Laboratory Note

Genetic Epistasis
I - Materials and methods
LN-1-2014

Ricardo Pinho and Rui Camacho
FEUP
Rua Dr Roberto Frias, s/n,
4200-465 PORTO
Portugal
Fax: (+351) 22 508 1440
e-mail: ei09045@fe.up.pt
www : http://www.fe.up.pt/∼ei09045 rcamacho@fe.up.pt
www : http://www.fe.up.pt/∼rcamacho

May 2014

Abstract

Based on literature results, we have selected 7 epistatic detection methods. The selected methods were empirically evaluated and compared using generated data from genomeSimla to simulate a smaller scale of genome wide studies. The simulated data includes 270 different configurations of datasets to simulate a wide array of disease models. The selected algorithms are BEAM 3.0, BOOST, MBMDR, Screen and Clean, SNPRuler, SNPHarvester, and TEAM. These algorithms are evaluated according to their Power, scalability and Type I Error Rate.

1 Introduction

The search for genetic predisposition to diseases has been researched for a long time. However, most early studies only focused on single SNP studies to determine disease predisposition. This is not the case in most complex diseases. Generally, most disease involve thousands or milions of SNPs, interacting between them in a large scale. Due to the complexity of these interactions, the computational costs for epistasis detection were infeasible until recently.
The main objective of the following experiments is to empirically evaluate the following algorithms: BEAM 3.0 [Zha12], BOOST [WYY+10a], MBMDR [MVV11], Screen and Clean [WDR+10], SNPRuler [WYY+10b], SNPHarvester [YHW+09], and TEAM [ZHZW10].
These algorithms will be evaluated according to their **Power**, **scalability**, and **Type I Error Rate**. Each algorithm will be executed with many data sets that simulate diseases with many different parameters.
These data sets are generated with **genomeSimla**, an open source data generator that contains many useful parameters to realistically simulate complex diseases.
The structure of the rest of the lab note consists of a brief description of the data sets that were used in the experiments, including the application used to generate them in Section 2. Section 3 includes a description of the evaluation measures used in these experiments. Section 4 contains the experimental methodology followed in these experiments. Section 5 is the summary of the experiments that will be detailed in the next lab notes.

2 The Data sets for the Experiments

The data sets for the experiments were created specifically for these experiments. The program used for the generation of the datasets was genomeSimla [EBT+08]. In total, 270 different configurations were generated. Each configuration consists of 100 data sets, which means that each algorithm was executed 27000 times.

Data Generation Application

The data generation application used for these experiments was genomeSimla. Due to its ability to evolve a population and achieve the desired allele frequencies, with any amount of SNPs, distributed by as many chromosomes as desired, genomeSimla is an adequate application for these kind of exper-

iments. The evolution of the population can follow a linear, exponential or logistic growth, the last one being the most preferred model.
Aside from generating and evolving a population to any amount of iterations as required, genomeSimla allows for an observation of the allele frequencies along the population and chose, based on those frequencies, which SNPs should be allocated, choosing how many chromosomes and block of SNPs per chromosome for each individual a priori.
After the generation of the population according to the selected parameters, genomeSimla can then be used to generate datasets, sampling from the population pool with as many individuals as necessary. The disease model can be further customized, with the desired odds ratio, prevalence of the disease, and type of disease model. Based on these values, a penetrance table is generated for each desired parameter.

- Allele Frequency - The frequency of the minor allele of the disease SNPs.
- Population - Number of individuals sampled in the data set.
- Disease Model - Type of disease model: main effect, epistasis interaction, and full effect.
- Odds ratio - Relation between disease SNPs. Probability of one disease SNP being present, given the presence of the other disease SNP.
- Prevalence - The proportion of a population with the disease. Affects the number of cases and controls in a data set.

With this data, data sets can be generated, using a configuration file, embedding the disease model into the desired alleles.

