

CERAMIC: a C Software for Mixed-Model Association Analysis of Binary Traits.

Version 1.0
October 23, 2016

Sheng Zhong¹, Duo Jiang², Mary Sara McPeck^{1,3}

Departments of Statistics¹ and Human Genetics³
The University of Chicago, Chicago, IL 60637, USA
Department of Statistics Oregon State University, Corvallis OR 97331, USA

CERAMIC (Binary Association Testing with Mixed-model ANalysis)

A C software for mixed-model case-control association analysis with covariates, related individuals and missing data.

Copyright(C) 2016 Sheng Zhong, Duo Jiang and Mary Sara McPeck

Homepage: <http://www.stat.uchicago.edu/~mcpeek/software/index.html>

Release 1.0 October 23, 2016

License

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY of FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program (see file `gpl.txt`); if not, write to the Free Software Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.

This program includes code provided by others under licenses compatible with GNU GPL as free software:

LAPACK, which is under the 3-clause BSD license.

Numerical recipes utility functions, which are public domain.

GNU Scientific Library, which is under GNU GPL version 3 (or later)

MQLS-XM, which is under GNU General Public License (version 3 or later)

We request that use of this software be cited in publications as follows:

Zhong, Sheng, Duo Jiang, and Mary Sara McPeck. "CERAMIC: Case-Control Association Testing in Samples with Related Individuals, Based on Retrospective Mixed Model Analysis with Adjustment for Covariates." *PLoS Genet* 12.10 (2016): e1006329.

Contents

1	Overview of CERAMIC	4
2	Installation of CERAMIC	4
2.1	Installation Instructions	4
2.2	Software Requirements for Installing CERAMIC	5
3	Using CERAMIC	5
3.1	Preliminary Steps	5
3.2	Using CERAMIC for Analysis of Phenotype Data under Null Hypothesis	6
3.3	Using CERAMIC for Genome-Wide Association Analysis with Complete Data	7
3.4	Using CERAMIC for Genome-Wide Association Analysis with Missing Data	7
4	Input Files	9
5	Output Files	12
6	Examples	13
7	Acknowledgments	14

1 Overview of CERAMIC

The CERAMIC software is a C program, created primarily for an implementation of CERAMIC, a novel binary trait association testing method for related individuals adjusted for covariates and incorporating missing data. CERAMIC tests for the association between a binary trait and a SNP using a mixed-model approach.

The CERAMIC testing method, developed based on an estimating equation approach, can be viewed as a hybrid of logistic regression and the linear mixed model (LMM) approaches. CERAMIC is a generalization of a previously reported association test for binary traits, MQLS [3], which does not adjust phenotypes for covariate and additive polygenic effects. Main features of the CERAMIC test are summarized as follows: (1) it is applicable to samples of arbitrarily related individuals, including unrelated individuals as well as individuals related by large, complex inbred pedigrees; (2) it incorporates information on individuals with partially missing data while correctly accounting for dependence; (3) it corrects binary phenotypes for covariates and additive polygenic effects with a quasi-likelihood model; and (4) it is computationally feasible (by avoiding the full generalized linear mixed model (GLMM) framework) for association studies with hundreds of thousands of markers.

In addition to the CERAMIC test, the CERAMIC software implements two other novel association tests for binary traits: MQLS-LIN and MQLS-LOG. Both tests take the retrospective approach similar to the CERAMIC test. However, MQLS-LIN adjusts a binary trait for covariates based on a classic linear regression model, while MQLS-LOG does so based on a classic logistic regression model ignoring familial correlation. Neither MQLS-LIN or MQLS-LOG accounts for relatedness among individuals when estimating phenotypic residuals. For details of the CERAMIC, MQLS-LIN, and MQLS-LOG methods, please refer to Zhong et al. (2016). [5].

CERAMIC also offers an implementation of MASTOR test [1], which takes a LMM approach, and is designed for association mapping of quantitative phenotypes.

