Received: 23 April 2019 | Revised: 13 August 2019 | Accepted: 16 September 2019

DOI: 10.1002/gepi.22265

Genetic Epidemiology

OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

WILEY

RESEARCH ARTICLE

Powerful rare variant association testing in a copula-based joint analysis of multiple phenotypes

Stefan Konigorski[1,2] | Yildiz E. Yilmaz[3,4,5] | Jürgen Janke[1] | Manuela M. Bergmann[6] | Heiner Boeing[6] | Tobias Pischon[1,7,8]

[1]Molecular Epidemiology Research Group, Max Delbrück Center (MDC) for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[2]Digital Health and Machine Learning Research Group, Hasso Plattner Institute for Digital Engineering, Potsdam, Germany

[3]Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, Canada

[4]Discipline of Genetics, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

[5]Discipline of Medicine, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

[6]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke (DIfE), Nuthetal, Germany

[7]Charité—Universitätsmedizin Berlin, Berlin, Germany

[8]DZHK (German Center for Cardiovascular Research), partner site Berlin, Berlin, Germany

Correspondence
Stefan Konigorski, Digital Health and Machine Learning Research Group, Hasso Plattner Institute for Digital Engineering, Rudolf-Breitscheid-Street 187, 14482 Potsdam, Germany.
Email: stefan.konigorski@hpi.de

Funding information
Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: 2014-04904; Memorial University of Newfoundland, Faculty of Medicine

Abstract

In genetic association studies of rare variants, the low power of association tests is one of the main challenges. In this study, we propose a new single-marker association test called C-JAMP (Copula-based Joint Analysis of Multiple Phenotypes), which is based on a joint model of multiple phenotypes given genetic markers and other covariates. We evaluated its performance and compared its empirical type I error and power with existing univariate and multivariate single-marker and multi-marker rare-variant tests in extensive simulation studies. C-JAMP yielded unbiased genetic effect estimates and valid type I errors with an adjusted test statistic. When strongly dependent traits were jointly analyzed, C-JAMP had the highest power in all scenarios except when a high percentage of variants were causal with moderate/small effect sizes. When traits with weak or moderate dependence were analyzed, whether C-JAMP or competing approaches had higher power depended on the effect size. When C-JAMP was applied with a misspecified copula function, it still achieved high power in some of the scenarios considered. In a real-data application, we analyzed sequencing data using C-JAMP and performed the first genome-wide association studies of high-molecular-weight and medium-molecular-weight adiponectin plasma concentrations. C-JAMP identified 20 rare variants with p-values smaller than 10^{-5}, while all other tests resulted in the identification of fewer variants with higher p-values. In summary, the results indicate that C-JAMP is a powerful, flexible, and robust method for association studies, and we identified novel candidate markers for adiponectin. C-JAMP is implemented as an R package and freely available from https://cran.r-project.org/package=CJAMP.

KEYWORDS

adipokines, adiponectin, copula models, genetic association study, joint modeling, multiple phenotypes, obesity, rare variant analysis

1 | INTRODUCTION

Technological advancements and collaborative efforts have revolutionized the way we study the genetic underpinnings of complex traits including molecular and disease phenotypes (Evangelou et al., 2018; PCAWG Transcriptome Core Group et al., 2018). High-throughput sequencing allows now to investigate rare variants in greater depth, which are abundant and have high interindividual variability (Telenti et al., 2016), often unknown phenotypic effects (Schork, Murray, Frazer, & Topol, 2009), and potentially pivotal roles in diseases (Mancuso et al., 2016). One of the main challenges in rare-variant association studies is the low power of tests, which limits the identification of novel loci and contributes to the fact that a large part of the estimated heritability of most complex traits is still unexplained. The power can be increased by analyzing larger sample sizes, optimizing study designs, and using powerful statistical approaches. The latter is the focus of this study.

In one line of methodological research, multi-marker tests (MMTs) have been proposed to increase the power by aggregating the information of (rare) genetic markers in a given genomic region and then testing the association of the region with the trait of interest. Popular MMTs include burden-type tests, SKAT (M. C. Wu et al., 2011), SKAT-O (Lee, Wu, & Lin, 2012), and other kernel-based variance-component tests (Asimit & Zeggini, 2010; Lee, Abecasis, Boehnke, & Lin, 2014; Li & Leal, 2008; Listgarten et al., 2013). While MMTs have larger power compared to single-marker tests (SMTs), which test each genetic marker separately, for the analysis of binary traits (Asimit & Zeggini, 2010; Konigorski, Yilmaz, & Pischon, 2017; Lee et al., 2014), this is not generally the case for the analysis of quantitative traits. Here, SMTs also have lower power compared to MMTs when single nucleotide variants (SNVs) have small effect sizes but equal or larger power compared to MMTs when SNVs have moderate or large effect sizes (Konigorski et al., 2017).

Another promising statistical approach to increase the power of association tests in genome-wide association studies (GWAS) is to jointly model multiple traits (Schillert & Konigorski, 2016), which is intuitive when different traits of a phenotype are available, such as different obesity measures or different isoforms of a gene or protein. Many previously proposed approaches are based on multivariate generalized linear or linear mixed models such as MURAT (Sun et al., 2016), aSPU, aSPUset, aSPUset-Score (Kim, Zhang, Pan, & Alzheimer's Disease Neuroimaging Initiative, 2016; Y. Zhang, Xu, Shen, Pan, & Alzheimer's Disease Neuroimaging Initiative, 2014), MultiPhen (O'Reilly et al., 2012), and others (Kaakinen et al., 2017; Lippert, Casale, Rakitsch, & Stegle, 2014; Maity, Sullivan, & Tzeng, 2012; Schifano, Li, Christiani, & Lin, 2013; Wang, Wang, Sha, & Zhang, 2016), or use dimension reduction methods (Aschard et al., 2014; Yang & Wang, 2012), structural equation modeling methods (Momen et al., 2018; Song, Morris, & Stein, 2016; Verhulst, Maes, & Neale, 2017), methods combining results from univariate analyses (Liang, Wang, & Zhang, 2016; Liu & Lin, 2018; O'Brien, 1984; van der Sluis, Posthuma, & Dolan, 2013), or others (Aschard et al., 2017; Jiang et al., 2015), see the review by Yang and Wang (2012) for more details. Most of these approaches assume a very specific and restricted Gaussian dependence structure between traits while there exist many empirical data sets with non-Gaussian dependencies. There exist more recent kernel-based approaches including GAMuT (Broadaway, Cutler, & Duncan, 2016), MSKAT (B. Wu & Pankow, 2016), DKAT (Zhan et al., 2017), and Multi-SKAT (Dutta, Scott, Boehnke, & Lee, 2019) that allow a more flexible modeling of the multivariate dependence structure. Most of the above tests are multi-marker and "multivariate" tests (i.e., testing the association of all variants in a region with all traits jointly), hence they may not have optimal power for testing variants with large effect sizes and when the tested variants are only associated with a few of the tested traits. There exist some empirical comparisons of multivariate tests for the analysis of common (Kim et al., 2016; Liang et al., 2016; Zhu, Zhang, & Sha, 2015) and rare genetic variants (Broadaway et al., 2016; Dutta et al., 2019; B. Wu & Pankow, 2016; Zhan et al., 2017), however, most comparisons have been limited to a few selected methods of similar type.

Here, we propose a novel SMT based on joint copula models of multiple phenotypes and call the approach C-JAMP (Copula-based Joint Analysis of Multiple Phenotypes). The use of copula functions has been proposed in some recent GWAS (Dutta et al., 2019; He, Li, Edmondson, Rader, & Li, 2012; Konigorski, Yilmaz, & Bull, 2014; Konigorski, Yilmaz, & Pischon, 2016; Lakhal-Chaieb, Oualkacha, Richards, & Greenwood, 2016; Ray & Basu, 2017; Rosen & Thompson, 2015; H. Zhang, Qin, Landi, Caporaso, & Yu, 2013; Zhao & Zhang, 2016), however, only for the analysis of common variants or with a different focus than to increase the power through a test based on the modeling of the dependence of multiple phenotypes. To our knowledge, the only previous application of testing rare-variant effects in the marginals of joint copula models is in our pilot study in Konigorski et al. (2016) and further, there is no available efficient and robust implementation of association tests in joint copula models for genome-wide analyses. The general goal of C-JAMP is to jointly model

two (or more) traits conditional on a genetic marker of interest using copula functions, in order to estimate and test the association of the marker with either trait. Copula functions are used to construct a joint distribution function of multiple traits by combining the marginal distributions of traits with a dependence structure. They provide a very flexible tool to model different multivariate distributions with appropriate marginal distributions, and we hypothesize that they allow increasing the power of association tests.

After describing C-JAMP in more detail in the next section, we assess its finite-sample properties as a SMT in extensive rare-variant simulation studies of quantitative traits, and compare its performance to a standard univariate SMT (i.e., SMT under the univariate model of a single trait; linear regression), univariate MMTs (a burden test, SKAT, and SKAT-O), multivariate SMTs (MultiPhen, aSPU), as well as to multivariate MMTs (MURAT, aSPUset, aSPUset-Score, MSKAT, GAMuT, and DKAT). Finally, we apply C-JAMP in a substudy of the large European Prospective Investigation into Cancer and Nutrition (EPIC) Potsdam cohort study (Boeing, Wahrendorf, & Becker, 1999; Riboli & Kaaks, 1997) with 200 probands, and perform—to our knowledge—the first GWAS of high-molecular-weight (HMW) and medium-molecular-weight (MMW) adiponectin plasma concentrations, which are different biologically active isoforms of the hormone adiponectin and have been associated with obesity and a number of chronic diseases (Aleksandrova et al., 2012; Pischon, 2009). An R package with the implementation of C-JAMP is freely available from https://cran.r-project.org/package=CJAMP.

