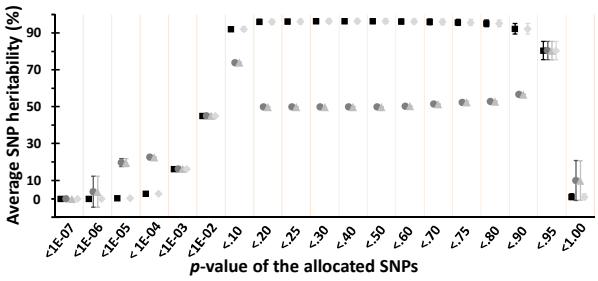


Supplementary Fig. 3 continued.

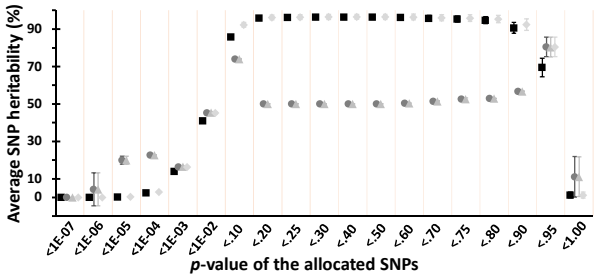
Analysis of 10 artificial random phenotypes on a panel of 10,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Random phenotypes	Range (N)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-100 (10,000)	■	●	▲	◆
2nd	40-100 (10,000)	■	●	▲	◆
3rd	40-100 (10,000)	■	●	▲	◆
4th	40-100 (10,000)	■	●	▲	◆
5th	40-100 (10,000)	■	●	▲	◆
6th	40-100 (10,000)	■	●	▲	◆
7th	40-100 (10,000)	■	●	▲	◆
8th	40-100 (10,000)	■	●	▲	◆
9th	40-100 (10,000)	■	●	▲	◆
10th	40-100 (10,000)	■	●	▲	◆

Linkage Equilibrium; 300kb distance between adjacent SNPs



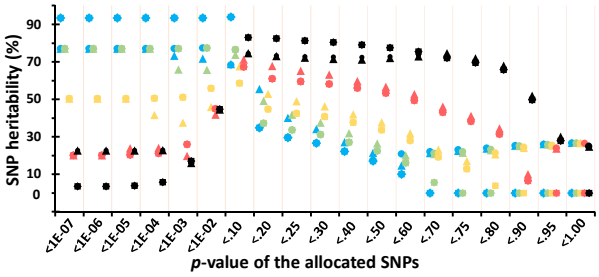
Linkage Disequilibrium; no distance between adjacent SNPs



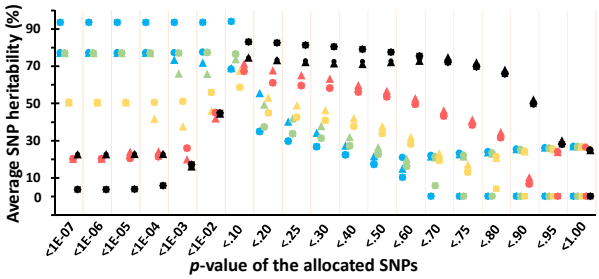
Analysis of 5 artificial non-random phenotypes on a panel of 10,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–10%)

Phenotypes	The introduced non-random SNP position: Ref/Ref Genotype range (N)	Alt/Alt Genotype range (N)	σ_e^2 (%)	σ_g^2 (%)	EMMAX	LDAK_GCTA	LDAK_Thin	GCTA_GREML
1st	40-50 (9,000)	80-100 (1,000)	6.605	93.395	■	●	▲	◆
2nd	40-60 (9,000)	75-100 (1,000)	22.738	77.262	■	●	▲	◆
3rd	40-70 (9,000)	70-100 (1,000)	48.937	51.063	■	●	▲	◆
4th	40-80 (9,000)	60-100 (1,000)	79.675	20.325	■	●	▲	◆
5th	40-90 (9,000)	50-100 (1,000)	96.309	3.691	■	●	▲	◆

