

ABSTRACT: We introduce software, Numericware i, to compute identical by state (IBS) matrix based on genotypic data. Calculating an IBS matrix with a large dataset requires large computer memory and takes lengthy processing time. Numericware i addresses these challenges with 2 algorithmic methods: multithreading and forward chopping. The multithreading allows computational routines to concurrently run on multiple central processing unit (CPU) processors. The forward chopping addresses memory limitation by dividing a dataset into appropriately sized subsets. Numericware i allows calculation of the IBS matrix for a large genotypic dataset using a laptop or a desktop computer. For comparison with different software, we calculated genetic relationship matrices using Numericware i, SPAGeDi, and TASSEL with the same genotypic dataset. Numericware i calculates IBS coefficients between 0 and 2, whereas SPAGeDi and TASSEL produce different ranges of values including negative values. The Pearson correlation coefficient between the matrices from Numericware i and TASSEL was high at .9972, whereas SPAGeDi showed low correlation with Numericware i (.0505) and TASSEL (.0587). With a high-dimensional dataset of 500 entities by 10 000 000 SNPs, Numericware i spent 382 minutes using 19 CPU threads and 64 GB memory by dividing the dataset into 3 pieces, whereas SPAGeDi and TASSEL failed with the same dataset. Numericware i is freely available for Windows and Linux under CC-BY 4.0 license at <https://figshare.com/s/f100f33a8857131eb2db>.

KEYWORDS: Identical by state matrix, Numericware i, Genetic relationship matrix, Forward chopping, Multithreading

RECEIVED: October 5, 2016. **ACCEPTED:** December 7, 2016.

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totaled 1357 words, excluding any confidential comments to the academic editor.

TYPE: Software or Database Review

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding was provided by the North Central Soybean Research Program and by the Department of Agronomy at Iowa State University.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: William D Beavis, Department of Agronomy, Iowa State University, Ames, IA 50011, USA. Email: wdbbeavis@iastate.edu

Background

The inbreeding, identical by descent (synonymous to IBD, kinship, and coancestry), and identical by state (IBS) coefficients are central parameters in population genetics.¹ By definition, (1) the inbreeding coefficient refers to a proportion that a pair of alleles in an entity is identical in origin and state,² (2) the IBD coefficient between 2 entities equals twice the inbreeding coefficient for their virtual offspring,³ and (3) the IBS coefficient between 2 entities equals twice a proportion that a pair of alleles in their virtual offspring is identical in state. The IBD matrix is a conventional indicator to represent genetic relationship among entities in a population, for which pedigrees are available. Emik and Terrill³ suggested a systematic method for calculating a numerator relationship matrix (NRM) that displays IBD coefficients among every pair of entities in a population. Because the NRM is based on pedigrees, it represents genetic relationship from the genealogical perspective.

High-throughput genotyping technologies provide abundant DNA profile that is useful to calculate the IBS matrix as a genetic relationship matrix. Some references^{4,5} introduced a method for computing the IBS matrix. Although the concept about the IBS matrix is general and simple, the IBS matrix is not widely used. Presumably, it might be due to the notoriously heavy computing burden. In this paper, we present

software referred to as Numericware i. In order to deal with heavy workload, Numericware i supports parallelization and data management to avoid low memory.

Implementation

IBS coefficient

The IBS coefficients^{4,5} can be calculated:

$$IBS_{A,B} = \frac{1}{2} \left(\begin{array}{l} P(a1 \equiv b1) + P(a1 \equiv b2) \\ + P(a2 \equiv b1) + P(a2 \equiv b2) \end{array} \right) \quad (1)$$

where $IBS_{A,B}$ = the IBS coefficient between A and B; $a1, a2$ = a pair of chromosomes for A; $b1, b2$ = a pair of chromosomes for B; $P(a1 \equiv b1)$ = the probability that $a1$ and $b1$ are homozygous; $P(a1 \equiv b2)$ = the probability that $a1$ and $b2$ are homozygous; $P(a2 \equiv b1)$ = the probability that $a2$ and $b1$ are homozygous; and $P(a2 \equiv b2)$ = the probability that $a2$ and $b2$ are homozygous.