2 | METHODS

2.1 | C-JAMP

In C-JAMP, we propose to jointly model two (or more) quantitative traits, Y_1 and Y_2, using copula functions. For Y_1, Y_2, and a covariate vector $\boldsymbol{Z}$, the joint distribution F of Y_1 and Y_2, conditional on $\mathbf{Z}=\mathbf{z}$, can be constructed by combining the marginal distributions of Y_1 and Y_2, F_1 and F_2, conditional on $\mathbf{Z}=\mathbf{z}$, using a copula function C_ψ with a parameter vector $\boldsymbol{\psi}$ modeling the dependence structure between $Y_1|\boldsymbol{z}$ and $Y_2|\boldsymbol{z}$:

$$F(y_1, y_2|\boldsymbol{z}) = C_{\boldsymbol{\psi}}(F_1(y_1|\boldsymbol{z}), F_2(y_2|\boldsymbol{z})). \quad (1)$$

The copula model allows marginal distributions to come from any family of distributions, and also to differ between phenotypes. In the following, we illustrate C-JAMP for a covariate vector $\boldsymbol{Z}=(X_1, X_2, G)^T$ including two nongenetic covariates X_1, X_2, and one genetic marker g, and the marginal models

$$Y_1 = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \beta g + \varepsilon, \quad (2)$$

$$Y_2 = \gamma_0' + \gamma_1' x_1 + \gamma_2' x_2 + \beta' g + \varepsilon', \quad (3)$$

with normally distributed errors $\varepsilon \sim N(0, \sigma_1^2)$ and $\varepsilon' \sim N(0, \sigma_2^2)$. The dependence parameters in $\boldsymbol{\psi}$ do not appear in the marginal distributions which allows to estimate the effect of a SNV G on Y_1 or Y_2, while considering the dependence between Y_1 and Y_2 adjusted for $\boldsymbol{Z}$.

Popular copula functions include the Clayton family (Clayton, 1978)

$$C_\varphi(u_1, u_2, \ldots u_p, \varphi) = \left(\left(\sum_{l=1}^{p} u_l^{-\varphi}\right) - (p-1)\right)^{-1/\varphi} \quad (4)$$

with $\varphi > 0$, which is illustrated in Figure 1, and the Gumbel-Hougaard family (Gumbel, 1960)

$$C_\theta(u_1, u_2, \ldots u_p, \theta) = \exp\left\{-\left[\sum_{l=1}^{p}(-\log u_l)^\theta\right]^{1/\theta}\right\} \quad (5)$$

with $\theta > 1$. A third family which includes both (4) and (5) for $\theta = 1$ and $\varphi \to 0$, respectively, is the 2-parameter copula family

$$C_{\boldsymbol{\psi}}(u_1, u_2, \ldots u_p, \varphi, \theta) = \left\{\left[\sum_{l=1}^{p}(u_l^{-\varphi} - 1)^\theta\right]^{1/\theta} + 1\right\}^{-1/\varphi} \quad (6)$$

with $0 \le u_1, u_2 \le 1$, and the copula parameter vector $\boldsymbol{\psi} = (\varphi, \theta)^T$, $\varphi > 0$, $\theta \ge 1$, which allows to model both the lower- and upper-tail dependence (Joe, 1997) and hence a flexible modeling of a large class of dependence structures. There is a one-to-one relation between the overall dependence measure Kendall's tau and the copula parameter vector $(\varphi, \theta)^T$ in the model (6); that is $\tau_{\varphi,\theta} = 1 - 2/(\theta(\varphi + 2))$. A stronger dependence in the tails of the distribution can often be found in real data, see for example the application described in this study.

Focusing on fully parametric models, maximum likelihood estimates of the parameters $(\varphi, \theta, \gamma_0, \gamma_1, \gamma_2, \beta, \gamma_0', \gamma_1', \gamma_2', \beta', \sigma_1, \sigma_2)^T$ are obtained by maximizing the likelihood function of the data using the quasi-Newton, variable metric BFGS (Broyden-Fletcher-Goldfarb-Shanno) method (Gentle, 2009). Standard error estimates of the parameter estimates

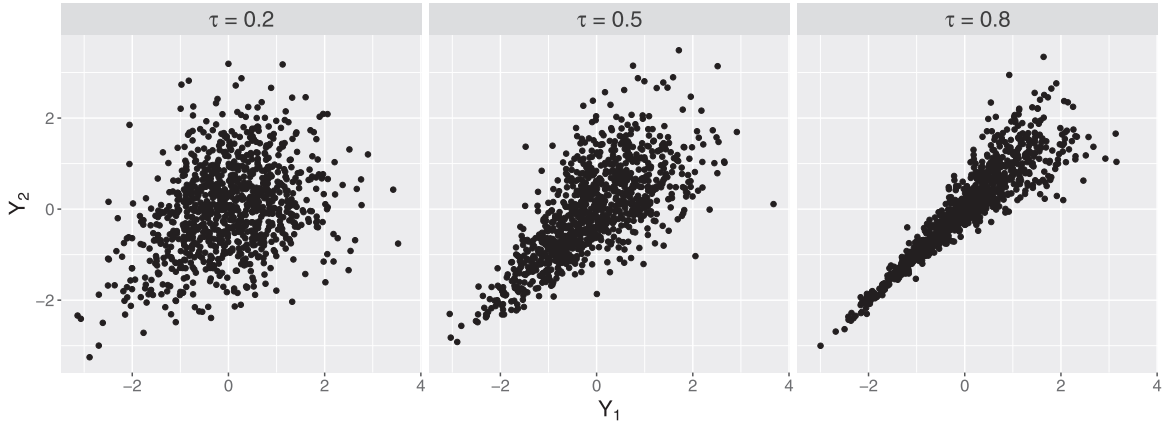

FIGURE 1 Scatterplots of bivariate data (Y_1 and Y_2) under the Clayton copula. Y_1 and Y_2 have standard normal marginal distributions and their dependence was set to Kendall's $\tau = 0.2$, 0.5, and 0.8. The data was generated using the *generate_clayton_copula()* function in the *CJAMP* R package using the relation $\varphi = 2\tau/(1-\tau)$ between Kendall's τ and the Clayton copula parameter φ, see the vignette of the *CJAMP* R package for further details. The Clayton copula allows modeling of the lower-tail dependence, while the Gumbel copula could be used to model upper-tail dependence and the 2-parameter copula in (6) could be used to model both tail dependencies

are obtained from the inverse of the observed information matrix. To identify SNVs associated with Y_1 or Y_2 using C-JAMP, we conduct SMTs and test the null hypotheses $H_0 : \beta = 0$ (vs. $H_A : \beta \neq 0$) and $H_0 : \beta' = 0$ (vs. $H_A : \beta' \neq 0$) for each SNV G, respectively, by using the Wald test statistics (Konigorski et al., 2017). For more statistical and computational details, see the Supporting Information Note 1 and the R package documentation at https://cran.r-project.org/web/packages/CJAMP/.

Preliminary empirical investigations of C-JAMP showed that the asymptotic distribution assumption for the Wald test statistic may not hold when testing the association of SNVs that have very few minor alleles

Box 1 Algorithm to obtain adjusted Wald test statistics. This algorithm was used to obtain adjusted Wald test statistics for our simulation study in the following in the scenarios with $\tau = 0.5$ and 0.8. We considered $k = 251$ single nucleotide variants (SNVs) with minor allele count (MAC) between 1 and 958 (which equal to MAFs of 0.0005, 0.001, 0.0015, ..., 0.4790, respectively), which were observed in the genetic data used in the simulation study in the main analysis, $m = 1{,}000$, for $n = 1{,}000$ individuals. We generated data from the Clayton copula in Step 2 using the *generate_clayton_copula*() function in the *CJAMP* R package and the relationship $\tau = \varphi/(\varphi + 2)$ between Kendall's τ and the Clayton copula parameter φ. In Step 3, the copula model was fitted using the *cjamp*() function in the *CJAMP* R package.

1. Generate data with n observations of k SNVs, for each of the MAC of the SNVs present in the data that will be analyzed.

Do the following for each of the k SNVs:

2. Generate m sets of phenotypes Y_1 and Y_2, of size n from the bivariate distribution of interest under the null hypothesis.
3. Fit the copula model separately for each set of traits conditional on the SNV and obtain the standard Wald test statistic W for testing the association of the SNV with the trait of interest.
4. Generate m standard normal quantiles Z.
5. Fit a linear regression model of the sorted Z and sorted W, $Z = \alpha + \beta W + \varepsilon$, and obtain estimates $\hat{\alpha}$ and $\hat{\beta}$.

In the data analysis, for each of the k SNVs, compute Wald test statistics $\widehat{W}$ and use the estimates $\hat{\alpha}$ and $\hat{\beta}$ to obtain adjusted Wald test statistics $\widehat{W}_{\text{adj}}$ as $\widehat{W}_{\text{adj}} = \hat{\alpha} + \hat{\beta}\widehat{W}$.