Data Set

The data sets were created using many different parameters, to maximize the diversity of disease models, to assert which algorithms are best for which scenarios. The data consists of a simulation of genotypes and phenotypes. For each individual, the attributes consist of genotypes associated with each SNP, for a total of 3 states: Homozygotic dominant, heterozygotic and homozygotic recessive. The label is binary, corresponding to an affected or not affected individual.
In each data set, a total of 2 pairs of chromosomes where generated. The first chromosome contains 20 blocks of 10 SNPs and the second contains 10

blocks of 10 SNPs, having 300 SNPs in total. There are two disease alleles placed in different chromosomes, according to the desired allele frequency.
The generated data sets contain 3 different number of individuals: 500, 1000, and 2000 individuals. The disease alleles contains 5 different minor allele frequencies: 0.01, 0.05, 0.1, 0.3, and 0.5. Three different disease models are used: data sets with marginal effects and no epistatic relations, without marginal effects and with epistatic relations and with marginal effects and epistatic relations. The odds ratio associated with both disease related alleles is 1.1, 1.5, or 2.0. The prevalence of the disease can is also configurated to either 0.0001 or 0.02, which also influances the amount of cases and controls.

3 Evaluation Measures

The evaluations measures used for these experiments consist of **Power**, **scalability**, and **Type I Error Rate**.
To evaluate the Power of the algorithms, for each configuration, the number of data sets were the ground truth is a statistically relevant interaction, measured using the χ^2 test, out of 100 data sets. Calculating the amount of datasets, within each configuration, how many data sets correctly identify the ground truth of the disease as the most significant SNP pair, considering that SNPs are ranked according to their importance to the phenotype, using statistical hypothesis tests.
To evaluate the scalability of each dataset, the average time for each dataset is calculated in each configuration.
In the Type I Error Rate, the proportion out of 100 data sets where non-disease related SNPs in each configuration are classified as a statistically relevant SNP pair, using χ^2 test.

4 Experimental Methodology

Initially, the population for the datasets is generated using genomeSimla. The population is generated using a logistic growth rate, with an initial population of 10000 and a maximum capacity of 1000000. The population chosen for the datasets is picked from reported generations, based on the allele frequencies desired for the experiment. The generation 1750 was selected for this purpose. 2 SNPs are selected for each configuration. The SNPs selected according to their minor allele frequency (MAF) were as follows:

- MAF 0.01 - $SNP112$ and $SNP267$

- MAF 0.05 - $SNP4$ and $SNP239$
- MAF 0.1 - $SNP135$ and $SNP230$
- MAF 0.3 - $SNP197$ and $SNP266$
- MAF 0.5 - $SNP80$ and $SNP229$

The first 200 SNPs belong to chromosome 1, where as the last 100 correspond to chromosome 2 SNPs.The table with all the allele frequencies can be seen in the annexes. Table 1 are the chromosome 1 allele frequencies and table 2 are the chromosome 2 allele frequencies
The penetrance tables are created from the allele frequencies in the population, following the configurations that were discussed earlier. The datasets are created, using each unique configuration file to create 100 datasets, generating all the configurations mentioned before.
With the data sets generated, the algorithms are tested for the most extreme configurations (minimum and maximum MAF) to find if the results are valid. Upon asserting the validity of the experiment, all algorithms are then executed for all configurations to analyze the potential of each algorithm.
For each algorithm in each dataset, a file containing the ranked SNPs, according to statistical relevancy, is generated, together with information about the time and memory used in the execution for each test. The Power and Type I Error Rates are are taken from the results that present a statistic relevance of $\alpha < 0.05$.
The computer used for this experiments used the 64-bit Ubuntu 13.10 operating system, with an Intel(R) Core(TM)2 Quad CPU Q6600 2.40GHz processor and 8,00 GB of RAM memory.

A Loci Frequencies

Chromosome 1

Table 1: Allele frequencies of the generated population for chromosome 1.

Label	**Freq Al1**	**Freq Al2**	**Map Dist.**	**Position**
RL0-1	0.704448	0.295552	0.0002523144	253
RL0-2	0.467747	0.532253	3.65488E-006	256
RL0-3	0.856627	0.143373	3.86582E-006	259