Linkage Equilibrium; 300kb distance between adjacent SNPs



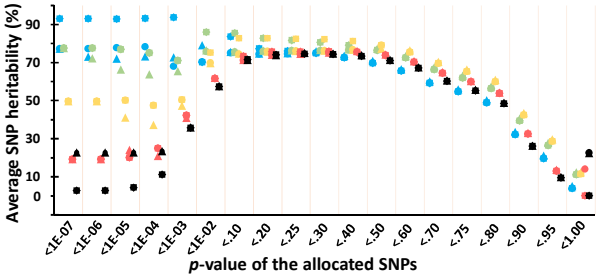
Linkage Disequilibrium; no distance between adjacent SNPs



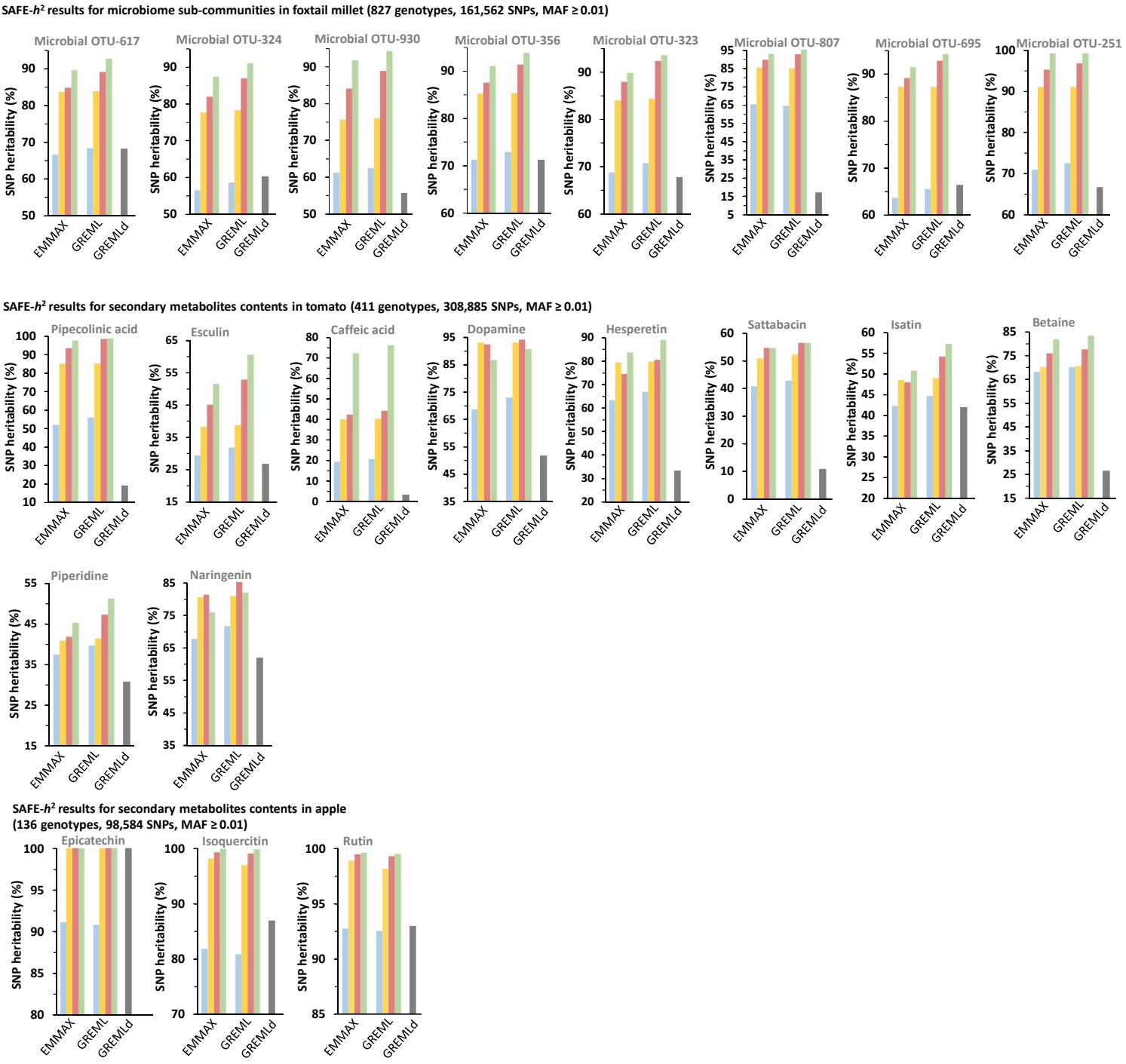
Supplementary Fig. 3: The SNP heritability profiles for simulated random and non-random quantitative phenotypes and SNPs, with a comparison of linkage disequilibrium and linkage equilibrium. The analysis compares linkage disequilibrium and linkage equilibrium. Initial GWAS was performed using EMMAX to realize association levels for the SNPs. The four different approaches of EMMAX, LDAKGCTA, LDAK-Thin, and GCTA-GREML were then applied to estimate SNP heritability for the clusters of SNP hits captured using the association p-values. N represents the sample sizes, i.e., number of genotypes. MAF stands for the minor allele frequency. The average SNP heritability estimates were obtained for 10 independent random phenotypes within identical ranges. The error bars represent 99% confidence intervals. The SNP heritability estimates were obtained for five artificial non-random phenotypes which were associated with a simulated non-random SNP position introduced within the simulated genotypes. By ensuring additive-only effects, the non-random phenotypes were defined at different quantitative ranges among the genotypic groups of the introduced non-random SNP position, i.e., Ref/Ref and Alt/Alt for the simulated inbred genotypes. The intragenotype assignment of the phenotypes, while being within their specific ranges, was still random. The inter-genotype variances of the phenotypes are shown as σ_g^2. σ_e^2 represents the intra-genotype variations. The analyses of inbred genotypes using GCTA-GREML were performed based on the "inbred" algorithm of GCTA.