The IBS coefficient for parents equals twice the homozygote coefficient (H) for their offspring. Thus, the H can be calculated:

$$H_C = \frac{1}{4} \left(\begin{array}{l} P(a1 \equiv b1) + P(a1 \equiv b2) \\ + P(a2 \equiv b1) + P(a2 \equiv b2) \end{array} \right) \quad (2)$$



where C = the offspring of A and B ; H_C = the homozygote coefficient for C ; a_1 , a_2 = a pair of chromosomes for A ; b_1 , b_2 = a pair of chromosomes for B ; $P(a_1 \equiv b_1)$ = the probability that a_1 and b_1 are homozygous; $P(a_1 \equiv b_2)$ = the probability that a_1 and b_2 are homozygous; $P(a_2 \equiv b_1)$ = the probability that a_2 and b_1 are homozygous; and $P(a_2 \equiv b_2)$ = the probability that a_2 and b_2 are homozygous.

As the cost for producing genotypic data is becoming less expensive, the dimensions of genotypic datasets are rapidly growing. The amount of computing workload can be represented:

$$w = n^2 m \quad (3)$$

where w = the amount of computational workload; n = the number of entities in a population; m = the number of markers.

According to equation 3, the growing dimension of genotypic dataset causes 2 computational challenges: (1) lengthy computational time and (2) low memory.

Functionality of Numericware i

Numericware i, written in C++, has a simple user interface. The software provides 2 special functionalities:

multithreading and forward chopping. The multithreading enables the computer to distribute the workload into multiple CPU threads. The forward chopping chops a dataset into multiple pieces that will not overextend memory capacity. Algorithm 1 shows the forward chopping algorithm.

Numericware i provides users with more conveniences:

1. Imputation not needed: The IBS computation is counting based. Numericware i skips counting missing genotypic data, assuming that the remaining genotypic data are of sufficiently large amount.
2. IBS computation for a haplotype: This allows computation of the IBS matrix for a partial genomic block.
3. Dataset integrity checking: This helps prevent a failure in the middle of analysis by checking the integrity of dataset at the beginning of work.
4. Dataset summary: This provides users with overview of genotypic data.
5. Supporting multiple types of datasets: This significantly reduces extra works for formatting the dataset. Numericware i accepts: alphanumeric, a pair of single-nucleotide polymorphisms (SNPs) and International Union of Pure and Applied Chemistry (IUPAC) formats.

Algorithm 1. Forward Chopping algorithm.

```

1: start_point=0
2: for (j=1; j <= num_pieces; j++){ // num_pieces = the total number of chopped pieces
3:   if (j <= width % num_pieces) { // width = the total number of columns
4:     chopped_width = ceil(width / num_pieces) // chopped_width = the width of a chopped piece
5:   } else{
6:     chopped_width = floor(width / num_pieces)
7:   }
8:   start_point = start_point + 1 // start_point = the first column coordinate of a chopped dataset
9:   end_point = start_point + chopped_width - 1 // end_point = the last column coordinate of a chopped dataset
10:  for (string line; getline(data, line)) {
11:    count = 1
12:    while (getline(line, temp, ",")) {
13:      if (count >= start_point) {
14:        row.push_back(temp)
15:      }
16:      if (count == end_point) { break }
17:      count++
18:    }
19:    table.push_back(row)
20:    row.clear()
21:  }
22:  //////////////////////////////////////
23:  IBS matrix computation with 'table'
24:  //////////////////////////////////////
25:  table.clear()
26:  start_point = end_point
27: }

```

6. Transposing the dataset.

Regarding item 5, details about dataset formats are described in the user manual, and example datasets are included in the software package. Numericware *i* completes the IBS matrix by copying the upper diagonals to the lower diagonals based on its symmetric property to reduce workload.

Application of the IBS matrix

Homozygote coefficient index. A diagonal value for an entity A ($IBS_{A,A}$) in the IBS matrix implies the H_A as $IBS_{A,A} = 1 + H_A$, and $IBS_{A,B}$ represents twice the H_C , in which C = the offspring of A and B. These principles can be useful in controlling homozygote level of progenies in a breeding program.

