Hum Genet (2014) 133:1343–1358
DOI 10.1007/s00439-014-1480-y

REVIEW PAPER

Practical aspects of genome-wide association interaction analysis

Elena S. Gusareva · Kristel Van Steen

Received: 21 May 2014 / Accepted: 18 August 2014 / Published online: 28 August 2014

Abstract Large-scale epistasis studies can give new clues to system-level genetic mechanisms and a better understanding of the underlying biology of human complex disease traits. Though many novel methods have been proposed to carry out such studies, so far only a few of them have demonstrated replicable results. Here, we propose a minimal protocol for genome-wide association interaction (GWAI) analysis to identify gene–gene interactions from large-scale genomic data. The different steps of the developed protocol are discussed and motivated, and encompass interaction screening in a hypothesis-free and hypothesis-driven manner. In particular, we examine a wide range of aspects related to epistasis discovery in the context of complex traits in humans, hereby giving practical recommendations for data quality control, variant selection or prioritization strategies and analytic tools, replication and meta-analysis, biological validation of statistical findings and other related aspects. The minimal protocol provides guidelines and attention points for anyone involved in GWAI analysis and aims to enhance the biological relevance of GWAI findings. At the same time, the protocol improves a better assessment of strengths and weaknesses of published GWAI methodologies.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-014-1480-y) contains supplementary material, which is available to authorized users.

E. S. Gusareva (✉) · K. Van Steen
Systems and Modeling Unit, Montefiore Institute, University of Liege, 10 Grande Traverse, Sart-Tilman, 4000 Liège, Belgium
e-mail: egusareva@ulg.ac.be

E. S. Gusareva · K. Van Steen
Bioinformatics and Modeling, GIGA-R, University of Liege, Liège, Belgium

Introduction

Genome-wide association (GWA) studies have been very successful in identifying predisposing genetic variants to a variety of complex traits (e.g., GWAS Diagram Browser for exploring GWA studies at http://www.ebi.ac.uk/fgpt/gwas/ and the Catalog of Published Genome-Wide Association Studies at http://www.genome.gov/page.cfm?pageid=26525384&clearquery=1#result_table). Still, yet to identify structural variants such as deletions, duplications, copy-number variants, insertions, inversions or translocations are likely to be associated with these traits. These may jointly contribute to the so-called "missing heritability". The proportion of heritability explained by a set of genetic variants is the ratio of the heritability due to these variants, estimated directly from their observed effects, to the total heritability, inferred indirectly from population data. This proportion is surprisingly low for most complex traits and corresponding identified genetic variants via GWA techniques. Multiple causes of consequential missing heritability have been proposed (Eichler et al. 2010), one of these being an overestimation of the total heritability by commonly adopted estimation methods. When these estimation methods do not account for non-additive genetic affects or genetic interactions, the resulting total heritability may be too high and thus the proportion of heritability explained may be smaller than in reality (Zuk et al. 2012). Another possible cause for missing heritability is related to the under-representation of the so-called "high hanging fruit", such as complex interplays within and between sets of rare and/or common variants. The study of these synergies, typically the subject of an interaction analysis, is complicated by an intrinsic deficit in power and the unavailability of a clear road map that points out the pitfalls and advantages of possible interaction analysis strategies.

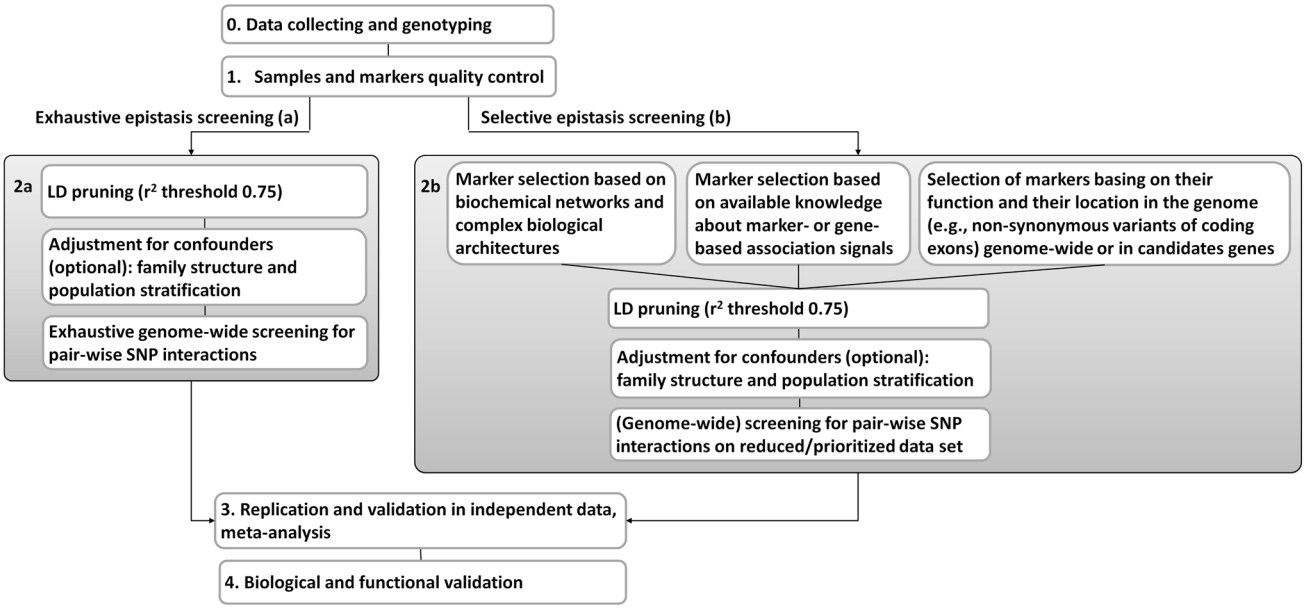

Fig. 1 Protocol for genome-wide association interaction (GWAI) analysis. The analytical blocs are *highlighted*

Genetic and biological epistasis is omnipresent (Miko 2008). However, in contrast to model organisms, the proportion of heritability that is due to epistasis for complex traits in humans is hard to estimate, in part because singular common and rare variants, structural variations, epistasis groups of genetic variations, epigenetic regulation, and the environment, all impact complex trait susceptibility in highly complex entangled ways. In yeast, the heritability of 46 quantitative trait loci (QTL) attributed to epistasis was estimated to vary from 2 to 54 % for the traits considered in the controlled experiments (Bloom et al. 2013). For selected chicken and mouse models, the proportion of variants caused by epistasis was estimated to be even higher, ranging from 0 to 81 %, for 18 QTL traits studied (Carlborg and Haley 2004). These and other studies in model organisms suggest not to ignore gene–gene interactions in explaining multifactorial trait variation in humans, and indicate that potentially complex population-dependent modes of inheritance need to be assumed (Carlborg and Haley 2004).

The discovery of genetic and biological epistasis via statistical methods remains a big challenge (Moore 2005), especially in the absence of prior hypotheses. Depending on the analysis protocol used, biological epistasis (occurring at the individual level) and statistical epistasis (occurring at a population level) may be closely connected or very far apart (Moore 2005). Despite the abundance of statistical methods and tools for epistasis analysis (Van Steen 2012), the success rate of GWA studies cannot be matched with a comparable success rate of GWAI studies. Taking into account two major obstacles in GWAI analysis, namely the unavailability of sufficiently large sample sizes and the dramatically increased multiple testing burden, genome-wide screenings for epistasis usually end with no statistically significant findings at all. At best, top results (not necessarily reaching genome-wide significance) are followed up by pathway or enrichment analyses. In this paper, we describe a minimal protocol for genome-wide association interaction (GWAI) analysis with genome-wide SNP panels, to maximize the detection of multiple epistasis signals with biological relevance. The different steps of this protocol are summarized in Fig. 1 and discussed and motivated in subsequent sections. Without being exhaustive, each step of the protocol includes example software tools. A full explanation of abbreviations for these, with references, is given as a supplementary document (Supplementary Note 1). Our experience has shown that by taking advantage of various methodologies and by examining data from different angles, it is feasible to reveal strong biological gene-based relationships and synergies via SNP data.

Data cleaning and quality control (QC)

General QC principles apply as for genome-wide association studies. For a review and practical guidelines, see Anderson et al. (2010), de Bakker et al. (2008). In the context of GWAI screening, not only standard sample level QC such as those based on call rate, heterozygosity, relatedness, ethnicity and gender discrepancies need to be passed, but also additional sample and marker quality control considerations have to be made (Fig. 1, block 1).

Clearly, the information load of the data affects the power to detect epistasis and therefore we recommend to pay extra attention to identifying and dealing with genotyping errors, and to use stringent HWE tests based on Chi-square test statistics, or Fisher's exact tests when appropriate. These tests are carried out in the available control populations (Xu et al. 2002). Note that deviations from HWE in cases may actually indicate disease associations. Apart from genotyping errors, population genetic causes (e.g., inbreeding caused by consanguinity, assortative mating, selection and migration) may also lead to deviations from HWE. These generally play a minor role in well-designed genetic association studies, whether focusing on main effects, or combinations of these (Minelli et al. 2008), with the exception of population stratification. The latter is an important observation since statistical power can be increased by increasing sample sizes and hence pooling data collections from large consortiums. As a consequence, the available data for GWAI analysis may be highly heterogeneous, a data property that will be discussed in more detail in a subsequent section on meta-analysis. In addition, principal components, constructed as continuous axes of genetic variation, are not always easy to integrate with available GWAI software tools or methodologies. When population stratification is the reason for Hardy–Weinberg disequilibrium observations, approaches to deal with this type of confounding are much more subtle in GWAI studies compared to main effects analyses (see later). To maximize power and to avoid inadequate use of large-sample statistics on extremely sparse multilocus genotype combinations, we also recommend setting the minimum allowable minor allele frequency (MAF) in each study to 5 %. Our experience has learned that analyzing lower MAFs leads to highly elevated type I errors, unless similar strategies are adopted as currently being used and developed by rare variants analytic groups. PLINK is a free, open-source whole-genome association analysis toolset, designed to perform a wide range of basic GWA analyses, including the implementation of well-established QC protocols (Anderson et al. 2010). SVS is part of the commercial software Golden Helix and provides a comprehensive set of conventional and state-of-the-art data quality assurance tools with user-friendly visualization options. Finally, we recommend a post-analytic visual inspection of signal intensity plots for the identified interacting SNPs to ensure that no epistasis signals were caused by genotyping errors.