RL0-4	0.94761	0.05239	1.18175E-006	260
RL0-5	0.747191	0.252809	1.23056E-006	261
RL0-6	0.868644	0.131356	5.41858E-006	266
RL0-7	0.869881	0.130119	9.49181E-006	275
RL0-8	0.634084	0.365916	1.72337E-006	276
RL0-9	0.616899	0.383101	4.81936E-006	280
RL0-10	0.603205	0.396795	8.88582E-006	288
RL0-11	0.951322	0.048678	0.000118908	406
RL0-12	0.928004	0.071996	9.53558E-006	415
RL0-13	0.7257	0.2743	3.20447E-006	418
RL0-14	0.547945	0.452055	3.96875E-006	421
RL0-15	0.735312	0.264688	5.03938E-006	426
RL0-16	0.983344	0.016655	4.10188E-006	430
RL0-17	0.809402	0.190598	7.11582E-006	437
RL0-18	0.908173	0.091827	2.25726E-006	439
RL0-19	0.628892	0.371108	2.13406E-006	441
RL0-20	0.824863	0.175137	9.99491E-006	450
RL0-21	0.640543	0.359457	0.000229233	679
RL0-22	0.542639	0.457361	5.61457E-006	684
RL0-23	0.776321	0.223679	5.05623E-006	689
RL0-24	0.925422	0.074578	4.39722E-006	693
RL0-25	0.596454	0.403546	9.48707E-006	702
RL0-26	0.80071	0.19929	7.38516E-006	709
RL0-27	0.712163	0.287837	3.95139E-006	712
RL0-28	0.91426	0.08574	5.07943E-006	717
RL0-29	0.902589	0.097411	0.000006668	723
RL0-30	0.933652	0.066348	2.4885E-006	725
RL0-31	0.486126	0.513874	0.000296081	1021
RL0-32	0.553701	0.446299	8.33422E-006	1029
RL0-33	0.887238	0.112762	4.95048E-006	1033
RL0-34	0.93165	0.06835	5.32692E-006	1038
RL0-35	0.887583	0.112417	2.23131E-006	1040
RL0-36	0.824546	0.175454	5.40611E-006	1045
RL0-37	1	0	7.03837E-006	1052
RL0-38	0.817039	0.182961	1.44855E-006	1053
RL0-39	0.762831	0.237169	9.89044E-006	1062
RL0-40	0.623942	0.376058	2.53856E-006	1064
RL0-41	0.886716	0.113284	0.0003574	1421

RL0-42	0.603873	0.396127	5.73344E-006	1426
RL0-43	0.708144	0.291856	7.18489E-006	1433
RL0-44	0.722182	0.277818	6.17693E-006	1439
RL0-45	0.59756	0.40244	6.57155E-006	1445
RL0-46	0.810217	0.189783	3.25347E-006	1448
RL0-47	0.679944	0.320056	8.28564E-006	1456
RL0-48	0.467092	0.532908	4.45383E-006	1460
RL0-49	0.518637	0.481363	2.97358E-006	1462
RL0-50	0.918397	0.081603	4.58774E-006	1466
RL0-51	0.979136	0.020864	0.0003772277	1843
RL0-52	0.571337	0.428663	3.32175E-006	1846
RL0-53	0.615734	0.384266	2.23233E-006	1848
RL0-54	0.695586	0.304414	2.98606E-006	1850
RL0-55	0.660442	0.339558	4.02315E-006	1854
RL0-56	0.910148	0.089852	5.9643E-006	1859
RL0-57	0.445087	0.554913	6.82648E-006	1865
RL0-58	0.470733	0.529267	9.5693E-006	1874
RL0-59	0.858588	0.141412	0.000004337	1878
RL0-60	0.681468	0.318532	2.15272E-006	1880
RL0-61	0.870466	0.129534	0.0003573156	2237
RL0-62	0.646194	0.353806	7.20147E-006	2244
RL0-63	0.763207	0.236793	3.17006E-006	2247
RL0-64	0.931087	0.068913	8.01624E-006	2255
RL0-65	0.7151	0.2849	6.35415E-006	2261
RL0-66	0.670911	0.329089	2.38872E-006	2263
RL0-67	0.888122	0.111878	2.52589E-006	2265
RL0-68	0.694165	0.305835	0.000008364	2273
RL0-69	0.864311	0.135689	7.35972E-006	2280
RL0-70	0.838895	0.161105	2.46709E-006	2282
RL0-71	0.823928	0.176073	0.0001992617	2481
RL0-72	0.583947	0.416053	6.33832E-006	2487
RL0-73	0.841979	0.158021	9.79685E-006	2496
RL0-74	0.6003	0.3997	7.07911E-006	2503
RL0-75	0.892639	0.107361	5.16523E-006	2508
RL0-76	0.761561	0.238439	2.85138E-006	2510
RL0-77	0.900447	0.099553	1.53824E-006	2511
RL0-78	0.599257	0.400743	3.89272E-006	2514
RL0-79	0.972086	0.027914	6.53018E-006	2520