Furthermore, our software CERAMIC can be used for association analysis of SNPs on X chromosomes, with a sample of related individuals and covariates, because it extends MQLS-XM [4] program by replacing the raw phenotypes by the phenotypic residuals adjusted for covariates and additive polygenic effects based on several different phenotype null models (i.e., the logistic-mixed model in the CERAMIC test, the LMM in MASTOR, the classic linear regression model in MQLS-LIN, and the classic logistic regression model in MQLS-LOG).

In the absence of genotype data, CERAMIC accommodates the option of only estimating parameters in the (null) phenotypic model with covariates and additive polygenic effects. This facilitates the analysis of phenotype data under the null hypothesis (of no association or linkage with a tested SNP).

2 Installation of CERAMIC

2.1 Installation Instructions

To install CERAMIC, you may proceed with the following steps:

1. Download the CERAMIC package. The package contains the documentation, source code, example files, binary executable, and the GNU GPL license.
2. Read the entire documentation (this document) carefully to understand the purpose of this software and how it works. It is highly recommended that users read the paper on the CERAMIC test [5].
3. Decompress the archive with GNU software `gzip`: `tar xvfz CERAMIC_v1.0.tar.gz`.
4. Change working directory to the newly created one (i.e, CERAMIC) by Step 3: `cd CERAMIC`.
5. This directory includes the GNU GPL license in file `gpl.txt` and four subdirectories:
 - `bin` contains a pre-compiled binary executable file (x86 64bit Linux);
 - `src` contains the source code;
 - `doc` contains this document `CERAMIC_v1.0_doc.pdf`;
 - `examples` contains example input and output files.
6. If the pre-compiled binary executable works on your system, no installation is necessary. Otherwise, follow Steps 7 and 8 to compile CERAMIC by yourself. To test if the pre-compiled binary works, you may switch to the `examples` directory and run (see `README.txt` for more details):


```
../bin/CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -N
```

In some cases, you may need to use `chmod a+x ../bin/CERAMIC` before using the binary executable.

When optimal computational efficiency is important (for example, for large data sets), it is recommended that the user compile CERAMIC on one's own computer system instead of using the pre-compiled executable.
7. If you want to compile MONSTER by yourself, make sure the compilation prerequisites explained in Section 2.2 are met.
8. Switch to the `src` directory. Type `make`. This will build an executable program called CERAMIC. Note that successful completion of this step requires a GNU gcc compiler.

2.2 Software Requirements for Installing CERAMIC

This section is relevant only when you want to compile the software by yourself. To ensure proper installation of CERAMIC, the LAPACK software package needs to be installed first. LAPACK is a free software package under the modified BSD license, and can be downloaded from netlib (see <http://www.netlib.org/lapack>). In addition, you will need a standard gcc compiler, such as GNU gcc (see <http://gcc.gnu.org/>) to build the source code.

3 Using CERAMIC

3.1 Preliminary Steps

To run CERAMIC, you may need to complete the following steps:

1. Create a directory for running CERAMIC, and copy the executable program CERAMIC to the directory.
2. Create input files with formats specified in Chapter 4, and copy them to the directory where the CERAMIC program is located. For the purpose of learning how to use the software, you may also use the example input files that come with the package.
3. Switch to the directory, and run CERAMIC by typing `./CERAMIC` with various options listed below. In some cases, you may need to run `chmod a+x CERAMIC` first.

3.2 Using CERAMIC for Analysis of Phenotype Data under Null Hypothesis

CERAMIC can be used to analyze the phenotype data only by fitting the phenotypic model under the null hypothesis without performing any association test. In this case, a genotype file is not required to run CERAMIC.