A note about the effect of genotyping errors on LD measures is given below. Simulations have shown that even small genotyping error rates can have profound consequences on LD measures, particularly when the frequency of the minor SNP allele is low (e.g., when MAF = 0.1 and there is a complete LD between a pair of SNPs ($D' = 1$, $r^2 = 1$) in the absence of genotyping errors, a 3 % error rate reduces D' and r^2 to 0.67 and 0.67, respectively, under a stochastic error model) (Akey et al. 2001). Therefore, in GWAs, LD structure of a trait-associated locus is frequently used to tease out genotyping artefacts that passed through QC procedure, for whatever reason. In GWAI analysis, we recommend examination of LD structures of a pair of interacting loci in the post-screening stage, i.e., when the genome-wide analysis has already been conducted.

Selecting an analytic strategy for GWAIS: exhaustive search or selective search in a prioritized (pre-filtered) marker set

It is important to achieve a balance between using as much information of the data as possible and keeping statistical limitations at a minimum. The use of information in the data can be maximized by adopting an exhaustive screening approach or a well-focused targeted screening that builds on reliable biological knowledge. An increased multiple testing burden with exhaustive searches and the presence of unknown confounders and unknown interaction effect modifiers require sophisticated statistical approaches that are characterized by optimal power and adequate false positive control under realistic scenarios.

Several criteria exist to categorize epistasis detection/modeling methods [reviewed in Van Steen (2012)]. Here, we use the categorization (a) exhaustive versus (b) non-exhaustive approaches (Fig. 1, analytical blocks 2a and 2b), since it reflects the current dilemma of choosing between a completely hypothesis-free search while analyzing all available data and a hypothesis-driven approach while analyzing only a limited set of markers. Often the latter choice is imposed by methodological (e.g., significance assessment) or IT infrastructure (e.g., computational burden) related aspects. The exhaustive search includes testing for all possible pair-wise interactions across all available genetic markers derived from genome-wide genotyping using micro-array technologies or exome/whole-genome sequencing efforts (usually involving between 0.5 and 1.5 million SNPs). Non-exhaustive approaches analyze a pre-selected set of markers.

Marker selection in non-exhaustive approaches is usually based on prior knowledge (e.g., biological knowledge about gene–gene interactions and network structures), and/or on desired properties of each single marker separately (e.g., functionality of a particular SNP or location of a SNP in an important candidate gene for a trait of interest). An interesting tool for gene annotation and prioritization, which allows system-level analysis and identification of disease-specific patterns of inheritance, and can in principle be used for a disease-specific pre-filtering of gene variants prior to gene–gene interaction testing or modelling is

the LYNX computational platform (Sulakhe et al. 2013). The annotation strategy of this tool is based on a large-scale integration of genomic and clinical data and various classes of biological information from over 35 public and private databases. These data are used for the identification of genes and molecular networks contributing to phenotypes of interest, as well as for the prediction of additional high-confidence disease genes. In addition, the analytical strategy of LYNX is based on (a) enrichment analysis of high-throughput genomic data by Bayes factor and *p* value estimates, (b) feature-based gene prioritization using support vector machines, and (c) the development of network-based disease models [using PINTA algorithms (Nitsch et al. 2011)]. Another interesting example of a marker pre-selection tool for epistasis screening is biofilter (Bush et al. 2009). This filtering tool allows for annotation of wide range of data (genes, SNPs, genomic locations, CNVs, etc.) and explicit detection and modeling of interactions between a large set of genetic markers, based on a wide spectrum of biological information related to gene–gene and gene–disease relationships (Bush et al. 2009). In particular, biofilter integrates twelve publicly available databases (including information about genomic locations of SNPs and genes, as well as relationships among genes and proteins such as interaction pairs, pathways and ontological categories) to produce multiple gene–gene models that have an established biological foundation. The knowledge-based support for the models is attributed by implication index, which reflects the number of data sources that provide evidence of gene–gene interaction or gene–disease relationship. Marker selection based on available knowledge about marker- or gene-based association signals can also be performed in eXtasy Variant Prioritization tool (Sifrim et al. 2013). It allows for ranking non-synonymous single-nucleotide variants given a specific phenotype. Another promising approach to prioritize genes with respect to their biological relevance is Data-driven Expression-Prioritized Integration for Complex Traits (DEPICT) (presented at ASHG 2013; http://www.ashg.org/2013meeting/abstracts/fulltext/f130123209.htm). It integrates expression microarrays, protein–protein interactions, gene-phenotype data from mouse knockout studies, and pathway databases to create a list of candidate causal factors (genes, pathways and related tissues) that may be indicative of a disease etiology.

An alternative to biological filtering statistical filtering approaches can be adopted. These data-filtering methods allow fast and exhaustive identification of potentially promising interacting SNPs. Examples include SIXPAC [based on LD-contrast between cases and controls (Prabhu and Pe'er 2012)], SNPRuler [based on a predictive rule learning approach (Wan et al. 2010b)], SNPHarvester [based on multiple path generation and a stochastic search for selection of SNP groups (Yang et al. 2009)], TEAM [based on a minimum spanning tree structure approach (Zhang et al. 2010a)], Screen and Clean [based on lasso procedure (Wu et al. 2010)], VariABEL [based on testing of variance heterogeneity of a phenotype among SNP genotypes (Struchalin et al. 2010)], and SURF or TuRF [based on ReliefF family algorithm for feature selection (Greene et al. 2009)].

It leaves no doubt that adopting a data pre-filtering/prioritization approach reduces the multiple testing burden and facilitates making biologically or clinically relevant interpretations. However, it is unclear to what extent pre-filtering hampers the detection of previously unreported SNP–SNP interactions. Therefore, more work is needed on selecting the most optimal criteria to avoid losing interesting data for subsequent analysis, and finding the optimal balance between data reduction and signal increase. Currently, when selecting a strategy for marker/gene prioritization, a good idea is to try a few of them taking into account capabilities, features and characteristics of every particular approach. For more information regarding pre-filtering prior to epistasis screening, we refer to the mini-review (Sun et al. 2014).

LD pruning

One of the underlying assumptions in genetic association studies is that there are some disease-causing loci which are in close proximity to some genetic markers in the genome, through which the disease causing signals can be detected and identified. Hence, the success of a genetic association study largely depends on the allelic associations between marker alleles and still to identify disease-causing alleles (Xu and George 2011). The impact of linkage disequilibrium (allelic association or LD) on epistasis is not as well investigated as it is in GWA studies. Whereas a single SNP association is dependent on strong LD between the marker SNP and causative variant, detection of epistasis requires strong LD for both loci (Wei et al. 2011). Motsinger et al. (2007) observed that strong patterns of LD actually increase the power of some epistasis detection methods, such as grammatical evolution neural networks. However, strong LD, whether it involves allelic associations between markers on the same or different chromosomes, induces strong correlations between variables in a dataset, which can complicate epistasis analysis [e.g., harmful multicollinearity (Van Steen et al. 2002)] and may lead to redundant interacting pairs (Moore et al. 2007) that unnecessarily increase the multiple testing burden. Therefore, we recommend performing LD pruning prior to (non-)exhaustive searches, to avoid an excess of redundant SNP–SNP interactions and thus to unnecessarily burden the computational aspects related to epistasis analysis. In our protocol, the LD r^2 filtering threshold is set to 0.75 (Fig. 1, analytical blocks

2a and 2b). This threshold is smaller than that is commonly used in main effects GWA contexts, to have a good compromise between information gain and control of computational burden [see for instance Gusareva et al. (2014a) for applications]. An extensive simulation study to identify the most optimal threshold for GWAI studies is underway.

When using data from different studies, SNP panels are often derived from different genotyping arrays. To obtain a common SNP panel in meta-GWAs, imputation has become the standard [reviewed in de Bakker et al. (2008)]. Imputation to a common set of SNPs has shown substantial bias though, which can be eliminated by imputing based on the intersection of genotyped SNPs (Johnson et al. 2013), with maintenance of adequate imputation quality. The magnitude of bias induced by statistical imputation of untyped SNP genotypes based on reference haplotype panels in the context of GWAI studies is largely unknown and is work in progress. Note that analytic epistasis detection strategies vary greatly in the way LD patterns influence increased false positives [redundant epistasis due to LD (Moore et al. 2007)], and that depending on the adopted imputation strategy, LD patterns may become more or less explicit. Until results from multiple studies regarding these issues become available, we recommend implementing a post-imputation filtering based on LD with an r^2 threshold of (only) 0.75, and interpreting meta-GWAI results based on imputed data (from different SNP panels) with extreme caution.