Best linear unbiased prediction. As a statistical model, the best linear unbiased prediction (BLUP)⁶ is widely used to estimate breeding values. The BLUP requires the genetic relationship matrix, for which Henderson⁶ suggested using the NRM. The IBS matrix is superior to the NRM in the following aspects. First, the IBS matrix fully has values greater than 0, whereas the NRM includes an identity matrix for a base population. The identity matrix results in underestimation of IBD coefficients within the NRM. Second, the IBS matrix provides an objective measure based on genotypic data whereas the NRM is based on statistical expectation. As the BLUP expands to genome-wide association study and genomic prediction, the IBS matrix can apply to these studies.⁷⁻¹¹ The IBS matrix will be useful especially for plants since plant pedigrees are often unknown and imprecise.¹² Previous studies reported that the BLUP outperforms with genome-based genetic relationship matrices than the NRM.¹³⁻¹⁸ In this context, the IBS matrix will be useful in improving the BLUP accuracy.

Results and Discussion

Negative effect of marker screen to IBS matrix

Marker collections generally consist of markers screened based on allele frequency. Marker screening secures allele diversity but causes an ascertainment bias in calculating the IBS matrix because the removed markers must be informative in representing the identical genomic state between entities. Thus, it is recommended to use all markers in calculating the IBS matrix.

Similarity between IBD and IBS

The IBD and IBS coefficients range between 0 and 2 in common. If it is assured that any identical alleles at the same locus from different entities were generated not independently but inherently, the IBS and IBD coefficients should be equal.⁴ We hypothetically assume numerous unique mutations might inherently have flourished genetic diversity because the probability of the same mutations coincidentally occurring on the

same loci of multiple genomes could be extremely low. If this assumption is true, the IBS and IBD matrices can be equal.

Comparison of results from Numericware i, SPAGeDi, and TASSEL

The IBS matrix can be expanded by computing the NRM based on pedigrees of entities within the IBS matrix. The expanded IBS matrix will have overall increased elemental values than the NRM solely based on pedigrees. The expansion of IBS matrix can be calculated using Numericware N.¹⁹ Other popular software, SPAGeDi²⁰ and TASSEL,⁸ implement different algorithms for calculating the genetic relationship matrix, causing their results to have different characteristics, such as negative values (SPAGeDi and TASSEL) or mono-diagonal values (SPAGeDi with 0's). Thus, the resulting matrices from SPAGeDi and TASSEL cannot be expanded by the NRM algorithm. For comparison, we applied Numericware *i*, SPAGeDi, and TASSEL to the same dataset of 20 entities by 2000 SNPs (included in Supplementary file). The algorithms used by SPAGeDi and TASSEL are the ones by Loiselle et al²¹ and Yang et al²², respectively, and widely used for calculating the genetic relationship matrices.^{7,23-26} Tables 1 to 3 are the resulting matrices. Pearson correlation coefficients among the 3 matrices (Table 4) indicate that the results from Numericware *i* and TASSEL are highly correlated at .9972, whereas the result from SPAGeDi shows low correlation coefficients with the results from Numericware *i* (0.0505) and TASSEL (0.0587). This illustrates that the results from Numericware *i* and TASSEL are substantially comparable, whereas SPAGeDi is not.

Performance

In our test, Numericware *i* took 382 minutes in computing an IBS matrix with a simulated dataset of 500 entities by 10 000 000 SNPs using 19 CPU threads (Intel Xeon processor E5-2600 v4) and 64 GB memory. For this test, the whole dataset was chopped into 3 pieces to circumvent the low memory, whereas SPAGeDi and TASSEL failed with the same dataset due to the low memory.

Conclusions

The IBS matrix can be useful as: (1) a foreseeing index about the homozygote coefficients for hybrid lines based on the IBS coefficient for parents being equal to twice the homozygote coefficient for an offspring, (2) an assessment of homozygote coefficient to an entity itself based on $IBS_{A,A}$ being equal to $1 + H_A$, and (3) a component of the BLUP. Thus, Numericware *i* can be an essential tool for breeding. The multithreading and forward chopping reduce computing time and allow processing of extremely large amount of data. In contrast, other software are often limited by the physical memory size, and only a

Table 1. Identical by state matrix calculated using Numericware i.

ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID8	ID9	ID10	ID11	ID12	ID13	ID14	ID15	ID16	ID17	ID18	ID19	ID20	
ID1	1.253	0.48525	0.4935	0.5045	0.493	0.4925	0.4865	0.4985	0.50025	0.521	0.49375	0.47575	0.50775	0.492	0.494	0.5145	0.4865	0.50625	0.50775	0.49675
ID2	0.48525	1.2555	0.50775	0.49275	0.497	0.50625	0.4905	0.505	0.504	0.50525	0.48225	0.4935	0.49375	0.514	0.50925	0.497	0.5015	0.48575	0.49025	0.4795
ID3	0.4935	0.50775	1.2365	0.51025	0.50175	0.49175	0.49175	0.49925	0.49475	0.50675	0.505	0.50325	0.4835	0.49775	0.49875	0.50525	0.502	0.50625	0.49575	0.499
ID4	0.5045	0.49275	0.51025	1.26	0.48075	0.48425	0.5025	0.48775	0.51225	0.50475	0.50225	0.49375	0.51225	0.515	0.4855	0.5035	0.51275	0.5035	0.511	0.48975
ID5	0.493	0.497	0.50175	0.48075	1.2365	0.48075	0.49825	0.50975	0.50325	0.505	0.506	0.507	0.51275	0.491	0.5065	0.5065	0.506	0.512	0.49775	0.51675
ID6	0.4925	0.50625	0.49175	0.48425	0.48075	1.241	0.49725	0.49175	0.49325	0.5085	0.5	0.5065	0.50275	0.509	0.50975	0.49575	0.499	0.49575	0.48775	0.4845
ID7	0.4865	0.4905	0.49175	0.5025	0.49825	0.49725	1.244	0.494	0.50975	0.489	0.48425	0.50675	0.4955	0.48825	0.526	0.50425	0.4985	0.49125	0.496	0.483
ID8	0.4985	0.505	0.49925	0.48775	0.50975	0.49175	0.494	1.2665	0.503	0.5075	0.507	0.50775	0.5015	0.49925	0.5005	0.49125	0.503	0.50425	0.485	0.50625
ID9	0.50025	0.504	0.49475	0.51225	0.50325	0.49325	0.49325	0.503	1.2435	0.50875	0.501	0.49925	0.5035	0.51475	0.504	0.49825	0.50525	0.48325	0.4885	0.49075
ID10	0.521	0.50525	0.50675	0.50475	0.505	0.5085	0.489	0.5075	0.50875	1.25	0.49175	0.50525	0.50625	0.507	0.48725	0.50725	0.49975	0.50375	0.50125	0.4765
ID11	0.49375	0.48225	0.505	0.50225	0.506	0.5	0.48425	0.507	0.501	0.49175	1.2485	0.493	0.5015	0.514	0.50525	0.5025	0.49525	0.51175	0.49775	0.5025
ID12	0.47575	0.4935	0.50325	0.49375	0.507	0.5065	0.50675	0.50775	0.49925	0.50525	0.493	1.229	0.5005	0.4875	0.50475	0.51475	0.4995	0.49875	0.515	0.49325
ID13	0.50775	0.49375	0.4835	0.51225	0.51275	0.50275	0.4955	0.5015	0.5035	0.50625	0.5015	0.5005	1.2445	0.5065	0.50175	0.50725	0.50125	0.5145	0.511	0.484
ID14	0.492	0.514	0.49775	0.515	0.491	0.509	0.48825	0.49925	0.51475	0.507	0.514	0.4875	0.5065	1.25	0.48775	0.501	0.48525	0.51	0.509	0.4595
ID15	0.494	0.50925	0.49875	0.4855	0.5065	0.50975	0.526	0.5005	0.504	0.48725	0.50525	0.50475	0.50175	0.48775	1.2345	0.5135	0.4945	0.51275	0.515	0.50525
ID16	0.5145	0.497	0.50525	0.5035	0.5065	0.49575	0.50425	0.49125	0.49825	0.50725	0.5025	0.51475	0.50725	0.501	0.5135	1.2335	0.48775	0.5015	0.51225	0.5005
ID17	0.4865	0.5015	0.502	0.51275	0.506	0.499	0.4985	0.503	0.50525	0.49975	0.49525	0.4995	0.50125	0.48525	0.4945	0.48775	1.252	0.5155	0.4985	0.51
ID18	0.50625	0.48575	0.50625	0.5035	0.512	0.49575	0.49125	0.50425	0.48325	0.50375	0.51175	0.49875	0.5145	0.51	0.51275	0.5015	0.5155	1.2445	0.5055	0.50175
ID19	0.50775	0.49025	0.49575	0.511	0.49775	0.48775	0.496	0.485	0.4885	0.50125	0.49775	0.515	0.511	0.509	0.515	0.51225	0.4985	0.5055	1.251	0.48925
ID20	0.49675	0.4795	0.499	0.48975	0.51675	0.4845	0.483	0.50625	0.49075	0.4765	0.5025	0.49325	0.484	0.4595	0.50525	0.5005	0.51	0.50175	0.48925	1.244