To demonstrate how to use CERAMIC in this scenario, we give some examples:

```
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -N
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -N -D
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -N -L
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -N -G
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -N -M
```

Comments:

- **-p yourphenocovfile.txt:** This option allows the user to specify the name of the phenotype data input file. This file includes phenotype and covariate information. The filename defaults to `phenofile.txt` if this flag is not used. To specify another filename, replace `yourphenocovfile.txt` with the appropriate filename. See Chapter 4 for format details.
- **-k yourkinfile.txt:** This option allows the user to specify the name of the kinship information input file. This file includes kinship coefficients and inbreeding coefficients. The filename defaults to `kinfile.txt` if this flag is not used. To specify another filename, replace `yourkinfile.txt` with the appropriate filename. See Chapter 4 for format details.
- **-N:** This option specifies that only phenotype data analysis (under the null hypothesis) is performed by CERAMIC and that no association test is conducted. The option is necessary if the user does not have genotype input files.
- CERAMIC software supports four (null) phenotype models: the logistic-mixed model in the CERAMIC test, the LMM in MASTOR, the classic linear regression model in MQLS-LIN, and the classic logistic regression model in MQLS-LOG.
 - The default option corresponds to the logistic-mixed model in the CERAMIC test.
 - **-L:** the classic linear regression model in MQLS-LIN.
 - **-G:** the classic logistic regression model in MQLS-LOG.
 - **-M:** the LMM in MASTOR.
 - **-D:** the logistic-mixed model in the CERAMIC test with an extra dispersion parameter for phenotypic variance.

– Note that only *one* option of above can be used for each run of CERAMIC.

- In the presence of missing data, CERAMIC uses all phenotyped individuals with non-missing covariates when fitting a phenotypic model.

3.3 Using CERAMIC for Genome-Wide Association Analysis with Complete Data

The primary purpose of CERAMIC is to perform genome-wide association testing. With complete data, we include the same set of individuals when analyzing different SNPs. This can greatly reduce computation burden, because the kinship matrix needs to be inverted only once for all SNPs.

Here are some example command lines:

```
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -g yourgenofile.txt  
-n yourSNPnamefile.txt -c  
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -g yourgenofile.txt  
-n yourSNPnamefile.txt -M -c
```

Comments:

- **-g yourgenofile.txt:** This option allows the user to specify the name of the genotype data input file. This file includes genotype data information. The filename defaults to `genofile.txt` if this flag is not used. To specify another filename, replace `yourgenofile.txt` with the appropriate filename. See Chapter 4 for format details.
- **-n yourSNPnamefile.txt:** This option allows the user to specify the name of the file that contains the names of SNPs included in analysis. If this option is omitted, SNPs is numbered from 1 to the total number of rows in `yourgenofile.txt`. See Chapter 4 for format details.
- **-c:** This option is necessary if the user needs CERAMIC to perform faster analysis for complete data. If the data actually contain missing values, the program may crash.

3.4 Using CERAMIC for Genome-Wide Association Analysis with Missing Data

In the presence of missing data, CERAMIC supports various command options to improve flexibility of genome-wide association analysis. To facilitate the understanding of the various command options, we first introduce some notation to denote four subsets of individuals.

- W_0 : the set of individuals who have non-missing phenotypes and non-missing covariates.
- W_1 : an individual in W_1 has non-missing phenotypes and non-missing covariates, and satisfies at least one of the following three conditions: (1) the individual is genotyped; (2) the individual has a genotyped relative; and/or (3) the individual is in the same pedigree with an individual with non-missing phenotype and covariates who either has non-missing genotype or has a relative with non-missing genotype at the test locus.

- W_2 : the set of individuals who have non-missing phenotypes and non-missing covariates, and are genotyped or have a genotyped relative.
- Q : the set of genotyped individuals with non-missing covariates.

Below is the default command for association analysis with missing data:

```
./CERAMIC -p yourphenocovfile.txt -k yourkinfile.txt -g yourgenofile.txt
-n yourSNPnamefile.txt
```

Under the default command, CERAMIC runs the CERAMIC test. To invoke MQLS-LIN test, the option `-N` should be added. See subsection 3.2 for other options. Note that only *one* method option can be chosen for each run of CERAMIC.