Adjustment for population stratification

Genetic ancestry may confound association results between genetic markers and the trait under investigation, especially in studies with unrelated individuals (e.g., case–control designs in which disease status is taken as a binary trait) (Fig. 1, analytical blocks 2a and 2b). Family-based association testing strategies such as FBAT (Horvath et al. 2001) offer protection against population stratification. Other family-based strategies may also provide such protection but—when genotypes are overmatched—at the expense of some loss of power (Thomas et al. 2005).

Whereas rigorously studied in the context of main effects or single gene/genetic marker associations, the effects of population stratification or admixture in large-scale epistasis screening efforts are not so well understood. For GWAI studies, simulation results seem to indicate an even larger impact of population stratification and admixture on generating spurious epistasis findings (study in progress). At the same time, it was shown that the likelihood of detecting true association between genetic variants and complex traits increases when studied in homogenous populations, thus suggesting that interactions are masked in heterogeneous populations (Fenger et al. 2008). In terms of population substructure handling, adjusting the analysis by a selection of principal components (PCs) is a widely used and accepted technique in main-effect studies (e.g., Heath et al. 2008; Novembre et al. 2008; Patterson et al. 2006; Reich et al. 2008). One of the reasons is that PCs are easily inserted as extra covariates in regression models, the most widely used models in GWA settings. However, choosing the number of PCs to retain is a non-trivial problem [see Peres-Neto et al. (2005) for a review], as well as deciding upon a fixed or variable number of PCs for the marker-trait combinations of interest (Peloso and Lunetta 2011). Moreover, not every regression-based epistasis detection software accommodates the inclusion of covariates (described in more details in the next section about the analytic tools). In fact, there are many more epistasis detection tools available that do not, due to their non-parametric nature (Van Steen 2012). And even when they do, it is not clear whether components capturing nonlinear relationships between genetic markers should be preferred over classic PCs. To date, linear PCs are believed to capture "confounding by population stratification" and represent continuous axes of genetic variation. They are either computed in the entire sample of individuals, or first computed in the control samples and second applied to (projected on) the case samples. But for a factor to act as a confounder in GWAI studies, it should be related to both the trait and the genetic exposures under investigation. Hence, it remains an unresolved question how to best capture confounding by population substructures in GWAI studies.

As mentioned before, a richer set of analytic approaches exist for GWAI analysis than for main effects GWA studies, the majority of these being data mining driven or non-parametric in nature. Hence, it may be more generic to combine features of the genome-wide rapid association using mixed model and regression approach [GRAMMAR (Aulchenko et al. 2007)] with the epistasis analytic tool of choice. In particular, instead of using the original trait of interest, residuals derived from a mixed regression model, in which relatedness is estimated via genomic kinship, are used for GWAI analysis. This approach has already been implemented in combination with MB-MDR (Cattaert et al. 2010). Although the mixed model estimation can be computationally challenging for large datasets, some efforts were already undertaken to increase its computational efficiency (called "compressed MLM") (Zhang et al. 2010b). Another disadvantage of the method includes the potential over-correction that may result in power loss and conservatism, as was observed by Cattaert et al. (2010). However, GWAI applications of this idea on admixed synthetic data give rise to spurious epistasis signals in the absence of gene–gene interactions (results not shown). These preliminary findings show that the problem of population substructure induced confounding to epistasis signals is highly

underappreciated. More work is needed to first understand the contribution of nonlinearity and interactions in characterizing admixed or structured populations, and second to propose optimal strategies to account for their potential confounding effects.

Statistical methods to screen for epistasis in large-scale genomic data

The growing interest in epistasis detection has caused a boom in methods development to identify pair-wise or higher-order SNP–SNP interactions. Although theoretically possible, only a few of these methods have resulted in software tools that make it practically feasible to screen among over at least half a million of genetic markers within a reasonable time-span (Table 1) (Van Steen 2012). Rigorous studies that compare the performance of promising (new) methods in relation to older methods, using several realistic scenarios of epistasis models and biologically complex reference data are largely lacking. In addition, honest and practically useful comparisons are hampered by the unavailability of standardized evaluation or performance criteria (e.g., the definition of "power" or "false positive") and the incomplete reporting of assumptions that implicitly underlie the method [see also argumentations in for instance Mahachie John et al. (2013), Van Steen (2012)]. Fortunately, several teams worldwide are putting their efforts together to create data simulators that will allow a consistent evaluation of methods on the basis of some minimal criteria. In what follows, we describe a few promising epistasis detection tools. For details about their performance, we refer to the initial references. The list is by no means exhaustive, but contains representatives of both parametric and non-parametric analytic tools. For excellent reviews in the context of large-scale epistasis analyses see for instance Cordell (2002), Miko (2008).

In general, the regression framework is seen as the most natural first-line approach when modeling or testing for genetic interactions (Cordell 2002) (Fig. 1, analytical block 2a and 2b), despite some associated and well-known difficulties it may generate. These difficulties may be technical in nature (e.g., when too many variables, possibly correlated, variables are included in the model or important confounders are left out from the model), computational in nature (e.g., when closed forms for test statistics do not exist) or may be more related to the interpretation of the regression-based signals [e.g., the interaction term has the interpretation of an effect modification; "interaction" and "modification" are not the same concepts (VanderWeele 2009)]. When adopting a particular coding scheme for genetic effects, an inherent assumption is being made about the corresponding modes of inheritance. In GWA

Table 1 Popular analytical tools for genome-wide exhaustive epistasis detections

Software	Core of the method	Traits	Covariates	Multiple testing correction	Lower-order effects adjustment	Missing genotype data accommodation	References
BOOST	Regression modeling (8 d.f. test)	Binary	–	–	+	–	Wan et al. (2010a)
BiForce	Regression modeling (4 d.f. test)	Binary and continuous	–	Bonferroni correction	+	+	Gyenesei et al. (2012)
epiGPU	*F* test-based and regression modeling (4 d.f. and 8 d.f. test)	Continuous	–	Permutations	+	+	Hemani et al. (2011)
EpiBlaster	Contrast of Pearson correlation	Binary and continuous	–	–	–	+	Kam-Thong et al. (2011a, b)
GLIDE	Regression modeling (4 d.f. test)	Binary and continuous	+	–	+	+	Kam-Thong et al. (2012)
MDR	Non-parametric, dimensionality reduction method	Binary, continuous, censored	–	Permutations	+	+	Ritchie et al. (2001)
MB-MDR	Non-parametric, dimensionality reduction method	Binary, continuous, censored, multivariate	–	Permutation-based maxT algorithm	+	+	Cattaert et al. (2010, 2011)

studies, biallelic markers are typically encoded using an additive model that has a reasonable power to detect both additive and dominant marginal effects (Bush and Moore 2012). However, such an additive coding, that implies a linear increase/decrease in disease risk or mean trait for each copy of the minor allele, seems to be too simplistic when interacting SNPs are envisaged. Adhering to a codominant coding scheme for genetic variants in GWAIS is believed to better agree with potentially complex biological interaction models. However, a clear mathematical argumentation still has to be performed (work in progress). One of the popular regression-based methods to search for epistasis associated with a binary outcome is BOolean Operation-based Screening and Testing (BOOST) (Wan et al. 2010a). This is a fast two-stage (screening and testing) approach that successfully uses algorithm of Boolean representation of the genotype data allowing for quick Boolean operations and thus efficiently speeds up calculations on the standard central processing units (CPU). The limitations of this method relate to its requirement to eliminate even moderate LD patterns in the data so as to keep the type I error and false positives under control, to its inability to accommodate missing data (hence input data are assumed to be imputed—hereby inducing additional correlations between markers) and the necessity to perform a multiple testing correction outside the software package. More recent methods that also exploit efficient bitwise data restructuring, Boolean bitwise operations and multithreaded (and/or multi-core) parallelization on standard CPUs include BiForce (Gyenesei et al. 2012). This regression-based tool is applicable to both binary and continuous outcomes, can efficiently accommodate data with missing genotypes (thus does not assume imputed input marker data), and has a built-in multiple testing strategy (the Bonferroni correction; although far from ideal in the context of exhaustive genome-wide epistasis screening). Based on a selective set of simulation scenarios, BiForce seems to outperform BOOST or the heavily used PLINK software using standard regression modeling (Purcell et al. 2007) in terms of statistical power and speed.

With increasing SNP densities, investments in speed become more relevant. It is therefore not surprising that novel epistasis detection methodologies take advantage of what parallel computing and/or graphical processing units (GPU) can offer regarding computational burden reduction. For instance, MB-MDR has been adapted to accommodate parallelized operations, hereby allowing larger SNP sets to be analyzed than with earlier sequential versions (Van Lishout et al. 2013). The more novel software tool GLIDE (Kam-Thong et al. 2012), which involves a linear-regression models and implements the same algorithm as PLINK epistasis analysis (Purcell et al. 2007), is faster than PLINK by a factor of 2,000 due to adoption of the parallel computations on GPU. Although not all facilities are able to run GPU-dependent software, these efforts make exhaustive epistasis screening on a genome-wide scale possible. The performance GLIDE is slightly lower compared to BOOST. However, the important advantage of the GLIDE approach is its possibility to perform quantitative trait epistasis studies (apart from binary trait studies—the only possible outcome type for BOOST) and its ability to adjust analyses for various continuous predictors or confounders (genetic or environmental). The advantages of GPU were also successfully implemented in epiGPU (Hemani et al. 2011) (an *F* test based method, currently restricted to continuous traits) and EpiBlaster [a Pearson's correlation coefficients-based method, applicable to binary outcomes (Kam-Thong et al. 2011a) or continuous traits (Kam-Thong et al. 2011b)]. Recently, another hardware technology, field-programmable gate array (FPGA), was used for problems related to bioinformatics. This technology was applied to outperform the iLOCi method (interaction prioritization algorithm based on LD differences between cases and controls) (Piriyapongsa et al. 2012; Wienbrandt et al. 2014), but in principle it can be adapted to many existing methodologies for exhaustive association interaction analysis.