Table 2. Genetic relationship matrix calculated based on the method of Loiselle et al. (1995) using SPAGeDI.

ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID8	ID9	ID10	ID11	ID12	ID13	ID14	ID15	ID16	ID17	ID18	ID19	ID20	
ID1	0	-0.0068	-0.0024	0.0039	-0.004	-0.0012	-0.0053	-0.0002	0.0012	0.143	-0.0027	-0.0146	0.0053	-0.0037	-0.0043	0.0097	-0.0079	0.0037	0.0062	0.0039
ID2	-0.0068	0	0.0076	-0.0039	-0.001	0.0084	-0.0023	0.0044	0.004	0.0038	-0.0103	-0.0023	-0.0041	0.0116	0.0064	-0.002	0.0026	-0.0101	-0.0055	-0.0077
ID3	-0.0024	0.0076	0	0.0068	0.0011	-0.0027	-0.0027	-0.0007	-0.0035	0.0036	0.0041	0.0032	-0.0122	-0.0007	0.0024	0.0017	0.0028	-0.0029	0.0045	0.0045
ID4	0.0039	-0.0039	0.0068	0	-0.0146	-0.0091	0.0034	-0.0098	0.0072	0.001	0.0009	-0.0046	0.0061	0.0098	-0.0123	0.0078	-0.0004	0.0062	-0.0031	-0.0031
ID5	-0.004	-0.001	0.0011	-0.0146	0	-0.0115	0.0005	0.0052	0.001	0.0012	0.0035	0.0045	0.0065	-0.0066	0.002	0.0032	0.0054	-0.0029	0.0153	0.0153
ID6	-0.0012	0.0084	-0.0027	-0.0091	-0.0115	0	0.0029	-0.004	-0.0027	0.0066	0.0024	0.0072	0.0027	0.0088	0.0074	-0.0023	0.0015	-0.0026	-0.0066	-0.0036
ID7	-0.0053	-0.0023	-0.0027	0.0034	0.0005	0.0029	0	-0.0024	0.0086	-0.0067	-0.0083	0.0074	-0.0022	-0.0053	0.0185	0.0036	0.0011	-0.0057	-0.001	-0.0046
ID8	-0.0002	0.0044	-0.0007	-0.0098	0.0052	-0.004	-0.0024	0	0.0008	0.0029	0.0041	0.005	-0.0012	-0.0009	-0.0021	-0.0084	0.0011	0.0001	-0.0116	0.0081
ID9	0.0012	0.004	-0.0035	0.0072	0.001	0.0027	0.0086	0.0008	0	0.0003	0.0003	-0.0006	0.0004	0.0099	0.0006	-0.0034	0.0029	-0.014	-0.009	-0.0022
ID10	0.0143	0.0038	0.0036	0.001	0.0012	0.0066	-0.0067	0.0029	0.004	0	-0.0071	0.0025	0.0012	0.0035	-0.0119	0.0017	-0.0019	-0.0011	-0.0013	-0.013
ID11	-0.0027	-0.0103	0.0041	0.0009	0.0035	0.0024	-0.0083	0.0041	0.0003	-0.0071	0	-0.0043	-0.0004	0.0099	0.002	0	-0.0034	0.006	-0.0021	0.0064
ID12	-0.0146	-0.0023	0.0032	-0.0046	0.0045	0.0072	0.0074	0.005	-0.0006	0.0025	-0.0043	0	-0.0008	-0.0078	0.002	0.0088	-0.0001	-0.0025	0.0101	0.0004
ID13	0.0053	-0.0041	-0.0122	0.0061	0.0065	0.0027	-0.0022	-0.0012	0.0004	0.0012	-0.0004	-0.0008	0	0.0032	0.0017	-0.0009	0.0063	0.0054	-0.0079	-0.0079
ID14	-0.0037	0.0116	-0.0007	0.0098	-0.0066	0.0088	-0.0053	-0.0009	0.0099	0.0035	0.0099	0.0032	0	-0.0098	-0.0007	-0.01	0.005	0.0058	-0.0228	-0.0228
ID15	-0.0043	0.0064	-0.002	-0.0123	0.002	0.0074	0.0185	-0.0021	0.0006	-0.0119	0.002	0.002	-0.002	0	0.0058	-0.0057	0.0049	0.008	0.0065	0.0065
ID16	0.0097	-0.002	0.0024	-0.0001	0.002	-0.0023	0.0036	-0.0084	-0.0034	0.0017	0	0.0088	0.0017	0.0058	0	-0.0103	-0.0028	0.006	0.0032	0.0032
ID17	-0.0079	0.0026	0.0017	0.0078	0.0032	0.0015	0.0011	0.0011	0.0029	-0.0019	-0.0034	-0.0001	-0.0009	-0.01	-0.0057	-0.0103	0	0.0083	-0.0019	0.0112
ID18	0.0037	-0.0101	0.0028	-0.0004	0.0054	-0.0026	-0.0057	0.0001	-0.014	-0.0011	0.006	-0.0025	0.0063	0.005	0.0049	-0.0028	0.0083	0	0.0011	0.0037
ID19	0.0062	-0.0055	-0.0029	0.0062	-0.0029	-0.0066	-0.001	-0.0116	-0.009	-0.0013	-0.0021	0.0101	0.0054	0.0058	0.006	-0.0019	0.0011	0	-0.0034	-0.0034
ID20	0.0039	-0.0077	0.0045	-0.0031	0.0153	-0.0036	-0.0046	0.0081	-0.0022	-0.013	0.0064	0.0004	-0.0079	-0.0228	0.0065	0.0032	0.0112	0.0037	-0.0034	0

