Results for the real Rheumatoid Arthritis dataset

1. Performances

Table 1. Number of epistatic interactions detected over ten runs of SMMB, BEAM, DASSO-MB and AntEpiSeeker respectively, for the 23 chromosomes of the WTCCC RA dataset. Parameters for SMMB: $t = 100$, $r = 4\times10^4$, $m = 30$, $K = 180$, $k = 3$, $\alpha = 0.05$. Parameters for BEAM: *nb iterations for burn-in phase:* 10^6, *nb iterations for stationary phase:* 10^7. Parameters for DASSO-MB: $\alpha = 0.05$. Parameters for AntEpiSeeker: 3×10^4 *iterations,* 5×10^4 *ants,* $\alpha = 0.01$.

Number of interactions detected in a run	Percentage of runs outputing the same number of interactions (over ten runs of the same software)			
	SMMB	BEAM	DASSO-MB	AntEpiSeeker
4	20%	0%	0%	0%
5	10%	0%	100%	20%
6	30%	0%	0%	30%
7	40%	100%	0%	50%

For instance, in 40% of the runs of SMMB (*i.e.* 4 out of 10 runs), seven interactions were identified, compared to 100%, 0% and 50% for BEAM, DASSO-MB and AntEpiSeeker, respectively.

2. Interactions detected

Table 2. Epistatic interactions identified by SMMB on the WTCCC RA dataset. NR stands for non-coding region. The p-values were computed using logistic regression (the phenotype was regressed against the SNPs in each 2-way pattern).

Epistasis interaction	Chromosomal location	Related genes	P-value
rs3890745-rs41348151	1p36.32-1p21.3	MMEL1-NR	$< 10^{-16}$
rs6457617-rs41443144	6p21.32-6p25.1	MHC-RP3	$< 10^{-16}$
rs6457620-rs41454544	6p21.32-6q21	HLA-DRB1-OSTM1	$< 10^{-16}$
rs6920220-rs17165379	6q23.3-6p25.3	TNFAIP3-NR	$< 10^{-16}$
rs4810485-rs2748666	20q13.12-20q11.23	CD40-NR	$< 10^{-16}$
rs2837960-rs41492246	21q22.3-21q22.12	NR-NR	$< 10^{-16}$
rs743777-rs9627642	22q12.3-22q13.31	NR-NR	$< 10^{-16}$

3. Running times

Table 3. Running times for SMMB, BEAM, DASSO-MB and AntEpiSeeker, for the 23 chromosomes of the WTCCC RA dataset (cumulative times), on XEON biprocessors 5462 2.66 GHz, 6 cores. Averages over 10 runs are provided for the non-deterministic software programs SMMB, BEAM and AntEpiSeeker.

Software	SMMB	BEAM	DASSO-MB	AntEpiSeeker
Running times	34h	59h	12h	69h

The number of SNPs of the largest chromosome is 38,867 (Chromosome 2). The ranking of the four methods by increasing running time was DASSO-MB, SMMB, BEAM and AntEpiSeeker for the datasets simulated for the multiplicative model (Model 1) (additional file 4, Supplementary data). This observation stands for the real dataset.

4. Epistatic interactions published for rheumatoid arthritis

A compilation of published interactions is presented in Table 4, together with the common results output by SMMB, BEAM and AntEpiSeeker. Eight genes are shown to be each involved in two epistatic interactions (*GPR133, SULF1, AKAP1, RTN4, GLIS3, MMEL1, BLK, PADI4).* There is no overlapping result between the different studies shown in Table 4, except for BEAM, SMMB and AntEpiSeeker, which used the same WTCCC dataset.

Julià and co-workers performed an exhaustive search of 2-order epistatic interactions for a Spanish cohort (Julià *et al.* 2008). They analyzed 317,503 SNPs in 400 cases and 400 controls. They selected a group of candidate SNPs for replication in an independent group of 410 cases and 394 controls. An important remark is that no known main-effect SNP was observed in the group of 10 top SNPs ($P < 1 \times 10^{-10}$), which defines pure epistasis. None of the top SNP pairs belong to genes from a known common biologic pathway. Protein phosphatase 1B is a regulator of NF-κB transcription factor. The study of Julià and co-workers shows that this protein is likely to have a strong interaction with protein G-coupled receptor 133. To date, the later protein has no associated biologic function. The human sulfatase 1 gene (*SULF1*) is a heparan sulfatase involved in tumor progression and inflammation (Kodaira,Y. *et al.*, 2000). It is shown to interact with 2 SNPs from the A kinase anchor protein 1 gene (*AKAP1*). The latter gene has been associated with cAMP-mediated signal transduction and messenger RNA trafficking (Furusawa,M. *et al.*, 2002).

Deshmukh and co-workers focused on 19 SNPs genotyped in 353 cases and 368 controls. They evidenced gene-gene interaction between SNPs in *MMEL1* (rs3890745) and *C80rf13-BLK* (rs13277113) (Deshmukh et *al.*, 2011).

Génin and collaborators detected an epistatic interaction between *BANK1* rs3733197 and *BLK* rs13277113, through a large meta-analysis (1,915 cases and 1,915 controls) (Génin *et al.*, 2013). Interestingly, none of the SNPs tested individually was significantly associated with RA in this study, which again indicates pure epistasis.