Parametric model (mis)specification is of major concern, especially in the presence of high-dimensional confounders (Vansteelandt et al. 2012). In addition, sensitivity to harmful multicollinearity (Sithisarankul et al. 1997; Slinker and Glantz 1985; Van Steen et al. 2002) induced by strong LD patterns may cause increased numbers of false positives or type I errors. Also, "small *n* big *p*" (*n*: number of subjects, *p*: number of variables/genetic markers) problems may give rise to curse of dimensionality problems (Bellman and Kalaba 1959). This led to the exploration of a variety of data mining approaches for epistasis searches (Fig. 1, analytical block 2b), including multifactor dimensionality reduction (MDR) (Hahn et al. 2003; Ritchie et al. 2001, 2003). These and similar non-parametric methods are good alternatives to traditional regression-based methods in that no assumption needs to be made about the epistatic mode of inheritance (which is extremely useful when there is no a priori knowledge of the genetic system), while at the same time maintaining power. The MDR method was initially designed to identify interactions among discrete variables in relation to binary outcomes. The basis of the MDR is a constructive induction algorithm [described in Ritchie et al. (2001), and inspired by a combinatorial partitioning method (Nelson et al. 2001)] that converts two or more variables or attributes into a single lower-dimensional attribute. This process of constructing a new attribute, changes the representation space of the data. The end goal is to create or to identify a representation that facilitates the detection of non-linear or non-additive interactions among the attributes such that prediction of the class variable is improved over that of the original representation of the

data. Data mining algorithms are good at finding patterns in completely random data, however, it is often difficult to determine whether a reported pattern is an important signal or just chance finding. Over-fitting issues in MDR are solved via cross-validation and permutations. Since the conception of MDR, several adaptations have been made (Van Steen 2012), extending the initial approach to accommodate continuous traits, censored traits, covariate adjustments, and gene-based analysis. In contrast, MB-MDR is another data dimensionality approach that breaks with the tradition of cross-validation and rather invests computational efforts in permutation-based multiple multilocus significance assessments and the implementation of the most appropriate association test for the data at hand (e.g., accommodating binary, continuous, censored, and multivariate traits within a single framework). Its unique way to create lower-dimensional constructs via selected robust association tests, hereby acknowledging that not all multilocus genotype combinations are informative, has proven to give optimal performance compared to MDR, especially in the presence of genetic heterogeneity (Cattaert et al. 2011). Different from MDR, both a global and a targeted epistasis specific test can be performed in MB-MDR. The latter is accomplished in MB-MDR by appropriately adjusting the association tests for lower-order effects. This explains the model-based (MB) component in MB-MDR. This strategy has been shown to avoid main effects signals giving rise to false epistasis effects (Mahachie John et al. 2012). Although the MB-MDR software (Van Lishout et al. 2013) has a built-in significance assessment strategy based on a permutation-based maxT algorithm (Westfall and Young 1993b) other multiple testing correction strategies can be incorporated as well. One of these, implemented in MB-MDR-4.2.2 version, upscales the MB-MDR approach to genome-wide epistasis screening studies. Currently, MB-MDR uses all available genotype data and accommodates a variety of study designs: unrelated or related individuals, or mixtures thereof, as well as apparently unrelated samples with cryptic relationships (Cattaert et al. 2010). Also, adjustments for covariates or important confounders (such as those capturing population substructure) can be incorporated in the MB part of MB-MDR (Cattaert et al. 2010).

Whereas regression and data mining approaches belong to the most commonly used analytic tools for large-scale epistasis detection, information-theoretic measures could never really reach the same popularity levels. Yet, interesting measures exist such as the k-way interaction information (KWII) (Chanda et al. 2007), the phenotype-associated information (PAI) (Chanda et al. 2008) and the interaction index (IID) (Chanda et al. 2009). As these and other entropy-based approaches may point towards interesting interactions for follow-up, they often do not provide levels of significance. However, we do not think that this is the main explanation for their limited use in large epistasis screening. Indeed, the heavily used Random Forests as a data mining approach (Schwarz et al. 2010) also does not assess significance for sets of variables, but provides individual variable importance scores and threshold above which to retain variables (r2VIM—recurrent relative variable importance scores; personal communication with Silke Szymczak). Notably, patterns obtained via Random Forests may be the result of random variations or of the recursive nature of the tree building algorithms, or of true interactions. Currently, it is not obvious how to make a distinction between these scenarios. Hence, a potential explanation for the limited use of entropy-based measures in large-scale epistasis screening is that they are often not readily implemented in easy-to-use computationally efficient software tools that in addition accommodate a variety of trait types. An interesting information-theoretic method, although limited to relatively small numbers of markers and binary traits, is the synergy disequilibrium (SD) plot (Anastassiou 2007; Watkinson and Anastassiou 2009). In such a plot, the synergy between two SNPs Si and Sj with respect to a disease C (or any phenotype or trait) refers to the amount of information conveyed by the pair of SNPs about the presence of the disease, minus the sum of the corresponding amounts of information conveyed by each SNP: I(Si,Sj;C) − [I(Si;C) + I(Sj;C)] (Anastassiou 2007; Watkinson and Anastassiou 2009). Large positive synergy (depicted by red dots in a standard SD plot) quantifies the amount of association between two SNPs and a phenotype that is due to purely cooperative effects among the factors. A blue dot for a SNP in a standard SD plot indicates that the corresponding SNP in itself is associated with disease (i.e., exhibits a main effect). Negative synergy for a pair of SNPs (depicted by blue dots in a standard SD plot), indicates that both SNPs in combination do not significantly enhance the association with the trait; the two SNPs are redundantly associated with disease.

Data pooling and meta-analysis in the context of epistasis screening

Common variants for complex diseases usually only have individual modest effects, and often involve odds ratios of <1.2 for dichotomous traits, or explained variances of <1 % for quantitative traits (de Bakker et al. 2008; Fellay et al. 2009). The power of a GWAI analysis using a single study is conceptually smaller than a corresponding main effects GWA analysis using the same data. In the pure epistasis models (when no individual marginal effects present for a pair of genes), simulations showed that more than 4,000 cases and controls or about 1,200 family trios are required to achieve 80 % power to detect an epistasis signal with an

odds ratio of 1.5 (assuming that the tested genetic markers are frequent in a population) (Gauderman 2002). For rare-allele markers (with MAF = 0.01), sample sizes have to be orders of magnitude larger (i.e., about 140,000 cases and controls or 47,000 family trios to achieve 80 % power to detect an epistasis signal with an odds ratio of 3.0) (Gauderman 2002). Therefore, the development of sound meta-analytic methodologies that would allow combining evidence across independent gene–gene interaction studies, taking into account the specific characteristics of individual epistasis studies, as well as protocols that outline good standard practice when performing such meta-GWAI studies, are needed (Fig. 1, block 3).

Meta-analysis of main effects GWA studies use two major methodologies: combination of effect sizes based on fixed effects or random effects models (Fleiss 1993) and combination of p values (or weighted z-scores) (Fisher 1948). Whereas the latter is suboptimal, it may be the first choice in meta-GWAI studies when the epistasis analytic method does not readily provide effect sizes or effect estimates (as is the case for MB-MDR). Moreover, methods such as Random Forests provide a ranking rather than a p value in which case it is not clear how to best perform a Random Forests meta-GWAI analysis. In fact, due to the diversity in adopted epistasis analytic tools, study heterogeneity in meta-GWAIs not only relates to different populations or study design protocols, but also to different analytic tools, that may vary in the way population substructure is accounted for, or how multiple testing is dealt with, or whether (perhaps more powerful, but certainly more focused) prior knowledge-based targeted searchers were performed. In addition, whereas in GWA settings, each SNP is usually assumed to have an additive genetic effect, in the absence of clear biological hypotheses (Ziegler and König 2006), a multitude of coding schemes exist and may be competing in GWAI contexts. Hence, also differential coding schemes between different GWAI studies (e.g., study 1 is analyzed with BOOST which assumes a codominant genotype coding; study 2 is analyzed with EpiBlaster which assumes an additive coding as well as a different methodology) further complicate meta-GWAI studies.

When the joint action of multiple genetic markers is of interest, it is unrealistic to assume a particular mode of action and thus model misspecification is a major concern (Pereira et al. 2011, 2009). An increasing number of authors promote a codominant coding scheme for SNPs in GWAI analysis [for instance, Mahachie John et al. (2012), Wan et al. (2010a)]. Hence, multilocus genotype (MLG) regression-based association interaction analyses will generate for each SNP pair eight multilocus genotypes (while the reference category is homozygous for the major alleles at both SNPs). As a consequence, when several such MLG studies are performed, the study-specific 8 effect sizes per SNP pair can be meta-analyzed (e.g., Gusareva et al. 2014a). Alternatively, p values from the appropriate 4 degrees of freedom likelihood ratio tests are combined, at the risk of losing refined information on the within multilocus architecture. However, such an approach is motivated and preferred in the belief that epistasis replication at the level of particular multilocus genotype combinations makes little sense. Nevertheless, awareness that suboptimal genetic models of analysis may cause dramatic loss of power has led to the development of more sophisticated GWA approaches (Minelli et al. 2005). Unfortunately, due to often unverifiable assumptions or increased computational intensity, these more advanced meta-analytic methods are not widely used in GWAs, let alone in GWAIs. In the context of meta-GWAIs, there is a clear need for model-free approaches that require no assumptions about the genetic models of action.