Table 3. Normalized identical by state matrix calculated based on the method of Yang et al. (2011) using TASSEL.

ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID8	ID9	ID10	ID11	ID12	ID13	ID14	ID15	ID16	ID17	ID18	ID19	ID20	
ID1	0.990677	-0.07475	-0.07468	-0.04347	-0.07634	-0.04866	-0.07975	-0.05198	-0.04992	0.002185	-0.03943	-0.09083	-0.02087	-0.01028	-0.07754	-0.03706	-0.0711	-0.05719	-0.04331	-0.04571
ID2	-0.07475	0.999107	-0.02417	-0.05987	-0.0615	-0.054	-0.06972	-0.0388	-0.01471	-0.055	-0.0717	-0.04805	-0.08701	-0.0283	-0.04932	-0.02896	-0.04279	-0.07973	-0.07387	-0.03686
ID3	-0.07468	-0.02417	0.933578	-0.02535	-0.04241	-0.0831	-0.07476	-0.0629	-0.05797	-0.04924	-0.02877	-0.05716	-0.0907	-0.04558	-0.05198	-0.03182	-0.04533	-0.03638	-0.03265	-0.01861
ID4	-0.04347	-0.05987	-0.02535	0.958366	-0.10577	-0.0528	-0.04032	-0.09919	-0.00316	-0.04468	-0.06807	-0.05746	-0.0491	-0.0297	-0.07252	-0.03923	-0.01946	-0.02074	-0.03641	-0.09108
ID5	-0.07634	-0.0615	-0.04241	-0.10577	0.953172	-0.06293	-0.02142	-0.05554	-0.03849	-0.06151	-0.03827	-0.053	-0.02515	-0.06697	-0.03413	-0.07029	-0.04079	-0.03478	-0.06206	-0.00182
ID6	-0.04866	-0.054	-0.0831	-0.0528	-0.06293	0.958312	-0.02279	-0.06428	-0.05163	-0.05079	-0.03247	-0.05604	-0.05783	-0.00742	-0.02654	-0.04114	-0.06854	-0.05117	-0.07728	-0.0489
ID7	-0.07975	-0.06972	-0.07476	-0.04032	-0.02142	-0.02279	0.9916	-0.03756	-0.0466	-0.07431	-0.06282	-0.04842	-0.0444	-0.04276	0.0015	-0.03824	-0.05802	-0.07736	-0.07846	-0.07538
ID8	-0.05198	-0.0388	-0.0629	-0.09919	-0.05554	-0.06428	-0.03756	0.997001	-0.05415	-0.06069	-0.03205	-0.03071	-0.03663	-0.03451	-0.05593	-0.0677	-0.04952	-0.06389	-0.06918	-0.03179
ID9	-0.04992	-0.01471	-0.05797	-0.00316	-0.03849	-0.05163	-0.0466	-0.05415	0.952509	-0.04427	-0.07858	-0.04378	-0.01943	-0.06021	-0.05587	-0.06662	-0.03338	-0.11517	-0.05965	-0.05892
ID10	0.002185	-0.055	-0.04924	-0.04468	-0.06151	-0.05079	-0.07431	-0.06069	-0.04427	0.980709	-0.05307	-0.05985	-0.0829	-0.02231	-0.07237	-0.05284	-0.04877	-0.04125	-0.03556	-0.07347