Analysis of 5 artificial non-random phenotypes on a panel of 1,000 simulated inbred genotypes (200,000 SNPs; MAF = 1.6%–50%)

Phenotypes	The introduced non-random SNP position				
	Ref/Ref Genotype	Alt/Alt Genotype			EMMAX / LDAK_GCTA / LDAK_Thin / GCTA_GREML
	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	
1st	40-50 (900)	80-100 (100)	6.995	93.005	■ ● ▲ ◆
2nd	40-60 (900)	75-100 (100)	22.389	77.611	■ ● ▲ ◆
3rd	40-70 (900)	70-100 (100)	50.267	49.733	■ ● ▲ ◆
4th	40-80 (900)	60-100 (100)	80.690	19.310	■ ● ▲ ◆
5th	40-90 (900)	50-100 (100)	97.170	2.830	■ ● ▲ ◆



Supplementary Fig. 4: The SNP heritability profiles for simulated random and non-random quantitative phenotypes and SNPs with minor allele frequencies ranging from 1.6%–50%. Initial GWAS was performed using EMMAX to realize association levels for the SNPs. The four different approaches of EMMAX, LDAKGCTA, LDAK-Thin, and GCTA-GREML were then applied to estimate SNP heritability for the clusters of SNP hits captured using the association *p*-values. N represents the sample sizes, i.e., number of genotypes. MAF stands for the minor allele frequency. The SNP heritability estimates were obtained for five artificial non-random phenotypes which were associated with a simulated non-random SNP position introduced within the simulated genotypes. By ensuring additive-only effects, the non-random phenotypes were defined at different quantitative ranges among the genotypic groups of the introduced non-random SNP position, i.e., Ref/Ref and Alt/Alt for the simulated inbred genotypes. The intragenotype assignment of the phenotypes, while being within their specific ranges, was still random. The inter-genotype variances of the phenotypes are shown as σ_g^2. σ_e^2 represents the intra-genotype variations. The analyses of inbred genotypes using GCTA-GREML were performed based on the "inbred" algorithm of GCTA.



Supplementary Fig. 5: SAFE-h^2 estimates of SNP heritability using additive-only model and the combined additive and non-additive models, compared to GREMLd estimates. The analyses of tomato and foxtail millet genotypes using GCTA-GREML model were performed based on the "inbred" algorithm of GCTA. The average inbreeding coefficients (Fhat) of tomato and foxtail millet genotypes were > 0.8. Add: heritability based on additive effects, AddDom: heritability based on additive and dominance effects, AddDomOD: heritability based on additive, dominance and overdominance effects, AddDomODHet: heritability based on additive, dominance, overdominance, and heterosis-like effects, GREMLd: heritability estimated by GREMLd model, MAF: minor allele frequency, OUT: operational taxonomic unit.