Under the default command, CERAMIC estimates the heritability once per genome screen using individuals in W_0 , and then computes, for each SNP, the regression parameters (and hence phenotypic residuals) separately using individuals in W_1 . CERAMIC also supports several other options:

- **-a**: Allows the CERAMIC test (or the MASTOR test) to estimate the phenotypic residuals once per genome screen using individuals in W_0 for analysis of *all* SNPs. This option is made available for the CERAMIC (default) and MASTOR (`-M`) methods only in order to speed up computation. When MQLS-LIN (`-L`) or MQLS-LOG (`-G`) is used, however, the option `-a` is not allowed by the current version of the software because computational burden is rarely a concern for MQLS-LIN and MQLS-LOG.
- **-s**: Allows the CERAMIC test (or the MASTOR test) to estimate the phenotypic residuals for each SNP separately using individuals in W_1 .
- **-f**: Allows the set W_2 to be used in replace of W_1 for the default option, or the option `-s`.

Under the default command, CERAMIC estimates the genotypic variance based on the generalized least square regression of genotypes on covariates, using individuals in Q . CERAMIC also supports two other options:

- **-i**: Allows CERAMIC to estimate the genotypic variance based on the generalized least square regression of genotypes with an intercept as the only covariate.
- **-h**: Allows CERAMIC to estimate the genotypic variance by plugging in the best linear unbiased estimator of the allele frequency [2] of a tested SNP. This option assumes Hardy-Weinberg Equilibrium for the tested SNP.

CERAMIC has some additional useful options, which are listed below:

- **-x**: Allows CERAMIC to perform association analysis for SNPs on chromosome X.
- **-u**: Drops from the analysis all genotyped individuals with either missing phenotypes and/or missing covariates.
- **-m**: Drops from the analysis all individuals who have non-missing phenotypes and/or non-missing covariates, but are not genotyped. This option should not be used together with the option `-a`.

- **-r**: Allows CERAMIC to use a slow genotype reader to read in a text genotype file. The slow reader performs some format checks.
- **-b**: Allows CERAMIC to read in binary genotype files.

4 Input Files

To perform phenotype data analysis under the null hypothesis, the user only needs to have an input file for phenotype and covariate information, and an input file for kinship information. To perform association analysis, a genotype file is required additionally. A SNP name file is not necessary, and if it is not provided, in the output file(s), the SNPs will be numbered from 1 to the total number of rows in the genotype data file. The formats for the input files are described below.

1. The phenotype and covariate data file (as in `-p yourphenocovfile.txt`):

The phenotype and covariate data file has at least three columns, which are as follows:

```
family ID (positive integer)
individual ID (positive integer)
phenotype (three values: 0, 1, 2)
covariate1 (numeric, optional)
covariate2 (numeric, optional)
⋮
```

The phenotype and covariate data file must satisfy the following format requirements:

- Tab or space delimited.
- Families must be numbered consecutively from 1 to F , where F is the total number of families in the sample.
- Individuals in a family must be grouped together in the file (order within the family does not matter). Individual IDs must be positive integers.
- An individual who is related to no other individuals in the sample must be included as a family with the individual as the only individual in the family, and a family ID must be assigned.
- The phenotypes in the file are coded as follows: 1 for unaffected, 2 for affected, 0 for unknown.
- An Intercept should *not* be included in this file as a covariate. In fact, CERAMIC automatically adds an intercept to the phenotype model.
- The missing value for an unobserved covariate should be encoded as -9999.
- The maximum number of covariates allowed is currently 10. This can be changed by modifying `MAXCOV` in the file `CERAMIC.h` in the directory `/src`.
- The maximum number of families allowed is currently 10000. This can be changed by modifying `MAXFAM` in the file `CERAMIC.h` in the directory `/src`.

To illustrate the format given above, consider a sample of 3 families with 6 individuals, with 2 observed covariates. The phenotype and covariate data file may look like the following:

```
1 8456 1 1 34.24
1 3379 1 2 30.34
1 38 2 1 35.66
2 439 0 2 34.10
2 543 1 1 -9999.0
3 1835 2 -9999 38.23
```

2. The kinship data file (as in `-k yourkinfile.txt`):

The kinship data file gives the kinship coefficient for a pair of individuals, and the inbreeding coefficient for an individual in the sample. The kinship data file has four columns, which are as follows:

```
family ID
individual1's ID
individual2's ID
relationship coefficient
```

If individual1's ID is different from individual2's ID, then the corresponding relationship coefficient is their kinship coefficient. If individual1's ID is the same as individual2's ID, then the corresponding relationship coefficient is individual1's (or equivalently individual2's) **inbreeding coefficient** (not the self-kinship coefficient, so if the individual is outbred, this value should be 0, not 1).