In the context of meta-analysis of GWAI studies, a special word is needed towards how to deal with uncertainty related to by-study imputations. Aulchenko et al. (2010) pointed out that an association method, which does not incorporate the uncertainty of imputed genotypes in the model, is biased and underpowered (Aulchenko et al. 2010). In GWA studies, a number of association methods that directly use genotype probabilities obtained after an imputation procedure have been developed. For instance, those based on approximate population genetics model for binary outcomes (Marchini et al. 2007), methods based on linear, logistic, and Cox proportional hazards models for quantitative, binary, and time-till-event outcomes, respectively [ProbABEL R package (Aulchenko et al. 2010)], approaches based on generalization of the Kruskal–Wallis test for binary outcomes (Acar and Sun 2013), and methods based on incorporation of uncertainty in the likelihood function when performing association analysis for longitudinal outcomes (Subirana and Gonzalez 2013). Accounting for imputation uncertainty in meta-analysis of GWAI studies has not been studied that extensively.

Notably, for GWAI settings, any regression-based technique for epistasis analysis can in principle be extended to incorporate the uncertainty of imputed genotypes, by regressing an outcome of interest onto posterior probability distributions in a manner similar to ProbABEL implementations. Some non-parametric methods, for instance MB-MDR extended to genetic regions of interest, can also incorporate estimated genotypic probabilities to perform association interaction analyses. Zhang (2011) proposed an interesting method based on a new Bayesian model for joint SNP imputation and epistasis association mapping. This method performs imputation conditional on the disease association status of SNPs, which is claimed to be more appropriate than imputing SNPs around disease loci.

In particular, SNP blocks are used to account for variable LD patterns among SNPs and untyped SNPs are imputed iteratively using Markov chain Monte Carlo (MCMC) algorithms. Non-ignorable power gains could be shown for detecting epistasis with the joint imputation and mapping strategy, compared to detecting single SNP effects, despite the elevated multiple testing burden due to imputation.

In summary, accounting for possible study-specific population stratification and LD patterns, for different multiple testing corrections, for different modes of low-order genetic effects (e.g., main effects), for possible study-specific confounder adjustments, for differential study designs, and hence different epistasis detection analytic tools, will be hard in a meta-GWAI context, if not impossible. When individual studies are weighted in a meta-GWAI analysis, weights ideally not only refer to sample size, but also to the expected power of the association test, which may depend on data quality (e.g., imputation quality scores for GWA studies), study design, the selected association test, and whether or not confounders were accounted for.

Biological and methodological replication of epistasis

It is widely accepted that association results from GWA studies should be supported by replication analysis in an independent sample (using the same analysis protocol; "replication by methodology") (Fig. 1, block 3) and/or by physiologically meaningful data, confirming a functional role of the polymorphism in question ("biological or functional validation") (Fig. 1, block 4). In the same spirit, current expectations require that positive findings from GWAI studies are replicated in independent data sets. Although the motivation of such a replication effort is clear (to ensure that the results are not artifacts caused by publication bias, selection bias, nor population stratification, etc.), it is less clear what the level of detail in a GWAI replication analysis should be. In the presence of highly complex networks of biochemical processes and genetic heterogeneity, is it reasonable to assume replication at the marker level? Should replication not be established at the gene/pathway level instead? Indeed, considering that most of the genotyping arrays are comprised with common tagging SNPs/genetic markers (i.e., most of them are "non-causal" and have no functional consequences) it is highly unlikely that the same combination of tagging genetic markers is associated with a trait of interest at the same level and in the same statistical model. We promote to append every GWAI analysis with a replication analysis on independent data, where "replication" is performed at the level of regions of interests that make sense for the genetic markers that gave significant signals in the first GWAI analysis. But which method to use? In theory, the replication analysis should involve the same analysis method and analysis conditions. Various statistical methods for discovering gene–gene interactions exist, each possibly address quite different aspects of the true biological epistasis mechanism. Hence, "replication by methodology" only makes sense when the behaviors of these methodologies are fully understood under a variety of theoretical and real-life settings: why would one replicate with a suboptimal method? This requires the development of realistic synthetic data sets, with complex genetic architectures, that can be promoted as reference data for methods comparisons. In addition, since simulation settings may be far off from the complexity of human disease models, extra work is needed in which the robustness of results is shown by using different methods (or options within the same method) on the same real-life data set, keeping in mind the strengths and weaknesses of each method as they are known to date.

Replication and validation efforts in independent data may be desirable, biological validation procedures (Fig. 1, block 4) are crucial to make important contributions in disease diagnostics and disease management. These procedures may rely in part on a systematic epistasis literature review. Most of the time, structured knowledge from databases (e.g., functional annotations, immunological pathways, eQTLs, DNA transcription factor binding sites, composite elements binding sites, etc.) is used to integrate data from a variety of experimental platforms. By doing so, additional insight is gained into underlying molecular and chemical interactions, cellular phenotypes, and disease processes relevant to the study. There are a number of platforms and tools, both commercial and publically available, which integrate a vast amount of biological knowledge and can facilitate the biological interpretation of statistical findings of epistasis. For instance, the IPA software (Ingenuity Systems, Inc.) integrates data from a variety of experimental platforms (transcriptomics, biomarkers, microRNAs, toxicogenomics, metabolomix, pharmacogenomics, and proteomics) providing insight into the molecular and chemical interactions, cellular phenotypes, and disease processes thus helping to understand connections between diseases and gene networks. Another similar publically available tools that allow interpreting experimental results in the context of a large cross-organism compendium of functional predictions and networks are Cytoscape (Shannon et al. 2003), CPDB (ConsensusPathDB-human) (Kamburov et al. 2009, 2011, 2013; Pentchev et al. 2010), GeneMANIA (Warde-Farley et al. 2010), BioGraph (Liekens et al. 2011), Integrative Multi-species Prediction (IMP) web server (Wong et al. 2012), etc. These tools integrate wide variety of data about protein and genetic interactions, pathways, co-expression, co-localization, protein domain similarity, microRNA, gene ontology, functional annotation, etc. Understanding mechanisms of gene regulation at the

level of transcription and information about transcriptional regulatory elements can also be very helpful to infer mechanisms of gene–gene interactions. Positions of transcription factor (TF) binding sites in human genome as well as in some other model organisms can be retrieved from the TRANSFAC® Professional database (one of the most comprehensive collection of experimentally determined TF binding sites) by P-Match™ tool (Combined Pattern-Matrix Search for Transcription Factor Binding Sites) (Chekmenev et al. 2005). Potential composite elements (CEs) for TFs in any DNA sequence can be predicted using the MatrixCatch tool (Deyneko et al. 2013).

Alternatively, a statistical epistasis network is built directly from GWAI results, in which nodes represent genes and edges capture the strength of statistical gene–gene interactions. Such an approach assumes having aggregated information about gene–gene interactions (based on SNP–SNP interactions). Important modules derived from the statistical epistasis network can be analyzed for their biological relevance, biological conservation or pathway enrichment. In addition, such a global interaction map facilitates the search for higher-order (>2) interactions by prioritizing genetic attributes clustered together in the network (Hu et al. 2013).

Finally, biological validation is performed in the lab via biological experiments (knockouts) in animal models. The use of model organisms overcomes some limitations of human genetic studies. The availability of genetically homogenous model organisms, the possibility to manipulate their genomes through selective breeding strategies, along with direct gene-targeting approaches and the ability to control environment to reduce phenotypic variance give model organisms considerable power to predict complex traits/disease susceptibility genes in humans (Gregersen et al. 2006; Gusareva et al. 2014b).

Perspective for rare variants in epistasis studies

Rare variants are those genetic variants that appear in <1 % of the population. These variants are sometimes private mutations that only appear in a few individuals or families. Several rare variants are autosomal and fully penetrant and have been associated to Mendelian disorders (Touitou et al. 2013; Wain 2014). Nevertheless, the abundance of rare variants identified via next-generation sequencing (NGS) efforts raises the question about whether or not also rare variants are involved in synergetic interactions. Epistasis detection with rare variants is even more challenging than with common SNPs, primarily due to the very low statistical power to identify these interactions, and the lack of methods for interaction analysis with rare variants. It remains to be investigated, whether the actual effect size of epistasis involving rare variants is higher or has a larger contribution to disease risk compared to SNP–SNP interactions. If so, provided adequate analytic tools are developed, epistasis involving rare variants or genes involving collections of rare and common variants will become detectable. If not, its detection via statistical methods for interactions and its replication will be extremely difficult (cfr. the large number of statistical challenges described in the previous sections; challenges that may potentially have an even more severe impact with rare variants than with SNPs). This also holds for the interpretation, especially when interactions refer to non-coding genomic regions or positions distant from well-known regulatory loci. Despite this skepticism, research in the field of rare variant epistasis analysis has already started. To facilitate interaction analysis with NGS data, Zhang et al. (2014) proposed to shift the paradigm of interaction analysis from pair-wise testing between genetic markers to genomic regions as a basic unit of interaction analysis and use high-dimensional data reduction and functional data analysis techniques to develop a novel functional regression model to collectively test for interactions between all possible pairs of SNPs within pairs of genome regions (presented at ASHG 2013; http://www.ashg.org/2013meeting/abstracts/fulltext/f130120242.htm). A region-based approach for NGS data epistasis analysis is also adopted by genomic MB-MDR (presented at ERCIM 2013 and HGV 2014; http://www.cmstatistics.org/ERCIM2013/docs/BoA.pdf), building on the MB-MDR framework. Another study, which assesses individual and cumulative effects of rare variants in families with atrial fibrillation, also supports the idea of adopting a global perspective in interpreting human genome sequence data and looking beyond a single-variant analysis (Mann et al. 2012). The latter study clearly demonstrates how instrumental an elaborate application of different biological and statistical approaches (i.e., comprehensive genetic screening, cellular electrophysiological data, atrial cell modeling, and systems biology approaches) can be in identifying interactions between rare variants.