Supplementary Table 1. Heritability profiling by SAFE-h^2 reveals negative contributions of SNPs with large association *p*-values

	Phenotypes	Number of Genotypes/Variants	[SAFE-h^2_{SNP}] - [h^2_{SNP} based on all SNPs]	
			EMMAX	GCTA-GREML[a]
Tomato	Cracked fruit	96 / 16,782	3.95%	2.37%
	Day-to-flowering	96 / 16,782	0.00%	0.00%
	Small fruit	96 / 16,782	17.96%	18.66%
	Iregular shape	96 / 16,782	2.71%	2.52%
	Caffeic acid	411 / 308,885	12.16%	12.80%
	Catechin	411 / 308,885	22.25%	21.92%
	Dopamine	411 / 308,885	68.64%	72.95%
	Esculin	411 / 308,885	19.53%	21.04%
	Hesperetin	411 / 308,885	61.11%	66.61%
	Naringenin	411 / 308,885	51.78%	54.29%
	Piperidine	411 / 308,885	8.53%	8.64%
	Viscumamide	411 / 308,885	2.83%	2.82%
	Betaine	411 / 308,885	49.64%	50.07%
	Isatin	411 / 308,885	1.42%	1.28%
	Pipecolinic acid	411 / 308,885	52.12%	56.12%
	Riboflavin	411 / 308,885	48.99%	50.73%
	Sattabacin	411 / 308,885	26.69%	27.50%
	Quercetin	411 / 308,885	62.37%	67.00%
Apple	Quercetin	136 / 98,584	30.78%	32.19%
	Catechin	136 / 98,584	13.03%	11.99%
	Avicularin	136 / 98,584	72.53%	71.99%
	Phloridzin	136 / 98,584	53.78%	52.99%
	Rutin	136 / 98,584	90.21%	90.07%
	Isoquercetin	136 / 98,584	46.78%	46.85%
	Epicatechin	136 / 98,584	60.93%	61.52%
	Fruit-weight	368 / 3,055,912	2.10%	2.21%
	Sucrose	361 / 2,927,501	20.20%	20.60%
	Fructose	448 / 2,682,418	15.52%	15.95%
	Tartrate	461 / 2,767,199	19.39%	10.73%
Foxtail millet	OUT-327	827 / 161,562	29.10%	28.84%
	OUT-403	827 / 161,562	44.80%	44.31%
	OUT-487	827 / 161,562	15.17%	14.61%
	OUT-564	827 / 161,562	19.23%	19.09%
	OUT-323	827 / 161,562	49.12%	49.73%
	OUT-356	827 / 161,562	41.99%	41.79%
	OUT-930	827 / 161,562	35.71%	35.13%
	OUT-324	827 / 161,562	29.64%	29.98%
	OUT-617	827 / 161,562	26.30%	25.78%
	OUT-807	827 / 161,562	53.93%	52.17%
	OUT-695	827 / 161,562	39.36%	39.63%
	OUT-251	827 / 161,562	49.04%	49.31%
Cats	Dibetes mellitus	1,022 / 826	12.07%	11.39%
Goats	Height	752 / 40,639	2.32%	2.37%
	Body-length	752 / 40,639	1.76%	1.85%
	Chest length	752 / 40,639	0.23%	0.23%
	Heart-girth	752 / 40,639	2.60%	2.75%
	Pubic bone-length	752 / 40,639	0.00%	0.03%
	Weight	752 / 40,639	4.86%	5.03%
Wild Avians	Clutch-size	314 / 40,766	52.92%	57.16%
Broiler lines	Lean vs Fat	475 / 47,402	0.00%	---
Average			29.27%	29.91%

[a]The "inbred" algorithm of GCTA was used to analyze the tomato and foxtail millet genotypes. The average inbreeding coefficients (Fhat) for these gentypes were > 0.8.

Supplementary Table 2: The SNP heritability estimates for simulated quantitative phenotypes and 200K SNPs in LE and LD.