It is a requirement that for each family, every pair (including same-individual pairs) of individuals included in the phenotype and covariate data file should appear in the kinship data file, and an individual that appears in the kinship data file should also appear in the phenotype and covariate data file. The kinship data file is tab or space delimited.

For the sample of 3 families with 6 individuals, the kinship data file looks like:

```
1 8456 8456 0
1 8456 3379 .25
1 8456 38 .25
1 3379 3379 0
1 3379 38 .25
1 38 38 0
2 439 439 0
2 439 543 .25
2 543 543 0
3 1835 1835 .25
```

There is no requirement on the order different rows appear in the kinship data file. However, the order in the example above is recommended for optimal computation time. The program KinInbCoef.c (available at <http://galton.uchicago.edu/~mcpeek/software/KinInbcoef/index.html>) can output the format required for CERAMIC. Note that this program only handles numerical IDs.

3. The genotype data file (as in `-g yourgenofile.txt`):

The genotype data file gives the genotypes of all individuals at all SNPs to be included in genome-wide analysis. Each row in the file encodes the genotypes (for all individuals) at a single SNP to be tested, and each column corresponds to an individual's genotypes at all SNPs included for analysis. The ordering of the columns should be corresponds exactly to the phenotype and covariate data file.

The genotype data file allows bi-allelic SNPs only. Suppose the two types of alleles for a SNP are named type 1 and type 0. Then three genotypic values are allowed: 0 for zero copy of the type 1 allele, 1 for one copy of the type 1 allele, and 2 for tow copies of the type 1 alleles. A missing value should be coded as -9. The genotype data file is tab or space delimited.

For a sample of 3 families with 6 individuals and 4 tested SNPs, the genotype data file may look like :

2	1	0	0	-9	0
-9	1	2	0	1	2
-9	-9	0	1	1	1
1	1	1	2	2	2

Again, each column represents an individual, and each row represents a SNP.

4. The SNP name file (as in `-n yourSNPnamefile.txt`):

The SNP name file is optional, and it contains only one column of SNP names which appear in the same order as SNPs appear in the genotype data file. A SNP name is a string with less than 99 characters. If a SNP name file is not provided, SNPs in the genotype file will be named consecutively from 1 in the output file(s).

Here is an example of a SNP name file with 4 SNPs in the above genotype data file:

(Note that the comments in parentheses in the example below are added for explaining purpose, and hence should be omitted.)

```
SNP_A-4294555 (Name for SNP1)
SNP_A-2291718 (Name for SNP2)
SNP_A-4205622 (Name for SNP3)
SNP_A-2101610 (Name for SNP4)
```

5 Output Files

The CERAMIC software outputs up to 4 files: `CERAMICtest.phenoestimates`, which gives the estimates (and their standard errors) of the applicable parameters in a phenotypic null model, `CERAMICtest.pvalues`, which gives the p-values for all SNPs present in the genotype data file, `CERAMICtest.top`, which gives the p-values for the top 20 most significant SNPs, and `CERAMICtest.err`, which gives the information on any errors occurred in the process of running CERAMIC. A description for each output file is given below.

1. **CERAMICtest.phenoestimates** File:

This output file gives the estimates (and their standard errors) of the coefficients on covariates and/or the variance parameters (only for the CERAMIC test and the MASTOR test) in the null phenotype model. Depending on the options (`-a`, `-s`, or the default option) the user has specified, `CERAMICtest.phenoestimates` outputs the global estimates based on all individuals with non-missing phenotypes and covariates, and/or the SNP-specific estimates based on a subset of individuals who are included depending on the genotype missing pattern of the tested SNP under consideration. Ample comments are made in the file to help the user to find out which option has been invoked, and what various portions of results are intended for.