Conclusion

The identification and detection of epistasis is a challenging task that, when successful, is believed to give new clues to systems-level genetics and a better understanding of the underling biology of human complex traits. Though many novel methods for detecting epistasis have been proposed and many studies for epistasis detection have been conducted, so far only a few studies can demonstrate replicable epistasis. In the present work, we described a comprehensive GWAI analysis protocol that involves screening for epistasis over large-scale genomic SNP panels, hereby

combining strengths of different methods and statistical tools. Issues regarding each of these steps were described in full detail and a schematic overview of the protocol was provided in Fig. 1. In summary, a rigorous quality control step is followed by an exhaustive or non-exhaustive screening methodology. Several strategies exist to incorporate functional and biological knowledge or prior information about complex biological architectures to narrow down the test space of alternatives and thus to reduce the multiple testing burden. The selection of the analytic tool is often driven by availability of efficient IT infrastructures and bioinformaticians who understand and are able to adapt or customize existing software tools to the needs of the user, the study design and the specific characteristics of the data. Once an analytic tool is selected, the results need to be interpreted at a more global level than the SNP level, and at the background of the tool's characteristics, strengths and weaknesses. Maintaining a genome-wide significance level of 0.05 remains an issue and no recommendation can be given towards the most optimal multiple testing corrective method, due to the variety of epistasis tests, modeling approaches and contexts (e.g., exhaustive or non-exhaustive, discovery or replication stage, presence of no, weak or moderate LD between genetic markers, etc.). The Bonferroni method is notorious for being suboptimal (Gusareva et al. 2014b), but usually, analytic tools come with a particular method for multiple testing correction and are not flexible in choosing an alternative methodology. Some alternatives look for the effective number of tests [e.g., Nyholt (2004), use higher criticism thresholding (Anderson et al. 2010; Donoho and Jin 2009), or implement permutation-based significance assessments (maxT method) (Westfall and Young 1993a), etc.]. The same level of scrutiny is adopted when setting up and carrying out an epistasis replication study with independent data. As a last step of the protocol, efforts are combined to support the statistical findings with biological and functional data.

It is highly unlikely that one method will fit all scenarios in GWAI studies. However, being able to asses which method performs well under which circumstance will accelerate the field. Currently, novel methods are evaluated using privately generated synthetic data, exhibiting overly simplified characteristics. We promote the creation of complex consensus GWAI synthetic data. This will greatly facilitate making honest comparisons between methods and/or identifying the true context-dependent benefits of each method. Combining multiple classification or regression models typically gives improved results compared to using only a single such model (Schwarz et al. 2010). Along the same line, each analytic epistasis detection tool can be envisaged to partition the (SNP–SNP) interaction space into "interesting" regions, according to some prespecified criteria or variables (which could include power to detect the interaction with the tool, biological interaction evidence, etc.). Similar to ensemble clustering (Strehl and Ghosh 2003), multiple tool-dependent partitions of the same interaction space can be combined so as to give a single partitioning solution of improved quality, borrowing strengths from several epistasis detection strategies.

Finally, despite the increasing number of investigators performing an epistasis analysis, the importance of epistasis in complex disease genetics still is the subject of heavy debate. Model organisms show that gene–gene interactions are important in explaining biological processes and in the ability to reveal the genetic secrets underlying complex traits. There is no reason to assume that this would not be the case for humans. Quite on the contrary, the increased complexity of human biology compared to the biology of model organisms requires investing in sophisticated epistasis detection methods, creating consensus criteria for their evaluation, and bringing awareness about pros and cons of each method. The presented minimal protocol and its discussion aim to contribute to creating such awareness, and promote viewing the epistasis problem from different angles. Only then we will be able to show what the impact of epistasis is on personalized medicine, disease risk prediction, and evolutionary genetics and will we obtain a more thorough understanding of biological and biochemical human complex disease mechanisms.

Acknowledgments The research was funded by the Belgian Science Policy Office Phase VII IAP network "Dynamical systems, control and optimization" (DYSCO II) and the Fonds de la Recherche Scientifique (FNRS). We thank our collaborators Jean-Charles Lambert and Céline Bellenguez (INSERM U744, Institut Pasteur de Lille, Universite de Lille Nord de France, Lille, France), Nilüfer Ertekin-Taner (Mayo Clinic Florida, Department of Neuroscience, Jacksonville, FL, USA; Mayo Clinic Florida, Department of Neurology, Jacksonville, FL, USA), Denise Harold and Julie Williams (GERAD1 Consortium; Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University School of Medicine, Cardiff, UK) and our colleagues François Van Lishout, Jestinah M. Mahachie John and Kyrylo Bessonov from the Systems and Modeling Unit, Montefiore Institute, University of Liege, Belgium without whom the development of the GWAI protocol would not have been possible.

Conflict of interest The authors declare that they have no competing interests.

References

Acar EF, Sun L (2013) A generalized Kruskal–Wallis test incorporating group uncertainty with application to genetic association studies. Biometrics 69:427–435. doi:10.1111/biom.12006

Akey JM, Zhang K, Xiong M, Doris P, Jin L (2001) The effect that genotyping errors have on the robustness of common linkage-disequilibrium measures. Am J Hum Genet 68:1447–1456. doi:10.1086/320607

Anastassiou D (2007) Computational analysis of the synergy among multiple interacting genes. Mol Syst Biol 3:83. doi:10.1038/msb4100124

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT (2010) Data quality control in genetic case–control association studies. Nat Protoc 5:1564–1573. doi:10.1038/nprot.2010.116

Aulchenko YS, de Koning DJ, Haley C (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 177:577–585. doi:10.1534/genetics.107.075614

Aulchenko YS, Struchalin MV, van Duijn CM (2010) ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinform 11:134. doi:10.1186/1471-2105-11-134

Bellman R, Kalaba R (1959) A mathematical theory of adaptive control processes. Proc Natl Acad Sci USA 45:1288–1290

Bloom JS, Ehrenreich IM, Loo WT, Lite TL, Kruglyak L (2013) Finding the sources of missing heritability in a yeast cross. Nature 494:234–237. doi:10.1038/nature11867

Bush WS, Moore JH (2012) Chapter 11: genome-wide association studies. PLoS Comput Biol 8:e1002822. doi:10.1371/journal.pcbi.1002822

Bush WS, Dudek SM, Ritchie MD (2009) Biofilter: a knowledge-integration system for the multi-locus analysis of genome-wide association studies. Pac Symp Biocomput 368–379

Carlborg O, Haley CS (2004) Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618–625. doi:10.1038/nrg1407

Cattaert T, Urrea V, Naj AC, De Lobel L, De Wit V, Fu M, Mahachie John JM, Shen H, Calle ML, Ritchie MD, Edwards TL, Van Steen K (2010) FAM-MDR: a flexible family-based multifactor dimensionality reduction technique to detect epistasis using related individuals. PLoS One 5:e10304. doi:10.1371/journal.pone.0010304

Cattaert T, Calle ML, Dudek SM, Mahachie John JM, Van Lishout F, Urrea V, Ritchie MD, Van Steen K (2011) Model-based multifactor dimensionality reduction for detecting epistasis in case–control data in the presence of noise. Ann Hum Genet 75:78–89. doi:10.1111/j.1469-1809.2010.00604.x

Chanda P, Zhang A, Brazeau D, Sucheston L, Freudenheim JL, Ambrosone C, Ramanathan M (2007) Information-theoretic metrics for visualizing gene–environment interactions. Am J Hum Genet 81:939–963. doi:10.1086/521878

Chanda P, Sucheston L, Zhang A, Brazeau D, Freudenheim JL, Ambrosone C, Ramanathan M (2008) AMBIENCE: a novel approach and efficient algorithm for identifying informative genetic and environmental associations with complex phenotypes. Genetics 180:1191–1210. doi:10.1534/genetics.108.088542

Chanda P, Sucheston L, Zhang A, Ramanathan M (2009) The interaction index, a novel information-theoretic metric for prioritizing interacting genetic variations and environmental factors. Eur J Hum Genet 17:1274–1286. doi:10.1038/ejhg.2009.38

Chekmenev DS, Haid C, Kel AE (2005) P-Match: transcription factor binding site search by combining patterns and weight matrices. Nucleic Acids Res 33:W432–W437. doi:10.1093/nar/gki441

Cordell HJ (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 11:2463–2468

de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF (2008) Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet 17:R122–R128. doi:10.1093/hmg/ddn288

Deyneko IV, Kel AE, Kel-Margoulis OV, Deineko EV, Wingender E, Weiss S (2013) MatrixCatch—a novel tool for the recognition of composite regulatory elements in promoters. BMC Bioinform 14:241. doi:10.1186/1471-2105-14-241