(with minor allele count [MAC] = 1, 2, or 4) with traits of moderate/high dependence in a finite-sample setting (see Supporting Information Note 2, Tables S1–S3, and Figure S1). As a consequence, we obtained adjusted Wald test statistics for traits with dependence $\tau \geq 0.5$ based on a Monte Carlo simulation study, which is outlined in Box 1 and described in more detail with a generalization in Supporting Information Note 2, Table S4, and Figures S2–S5. These adjusted Wald test statistics were used for all analyses in the simulation study that are reported in the following.

2.2 | Simulation study

2.2.1 | Overview

In the simulation study, our aim was to evaluate the empirical type I error and power of C-JAMP as a SMT under a joint model of multiple traits, and to compare its performance with a standard univariate SMT (linear regression), univariate MMTs (a burden test, SKAT, and SKAT-O), multivariate SMTs (MultiPhen and aSPU), as well as to multivariate MMTs (MURAT, aSPUset, aSPUset-Score, MSKAT, GAMuT, and DKAT) for the analysis of rare variants. All approaches were evaluated in their power to identify a causal gene, which amounts to testing whether any of the SNVs in the gene has an effect on the trait for SMTs, using a multiple testing correction for all SNVs in the gene. It has to be noted that the univariate and multivariate tests were compared in testing different hypotheses (association with Y_1 $[H_0: \beta = 0]$ with C-JAMP vs. association with all traits $[H_0: \beta = \beta' = 0]$ in multivariate tests) which constitutes an advantage for the multivariate tests as we primarily analyzed data generated with effects on both traits. In addition, we considered scenarios where the genetic effects on the second trait are smaller or absent. Our primary focus was on the analysis of non-Gaussian dependencies between traits, which often occur in empirical datasets, but we also considered phenotypes that follow a bivariate normal distribution.

2.2.2 | Genetic data generation and general study set-up

The set-up of the simulation study is described in detail in the Supporting Information Note 3. In brief, we used the genetic data set provided in the SKAT package in R to obtain $m = 10{,}000$ genes as replicates with on average 58 SNVs (min = 36; max = 81 SNVs). The sample size was set to $n = 1{,}000$. Two traits, Y_1 and Y_2, were generated from the 1-parameter Clayton copula model in (4) conditional on two covariates x_1 and x_2 (and conditional on the causal SNVs g_j under the alternative hypotheses with additive genetic effects). We considered weak (Kendall's tau, $\tau = 0.2$), moderate ($\tau = 0.5$) and strong ($\tau = 0.8$) dependence levels under the Clayton copula model. Using the one-to-one relation $\tau = \varphi/(\varphi + 2)$ between Kendall's τ and the Clayton copula parameter φ, the corresponding copula parameter values are obtained from $\varphi = 2\tau/(1 - \tau)$. The marginal models of the two traits were considered as

$$Y_1 = 0.5x_1 + 0.5x_2 + \sum \beta_j g_j + \varepsilon, \quad (7)$$

$$Y_2 = 0.5x_1 + 0.5x_2 + \sum \beta'_j g_j + \varepsilon', \quad (8)$$

where $X_1 \sim Bin(0.5)$, $X_2 \sim N(0, 1)$, $\varepsilon, \varepsilon' \sim N(0, 1)$, $\beta_j = c_{Y_1} |log_{10}(\mathrm{MAF}_j)|$, $\beta'_j = c_{Y_2} |log_{10}(\mathrm{MAF}_j)|$, with different values for c_{Y_1} and c_{Y_2} and different percentages of causal variants (see Table 1; Konigorski et al., 2017; Lee et al., 2012). In each scenario, genetic effects were all in the same direction. This allowed evaluating C-JAMP when MMTs are more powerful (Konigorski et al., 2017; Lee et al., 2012). For Scenarios 0–12, for C-JAMP, all results regarding type I error and power estimates were similar for the two traits Y_1 and Y_2, and are reported only for testing the association with Y_1. Scenarios 13–15 were analyzed for C-JAMP to evaluate its type I error with respect to one trait (here: Y_2) when the SNVs are only associated with the other trait (Y_1). Also, Scenarios 13–17 were analyzed to investigate the power of C-JAMP, MURAT, aSPUset, aSPUset-Score, MSKAT, GAMuT, and DKAT when there is no or only a smaller genetic effect on the second trait Y_2. Here, the power of C-JAMP was reported with respect to testing the association with Y_1. In the evaluation, we did not assess the power of MultiPhen and aSPU since they did not yield valid empirical type I errors under the assumed model. To investigate scenarios with bivariate normally distributed phenotypes, Y_1 and Y_2 were generated using the same marginal models in (7) and (8) but with a bivariate Gaussian dependence structure (see Supporting Information Note 3 for details).

2.2.3 | Investigated rare variant tests

To evaluate the performance of C-JAMP, joint models of the generated phenotypes Y_1 and Y_2 given the SNV g and covariates X_1, X_2 were fitted using the Clayton copula in (4) with the marginal models in (2)–(3) using the *cjamp*() function in the R package *CJAMP*. Initially, we also fitted the 2-parameter copula family as a check and obtained the same results (data not shown), since the two-parameter copula family includes the Clayton model. Hence, for data generated from the Clayton copula, we

TABLE 1 Overview of the simulation study scenarios

Investigation	Scenario	% of causal variants	Effect sizes	Median (MAD) of explained variance in %
Type I error	0	0%	$c_{Y_1} = c_{Y_2} = 0$	–
Power	1	5%	$c_{Y_1} = c_{Y_2} = 0.6$	0.9% (0.6)
	2	5%	$c_{Y_1} = c_{Y_2} = 0.3$	0.2% (0.2)
	3	5%	$c_{Y_1} = c_{Y_2} = 0.2$	0.1% (0.1)
	4	10%	$c_{Y_1} = c_{Y_2} = 0.6$	1.9% (1.4)
	5	10%	$c_{Y_1} = c_{Y_2} = 0.3$	0.5% (0.3)
	6	10%	$c_{Y_1} = c_{Y_2} = 0.2$	0.2% (0.2)
	7	20%	$c_{Y_1} = c_{Y_2} = 0.6$	3.8% (2.1)
	8	20%	$c_{Y_1} = c_{Y_2} = 0.3$	1.0% (0.6)
	9	20%	$c_{Y_1} = c_{Y_2} = 0.2$	0.4% (0.2)
	10	50%	$c_{Y_1} = c_{Y_2} = 0.6$	9.1% (3.0)
	11	50%	$c_{Y_1} = c_{Y_2} = 0.3$	2.4% (0.9)
	12	50%	$c_{Y_1} = c_{Y_2} = 0.2$	1.1% (0.4)
Additional investigations of type I error and power	13	10%	$c_{Y_1} = 0.6; c_{Y_2} = 0$	As in Scenario 4
	14	20%	$c_{Y_1} = 0.3; c_{Y_2} = 0$	As in Scenario 8
	15	50%	$c_{Y_1} = 0.2; c_{Y_2} = 0$	As in Scenario 12
	16	10%	$c_{Y_1} = 0.6; c_{Y_2} = 0.1$	As in Scenario 4
	17	10%	$c_{Y_1} = 0.6; c_{Y_2} = 0.2$	As in Scenarios 1 and 3

Note: The 17 scenarios vary the percentage of causal variants and their effect sizes to investigate the empirical type I error and power. The percentage of causal rare variants is with respect to the total number of rare variants with MAF ≤ 0.03 in the gene. The effect size β of a SNV with a given MAF on the first trait Y_1 is $c_{Y_1} \cdot |log_{10}(\text{MAF})|$, and on the second trait Y_2 is $c_{Y_2} \cdot |log_{10}(\text{MAF})|$. The percentage of explained variance for a given gene is calculated as the sum of $2 \cdot \text{MAF}_i \cdot (1 - \text{MAF}_i) \cdot \beta_i^2 / \text{Var}(Y)$ over all variants i in the gene.

fitted C-JAMP based on the Clayton copula model since it is computationally faster. The unadjusted Wald test statistic was used for since it already has valid type I error, and adjusted Wald test statistics were computed for $\tau = 0.5$ and 0.8 due to the deflated and inflated type I errors for some nominal values. In the analysis of bivariate normally distributed phenotypes, C-JAMP was evaluated by fitting the (misspecified) 2-parameter copula without adjustment of the Wald test statistic. For evaluation of the existing methods, their available implementations in R were used. In short, the *lm*() function was used to fit linear regression models and to obtain SMT *p*-values, and the *p.adjust*() function was used to obtain the Benjamini-Hochberg (BH; Benjamini & Hochberg, 1995) corrected SMT *p*-values. Both SKAT and SKAT-O were computed using the default linear-weighted kernel in the test statistics as it returned the highest power estimates among all possible options provided in the *SKAT*() function in the *SKAT* package in R. The burden test was conducted under a linear regression model of each trait separately by using the sum of the minor alleles of all rare SNVs in the gene as predictor. MURAT is a multivariate generalization of the multi-marker SKAT and a data-adaptive variance-component test for the overall effect of all SNVs on all traits. It was computed using the *MURAT*() function in the *MURAT* R package with default settings. MultiPhen uses proportional odds logistic regression to predict each genotype by the multiple phenotypes and was computed using the *mPhen*() function in the *MultiPhen* R package with default settings, which tests the association of each SNV with each trait separately, as well as the association of each SNV with all traits. aSPU, aSPUset, and aSPUset-Score are data-adaptive powered score tests based on a multivariate marginal linear model of the traits conditional on one SNV (aSPU) or multiple SNVs (aSPUset and aSPUset-Score). aSPU was computed using the *GEEaSPU*() function in the *GEEaSPU* R package with default setting and evaluated for testing all SNVs within a gene to yield gene-level *p*-values. aSPUset and aSPUset-Score are multivariate gene-level association tests of the absence of effects of all SNVs on all traits and were performed using the *GEEaSPUset*() function in the *GEEaSPU* R package with default settings. Similar to MURAT, MSKAT is a multivariate extension of SKAT assuming a linear model between each trait and all SNVs, and constructs a score-based test statistic associating all SNVs in a gene with all traits. We used the *MSKAT*() function in the *MSKAT* package with default settings. GAMuT uses two kernels to model the similarities between phenotypes and between SNVs and constructs a test statistic based on the kernel distance covariance. We applied GAMuT by using the *TestGAMuT*() function with default settings. Finally, DKAT uses the dual-kernel framework similar to GAMuT but constructs a different test statistic for comparing the phenotype and SNV kernels, and we used the *DKAT*() function with default settings. See Supporting Information Note 4 for details.