RL0-80	0.560663	0.439337	8.62124E-006	2528
RL0-81	0.554206	0.445794	0.000199997	2727
RL0-82	0.93403	0.06597	8.61757E-006	2735
RL0-83	0.542574	0.457426	9.10087E-006	2744
RL0-84	0.837702	0.162298	1.23079E-006	2745
RL0-85	0.909783	0.090217	6.84162E-006	2751
RL0-86	0.91318	0.08682	4.48263E-006	2755
RL0-87	0.725569	0.274431	0.000001848	2756
RL0-88	0.90355	0.09645	2.79894E-006	2758
RL0-89	0.716186	0.283814	4.00443E-006	2762
RL0-90	0.612835	0.387165	6.94976E-006	2768
RL0-91	0.582162	0.417838	0.0003616833	3129
RL0-92	0.83582	0.16418	0.000009529	3138
RL0-93	0.558802	0.441198	9.02466E-006	3147
RL0-94	0.86217	0.13783	5.29547E-006	3152
RL0-95	0.617906	0.382094	7.09319E-006	3159
RL0-96	0.801595	0.198405	6.73657E-006	3165
RL0-97	0.676978	0.323022	6.97316E-006	3171
RL0-98	0.738348	0.261652	7.87644E-006	3178
RL0-99	0.591386	0.408614	3.67391E-006	3181
RL0-100	0.521751	0.478249	4.20054E-006	3185
RL0-101	0.508844	0.491156	9.09917E-005	3275
RL0-102	0.565387	0.434613	9.41043E-006	3284
RL0-103	0.479309	0.520691	7.40872E-006	3291
RL0-104	0.745518	0.254482	3.35237E-006	3294
RL0-105	0.532452	0.467548	4.28727E-006	3298
RL0-106	0.935416	0.064584	9.89425E-006	3307
RL0-107	0.662617	0.337383	8.74864E-006	3315
RL0-108	0.658306	0.341694	2.01241E-006	3317
RL0-109	0.712991	0.287009	5.8733E-006	3322
RL0-110	0.665501	0.334499	6.69027E-006	3328
RL0-111	0.568289	0.431711	8.718047E-005	3415
RL0-112	0.98671	0.01329	8.66949E-006	3423
RL0-113	0.79789	0.20211	5.05033E-006	3428
RL0-114	0.553154	0.446846	9.60618E-006	3437
RL0-115	0.667399	0.332601	6.92172E-006	3443
RL0-116	0.700185	0.299815	9.52134E-006	3452
RL0-117	0.610748	0.389252	5.60877E-006	3457

RL0-118	0.661102	0.338898	6.63784E-006	3463
RL0-119	0.820744	0.179256	3.09427E-006	3466
RL0-120	0.912926	0.087073	4.1968E-006	3470
RL0-121	0.68335	0.31665	0.0003871028	3857
RL0-122	0.707937	0.292063	5.00312E-006	3862
RL0-123	0.589477	0.410523	2.13525E-006	3864
RL0-124	0.745493	0.254507	9.8212E-006	3873
RL0-125	0.698088	0.301912	7.02674E-006	3880
RL0-126	0.424467	0.575533	5.18827E-006	3885
RL0-127	0.787719	0.212281	4.74483E-006	3889
RL0-128	0.860644	0.139356	5.22368E-006	3894
RL0-129	0.638396	0.361604	3.96526E-006	3897
RL0-130	0.731953	0.268047	8.71207E-006	3905
RL0-131	0.744233	0.255766	0.0002181738	4123
RL0-132	1	0	1.69539E-006	4124
RL0-133	0.771704	0.228296	9.71469E-006	4133
RL0-134	0.878927	0.121073	0.000002233	4135
RL0-135	0.90145	0.09855	4.28905E-006	4139
RL0-136	0.648369	0.351631	0.00000754	4146
RL0-137	0.80335	0.19665	8.70869E-006	4154
RL0-138	0.856866	0.143134	9.44719E-006	4163
RL0-139	0.615518	0.384482	3.60345E-006	4166
RL0-140	0.788087	0.211913	0.000002436	4168
RL0-141	0.678961	0.321039	0.0002748812	4442
RL0-142	0.771435	0.228565	5.86447E-006	4447
RL0-143	0.503258	0.496742	3.67578E-006	4450
RL0-144	0.795211	0.204789	2.75252E-006	4452
RL0-145	0.490144	0.509856	4.10642E-006	4456
RL0-146	0.488492	0.511508	4.30833E-006	4460
RL0-147	0.667302	0.332698	7.3961E-006	4467
RL0-148	0.643159	0.356841	2.3613E-006	4469
RL0-149	0.673992	0.326008	9.5407E-006	4478
RL0-150	0.788535	0.211465	5.39342E-006	4483
RL0-151	0.781059	0.218941	0.0002359844	4718
RL0-152	0.502629	0.497371	5.62238E-006	4723
RL0-153	0.466542	0.533458	2.22743E-006	4725
RL0-154	0.538982	0.461018	3.21068E-006	4728
RL0-155	0.841056	0.158944	2.43989E-006	4730