Donoho D, Jin J (2009) Feature selection by higher criticism thresholding achieves the optimal phase diagram. Philos Trans A Math Phys Eng Sci 367:4449–4470. doi:10.1098/rsta.2009.0129

Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH (2010) Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11:446–450. doi:10.1038/nrg2809

Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, Urban TJ, Zhang K, Gumbs CE, Smith JP, Castagna A, Cozzi-Lepri A, De Luca A, Easterbrook P, Gunthard HF, Mallal S, Mussini C, Dalmau J, Martinez-Picado J, Miro JM, Obel N, Wolinsky SM, Martinson JJ, Detels R, Margolick JB, Jacobson LP, Descombes P, Antonarakis SE, Beckmann JS, O'Brien SJ, Letvin NL, McMichael AJ, Haynes BF, Carrington M, Feng S, Telenti A, Goldstein DB, Immunology NCfHAV (2009) Common genetic variation and the control of HIV-1 in humans. PLoS Genet 5:e1000791. doi:10.1371/journal.pgen.1000791

Fenger M, Linneberg A, Werge T, Jorgensen T (2008) Analysis of heterogeneity and epistasis in physiological mixed populations by combined structural equation modelling and latent class analysis. BMC Genet 9:43. doi:10.1186/1471-2156-9-43

Fisher R (1948) Combining independent tests of significance. Am Stat 2:30

Fleiss JL (1993) The statistical basis of meta-analysis. Stat Methods Med Res 2:121–145

Gauderman WJ (2002) Sample size requirements for association studies of gene–gene interaction. Am J Epidemiol 155:478–484

Greene CS, Penrod NM, Kiralis J, Moore JH (2009) Spatially uniform relieff (SURF) for computationally-efficient filtering of gene–gene interactions. BioData Min 2:5. doi:10.1186/1756-0381-2-5

Gregersen JW, Kranc KR, Ke X, Svendsen P, Madsen LS, Thomsen AR, Cardon LR, Bell JI, Fugger L (2006) Functional epistasis on a common MHC haplotype associated with multiple sclerosis. Nature 443:574–577. doi:10.1038/nature05133

Gusareva ES, Carrasquillo MM, Bellenguez C, Cuyvers E, Colon S, Graff-Radford NR, Petersen RC, Dickson DW, Mahachie John JM, Bessonov K, Van Broeckhoven C, the GC, Harold D, Williams J, Amouyel P, Sleegers K, Ertekin-Taner N, Lambert JC, Van Steen K et al (2014a) Genome-wide association interaction analysis for Alzheimer's disease. Neurobiol Aging. doi:10.1016/j.neurobiolaging.2014.05.014

Gusareva ES, Kurey I, Grekov I, Lipoldova M (2014b) Genetic regulation of immunoglobulin E level in different pathological states: integration of mouse and human genetics. Biol Rev Camb Philos Soc 89:375–405. doi:10.1111/brv.12059

Gyenesei A, Moody J, Semple CA, Haley CS, Wei WH (2012) High-throughput analysis of epistasis in genome-wide association studies with BiForce. Bioinformatics 28:1957–1964. doi:10.1093/bioinformatics/bts304

Hahn LW, Ritchie MD, Moore JH (2003) Multifactor dimensionality reduction software for detecting gene–gene and gene–environment interactions. Bioinformatics 19:376–382

Heath SC, Gut IG, Brennan P, McKay JD, Bencko V, Fabianova E, Foretova L, Georges M, Janout V, Kabesch M, Krokan HE, Elvestad MB, Lissowska J, Mates D, Rudnai P, Skorpen F, Schreiber S, Soria JM, Syvanen AC, Meneton P, Hercberg S, Galan P, Szeszenia-Dabrowska N, Zaridze D, Genin E, Cardon LR, Lathrop M (2008) Investigation of the fine structure of European populations with applications to disease association studies. Eur J Hum Genet 16:1413–1429. doi:10.1038/ejhg.2008.210

Hemani G, Theocharidis A, Wei W, Haley C (2011) EpiGPU: exhaustive pairwise epistasis scans parallelized on consumer level

graphics cards. Bioinformatics 27:1462–1465. doi:10.1093/bioinformatics/btr172

Horvath S, Xu X, Laird NM (2001) The family based association test method: strategies for studying general genotype–phenotype associations. Eur J Hum Genet 9:301–306. doi:10.1038/sj.ejhg.5200625

Hu T, Andrew AS, Karagas MR, Moore JH (2013) Statistical epistasis networks reduce the computational complexity of searching three-locus genetic models. Pac Symp Biocomput 397–408

Johnson EO, Hancock DB, Levy JL, Gaddis NC, Saccone NL, Bierut LJ, Page GP (2013) Imputation across genotyping arrays for genome-wide association studies: assessment of bias and a correction strategy. Hum Genet 132:509–522. doi:10.1007/s00439-013-1266-7

Kamburov A, Wierling C, Lehrach H, Herwig R (2009) ConsensusPathDB—a database for integrating human functional interaction networks. Nucleic Acids Res 37:D623–D628. doi:10.1093/nar/gkn698

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R (2011) ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res 39:D712–D717. doi:10.1093/nar/gkq1156

Kamburov A, Stelzl U, Lehrach H, Herwig R (2013) The ConsensusPathDB interaction database: 2013 update. Nucleic Acids Res 41:D793–D800. doi:10.1093/nar/gks1055

Kam-Thong T, Czamara D, Tsuda K, Borgwardt K, Lewis CM, Erhardt-Lehmann A, Hemmer B, Rieckmann P, Daake M, Weber F, Wolf C, Ziegler A, Putz B, Holsboer F, Scholkopf B, Muller-Myhsok B (2011a) EPIBLASTER-fast exhaustive two-locus epistasis detection strategy using graphical processing units. Eur J Hum Genet 19:465–471. doi:10.1038/ejhg.2010.196

Kam-Thong T, Putz B, Karbalai N, Muller-Myhsok B, Borgwardt K (2011b) Epistasis detection on quantitative phenotypes by exhaustive enumeration using GPUs. Bioinformatics 27:i214–i221. doi:10.1093/bioinformatics/btr218

Kam-Thong T, Azencott CA, Cayton L, Putz B, Altmann A, Karbalai N, Samann PG, Scholkopf B, Muller-Myhsok B, Borgwardt KM (2012) GLIDE: GPU-based linear regression for detection of epistasis. Hum Hered 73:220–236. doi:10.1159/000341885

Liekens AM, De Knijf J, Daelemans W, Goethals B, De Rijk P, Del-Favero J (2011) BioGraph: unsupervised biomedical knowledge discovery via automated hypothesis generation. Genome Biol 12:R57. doi:10.1186/gb-2011-12-6-r57

Mahachie John JM, Cattaert T, Lishout FV, Gusareva ES, Steen KV (2012) Lower-order effects adjustment in quantitative traits model-based multifactor dimensionality reduction. PLoS One 7:e29594. doi:10.1371/journal.pone.0029594

Mahachie John JM, Van Lishout F, Gusareva ES, Van Steen K (2013) A robustness study of parametric and non-parametric tests in model-based multifactor dimensionality reduction for epistasis detection. BioData Min 6:9. doi:10.1186/1756-0381-6-9

Mann SA, Otway R, Guo G, Soka M, Karlsdotter L, Trivedi G, Ohanian M, Zodgekar P, Smith RA, Wouters MA, Subbiah R, Walker B, Kuchar D, Sanders P, Griffiths L, Vandenberg JI, Fatkin D (2012) Epistatic effects of potassium channel variation on cardiac repolarization and atrial fibrillation risk. J Am Coll Cardiol 59:1017–1025. doi:10.1016/j.jacc.2011.11.039

Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39:906–913. doi:10.1038/ng2088

Miko I (2008) Epistasis: gene interaction and phenotype effects. Nat Educ 1

Minelli C, Thompson JR, Abrams KR, Lambert PC (2005) Bayesian implementation of a genetic model-free approach to the meta-analysis of genetic association studies. Stat Med 24:3845–3861. doi:10.1002/sim.2393

Minelli CT, Thompson JR, Abrams KR, Thakkinstian A, Attia J (2008) How should we use information about HWE in the meta-analyses of genetic association studies? Int J Epidemiol 37:136–146

Moore JH (2005) A global view of epistasis. Nat Genet 37:13–14. doi:10.1038/ng0105-13

Moore JH, Barney N, Tsai CT, Chiang FT, Gui J, White BC (2007) Symbolic modeling of epistasis. Hum Hered 63:120–133. doi:10.1159/000099184

Motsinger AA, Reif DM, Fanelli TJ, Davis AC, Ritchie MD (2007) Linkage disequilibrium in genetic association studies improves the performance of grammatical evolution neural networks. Proc IEEE Symp Comput Intell Bioinform Comput Biol 2007:1–8

Nelson MR, Kardia SL, Ferrell RE, Sing CF (2001) A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. Genome Res 11:458–470. doi:10.1101/gr.172901

Nitsch D, Tranchevent LC, Goncalves JP, Vogt JK, Madeira SC, Moreau Y (2011) PINTA: a web server for network-based gene prioritization from expression data. Nucleic Acids Res 39:W334–W338. doi:10.1093/nar/gkr289

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD (2008) Genes mirror geography within Europe. Nature 456:98–101. doi:10.1038/nature07331

Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 74:765–769. doi:10.1086/383251

Patterson N, Price AL, Reich D (2006) Population structure and Eigen analysis. PLoS Genet 2:e190. doi:10.1371/journal.pgen.0020190