2.3 | Analysis of adiponectin levels

For an application of C-JAMP, a GWAS of rare SNVs with HMW and MMW adiponectin concentrations was performed. Adiponectin is a hormone predominantly synthesized and secreted by adipocytes in the adipose tissue. It is involved, among others, in the regulation of insulin sensitivity and energy homeostasis. It increases insulin sensitivity, decreases inflammation, and is inversely associated with obesity, risk of type 2 diabetes, coronary heart disease, and several types of cancer (Aleksandrova et al., 2012; Pischon, 2009). Adiponectin circulates in different molecular fractions, which differ in their metabolic roles and involvement in molecular processes (Pischon, 2009), yet they are correlated, so including their information in joint modeling of the different fractions is an intuitive way to increase the power of association tests.

The data stems from a substudy of $n = 200$ probands within the large EPIC Potsdam study (Konigorski et al., 2018). 2,109,385 SNVs were called from RNA-Seq experiments and after stringent quality control checks, 176,733 biallelic autosomal rare SNVs in 23,922 genes with MAF $\leq$ 0.03 were included in the analysis, excluding singletons with MAF = 0.0027. C-JAMP was applied as an SMT by estimating the association of each SNV separately with each adiponectin fraction, and performing both variant-level and gene-level tests for each gene assessing whether any of the SNVs in the gene (within 5 kb of the gene boundary) shows an association with either adiponectin fraction. In addition, SMT under a univariate linear regression model and SKAT-O, MURAT, MSKAT, and DKAT as MMTs were computed. In all approaches, age, sex, physical activity, education, and BMI were incorporated into the models as covariates, and a complete data analysis was performed yielding a sample size of $n = 188$. For further details, see Supporting Information Note 5.

3 | RESULTS

3.1 | Empirical type I error rates of C-JAMP and other approaches

First, we report empirical type I errors of C-JAMP and the competing approaches for testing the null hypothesis that the gene is not associated with the trait. We provide results for both the data generated from the Clayton copula and from the bivariate normal distribution (Table 2).

The empirical type I error of C-JAMP was generally close to the nominal level under each dependence level and multivariate distribution considered. The univariate SMT based on linear regression and the three univariate MMTs all had valid empirical type I errors. Similarly, MURAT, aSPUset, aSPUset-Score, MSKAT, GAMuT, and

TABLE 2 Empirical type I error estimates of the approaches

Dependence	C-JAMP	SMT	Burden	SKAT	SKAT-O	MURAT	MultiPhen Y_1	MultiPhen Joint	aSPU	aSPUset	aSPUset-Score	MSKAT	GAMuT	DKAT
Dependence between traits from Clayton copula														
$\tau = 0.2$	0.050					0.051	0.019	0.298	0.100	0.051	0.054	0.051	0.051	0.053
$\tau = 0.5$	0.046	0.049	0.050	0.049	0.052	0.053	0.002	0.278	0.101	0.052	0.058	0.049	0.050	0.051
$\tau = 0.8$	0.049					0.059	0.004	0.225	0.090	0.053	0.056	0.055	0.055	0.049
Dependence between traits from bivariate normal distribution														
$\tau = 0.2$	0.054					0.054	0.002	0.317	0.107	0.049	0.050	0.054	0.054	0.053
$\tau = 0.5$	0.053	0.049	0.050	0.049	0.052	0.054	0.002	0.306	0.108	0.053	0.055	0.051	0.052	0.051
$\tau = 0.8$	0.029					0.049	0.003	0.308	0.067	0.052	0.052	0.048	0.049	0.048

Note: Empirical type I error estimates are shown for the nominal level of $\alpha = 0.05$. Data was generated from the null model described in Scenario 0 in Table 1 for $n = 1{,}000$ individuals with $m = 10{,}000$ replicates, and under different dependence levels between the two traits Y_1, Y_2 (Kendall's $\tau = 0.2$, 0.5, and 0.8), from the Clayton copula (upper panel) or bivariate normal distribution (lower panel) as described in Supporting Information Note 3. Results of C-JAMP are based on the adjusted Wald test statistics described in the Supporting Information Note 2 for $\tau = 0.5$ and 0.8, and based on unadjusted Wald test statistics for $\tau = 0.2$. Adjustments for multiple testing of all SNVs in a gene with C-JAMP, MultiPhen, and aSPU were done using the Bonferroni correction, and for the SMT using the BH (Benjamini & Hochberg, 1995) correction. Type I error estimates are shown for testing the association with the first trait Y_1 in C-JAMP and "MultiPhen Y_1", (and SMT, Burden, SKAT, SKAT-O), and for testing the joint association with both Y_1 and Y_2 in all other tests. Type I error estimates of the SMT, Burden, SKAT, and SKAT-O are not affected by the dependence structure between Y_1 and Y_2 and are hence identical over all these scenarios.

DKAT had empirical type I errors close to the nominal level, but some of them showed a slight inflation in the analysis of traits with high dependence. The aSPU test and the MultiPhen joint test of all traits led to inflated or highly inflated type I errors—for example, the type I error of the MultiPhen test was six times as high as the nominal level. The MultiPhen test of the genetic association with one trait, on the other hand, provided highly deflated type I errors and was also not valid. Consequently, the aSPU and MultiPhen tests were not included in the power study.

The empirical type I errors of C-JAMP were also close to the nominal levels for smaller levels ($\alpha = 10^{-2}$, 10^{-3}, 10^{-4}, 10^{-5}; see Table S5), but slightly larger in a few scenarios for some of the smaller nominal levels. They can be further controlled if the analysis is restricted to SNVs with MAC > 2 or MAC > 4, that is removing singletons and doubletons from the analysis (Table S6). The Bonferroni correction provided empirical levels closer to the nominal levels compared to the BH correction (Table S5) and was, therefore, used in the power evaluation in the next section. Further investigations of C-JAMP confirmed that the empirical type I errors were also well-calibrated when evaluated on a SNV-level instead of the gene-level (Table S7). In the scenario that SNVs were only affecting one of the two traits in the joint model, the empirical type I errors were valid when there was a weak or moderate dependence between the traits. If the traits had a strong dependence, however, then the type I errors were inflated (Table S8). See Supporting Information Note 6 for further details.

3.2 | Empirical power of C-JAMP compared with other approaches

The results of the power comparisons are shown in Figure 2, comparing C-JAMP with the univariate SMT and MMTs, and in Figure 3, comparing C-JAMP with the multivariate MMTs, under the Scenarios 1–12 described in