RL0-156	0.462765	0.537235	7.40954E-006	4737
RL0-157	0.90605	0.09395	3.96506E-006	4740
RL0-158	0.681072	0.318928	2.10963E-006	4742
RL0-159	0.596135	0.403865	6.71541E-006	4748
RL0-160	0.855496	0.144504	0.00000768	4755
RL0-161	0.727272	0.272728	0.0002969833	5051
RL0-162	0.774272	0.225728	2.62789E-006	5053
RL0-163	0.791941	0.208059	6.76876E-006	5059
RL0-164	0.644252	0.355748	0.000005599	5064
RL0-165	0.549582	0.450418	8.32549E-006	5072
RL0-166	0.428749	0.571251	8.10471E-006	5080
RL0-167	0.376485	0.623515	9.96927E-006	5089
RL0-168	0.535948	0.464052	9.47661E-006	5098
RL0-169	0.514295	0.485705	3.16517E-006	5101
RL0-170	0.700045	0.299955	5.98168E-006	5106
RL0-171	0.571955	0.428045	0.0003862553	5492
RL0-172	0.586523	0.413477	2.88618E-006	5494
RL0-173	0.783275	0.216725	7.29982E-006	5501
RL0-174	0.610016	0.389985	9.43182E-006	5510
RL0-175	0.866664	0.133336	7.05865E-006	5517
RL0-176	0.75876	0.24124	7.56181E-006	5524
RL0-177	0.600093	0.399907	1.005344E-005	5534
RL0-178	0.577467	0.422533	2.42474E-006	5536
RL0-179	0.789476	0.210524	6.1728E-006	5542
RL0-180	0.590153	0.409847	5.99256E-006	5547
RL0-181	0.422633	0.577367	9.624393E-005	5643
RL0-182	0.526449	0.473551	1.007159E-005	5653
RL0-183	0.83354	0.16646	3.23814E-006	5656
RL0-184	0.737217	0.262783	8.58028E-006	5664
RL0-185	0.650092	0.349908	9.27841E-006	5673
RL0-186	0.56464	0.43536	5.87977E-006	5678
RL0-187	0.717536	0.282464	3.16557E-006	5681
RL0-188	0.961919	0.038081	2.93894E-006	5683
RL0-189	0.84241	0.15759	8.25314E-006	5691
RL0-190	0.817398	0.182602	4.0069E-006	5695
RL0-191	1	0	0.0002386956	5933
RL0-192	1	0	4.98276E-006	5937
RL0-193	0.709334	0.290666	0.000002811	5939

RL0-194	0.78411	0.21589	0.000008052	5947
RL0-195	0.932612	0.067388	2.89373E-006	5949
RL0-196	0.865947	0.134053	8.6839E-006	5957
RL0-197	0.725338	0.274662	5.21764E-006	5962
RL0-198	0.795964	0.204036	7.8731E-006	5969
RL0-199	0.583016	0.416984	4.61094E-006	5973
RL0-200	0.803726	0.196274	8.37366E-006	5981

Chromosome 2

Table 2: Allele frequencies of the generated population for chromosome 2.