Peloso GM, Lunetta KL (2011) Choice of population structure informative principal components for adjustment in a case–control study. BMC Genet 12:64. doi:10.1186/1471-2156-12-64

Pentchev K, Ono K, Herwig R, Ideker T, Kamburov A (2010) Evidence mining and novelty assessment of protein–protein interactions with the ConsensusPathDB plugin for Cytoscape. Bioinformatics 26:2796–2797. doi:10.1093/bioinformatics/btq522

Pereira TV, Patsopoulos NA, Salanti G, Ioannidis JP (2009) Discovery properties of genome-wide association signals from cumulatively combined data sets. Am J Epidemiol 170:1197–1206. doi:10.1093/aje/kwp262

Pereira TV, Patsopoulos NA, Pereira AC, Krieger JE (2011) Strategies for genetic model specification in the screening of genome-wide meta-analysis signals for further replication. Int J Epidemiol 40:457–469. doi:10.1093/ije/dyq203

Peres-Neto P, Jackson D, Somers K (2005) How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Comput Stat Data Anal 49:974–997

Piriyapongsa J, Ngamphiw C, Intarapanich A, Kulawonganunchai S, Assawamakin A, Bootchai C, Shaw PJ, Tongsima S (2012) iLOCi: a SNP interaction prioritization technique for detecting epistasis in genome-wide association studies. BMC Genom 13(Suppl 7):S2. doi:10.1186/1471-2164-13-S7-S2

Prabhu S, Pe'er I (2012) Ultrafast genome-wide scan for SNP–SNP interactions in common complex disease. Genome Res 22:2230–2240. doi:10.1101/gr.137885.112

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. doi:10.1086/519795

Reich D, Price AL, Patterson N (2008) Principal component analysis of genetic data. Nat Genet 40:491–492. doi:10.1038/ng0508-491

Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 69:138–147. doi:10.1086/321276

Ritchie MD, Hahn LW, Moore JH (2003) Power of multifactor dimensionality reduction for detecting gene–gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. Genet Epidemiol 24:150–157. doi:10.1002/gepi.10218

Schwarz DF, Konig IR, Ziegler A (2010) On safari to Random Jungle: a fast implementation of Random Forests for high-dimensional data. Bioinformatics 26:1752–1758. doi:10.1093/bioinformatics/btq257

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. doi:10.1101/gr.1239303

Sifrim A, Popovic D, Tranchevent LC, Ardeshirdavani A, Sakai R, Konings P, Vermeesch JR, Aerts J, De Moor B, Moreau Y (2013) eXtasy: variant prioritization by genomic data fusion. Nat Methods 10:1083–1084. doi:10.1038/nmeth.2656

Sithisarankul P, Weaver VM, Diener-West M, Strickland PT (1997) Multicollinearity may lead to artificial interaction: an example from a cross sectional study of biomarkers. Southeast Asian J Trop Med Public Health 28:404–409

Slinker BK, Glantz SA (1985) Multiple regression for physiological data analysis: the problem of multicollinearity. Am J Physiol 249:R1–R12

Strehl A, Ghosh J (2003) Cluster ensembles—a knowledge reuse framework for combining multiple partitions. J Mach Learn Res 3:583–617. doi:10.1162/153244303321897735

Struchalin MV, Dehghan A, Witteman JC, van Duijn C, Aulchenko YS (2010) Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. BMC Genet 11:92. doi:10.1186/1471-2156-11-92

Subirana I, Gonzalez JR (2013) Genetic association analysis and meta-analysis of imputed SNPs in longitudinal studies. Genet Epidemiol 37:465–477. doi:10.1002/gepi.21719

Sulakhe D, Balasubramanian S, Xie B, Feng B, Taylor A, Wang S, Berrocal E, Dave U, Xu J, Bornigen D, Gilliam TC, Maltsev N (2013) Lynx: a database and knowledge extraction engine for integrative medicine. Nucleic Acids Res. doi:10.1093/nar/gkt1166

Sun X, Lu Q, Mukheerjee S, Crane PK, Elston R, Ritchie MD (2014) Analysis pipeline for the epistasis search—statistical versus biological filtering. Front Genet 5:106. doi:10.3389/fgene.2014.00106

Thomas DC, Haile RW, Duggan D (2005) Recent developments in genomewide association scans: a workshop summary and review. Am J Hum Genet 77:337–345. doi:10.1086/432962

Touitou I, Galeotti C, Rossi-Semerano L, Hentgen V, Piram M, Kone-Paut I, CeReMai Frcfad (2013) The expanding spectrum of rare monogenic autoinflammatory diseases. Orphanet J Rare Dis 8:162. doi:10.1186/1750-1172-8-162

Van Lishout F, Mahachie John JM, Gusareva ES, Urrea V, Cleynen I, Theatre E, Charloteaux B, Calle ML, Wehenkel L, Van Steen K (2013) An efficient algorithm to perform multiple testing in epistasis screening. BMC Bioinform 14:138. doi:10.1186/1471-2105-14-138

Van Steen K (2012) Travelling the world of gene–gene interactions. Brief Bioinform 13:1–19. doi:10.1093/bib/bbr012

Van Steen K, Curran D, Kramer J, Molenberghs G, Van Vreckem A, Bottomley A, Sylvester R (2002) Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. Stat Med 21:3865–3884. doi:10.1002/sim.1358

VanderWeele TJ (2009) On the distinction between interaction and effect modification. Epidemiology 20:863–871. doi:10.1097/EDE.0b013e3181ba333c

Vansteelandt S, Bekaert M, Claeskens G (2012) On model selection and model misspecification in causal inference. Stat Methods Med Res 21:7–30. doi:10.1177/0962280210387717

Wain LV (2014) Rare variants and cardiovascular disease. Brief Funct Genomics. doi:10.1093/bfgp/elu010

Wan X, Yang C, Yang Q, Xue H, Fan X, Tang NL, Yu W (2010a) BOOST: a fast approach to detecting gene–gene interactions in genome-wide case–control studies. Am J Hum Genet 87:325–340. doi:10.1016/j.ajhg.2010.07.021

Wan X, Yang C, Yang Q, Xue H, Tang NL, Yu W (2010b) Predictive rule inference for epistatic interaction detection in genome-wide association studies. Bioinformatics 26:30–37. doi:10.1093/bioinformatics/btp622

Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 38:W214–W220. doi:10.1093/nar/gkq537

Watkinson J, Anastassiou D (2009) Synergy disequilibrium plots: graphical visualization of pairwise synergies and redundancies of SNPs with respect to a phenotype. Bioinformatics 25:1445–1446. doi:10.1093/bioinformatics/btp159

Wei W, Hemani G, Hicks AA, Vitart V, Cabrera-Cardenas C, Navarro P, Huffman J, Hayward C, Knott SA, Rudan I, Pramstaller PP, Wild SH, Wilson JF, Campbell H, Dunlop MG, Hastie N, Wright AF, Haley CS (2011) Characterisation of genome-wide association epistasis signals for serum uric acid in human population isolates. PLoS One 6:e23836. doi:10.1371/journal.pone.0023836

Westfall PH, Young SS (1993a) Resampling-based multiple testing: examples and methods for *p* value adjustment. Wiley, New York

Westfall PH, Young SS (1993b) Resampling-based multiple testing. Wiley, New York

Wienbrandt L, Kassens JC, Gonzalez-Domınguez J, Schmidt B, Ellinghaus D, Schimmler M (2014) FPGA-based acceleration of detecting statistical epistasis in GWAS. 14th International Conference on Computational Science, vol 29. Procedia Computer Science, pp 220–230

Wong AK, Park CY, Greene CS, Bongo LA, Guan Y, Troyanskaya OG (2012) IMP: a multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks. Nucleic Acids Res 40:W484–W490. doi:10.1093/nar/gks458

Wu J, Devlin B, Ringquist S, Trucco M, Roeder K (2010) Screen and clean: a tool for identifying interactions in genome-wide association studies. Genet Epidemiol 34:275–285. doi:10.1002/gepi.20459

Xu H, George V (2011) A Monte Carlo test of linkage disequilibrium for single nucleotide polymorphisms. BMC Res Notes 4:124. doi:10.1186/1756-0500-4-124

Xu J, Turner A, Little J, Bleecker ER, Meyers DA (2002) Positive results in association studies are associated with departure from Hardy–Weinberg equilibrium: hint for genotyping error? Hum Genet 111:573–574

Yang C, He Z, Wan X, Yang Q, Xue H, Yu W (2009) SNPHarvester: a filtering-based approach for detecting epistatic interactions in genome-wide association studies. Bioinformatics 25:504–511. doi:10.1093/bioinformatics/btn652

Zhang Y (2011) Bayesian epistasis association mapping via SNP imputation. Biostatistics 12:211–222. doi:10.1093/biostatistics/kxq063

Zhang X, Huang S, Zou F, Wang W (2010a) TEAM: efficient two-locus epistasis tests in human genome-wide association study. Bioinformatics 26:i217–i227. doi:10.1093/bioinformatics/btq186

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES (2010b) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–360. doi:10.1038/ng.546

Zhang F, Boerwinkle E, Xiong M (2014) Epistasis analysis for quantitative traits by functional regression model. Genome Res 24:989–998. doi:10.1101/gr.161760.113

Ziegler A, König I (2006) A Statistical Approach to Genetic Epidemiology: Concepts and Applications, with an e-Learning Platform. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci USA 109:1193–1198. doi:10.1073/pnas.1119675109