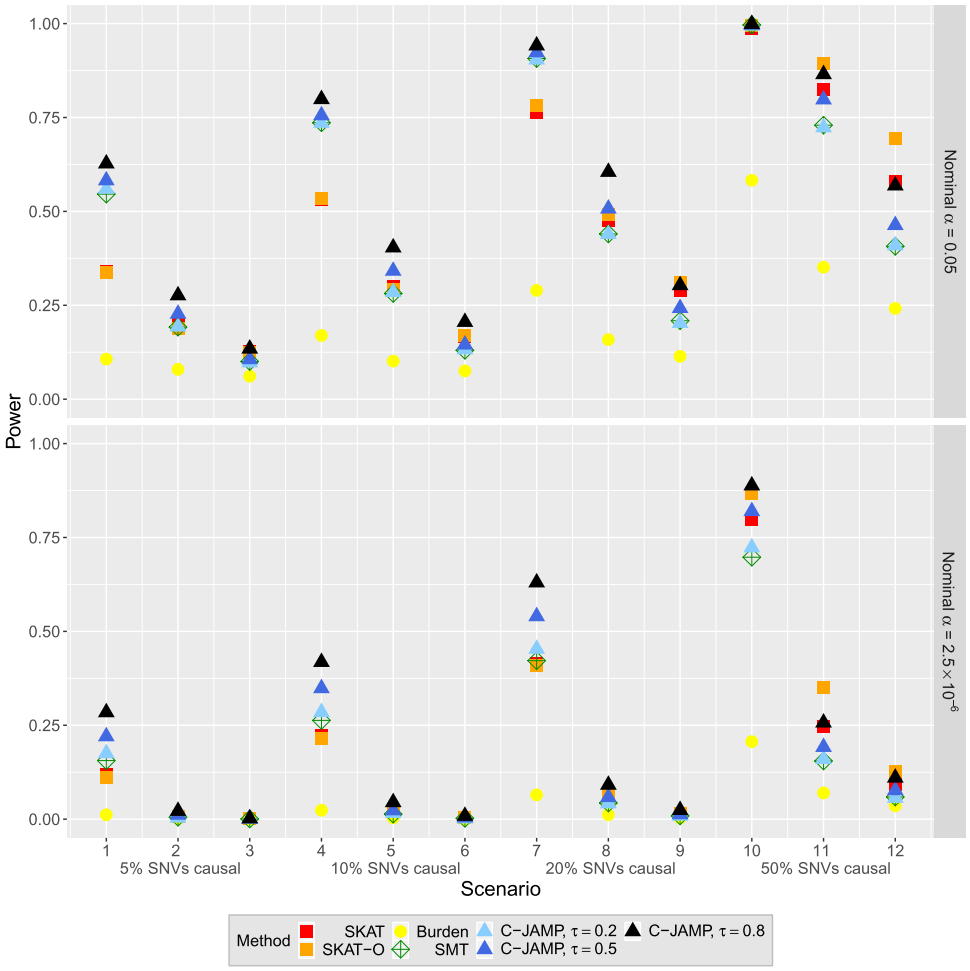

FIGURE 2 Empirical power estimates of C-JAMP and the univariate single-marker test (SMT) and multi-marker tests. Data was generated under an alternative-hypothesis model described in Scenarios 1–12 in Table 1 for $n = 1{,}000$ individuals with $m = 10{,}000$ replicates. The nominal α was set to 0.05 and 2.5×10^{-6}. Adjustments for multiple testing of all single nucleotide variants in a gene with C-JAMP were done using the Bonferroni correction and for testing with SMT using the BH-correction. Power estimates are shown for testing the association with the first trait Y_1.

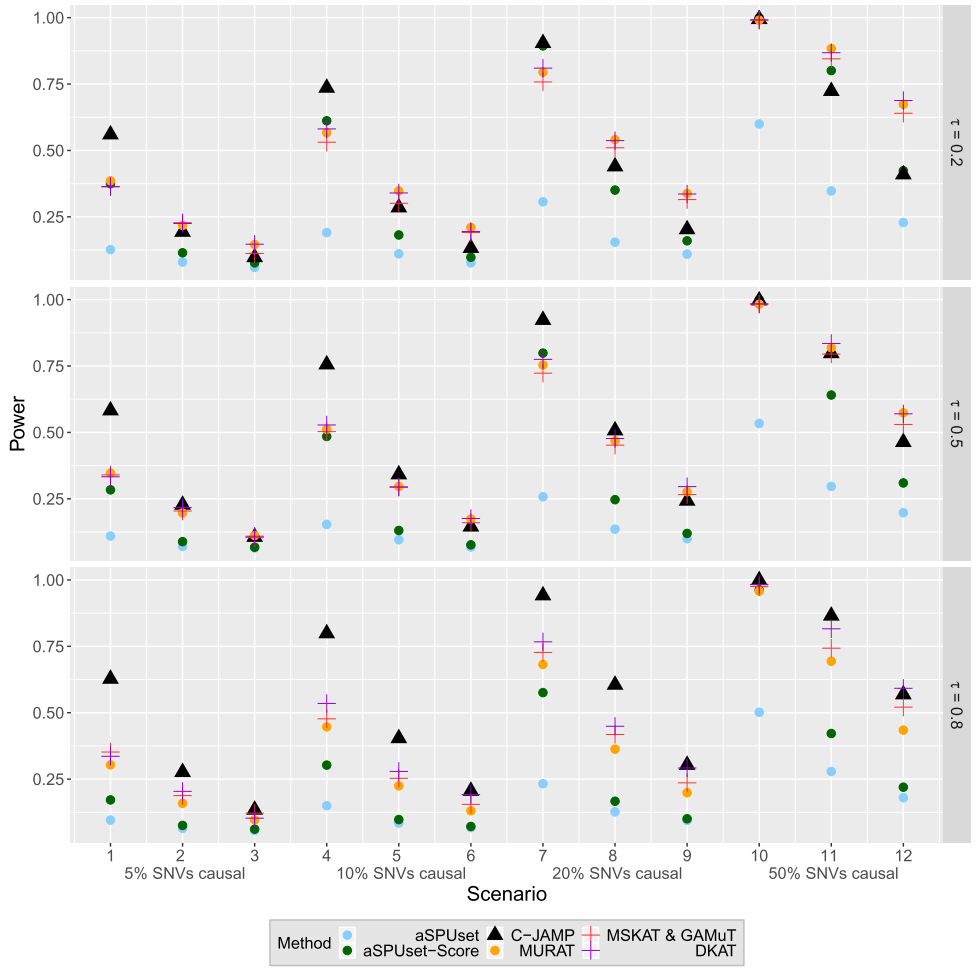

FIGURE 3 Empirical power estimates of C-JAMP and the multivariate multi-marker tests. Data was generated under an alternative-hypothesis model described in Scenarios 1–12 in Table 1 for $n = 1{,}000$ individuals with $m = 10{,}000$ replicates, and under different dependence levels between the two traits Y_1 and Y_2 (Kendall's $\tau = 0.2, 0.5,$ and 0.8). The nominal α was set to 0.05. Adjustments for multiple testing of all single nucleotide variants in a gene with C-JAMP were done using the Bonferroni-correction. Power estimates are shown for testing the association with the first trait Y_1 for C-JAMP and for testing the association with both traits with aSPUset, aSPUset-Score, MURAT, MSKAT, GAMuT, and DKAT. Since the power of MSKAT and GAMuT was identical up to three decimals in all scenarios, they are shown together in the figure.

Table 1, when the nominal type I errors were 0.05 or 2.5×10^{-6} (which would be the Bonferroni-corrected threshold for the analysis of 20,000 genes in applications) and data was generated from the Clayton copula (see Tables S9–S10 for more details).

As a first observation, the power of C-JAMP increased with increasing dependence between the traits (Figure 2). This was in contrast to all other investigated multivariate approaches, whose power was inversely affected by the dependence between traits: if the dependence between traits increased, their power decreased (Figure 3). Second, in comparison to the SMT based on a univariate model of a trait, C-JAMP led to consistently higher power, and the power gain was larger when there was a higher dependence between the two traits. Third, the burden test and aSPUset had the lowest power across all scenarios, and always had smaller power compared to C-JAMP. Fourth, regarding the multivariate tests, MSKAT and GAMuT had identical power up to three decimals and had lower power compared to DKAT in almost all scenarios. DKAT and MURAT had often similar power and in some scenarios DKAT had higher power, in other scenarios MURAT had higher power. Fifth, in a comparison of C-JAMP with all competing approaches, C-JAMP had the highest power in all scenarios when traits with high dependence were analyzed, except when 50% of variants were causal and all had moderate/small effect sizes (Scenarios 11 and 12), where SKAT-O or DKAT had slightly higher power. When traits with weak or moderate dependence were analyzed, whether C-JAMP or competing approaches had the highest power depended on the effect size and the

percentage of causal variants. In this case, C-JAMP still had the highest power compared to all competing approaches when effect sizes were high (Scenarios 1, 4, 7, and 10). When effect sizes were moderate (Scenarios 2, 5, 8, and 11), C-JAMP had similar power compared to SKAT, SKAT-O, MURAT, aSPUset-Score, MSKAT, and GAMuT if 5%, 10%, or 20% SNVs in a gene were causal (Scenarios 2, 5, and 8), and lower power if 50% SNVs in a gene was causal (Scenario 11). Finally, when effect sizes were small (Scenarios 3, 6, 9, and 12), C-JAMP had similar power for 5% or 10% causal SNVs and smaller power for 20% and 50% causal SNVs in a gene.

In the first sensitivity checks, the power of C-JAMP was investigated when SNVs with MAC ≤ 2 or with MAC ≤ 4 were removed from the analysis (Table S11) to control the type I error more stringently for small nominal values (i.e., $\alpha = 2.5 \times 10^{-6}$). The results interestingly showed that the power of C-JAMP was consistently higher compared to the analyses where all SNVs in a gene were included. This was even more apparent for low-powered scenarios.

Second, when bivariate normally distributed traits were analyzed by applying C-JAMP with a misspecified copula function (Table S12), C-JAMP still had the highest power in scenarios with large genetic effect sizes (Scenario 1) for every trait dependence level—in Scenario 10, all approaches except aSPUset had a power of 1. When there were 5% causal SNVs with moderate or weak effect sizes (Scenarios 2 and 3), C-JAMP had similar power to MURAT, MSKAT, GAMuT and DKAT for every trait dependence level except for Scenario 3 and weak dependence, where C-JAMP had lower power. Finally, for 50% causal SNVs with moderate or weak effect sizes (Scenarios 11, 12), C-JAMP had consistently lower power compared to MSKAT, GAMuT, and DKAT for every trait dependence level, and partly also compared to MURAT.