	The introduced non-random (associated) SNP position				On a 200kb chromosome (Linkage Disequilibrium)			300kb distance between two adjacent SNPs (Linkage Equilibrium)		
	Ref/Ref Genotype	Alt/Alt Genotype			BayesR			BayesR		Process time (BayesR/SAFE-h^2)
Phenotypes	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	h^2 (%)	Process time	Process time (BayesR/SAFE-h^2)	h^2 (%)	Process time	
1st	40-50 (900)	80-100 (100)	6.995	93.005	na	150 min	10.7	na	150 min	10.7
2nd	40-60 (900)	75-100 (100)	22.389	77.611	na	150 min	10.7	na	150 min	10.7
3rd	40-70 (900)	70-100 (100)	50.267	49.733	na	150 min	10.7	34.67	150 min	10.7
4th	40-80 (900)	60-100 (100)	80.69	19.31	6.04	150 min	10.7	8.33	150 min	10.7
5th	40-90 (900)	50-100 (100)	97.17	2.83	3.47	150 min	10.7	1.07	150 min	10.7

Ref/Ref and Alt/Alt for the introduced associated single SNP

σ_g^2: Inter-genotype variances of the phenotypes

σ_e^2: Intra-genotype variations

N: number of genotypes (individuals)

Supplementary Table 3: The SNP heritability estimates for simulated quantitative phenotypes and the chr. 22 SNPs from human HapMap phase 1 genotypes.

	The introduced non-random (associated) SNP position					Real genotypes based on Chr22 + introduced single associated SNP		
	Ref/Ref Genotype	Ref/Alt Genotype	Alt/Alt Genotype			BayesR		Process time (BayesR/SAFE-h^2)
Phenotypes	range (N)	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	h^2 (%)	Process time	
1st	40-50 (982)	80-100 (103)	120-150 (7)	5.704	94.296	na	120 min	8.6
2nd	40-60 (982)	75-100 (103)	112-150 (7)	21.037	78.963	na	120 min	8.6
3rd	40-70 (982)	70-100 (103)	105-150 (7)	41.501	58.499	na	120 min	8.6
4th	40-80 (982)	60-100 (103)	90-150 (7)	72.099	27.901	22.21	120 min	8.6
5th	40-90 (982)	50-100 (103)	75-150 (7)	91.416	8.584	6.17	120 min	8.6

Ref/Ref, Ref/Alt, and Alt/Alt for the introduced associated single SNP

σ_g^2: Inter-genotype variances of the phenotypes

σ_e^2: Intra-genotype variations

N: number of genotypes (individuals)

Supplementary Table 4: The SNP heritability estimates for real phenotypes using BayesR and SAFE-h^2.

		Heritability estimates (h^2 %)		
	Phenotype	BayesR	SAFE-h^2 by EMMAX	SAFE-h^2 by GCTA_GREML
Foxtail millet				
827 genotypes & 161k SNPs	OUT_25	27.84	71.73	72.05
827 genotypes & 161k SNPs	OUT_312	9.47	54.05	54.01
Tomato				
411 genotypes & 308k SNPs	Betaine	17.4	68.08	70.02
411 genotypes & 308k SNPs	Hesperetin	7.37	62.85	66.41
Arabidopsis				
345 genotypes & 2.8M SNPs	Root-length	6.3	71.5	71.4
936 genotypes & 2.8M SNPs	FBE	69.41	93.1	93.1
Apple				
136 genotypes & 98k SNPs	Avicularin	24.2	72.63	77.08
136 genotypes & 98k SNPs	Catechin	35.48	55.73	53.61

FBE stands for flowering bud emergence.

Supplementary Table 5. SNP heritability estimations on additive or additive + non-additive effects of a biallelic locus.