Label	Freq Al1	Freq Al2	Map Dist.	Position
RL1-201	0.893976	0.106024	0.0003986369	399
RL1-202	0.584141	0.415859	2.05934E-006	401
RL1-203	0.422083	0.577917	0.000005955	406
RL1-204	0.73351	0.26649	5.58855E-006	411
RL1-205	0.694034	0.305966	4.1723E-006	415
RL1-206	0.765355	0.234645	2.06415E-006	417
RL1-207	0.965014	0.034986	7.44318E-006	424
RL1-208	0.668517	0.331483	9.60649E-006	433
RL1-209	0.634885	0.365115	8.56251E-006	441
RL1-210	0.725027	0.274973	6.14954E-006	447
RL1-211	0.698398	0.301602	7.386583E-005	520
RL1-212	0.595985	0.404015	9.7547E-006	529
RL1-213	0.710597	0.289403	1.58667E-006	530
RL1-214	0.663247	0.336753	4.37889E-006	534
RL1-215	0.75663	0.24337	7.38782E-006	541
RL1-216	0.936743	0.063257	8.35938E-006	549
RL1-217	0.663784	0.336216	1.64064E-006	550
RL1-218	0.680104	0.319896	9.16445E-006	559
RL1-219	0.688756	0.311244	0.000007628	566
RL1-220	0.9333	0.0667	7.01934E-006	573
RL1-221	0.742415	0.257585	0.0003420352	915
RL1-222	0.799322	0.200678	4.01391E-006	919
RL1-223	0.709122	0.290879	0.000002737	921
RL1-224	0.565597	0.434403	6.28353E-006	927

RL1-225	0.863029	0.136971	9.64911E-006	936
RL1-226	0.752561	0.247439	6.74076E-006	942
RL1-227	0.676998	0.323002	1.004539E-005	952
RL1-228	0.840474	0.159526	1.71067E-006	953
RL1-229	0.49346	0.50654	0.000001589	954
RL1-230	0.910095	0.089905	7.41687E-006	961
RL1-231	0.960868	0.039132	0.0002261121	1187
RL1-232	0.933743	0.066257	1.91042E-006	1188
RL1-233	0.760953	0.239047	4.80473E-006	1192
RL1-234	0.748072	0.251928	7.04549E-006	1199
RL1-235	0.663473	0.336527	8.21959E-006	1207
RL1-236	0.964783	0.035217	2.82873E-006	1209
RL1-237	0.905525	0.094475	0.000007663	1216
RL1-238	0.691349	0.308651	5.04876E-006	1221
RL1-239	0.951645	0.048355	1.59639E-006	1222
RL1-240	0.989216	0.010784	8.73616E-006	1230
RL1-241	0.738781	0.261219	0.0003243203	1554
RL1-242	0.795527	0.204473	8.89964E-006	1562
RL1-243	0.795563	0.204437	2.02264E-006	1564
RL1-244	0.703822	0.296178	3.36477E-006	1567
RL1-245	0.57285	0.42715	9.19778E-006	1576
RL1-246	0.767369	0.232631	7.29139E-006	1583
RL1-247	0.645825	0.354175	2.43094E-006	1585
RL1-248	0.802402	0.197598	1.73925E-006	1586
RL1-249	0.944397	0.055603	9.46653E-006	1595
RL1-250	0.622399	0.377601	8.82309E-006	1603
RL1-251	0.630848	0.369152	0.0002217582	1824
RL1-252	0.818129	0.181871	1.91494E-006	1825
RL1-253	0.484804	0.515196	1.6334E-006	1826
RL1-254	0.676497	0.323503	1.59652E-006	1827
RL1-255	0.880815	0.119185	3.35782E-006	1830
RL1-256	0.959511	0.040489	2.75846E-006	1832
RL1-257	0.784072	0.215928	3.03069E-006	1835
RL1-258	0.52286	0.47714	6.06819E-006	1841
RL1-259	0.623466	0.376534	6.91131E-006	1847
RL1-260	0.874709	0.125291	7.25071E-006	1854
RL1-261	0.803013	0.196987	0.000331411	2185
RL1-262	0.545178	0.454822	4.47452E-006	2189