Xu and Yuan focused their work on four genes (*VEGFA, PADI4, C5, ITGAV*) to detect interactions with RA susceptibility (Xu,J. and Yuan,Z., 2016). Respectively 4, 6, 8 and 8 SNPs were examined in each of the previous genes. The GWAS cohorts consisted of 868 cases and 1,194 controls. These authors confirm the existence of two patterns of epistasis: *VEGFA-PADI4*, and *C5-PADI4*.

Table 4. Epistatic interactions published for rheumatoid arthritis, and epistatic interactions discovered by SMMB, BEAM and AntEpiSeeker software programs.

Study	SNP1	Chromosome location	*Gene 1*	SNP2	Chromosome location	*Gene 2*	p-value for the interaction
Julià *et al.* 2008	rs9752494	2	*PPM1B*	rs1569020	12	*GPR133*	1.22×10^{-12}
	rs10465885	1	*GJA5*	rs2302502	18	*PTPRM*	3.62×10^{-11}
	rs950675	2	*TPO*	rs1569020	12	*GPR133*	4.84×10^{-11}
	rs12755965	1	*GJA5*	rs6776932	3	*ACPP*	5.41×10^{-11}
	rs259401	6	*RAB32*	rs2322140	17	*DNAH9*	7.73×10^{-11}
	rs2244817	8	*SULF1*	rs3826296	17	*AKAP1*	8.56×10^{-11}
	rs2244817	8	*SULF1*	rs998113	17	*AKAP1*	9.52×10^{-11}
	rs10171653	2	*RTN4*	rs7033413	9	*GLIS3*	5.69×10^{-12}
	rs2580768	2	*RTN4*	rs7033413	9	*GLIS3*	2.63×10^{-11}
	rs4849025	2	*CNTNAP5*	rs2392829	8	*PXDNL*	9.07×10^{-11}
Deshmukh *et al.*, 2011	rs3890745	1	*MMEL1*	rs13277113		C80rf13-BLK	0.0002
Génin *et al*., 2013	rs3733197		*BANK1*	rs13277113		*BLK*	
Xu and Yuan, 2016			*VEGFA*			*PADI4*	
			C5			*PADI4*	
Zhang and Liu, 2007 (BEAM), SMMB AntEpiSeeker	rs3890745	l p36.32	*MMELl*	rs41348151	lp21.3	*NR*	$< 10^{-16}$
	rs6457617	6p21.32	*MHC*	rs41443144	6p25.1	*RP3*	$< 10^{-16}$
	rs6457620	6p21.32	*HLA-DRBl*	rs41454544	6q21	*0STM1*	$< 10^{-16}$
	rs6920220	6q23.3	*T NFAIP 3*	rsl 7165379	6p25.3	*NR*	$< 10^{-16}$
	rs 4810485	20q13.12	*CD40*	rs2748666	20ql l. 23	*NR*	$< 10^{-16}$
	rs2837960	21q22.3	*NR*	rs41492246	21q22.12	*NR*	$< 10^{-16}$
	rs7 43777	22q12 .3	*NR*	rs9627642	22q13.31	*NR*	$< 10^{-16}$

The first interaction evidenced by SMMB involves *MMEL1*, a gene of the membrane metallo-endopeptidase family. Genes in this family play important roles in pain perception, arterial pressure regulation, phosphate metabolism and homeostasis. The second interaction detected by SMMB involves rs6457617, located on the major histocompatibility complex region. *MHC* class I proteins trigger an immediate response from the immune system by displaying peptide fragments of non-self proteins from within the cell to cytotoxic T cells. The third interaction evidenced by SMMB points out a SNP located on the *HLA-DRB1* gene. This gene encodes the *DRB1* beta chain, which belongs to the *HLA* class II beta chain paralogues, and is a *HLA* class II histocompatibility antigen. This protein plays a key role in the immune system by presenting peptides derived from extracellular proteins to T helper cells. The second gene involved in the interaction detected is *OSTM1*, which encodes osteopetrosis-associated transmembrane protein, a protein required for osteoclast and melanocyte maturation and function. The fifth interaction detected by SMMB involves *CD40* gene, a member of the *TNF*-receptor superfamily. The *CD40* gene encodes a protein that is a receptor on antigen-presenting cells of the immune system. This protein is the key to mediating a broad variety of immune and inflammatory responses including T cell-dependent immunoglobulin class switching, memory B cell development, and germinal center formation.

Bibliographical references

Deshmukh,H.A. *et al.* (2011) Evaluation of 19 autoimmune disease-associated loci with rheumatoid arthritis in a Colombian population: evidence for replication and gene-gene interaction. *J. Rheumatol.*, 38(9),1866-1870.

Furusawa,M. et al. (2002) AMY-1 interacts with S-AKAP84 and AKAP95 in the cytoplasm and the nucleus, respectively, and inhibits cAMP-dependent protein kinase activity by preventing binding of its catalytic subunit to A-kinase-anchoring protein (AKAP) complex, *J. Biol. Chem*., 277,50885–50892.

Génin,A. *et al*. (2013) Epistatic interaction between *BANK1* and *BLK* in rheumatoid arthritis: results from a large trans-ethnic meta-analysis. *PLOS One*, 8(4): e1044. 10.1371/journal.pone.0061044.

Julià,A. *et al.* (2008) Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. *Arthritis Rheum*., 58(8), 2275-2286. doi: 10.1002/art.23623.

Kodaira,Y. *et al.* (2000) Phenotypic and functional maturation of dendritic cells mediated by heparin sulfate. *J. Immunol.*,165, 1599-1604.

Xu.J. and Yuan,Z. (2016) A powerful score-based test statistic for detecting gene-gene co-association, *BMC Genet*, 17: 31.