Phenotypes						SAFE-h^2 estimates of SNP-heritability							
	Ref/Ref Genotype	Ref/Alt Genotype	Alt/Alt Genotype			using EMMAX model				using GEMMA model			
Data_Genetic effect	range (N)	range (N)	range (N)	σ_e^2 (%)	σ_g^2 (%)	A[a]	AD[b]	ADO[c]	ADOH[d]	A[a]	AD[b]	ADO[c]	ADOH[d]
1st_Additice	40-50 (982)	80-100 (103)	120-150 (7)	5.7	94.3	94.29	94.29	94.29	94.29	94.28	94.28	94.28	94.28
1st_Dominance	40-50 (982)	120-150 (103)	120-150 (7)	2.16	97.84	92.8	▲ 97.83	▲ 97.83	▲ 97.83	92.8	▲ 97.83	▲ 97.83	▲ 97.83
1st_Overdominance	40-50 (982)	120-150 (103)	80-100 (7)	2.35	97.65	84.77	▲ 95.9	▲ 97.65	▲ 97.65	84.76	▲ 95.9	▲ 97.64	▲ 97.64
1st_Heterosis-like	40-50 (982)	120-150 (103)	40-50 (7)	6.11	93.89	70.51	▲ 86.47	▲ 91.72	▲ 93.86	70.49	▲ 86.46	▲ 91.71	▲ 93.86
2nd_Additive	40-60 (982)	75-100 (103)	112-150 (7)	21.04	78.96	78.86	78.86	78.86	78.86	78.85	78.85	78.85	78.85
2nd_Dominance	40-60 (982)	112-150 (103)	112-150 (7)	12.75	87.25	82.1	▲ 87.23	▲ 87.23	▲ 87.23	82.1	▲ 87.22	▲ 87.22	▲ 87.22
2nd_Overdominance	40-60 (982)	112-150 (103)	75-100 (7)	14.7	85.3	71.58	▲ 82.7	▲ 85.13	▲ 85.13	71.56	▲ 82.69	▲ 85.12	▲ 85.12
2nd_Heterosis-like	40-60 (982)	112-150 (103)	40-60 (7)	8.1	91.9	70.56	▲ 85.58	▲ 90.28	▲ 91.89	70.54	▲ 85.57	▲ 90.27	91.89
3rd_Additive	40-70 (982)	70-100 (103)	105-150 (7)	41.5	58.5	58.34	58.34	58.34	58.34	58.31	58.31	58.31	58.31
3rd_Dominance	40-70 (982)	105-150 (103)	105-150 (7)	15.47	84.53	79.48	▲ 84.51	▲ 84.51	▲ 84.51	79.47	▲ 84.5	▲ 84.5	▲ 84.5
3rd_Overdominance	40-70 (982)	105-150 (103)	70-100 (7)	15.16	84.84	72.05	▲ 82.65	▲ 84.76	▲ 84.76	72.04	▲ 82.64	▲ 84.74	▲ 84.74
3rd_Heterosis-like	40-70 (982)	105-150 (103)	40-70 (7)	16.44	83.56	62.59	▲ 76.85	▲ 81.56	▲ 83.52	62.57	▲ 76.84	▲ 81.55	▲ 83.5
4th_Additive	40-80 (982)	60-100 (103)	90-150 (7)	72.1	27.9	27.34	27.34	27.34	27.34	27.33	27.33	27.33	27.33
4th_Dominance	40-80 (982)	90-150 (103)	90-150 (7)	33.36	66.64	62.6	▲ 66.6	▲ 66.6	▲ 66.6	62.58	▲ 66.59	▲ 66.59	▲ 66.59
4th_Overdominance	40-80 (982)	90-150 (103)	60-100 (7)	32.11	67.89	56.78	▲ 65.73	▲ 67.72	▲ 67.72	56.76	▲ 65.71	▲ 67.7	▲ 67.7
4th_Heterosis	40-80 (982)	90-150 (103)	40-80 (7)	35.4	64.6	51.1	▲ 60.96	▲ 63.84	▲ 64.5	50.98	▲ 60.94	▲ 63.82	▲ 64.49
5th_Additive	40-90 (982)	50-100 (103)	75-150 (7)	91.42	8.58	7.7	7.7	7.7	7.7	7.69	7.69	7.69	7.69
5th_Dominance	40-90 (982)	75-150 (103)	75-150 (7)	61.67	38.33	36.61	▲ 38.24	▲ 38.24	▲ 38.24	36.59	▲ 38.22	▲ 38.22	▲ 38.22
5th_Overdominance	40-90 (982)	75-150 (103)	50-100 (7)	55.12	44.88	34.61	▲ 41.87	▲ 44.1	▲ 44.82	34.59	▲ 41.84	▲ 44.08	▲ 44.8
5th_Heterosis-like	40-90 (982)	75-150 (103)	40-90 (7)	56	44	34.25	▲ 41.23	▲ 43.33	▲ 43.93	34.23	▲ 41.21	▲ 43.3	▲ 43.91
6th_Additive	40-95 (982)	45-100 (103)	50-150 (7)	97.56	2.44	2.3	2.3	2.3	2.3	2.29	2.29	2.29	2.29
6th_Dominance	40-95 (982)	50-150 (103)	50-150 (7)	76.6	23.4	20.67	▲ 23.03	▲ 23.31	▲ 23.31	20.65	▲ 23.05	▲ 23.29	▲ 23.29
6th_Overdominance	40-95 (982)	50-150 (103)	45-100 (7)	76.64	23.36	19.5	▲ 22.56	▲ 23.24	▲ 23.24	19.49	▲ 22.55	▲ 23.22	▲ 23.22
6th_Heterosis-like	40-95 (982)	50-150 (103)	40-95 (7)	76.57	23.43	16.19	▲ 20.62	▲ 22.57	▲ 23.21	16.17	▲ 20.61	▲ 22.56	▲ 23.19

