

Supplementary data

Estimation of genomic prediction accuracy from reference populations with varying degrees of relationship

S. Hong Lee, Sam Clark, and Julius H.J. van der Werf

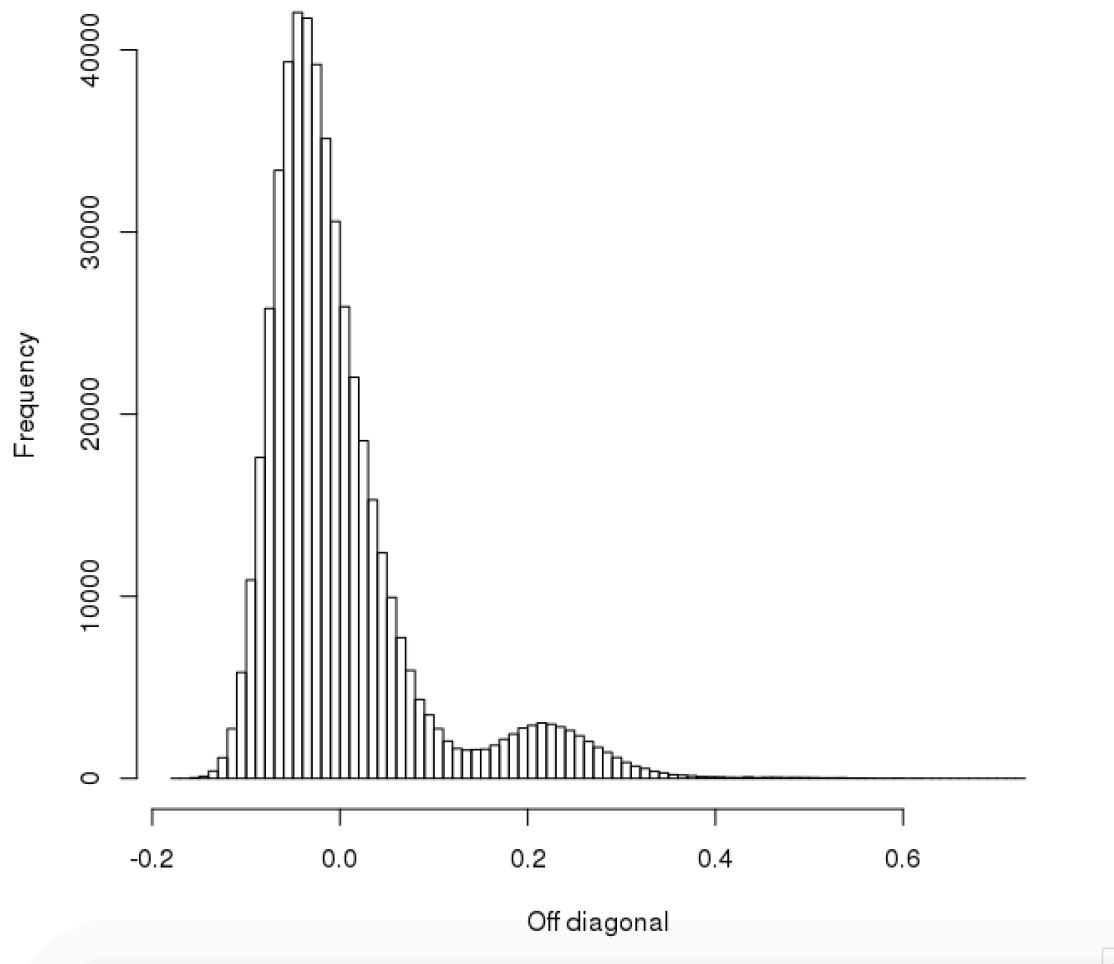


Figure S1. The distribution of off-diagonal of the genomic relationships matrix among 1000 individuals when effective population size is 50.

Although a substantial proportion includes close relationships with an effective population size of 50, the mean and variance of the genomic relationships is -0.001 and 0.0072, respectively, which agrees with the expected value from the theory (0 and 0.0077). This indicates that Eq. (5) is valid with any random sample even having close relationship.