ID11	-0.03943	-0.0717	-0.02877	-0.06807	-0.03827	-0.03247	-0.06282	-0.03205	-0.07858	-0.05307	0.943432	-0.0346	-0.0497	-0.0504	-0.049	-0.06394	-0.05464	-0.0468	-0.07238	-0.01673
ID12	-0.09083	-0.04805	-0.05716	-0.05746	-0.053	-0.05604	-0.04842	-0.03071	-0.04378	-0.05985	-0.0346	0.955251	-0.07892	-0.0773	-0.03479	-0.03686	-0.04771	-0.06049	-0.01148	-0.02779
ID13	-0.02087	-0.08701	-0.0907	-0.0491	-0.02515	-0.05783	-0.0444	-0.03663	-0.01943	-0.0829	-0.0497	-0.07892	0.988834	-0.03807	-0.05919	-0.06508	-0.04826	-0.03562	-0.02503	-0.05494
ID14	-0.01028	-0.0283	-0.04558	-0.0297	-0.06697	-0.00742	-0.04276	-0.03451	-0.06021	-0.02231	-0.0504	-0.0773	-0.03807	0.931201	-0.09514	-0.02666	-0.07504	-0.05541	-0.06462	-0.10051
ID15	-0.07754	-0.04932	-0.05198	-0.07252	-0.03413	-0.02654	0.0015	-0.05593	-0.05587	-0.07237	-0.049	-0.03479	-0.05919	-0.09514	0.953138	-0.04718	-0.05297	-0.03786	-0.0523	-0.03
ID16	-0.03706	-0.02896	-0.03182	-0.03923	-0.07029	-0.04114	-0.03824	-0.0677	-0.06662	-0.05284	-0.06394	-0.03686	-0.06508	-0.02666	-0.04718	0.945105	-0.0886	-0.03571	-0.03108	-0.07608
ID17	-0.0711	-0.04279	-0.04533	-0.01946	-0.04079	-0.06854	-0.05802	-0.04952	-0.03338	-0.04877	-0.05464	-0.04771	-0.04826	-0.07504	-0.05297	-0.0886	0.988198	-0.05425	-0.06342	-0.02561
ID18	-0.05719	-0.07973	-0.03638	-0.02074	-0.03478	-0.05117	-0.07736	-0.06389	-0.11517	-0.04125	-0.0468	-0.06049	-0.03562	-0.05541	-0.03786	-0.03571	0.975567	-0.02854	-0.04324	-0.04324
ID19	-0.04331	-0.07387	-0.03265	-0.03641	-0.06206	-0.07728	-0.07846	-0.06918	-0.05965	-0.03556	-0.07238	-0.01148	-0.02503	-0.06462	-0.03108	-0.06342	-0.02854	0.979937	-0.06266	-0.06266
ID20	-0.04571	-0.03686	-0.01861	-0.09108	-0.00182	-0.0489	-0.07538	-0.03179	-0.05892	-0.07347	-0.01673	-0.02779	-0.05494	-0.10051	-0.03	-0.07608	-0.02561	-0.04324	-0.06266	0.9201

Table 4. Pearson correlation coefficients among results from Numericware i (Table 1), SPAGeDi (Table 2), and TASSEL (Table 3) for the same dataset.