RL1-263	0.815965	0.184035	4.89193E-006	2193
RL1-264	0.818366	0.181634	1.90565E-006	2194
RL1-265	0.724692	0.275308	1.45521E-006	2195
RL1-266	0.68352	0.31648	1.001287E-005	2205
RL1-267	0.989999	0.010001	3.48414E-006	2208
RL1-268	0.985774	0.014226	9.2895E-006	2217
RL1-269	0.642113	0.357887	4.82072E-006	2221
RL1-270	0.464929	0.535071	0.000002507	2223
RL1-271	0.734131	0.265869	0.0003870134	2610
RL1-272	0.632632	0.367368	8.94209E-006	2618
RL1-273	0.553081	0.446919	0.000004175	2622
RL1-274	0.764977	0.235023	5.60863E-006	2627
RL1-275	0.464551	0.535449	9.08894E-006	2636
RL1-276	0.851137	0.148863	1.002911E-005	2646
RL1-277	0.739427	0.260573	9.30477E-006	2655
RL1-278	0.555538	0.444462	6.07683E-006	2661
RL1-279	0.551021	0.448979	1.71434E-006	2662
RL1-280	0.593129	0.406871	6.07637E-006	2668
RL1-281	0.79749	0.20251	0.0002724436	2940
RL1-282	0.848332	0.151668	7.00277E-006	2947
RL1-283	0.812696	0.187304	5.19928E-006	2952
RL1-284	0.715573	0.284426	6.3489E-006	2958
RL1-285	0.578981	0.421019	3.26024E-006	2961
RL1-286	0.786632	0.213368	0.000008282	2969
RL1-287	0.64689	0.35311	8.91268E-006	2977
RL1-288	0.600677	0.399323	2.59076E-006	2979
RL1-289	0.552264	0.447736	7.46941E-006	2986
RL1-290	0.836774	0.163226	7.75812E-006	2993
RL1-291	0.910408	0.089592	6.30604E-005	3056
RL1-292	0.705616	0.294384	9.07012E-006	3065
RL1-293	0.833055	0.166945	7.11207E-006	3072
RL1-294	0.55822	0.44178	9.56152E-006	3081
RL1-295	0.684736	0.315264	2.78967E-006	3083
RL1-296	0.973315	0.026685	9.43315E-006	3092
RL1-297	0.676965	0.323035	5.50475E-006	3097
RL1-298	0.698511	0.301489	4.26716E-006	3101
RL1-299	0.514109	0.485891	9.42184E-006	3110
RL1-300	0.895842	0.104158	9.44663E-006	3119

References

[EBT+08] Todd L Edwards, William S Bush, Stephen D Turner, Scott M Dudek, Eric S Torstenson, Mike Schmidt, Eden Martin, and Marylyn D Ritchie. Generating Linkage Disequilibrium Patterns in Data Simulations using genomeSIMLA. *Lecture Notes in Computer Science*, 4973:24–35, 2008.

[MVV11] Jestinah M Mahachie John, Francois Van Lishout, and Kristel Van Steen. Model-Based Multifactor Dimensionality Reduction to detect epistasis for quantitative traits in the presence of error-free and noisy data. *Eur J Hum Genet*, 19(6):696–703, June 2011.

[WDR+10] Jing Wu, Bernie Devlin, Steven Ringquist, Massimo Trucco, and Kathryn Roeder. Screen and clean: a tool for identifying interactions in genome-wide association studies. *Genetic epidemiology*, 34:275–285, 2010.

[WYY+10a] Xiang Wan, Can Yang, Qiang Yang, Hong Xue, Xiaodan Fan, Nelson L S Tang, and Weichuan Yu. BOOST: A fast approach to detecting gene-gene interactions in genome-wide case-control studies. *American journal of human genetics*, 87:325–340, 2010.

[WYY+10b] Xiang Wan, Can Yang, Qiang Yang, Hong Xue, Nelson L S Tang, and Weichuan Yu. Predictive rule inference for epistatic interaction detection in genome-wide association studies. *Bioinformatics (Oxford, England)*, 26:30–37, 2010.

[YHW+09] Can Yang, Zengyou He, Xiang Wan, Qiang Yang, Hong Xue, and Weichuan Yu. SNPHarvester: a filtering-based approach for detecting epistatic interactions in genome-wide association studies. *Bioinformatics (Oxford, England)*, 25:504–511, 2009.

[Zha12] Yu Zhang. A novel bayesian graphical model for genome-wide multi-SNP association mapping. *Genetic Epidemiology*, 36:36–47, 2012.

[ZHZW10] Xiang Zhang, Shunping Huang, Fei Zou, and Wei Wang. TEAM: efficient two-locus epistasis tests in human genome-wide association study. *Bioinformatics (Oxford, England)*, 26:i217–i227, 2010.