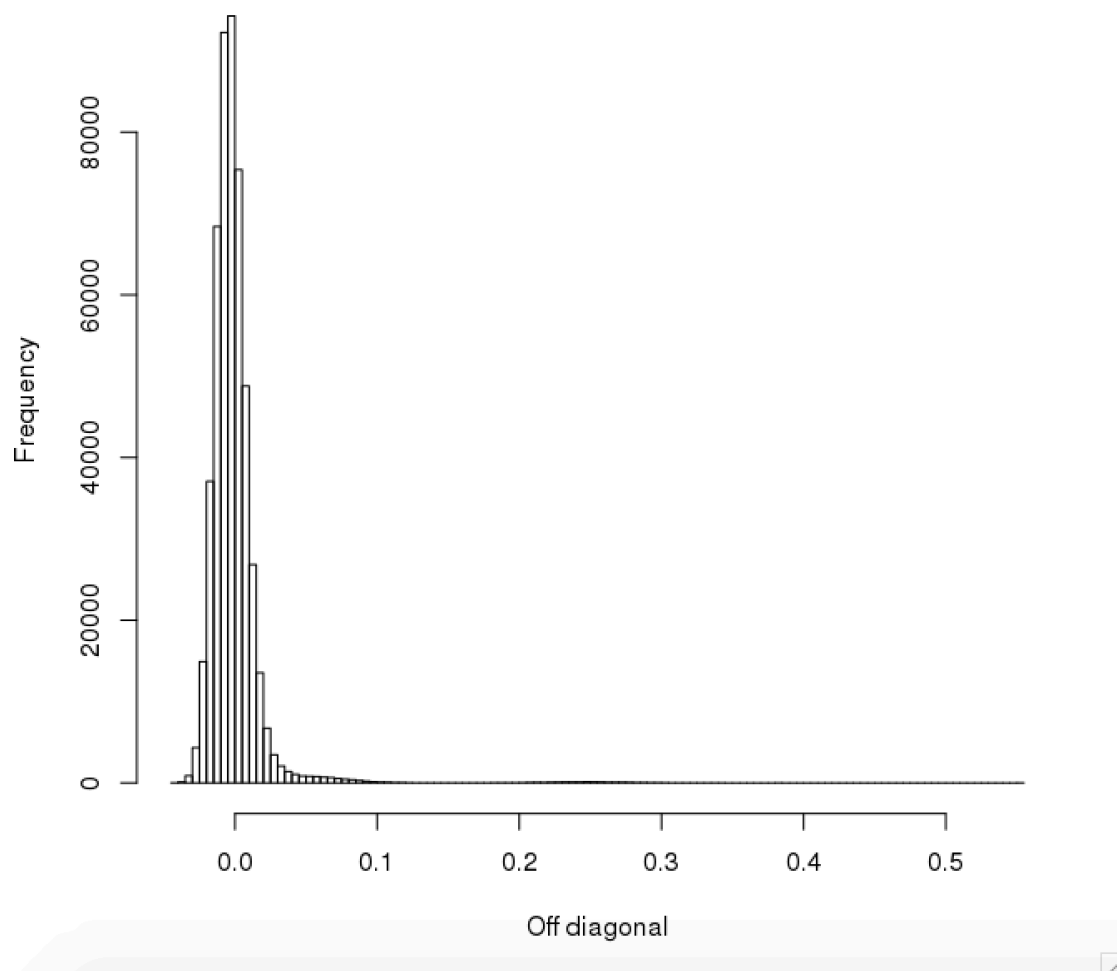


Figure S2. The distribution of off-diagonal of the genomic relationships matrix among 1000 individuals when effective population size is 1000.

With an effective population size of 1000 (having hardly close relationships as expected), the mean and variance of the genomic relationships is -0.001 and 0.0004 , respectively, which agrees with the expected value from the theory (0 and 0.0004). It is noted that whether using a high or low effective population size, the mean and variance of genomic relationships is generally matched with the expectation, supporting that any random samples from the population can be used for Eq. (5).

Supplementary Table 1. Approximated value for effective population size (N_e) for structured population of full sib families (kinship coefficient of 0.5)

Number of family	Number of members within family											
	2	5	10	20	30	40	50	60	70	80	90	100
2	5	6	6	6	6	6	6	6	6	6	6	6
5	15	10	10	9	9	9	9	9	9	9	9	9
10	31	20	18	17	17	17	17	17	17	17	17	17
20	63	40	36	34	33	33	33	33	33	33	33	33
30	96	60	53	51	50	49	49	49	49	49	49	49
40	129	80	72	68	67	66	66	66	65	65	65	65
50	162	101	90	85	84	83	82	82	82	82	82	82
60	196	122	108	102	101	100	99	99	99	99	98	98
70	230	143	127	120	118	117	116	116	115	115	115	115
80	264	164	145	137	135	134	133	133	132	132	132	132
90	298	185	164	155	152	151	150	150	149	149	149	149
100	333	206	183	173	170	168	167	167	166	166	166	166

Given the family structure, we estimated variance of relationships, which can be used to obtain M_e from Eq. (5). Assuming that the number of chromosomes is 30 chromosomes and the length of each chromosome is 1 Morgan (total 30 Morgan), the effective population size was estimated from Eq. (2).

Supplementary Table 2. Approximated value for effective population size (N_e) for structured population of half sib families (kinship coefficient of 0.25)

Number of family	Number of members within family											
	2	5	10	20	30	40	50	60	70	80	90	100
2	23	24	24	24	24	24	24	24	24	24	24	24
5	62	45	71	40	39	39	39	39	39	39	39	39
10	128	84	137	73	72	72	71	71	71	71	71	71
20	264	167	205	142	140	139	138	138	138	137	137	137
30	402	252	274	213	210	208	207	206	206	206	205	205
40	542	338	343	285	280	278	277	276	275	274	274	274
50	682	425	413	358	351	348	347	346	345	344	344	343
60	824	512	484	431	423	420	417	416	415	414	414	413
70	967	599	554	504	495	491	488	487	486	485	484	484
80	1110	688	625	578	568	563	560	558	557	556	555	554
90	1255	776	697	652	641	635	632	630	628	627	626	625
100	1399	865	767	726	714	707	704	701	700	698	697	697

Given the family structure, we estimated variance of relationships, which can be used to obtain M_e from Eq. (5). Assuming that the number of chromosomes is 30 chromosomes and the length of each chromosome is 1 Morgan (total 30 Morgan), the effective population size was estimated from Eq. (2).