Green triangles indicate improved estimation when compared to the analysis on the original genotypes (a). Red triangles indicate underestimations when compared to the analysis on the original genotypes (a). σ_e^2 & σ_g^2 represent the the enviromental and genetic variances, respectively. σ_g^2 is equivalent to the broad-sense heritability. [a]On original genotypes to capture additive effects; [b]On adjusted genotypes for dominance effects to capture additive and dominance effects; [c]On adjusted genotypes for dominance and overdominance effects to capture additive, dominance, and overdominace effects; [d]On adjusted genotypes for dominance,overdominance, and heterosis-like effects to capture additive, dominance, overdominace, and heterosis-like effects; [e]Overestimated.

Supplementary Table 6. Heritability estimates based on *p*-values, calculated using EMMAX and GEMMA.

		Heritability estimates (h^2 %)											
		Association analysis using EMMAX (*p*-value source)						Association analysis using GEMMA (*p*-value source)					
	Phenotype	EMMAX (on all SNPs)	SAFE-h^2 using EMMAX	GCTA_GREML (on all SNPs)	SAFE-h^2 using GCTA_GREML	GEMMA (on all SNPs)	SAFE-h^2 using GEMMA	EMMAX (on all SNPs)	SAFE-h^2 using EMMAX	GCTA_GREML (on all SNPs)	SAFE-h^2 using GCTA_GREML	GEMMA (on all SNPs)	SAFE-h^2 using GEMMA
Foxtail millet													
827 genotypes & 161k SNPs	OUT_25	32.42	71.73	34.26	72.05	32.94	74.41	32.43	94.97	34.26	95.17	32.94	97.97
827 genotypes & 161k SNPs	OUT_312	17.48	54.05	18.68	54.01	17.86	57.37	17.49	85.01	18.68	86.02	17.86	88.567
Tomato													
411 genotypes & 308k SNPs	Betaine	18.43	68.08	19.99	70.02	18.02	67.35	18.43	73.23	19.99	74.85	18.02	74.46
411 genotypes & 308k SNPs	Hesperetin	0.21	63.39	0.23	66.84	0.26	62.5	0.21	62.04	0.23	65.58	0.26	61.18
Arabidopsis													
345 genotypes & 2.8M SNPs	Root-length	0.01	71.47	0.01	71.42	0	71.41	0.01	71.58	0	71.52	0	71.52
936 genotypes & 2.8M SNPs	FBE	51.03	93.14	51	93.13	51	93.13	51.03	93.58	51	93.58	51	93.58
Apple													
136 genotypes & 98k SNPs	Avicularin	0.09	72.63	0.09	72.08	0.01	80.34	0.09	79.74	0.09	79.26	1.47	86.54
136 genotypes & 98k SNPs	Catechin	42.69	55.73	41.61	53.61	45.05	58.88	42.69	95.77	41.61	95.65	45.05	97.09