In further sensitivity checks, if the genetic effects were only affecting the first trait and were absent on the second trait (Scenarios 13–17), then the power of C-JAMP was not affected and did not decrease (see Table S13). The power of aSPUset, aSPUset-Score, MURAT, MSKAT, GAMuT, and DKAT, however, decreased markedly when the genetic effect was absent or smaller on the second trait.

3.3 | Analysis of adiponectin levels

For the analysis of adiponectin traits, Table S14 shows descriptive characteristics of the study population, and Figure S6 shows a scatterplot of HMW and MMW adiponectin, indicating that the 2-parameter copula was appropriate to model their bivariate distribution with higher upper-tail dependence. In support, we note that the Akaike information criterion (AIC) value under the 2-parameter copula model in (6) with the marginal models (2) and (3) was much lower than the AIC under a bivariate normal model, indicating that the copula model was a better fit. For example, the AIC value under the copula model conditioning on nongenetic covariates without conditioning on a particular SNV was 572.4 compared to an AIC value of 633.7 under a bivariate normal model with the same marginals. An inspection of quantile–quantile (Q–Q) plots of the p-values did not indicate a substantial inflation for variant-level tests using C-JAMP and SMT under univariate linear regression, or of gene-based tests using C-JAMP, the SMT under linear regression, SKAT-O, MURAT, MSKAT, or DKAT.

In the analysis using C-JAMP, in variant-level analyses, while none of the SNVs reached genome-wide significance with p-value $< 5 \times 10^{-8}$, C-JAMP identified 11 SNVs for HMW adiponectin and nine SNVs for MMW adiponectin with p-value $< 10^{-5}$, of which 7 (for HMW) and 0 (for MMW) had BH-adjusted p-value < 0.05, respectively. Gene-level analyses supported the results of the variant-level analyses and identified six genes for HMW adiponectin and one gene for MMW adiponectin with p-value $< 10^{-5}$, and 10 genes (for HMW) as well as no gene (for MMW) with BH-adjusted p-value < 0.05, respectively. All except two genes identified in the gene-level analyses were also identified in the variant-level analyses, while none of the SNVs or genes identified for HMW adiponectin overlapped with those for MMW adiponectin (see Tables 3 and 4 for detailed results).

In comparison to C-JAMP, SMTs based on linear regression yielded only 8 SNVs for HMW adiponectin and four SNVs for MMW adiponectin with p-value $< 10^{-5}$. All these SNVs were also identified by C-JAMP but had higher p-values in linear regression (see Table 3). In addition, none of these SNVs had a BH-adjusted p-value < 0.05 and gene-level tests also did not yield any gene with p-value $< 10^{-5}$ (Table 4). Gene-level tests with SKAT-O, MURAT, MSKAT, and DKAT did not yield any gene with p-value $< 10^{-5}$ or BH-adjusted p-value < 0.05 (Table 4). The conclusion for these methods did not change when other kernels were used or when singletons were included in the analysis. As a result, they provided higher p-values compared to C-JAMP in the real-data application.

4 | DISCUSSION

In this study, we introduce a novel rare-variant association test based on a joint copula model of multiple phenotypes called C-JAMP, evaluate it in comparison to different established approaches in extensive simulation studies, and apply C-JAMP in an empirical analysis of adiponectin traits. One of the main advantages of C-JAMP is that through the use of copula functions, many different dependence

TABLE 3 Results from variant-level association analyses of adiponectin traits

					C-JAMP		SMT	
rsID	**chr**	**Position**	**Gene**	**MAF**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**
HMW adiponectin								
–	17	35,097,367	RAD51D	0.011	3.2×10^{-7}	0.02	6.0×10^{-6}	0.17
rs11746883	5	151,676,482	SPARC	0.019	3.3×10^{-7}	0.02	2.7×10^{-6}	0.17
rs1029303	13	32,439,013	N4BP2L2	0.008	3.4×10^{-7}	0.02	4.1×10^{-6}	0.17
rs188769218	2	88,886,106	–	0.011	3.6×10^{-7}	0.02	6.7×10^{-7}	0.12
–	2	109,819,508	RGPD5	0.005	1.1×10^{-6}	0.03	7.7×10^{-6}	0.17
–	8	100,709,384	PABPC1	0.005	1.1×10^{-6}	0.03	7.7×10^{-6}	0.17
rs149200056	17	34,962,071	CCT6B, ZNF830	0.005	1.1×10^{-6}	0.03	7.7×10^{-6}	0.17
–	19	38,408,438	FAM98C	0.021	2.9×10^{-6}	0.06	5.9×10^{-6}	0.17
rs17055869	8	26,740,131	–	0.021	3.2×10^{-6}	0.06	1.7×10^{-5}	0.33
rs144213212	9	128,356,275	SLC27A4	0.019	3.9×10^{-6}	0.07	2.7×10^{-5}	0.48
rs75975249	15	74,513,902	–	0.013	9.8×10^{-6}	0.10	5.1×10^{-5}	0.54
MMW adiponectin								
rs36043647	2	201,277,115	CASP8	0.021	1.0×10^{-6}	0.07	4.8×10^{-6}	0.28
rs1367710075	15	32,185,288	–	0.011	1.2×10^{-6}	0.07	7.5×10^{-6}	0.33
rs74092385	14	105,463,408	MTA1	0.016	1.3×10^{-6}	0.07	2.1×10^{-6}	0.28
rs10513865	3	191,371,096	CCDC50	0.021	1.6×10^{-6}	0.07	4.7×10^{-6}	0.28
rs55656828	5	151,447,743	SLC36A1	0.013	2.8×10^{-6}	0.10	1.6×10^{-5}	0.41
rs113408613	12	112,912,550	OAS1	0.005	6.9×10^{-6}	0.15	2.8×10^{-5}	0.45
–	3	142,827,996	PCOLCE2	0.005	6.9×10^{-6}	0.15	2.8×10^{-5}	0.45
–	15	98,923,881	IGF1R	0.005	6.9×10^{-6}	0.15	2.8×10^{-5}	0.45
rs56213419	8	143,694,737	ZNF707, LOC101928160	0.008	9.8×10^{-6}	0.17	2.3×10^{-5}	0.45

Note: Shown are SNVs with *p*-values $< 10^{-5}$ in variant-level tests using C-JAMP, for HMW (upper panel) and MMW adiponectin (lower panel), and the respective *p*-values from variant-level tests using linear regression. Shown are the rs SNP ID, chromosome, position, gene symbol, minor allele frequency (MAF), unadjusted *p*-values ("*p*-value") and BH-adjusted *p*-values ("adj. *p*-value") of C-JAMP and linear regression ("SMT").

structures between traits can be modeled. Model selection approaches and goodness of fit tests (see Yilmaz, 2009, for a review) can be used to obtain a plausible copula model for the joint distribution of the considered traits in a given application. The 2-parameter copula function used in this study encompasses a wide range of dependencies, and as an Archimedean copula also encompasses random effect models (Joe, 1997). In addition, C-JAMP provides the flexibility to use different genetic models for the effect of each SNV, to use any marginal distribution of a given trait conditional on the SNV and other factors, to model other trait types such as time-to-event traits, and to jointly model traits of family members. As one of the challenges that was revealed in preliminary investigations, the large sample Wald test statistic in copula models needs an adjustment for the analysis of moderate or strong trait dependences and SNVs with very low frequency to avoid inflated type I errors. Here, we presented an empirical adjustment of the Wald test statistic, which yielded empirical type I errors close to the nominal levels. The restriction to analyze only SNVs with at least 2 or 4 copies of the minor allele can provide additional control of the type I error, which is important for real-data analyses.

Based on these adjusted Wald test statistics, our results indicate that the power of association tests can be increased through the joint modeling of multiple traits using C-JAMP. The power of C-JAMP increased when the traits had higher dependence—contrary to the performance of all competing multivariate approaches—and in some of the considered scenarios, the power of C-JAMP was 1.5× or 2× as high as the best competitor. More specifically, when traits with high dependence were analyzed, C-JAMP outperformed all competitors in all scenarios except when 50% of variants were causal and all had moderate/small effect sizes. When traits with weak or moderate dependence were analyzed, whether C-JAMP or competing approaches had the highest power depended on the effect size. SKAT-O, MURAT, MSKAT, GAMuT, and DKAT had the highest power gain compared to C-JAMP when traits were weakly dependent and many variants in a gene were causal but all had small effects sizes. Regarding the other investigated tests, aSPU and MultiPhen yielded invalid inference for the analysis of rare variants under the considered models, and the standard SMT, burden test and aSPUset always had less power compared to C-JAMP. In the simulation study, we also investigated C-JAMP when the multivariate distribution was misspecified and obtained similar results in the type I error evaluation and power comparison, except when there were a large number of causal SNVs with moderate or weak effect sizes then MSKAT, GAMuT, and DKAT had consistently higher power. These results indicated that C-JAMP with the