	NUMERICWARE I	SPAGeDi	TASSEL
Numericware i	1	0.0505	0.9972
SPAGeDi	0.0505	1	0.0587
TASSEL	0.9972	0.0587	1

single CPU is supported. Numericware i is freely available for Windows and Linux under CC-BY 4.0 license and can be downloaded from <https://figshare.com/s/f100f33a8857131eb2db>.

Author Contributions

BK built software. BK and WDB analyzed, discussed, and wrote.

REFERENCES

- Cockerham CC, Weir BS. Variance of actual inbreeding. *Theor Popul Biol.* 1983;23:85–109.
- Wright S. Coefficients of inbreeding and relationship. *Am Nat.* 1922;56:330–338.
- Emik LO, Terrill CE. Systematic procedures for calculating inbreeding coefficients. *J Hered.* 1949;40:51–55.
- Eding H, Meuwissen THE. Marker-based estimates of between and within population kinships for the conservation of genetic diversity. *J Anim Breed Genet.* 2001;118:141–159.
- Bernardo R. *Breeding for Quantitative Traits in Plants.* Woodbury, MN: Stemma Press; 2002.
- Henderson CR. Best linear unbiased estimation and prediction under a selection model. *Biometrics.* 1975;31:423–447.
- Yu J, Pressoir G, Briggs WH, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet.* 2006;38:203–208.
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* 2007;29:2633–2635.
- Lipka AE, Tian F, Wang Q, et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics.* 2012;28:2397–2399.
- Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* 2001;157:1819–1829.
- Endelman JB. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome.* 2011;4:250–255.
- Bauer A, Reets TC, Léon J. Estimation of breeding values of inbred lines using best linear unbiased prediction (BLUP) and genetic similarities. *Crop Sci.* 2006;46:2685–2691.
- Colleau J. An indirect approach to the extensive calculation of relationship coefficients. *Genet Sel Evol.* 2002;34:409–421.
- Habier D, Fernando RL, Dekkers JCM. The impact of genetic relationship information on genome-assisted breeding values. *Genetics.* 2007;177:2389–2397.
- VanRaden PM. Efficient methods to compute genomic prediction. *J Dairy Sci.* 2008;91:4414–4423.
- Legarra A, Aguilar I, Misztal I. A relationship matrix including full pedigree and genomic information. *J Dairy Sci.* 2009;92:4656–4663.
- Endelman JB, Jannink JL. Shrinkage estimation of the realized relationship matrix. *G3 (Bethesda).* 2012;2:1405–1413.
- Müller D, Technow F, Melchinger AE. Shrinkage estimation of the genomic relationship matrix can improve genomic estimated breeding values in the training set. *Theor Appl Genet.* 2015;128:693–703.
- Kim B, Beavis WD, Léon J. Numericware N: numerator relationship matrix calculator. *J Hered.* 2016;107:686–690.
- Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes.* 2002;2:618–620.
- Loiselle BA, Sork VL, Nason J, Graham C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot.* 1995;82:1420–1425.
- Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010;42:565–569.
- Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet.* 2013;45:613–620.
- Slavov GT, Nipper R, Robson P, et al. Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass *Miscanthus sinensis*. *New Phytol.* 2014;201:1227–1239.
- Zhang P, Liu X, Tong H, Lu Y, Li J. Association mapping for important agronomic traits in core collection of rice (*Oryza sativa* L.) with SSR markers. *PLoS ONE.* 2014;9:e111508.
- Sukumaran S, Dreisigacker S, Lopes M, Chavez P, Reynolds MP. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor Appl Genet.* 2015;128:353–363.