TABLE 4 Results from gene-level association analyses of adiponectin traits

		C-JAMP		SMT		SKAT-O		MURAT		MSKAT		DKAT	
Ensembl ID	**Gene**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**	***p*-value**	**adj. *p*-value**
HMW adiponectin													
ENSG00000198783	ZNF830	3.3×10^{-6}	0.03	3.9×10^{-5}	0.20	1.6×10^{-1}	1.00	5.8×10^{-2}	0.99	4.8×10^{-2}	1.00	8.0×10^{-2}	1.00
ENSG00000015568	RGPD5	5.5×10^{-6}	0.03	6.7×10^{-5}	0.23	6.2×10^{-2}	1.00	5.2×10^{-1}	0.99	5.5×10^{-1}	1.00	5.5×10^{-1}	1.00
ENSG00000132141	CCT6B	5.5×10^{-6}	0.03	3.9×10^{-5}	0.20	2.1×10^{-3}	1.00	9.0×10^{-4}	0.90	1.1×10^{-3}	0.99	7.1×10^{-3}	1.00
ENSG00000249035	CLMAT3	6.3×10^{-6}	0.03	2.3×10^{-5}	0.20	9.5×10^{-4}	1.00	6.7×10^{-5}	0.72	9.5×10^{-5}	0.51	1.7×10^{-3}	1.00
ENSG00000113140	SPARC	8.3×10^{-6}	0.03	5.1×10^{-5}	0.20	4.6×10^{-2}	1.00	3.5×10^{-1}	0.99	3.5×10^{-1}	1.00	3.8×10^{-1}	1.00
ENSG00000261775	–	9.8×10^{-6}	0.03	5.1×10^{-5}	0.20	2.7×10^{-5}	0.65	NA	NA	NA	NA	NA	NA
MMW adiponectin													
ENSG00000257270	–	8.1×10^{-6}	0.17	1.2×10^{-5}	0.29	4.1×10^{-4}	0.58	1.7×10^{-3}	0.99	3.7×10^{-4}	0.69	9.4×10^{-5}	0.97

Note: Shown are genes identified with p-values $< 10^{-5}$ in gene-level tests using C-JAMP, for HMW (upper panel) and MMW adiponectin (lower panel), respectively, and the corresponding p-values from gene-level tests using linear regression, SKAT-O, MURAT, MSKAT, and DKAT. Shown are the ensembl gene ID, gene symbol, unadjusted gene-level p-values ("p-value") using the Bonferroni correction for C-JAMP and BH-correction for linear regression ("SMT") to account for multiple testing of the SNVs within the gene, as well as BH-adjusted gene-level p-values ("adj. p-value"). For ENSG00000261775, MURAT, MSKAT and DKAT were not able to perform an association test since only one SNV was included in the test region.

assumed 2-parameter copula model is robust against misspecification of the dependence structure, and we recommend using this more general copula model in practical applications. Nonetheless, it is always important to assess the plausibility of the chosen copula function first in real data analyses and to use a plausible model.

A general downside of MMTs is that they rely on multiple assumptions (e.g., many SNVs in a gene are causal) and parameters (e.g., the choice of power sets), and even an integration of many different subtests does not always yield tests with optimal power, as shown by the results of the simulation studies. On the other hand, SMTs rely on much fewer assumptions. For example, since we investigated the situation where MMTs have optimal power (all causal variants have an effect in the same direction), the power of C-JAMP can be expected to compare even more favorably to MMTs when effects are not all in the same direction, since the power of C-JAMP as an SMT is not affected by this (Konigorski et al., 2017). Further, it should be noted that in the simulation study, the data was generated from the same model as in Lee et al. (2012) where the effect size of each SNV depends on the MAF of the corresponding SNV through a specified function (see Table 1). Since the weighted kernel functions in SKAT, SKAT-O, MURAT, MSKAT, GAMuT, and DKAT use weights dependent on the MAF of SNVs, this gives them an advantage and allowed us to evaluate C-JAMP in situations where these MMTs have optimal power. In addition, while multi-degree-of-freedom tests generally have the highest power when a genetic variant affects all tested traits, they provide less or no information on which traits are associated, and they lose power when only one or a few traits are associated. Therefore, in real applications, the power of C-JAMP might compare even more favorably to these MMTs.

Regarding the computational cost, C-JAMP is computationally intensive compared to standard regression approaches and the kernel-based tests due to the optimization of a more complex likelihood function, but still fast enough to be employed on a genome-wide scale. For example, the real-data association analysis of the 23,922 genes with adiponectin levels was computed in less than 4 hours on a cluster with 200 compute nodes (cf. Figure S7 for more general run times). In this regard, it would be interesting in future studies to investigate algorithmic improvements of C-JAMP to be computationally more efficient and scale to high-dimensional traits. Some competing tests such as the aSPUset and aSPUset-Score tests are computationally more intensive and not suitable for genome-wide analyses, where evidence in form of p-values smaller than 10^{-5} or 10^{-8} is needed. That is, since permutation or simulation approaches are employed to derive p-values of aSPUset and aSPUset-Score empirically, 10^k permutations are needed to be able to obtain p-values that are potentially as small as $(10^k + 1)^{-1}$.

The results from the data application to adiponectin levels supported the results from the simulation study that C-JAMP is a powerful approach for rare-variant analyses and can outperform competing approaches and yielded 20 SNVs with p-values $< 10^{-5}$. While the data application was limited to SNVs in coding regions of genes that are expressed in adipose tissue and limited by the sample size, it constitutes the first GWAS of rare SNVs with high- and medium-molecular-weight adiponectin fractions and yielded novel candidate markers for adiponectin. In our opinion, the most interesting marker for replication studies is rs11746883 in the *SPARC* gene, which has been previously implicated in obesity and chronic diseases (Kos & Wilding, 2010; Takahashi et al., 2001). This marker might help to explain the relation between the *SPARC* gene and adiponectin and their role in obesity. This illustrates the potential of applying appropriate statistical models to analyze the complex interplay of genetic factors with molecular as well as nonmolecular phenotypes to investigate more complex biological models. With the multitude of factors having a likely role in the development of complex traits (Solovieff, Cotsapas, Lee, Purcell, & Smoller, 2013), C-JAMP provides a powerful tool for association studies and beyond, for example, to investigate alternative definitions of pleiotropy (Konigorski et al., 2014).

ACKNOWLEDGMENTS

Y. E. Y. is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN 2014-04904) and the Faculty of Medicine of Memorial University of Newfoundland. The authors thank the participants of the EPIC Potsdam substudy and the staff of the study center EPIC Potsdam for their work. We are grateful to Ellen Kohlsdorf for data handling and to Sarah Moreno Garcia and Henning Damm for the sample handling and for performing the ELISA.

CONFLICT OF INTEREST

The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

All computations and visualizations were performed in R 3.3.1 and higher. The computer code used to obtain all results and to generate the simulated data is available from the corresponding author upon request. The study underlying the analysis of adiponectin levels was approved by

the Ethics Committee of the medical association of the state of Brandenburg (Germany) and all participants gave written informed consent. This data is not publicly available due to privacy or ethical restrictions. It is available on request from the corresponding author after approval of the request by the study committee.

ORCID

Stefan Konigorski http://orcid.org/0000-0002-9966-6819

REFERENCES

Aleksandrova, K., Boeing, H., Jenab, M., Bueno-de-Mesquita, H. B., Jansen, E., van Duijnhoven, F. J., ... Pischon, T. (2012). Total and high-molecular weight adiponectin and risk of colorectal cancer: The European prospective investigation into cancer and nutrition study. *Carcinogenesis*, *33*, 1211–1218.

Aschard, H., Guillemot, V., Vilhjalmsson, B., Patel, C. J., Skurnik, D., Ye, C. J., ... Zaitlen, N. (2017). Covariate selection for association screening in multi–phenotype genetic studies. *Nature Genetics*, *49*(12), 1789–1795.

Aschard, H., Vilhjálmsson, B. J., Greliche, N., Morange, P. E., Trégouët, D. A., & Kraft, P. (2014). Maximizing the power of principal-component analysis of correlated phenotypes in genome-wide association studies. *The American Journal of Human Genetics*, *94*, 662–676.

Asimit, J., & Zeggini, E. (2010). Rare variant association analysis methods for complex traits. *Annual Review of Genetics*, *44*, 293–308.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *57*, 289–300.

Boeing, H., Wahrendorf, J., & Becker, N. (1999). EPIC-Germany—A source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Annals of Nutrition and Metabolism*, *43*, 195–204.

Broadaway, K. A., Cutler, D. J., & Duncan, R. (2016). A statistical approach for testing cross-phenotype effects of rare variants. *The American Journal of Human Genetics*, *98*(3), 525–540.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*, 141–151.

Dutta, D., Scott, L., Boehnke, M., & Lee, S. (2019). Multi-SKAT: General framework to test for rare-variant association with multiple phenotypes. *Genetic Epidemiology*, *43*(1), 4–23.

Evangelou, E., Warren, H. R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., & Gao, H., ... Million Veteran Program. (2018). Genetic analysis of over one million people identifies 535 novel loci for blood pressure. *Nature Genetics*, *50*(10), 1412–1425.

Gentle, J. E. (2009). *Computational Statistics*. New York, NY: Springer.

Gumbel, E. J. (1960). Distribution des valeurs extrêmes en plusieurs dimensions. *Publications de l'Institut de statistique de l'Université de Paris*, *9*, 171–173.

He, J., Li, H., Edmondson, A. C., Rader, D. J., & Li, M. (2012). A Gaussian copula approach for the analysis of secondary phenotypes in case–control genetic association studies. *Biostatistics*, *13*(3), 497–508.

Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z., & Liu, J. F. (2015). Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity*, *115*(1), 29–36.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. London: Chapman & Hall.

Kaakinen, M., Mägi, R., Fischer, K., Heikkinen, J., Järvelin, M. R., Morris, A. P., & Prokopenko, I. (2017). A rare-variant test for high-dimensional data. *European Journal of Human Genetics*, *25*(8), 988–994.

Kim, J., Zhang, Y., & Pan, W., Alzheimer's Disease Neuroimaging Initiative. (2016). Powerful and adaptive testing for multi-trait and multi-SNP associations with GWAS and sequencing data. *Genetics*, *203*(2), 715–731.

Konigorski, S., Janke, J., Drogan, D., Bergmann, M. M., Hierholzer, J., Kaaks, R., ... Pischon, T. (2018). Prediction of circulating adipokine levels based on body fat compartments and adipose tissue gene expression. *Obesity Facts*, https://doi.org/10.1159/000502117

Konigorski, S., Yilmaz, Y. E., & Bull, S. B. (2014). Bivariate genetic association analysis of systolic and diastolic blood pressure by copula models. *BMC Proceedings*, *8*, S72.

Konigorski, S., Yilmaz, Y. E., & Pischon, T. (2016). Genetic association analysis based on a joint model of gene expression and blood pressure. *BMC Proceedings*, *10*(Suppl 7), 289–294.

Konigorski, S., Yilmaz, Y. E., & Pischon, T. (2017). Comparison of single-marker and multi-marker tests in rare variant association studies of quantitative traits. *PLOS One*, *12*, e0178504.

Kos, K., & Wilding, J. P. (2010). SPARC: A key player in the pathologies associated with obesity and diabetes. *Nature Reviews Endocrinology*, *6*(4), 225–235.

Lakhal-Chaieb, L., Oualkacha, K., Richards, B. J., & Greenwood, C. M. T. (2016). A rare variant association test in family-based designs and non-normal quantitative traits. *Statistics in Medicine*, *35*, 905–921.

Lee, S., Abecasis, G. R., Boehnke, M., & Lin, X. (2014). Rare-variant association analysis: Study designs and statistical tests. *The American Journal of Human Genetics*, *95*, 5–23.

Lee, S., Wu, M. C., & Lin, X. (2012). Optimal tests for rare variant effects in sequencing association studies. *Biostatistics*, *13*, 762–775.

Li, B., & Leal, S. M. (2008). Methods for detecting associations with rare variants for common diseases: Application to analysis of sequence data. *The American Journal of Human Genetics*, *83*, 311–321.

Liang, X., Wang, Z. Q., & Zhang, S. (2016). An adaptive Fisher's combination method for joint analysis of multiple phenotypes in association studies. *Scientific Reports*, *6*, 34323.

Lippert, C., Casale, F. P., Rakitsch, B., & Stegle, O. (2014). LIMIX: Genetic analysis of multiple traits. *BioRxiv*, 003905.

Listgarten, J., Lippert, C., Kang, E. Y., Xiang, J., Kadie, C. M., & Heckerman, D. (2013). A powerful and efficient set test for genetic markers that handles confounders. *Bioinformatics*, *29*, 1526–1533.

Liu, Z., & Lin, X. (2018). Multiple phenotype association tests using summary statistics in genome-wide association studies. *Biometrics*, *74*(1), 165–175.

Maity, A., Sullivan, P. F., & Tzeng, J. Y. (2012). Multivariate phenotype association analysis by marker-set kernel machine regression. *Genetic Epidemiology*, *36*, 686–695.

Mancuso, N., Rohland, N., Rand, K. A., Tandon, A., Allen, A., Quinque, D., ... Reich, D. (2016). The contribution of rare variation to prostate cancer heritability. *Nature Genetics*, *48*, 30–35.

Momen, M., Mehrgardi, A. A., Roudbar, M. A., Kranis, A., Pinto, R. M., Valente, B. D., ... Gianola, D. (2018). Including phenotypic causal networks in genome-wide association studies using mixed effects structural equation models. *BioRxiv*, 251421.

O'Brien, P. C. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics*, *40*(4), 1079–1087.

O'Reilly, P. F., Hoggart, C. J., Pomyen, Y., Calboli, F. C., Elliott, P., Jarvelin, M. R., & Coin, L. J. (2012). Multi-Phen: Joint model of multiple phenotypes can increase discovery in GWAS. *PLOS One*, *7*, e34861.

PCAWG Transcriptome Core Group, Calabrese, C., Davidson, N. R., Fonseca, N. A., He, Y., ... Zhang, Z. (2018). Genomic basis for RNA alterations revealed by whole-genome analyses of 27 cancer types. *BioRxiv*, 183889.

Pischon, T. (2009). Use of obesity biomarkers in cardiovascular epidemiology. *Disease Markers*, *26*, 247–263.

Ray, D., & Basu, S. (2017). A novel association test for multiple secondary phenotypes from a case-control GWAS. *Genetic Epidemiology*, *41*, 413–426.

Riboli, E., & Kaaks, R. (1997). The EPIC Project: Rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology*, *26*(Suppl 1), S6–S14.

Rosen, O., & Thompson, W. K. (2015). Bayesian semiparametric copula estimation with application to psychiatric genetics. *Biometrical Journal*, *57*, 468–484.

Schifano, E. D., Li, L., Christiani, D. C., & Lin, X. (2013). Genome-wide association analysis for multiple continuous secondary phenotypes. *The American Journal of Human Genetics*, *92*(5), 744–759.

Schillert, A., & Konigorski, S. (2016). Joint analysis of multiple phenotypes—Summary of results and discussions from the Genetic Analysis Workshop 19. *BMC Genetics*, *17*(Suppl 2), 67–73.

Schork, N. J., Murray, S. S., Frazer, K. A., & Topol, E. J. (2009). Common vs. rare allele hypotheses for complex diseases. *Current Opinion in Genetics & Development*, *19*, 212–219.

van der Sluis, S., Posthuma, D., & Dolan, C. V. (2013). TATES: Efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLOS Genetics*, *9*(1):e1003235.

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: Challenges and strategies. *Nature Review Genetics*, *14*, 483–495.

Song, Y. E., Morris, N. J., & Stein, C. M. (2016). Structural equation modeling with latent variables for longitudinal blood pressure traits using general pedigrees. *BMC Proceedings*, *10*(Suppl 7), 303–307.

Sun, J., Oualkacha, K., Forgetta, V., Zheng, H. F., Richards, B. J., Ciampi, A., & Greenwood, C. M. T., UK10K Consortium. (2016). A method for analyzing multiple continuous phenotypes in rare variant association studies allowing for flexible correlations in variant effects. *European Journal of Human Genetics*, *24*(9), 1344–1351.

Takahashi, M., Nagaretani, H., Funahashi, T., Nishizawa, H., Maeda, N., Kishida, K., ... Matsuzawa, Y. (2001). The expression of SPARC in adipose tissue and its increased plasma concentration in patients with coronary artery disease. *Obesity Research*, *9*(7), 388–393.

Telenti, A., Pierce, L. C., Biggs, W. H., di Iulio, J., Wong, E. H., Fabani, M. M., ... Venter, J. C. (2016). Deep sequencing of 10,000 human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 11901–11906.

Verhulst, B., Maes, H. H., & Neale, M. C. (2017). GW-SEM: A statistical package to conduct genome-wide structural equation modeling. *Behavior Genetics*, *47*, 345.

Wang, Z., Wang, X., Sha, Q., & Zhang, S. (2016). Joint analysis of multiple traits in rare variant association studies. *Annals of Human Genetics*, *80*, 162–171.

Wu, B., & Pankow, J. S. (2016). Sequence kernel association test of multiple continuous phenotypes. *Genetic Epidemiology*, *40*(2), 91–100.

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *The American Journal of Human Genetics*, *89*, 82–93.

Yang, Q., & Wang, Y. (2012). Methods for analyzing multivariate phenotypes in genetic association studies. *Journal of Probability and Statistics*, *2012*, 652569.

Yilmaz, Y.E. (2009). Estimation and goodness of fit for multivariate survival models based on copulas. UWSpace. Retrieved from http://hdl.handle.net/10012/4571

Zhan, X., Zhao, N., Plantinga, A., Thornton, T. A., Conneely, K. N., Epstein, M. P., & Wu, M. C. (2017). Powerful genetic association analysis for common or rare variants with high-dimensional structured traits. *Genetics*, *206*(4), 1779–1790.

Zhang, H., Qin, J., Landi, M., Caporaso, N., & Yu, K. (2013). A copula-model based semiparametric interaction test under the case-control design. *Statistica Sinica*, *23*(4), 1505–1521.

Zhang, Y., Xu, Z., Shen, X., Pan, W., & Alzheimer's Disease Neuroimaging Initiative. (2014). Testing for association with multiple traits in generalized estimation equations, with application to neuroimaging data. *Neuroimage*, *96*, 309–325.

Zhao, J., & Zhang, H. (2016). Modeling multiple responses via bootstrapping margins with an application to genetic association testing. *Statistics and its Interface*, *9*(1), 47–56.

Zhu, H., Zhang, S., & Sha, Q. (2015). Power comparisons of methods for joint association analysis of multiple phenotypes. *Human Heredity*, *80*, 144–152.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Konigorski S, Yilmaz YE, Janke J, Bergmann MM, Boeing H, Pischon T. Powerful rare variant association testing in a copula-based joint analysis of multiple phenotypes. *Genet. Epidemiol.* 2020;44:26–40. https://doi.org/10.1002/gepi.22265