In `CERAMICtest.phenoestimates` file, a typical table showing the estimation results of the CERAMIC method looks like:

Parameter	Estimate	SE
Heritability	0.409074	NA
Additive_Var	NA	NA
Error_Var	NA	NA
Intercept	-6.33767	0.412643
Covariate_1	-0.752426	0.109866
Covariate_2	0.235426	0.0136762

For the MASTOR test and the CERAMIC test with the extra dispersion parameter, additive variance and error variance estimates are available. For MQLS-LIN and MQLS-LOG, heritability estimate is not available.

2. **CERAMICtest.pvalues** File:

This output file has three columns: SNP label (from 1 to the total number of SNPs tested, in the same order as in the genotype data file), SNP name, and the association p-value given by the chosen test.

For an association study with 4 SNPs, `CERAMICtest.pvalues` may look like:

SNP	NAME	CERAMIC_Robust
1	SNP_A-4294555	0.67826
2	SNP_A-2291718	0.988989
3	SNP_A-4205622	0.416211
4	SNP_A-2101610	0.709789

In the header of the third column, “_Robust” implies that the genotypic variance is estimated without assuming Hardy-Weinberg equilibrium. If the option `-h` is used, the header of the third column becomes “CERAMIC_HWE”.

3. **CERAMICtest.top** File:

This output file gives the names and p-values for the top 20 most significant SNPs. To show more (or less) top p-values, modification can be made to the variable `MAXTOP` in the file `CERAMIC.h` in the directory `/src`.

4. **CERAMICtest.err** File:

This output file gives the information on any errors occurred during the execution of the CERAMIC program.

6 Examples

The directory `CERAMIC/examples/` provides example input files: `phenocovfile_ex.txt`, `genofile_ex.txt`, `SNPnamefile_ex.txt`, and `kinfile_ex.txt`. A number of example command lines (followed by comments) are given below for reference.

1. `./CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -g genofile_ex.txt -n SNPnamefile_ex.txt -a -i`

The CERAMIC test is performed with phenotypic residuals estimated once per genome using all phenotyped individuals with non-missing covariates. The genotype variance is estimated by the generalized least square regression of genotypes on an intercept.

2. `./CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -g genofile_ex.txt -n SNPnamefile_ex.txt -s -h -f`

The CERAMIC test is performed with phenotypic residuals estimated separately for each SNP using the set W_2 . The genotype variance is estimated by a plug-in estimator under the assumption of Hardy-Weinberg equilibrium.

3. `./CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -g genofile_ex.txt
-n SNPnamefile_ex.txt -m`

The CERAMIC test is performed. Heritability is estimated once per genome screen using the set W_0 , and the regression coefficients on the covariates (and hence phenotypic residuals) are estimated separately for each SNP using individuals with observed genotypes, phenotypes and covariates.

4. `./CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -g genofile_ex.txt
-n SNPnamefile_ex.txt -G -m -u`

MQLS-LOG test is performed. The phenotypic residuals are estimated separately for each SNP based on the set of individuals with non-missing phenotypes, non-missing covariates, and non-missing genotypes. We do not recommend the use of the option `-a` for MQLS-LOG.

5. `./CERAMIC -p phenocovfile_ex.txt -k kinfile_ex.txt -g genofile_ex.txt
-n SNPnamefile_ex.txt -L -u`

MQLS-LIN test is performed. The phenotypic residuals are estimated separately for each SNP based on the set W_1 . Then the set of genotyped individuals who have either missing phenotypes or covariates are dropped from the calculation of the MQLS-LIN test statistic. We do not recommend the use of the option `-a` for MQLS-LIN.

7 Acknowledgments

We gratefully acknowledge:

- MQLS-XM, based on which the CERAMIC software is developed,
- LAPACK, for the `dposv` routine,
- Numerical Recipes in C, for the utility functions in `nrutil.h`,
- Mark Abeny for sharing the Cholesky decomposition routine,
- GNU Scientific library, for the implementation of Brent's method,
- Valgrind and gdb, for software debugging.

References

- [1] Johanna Jakobsdottir and Mary Sara McPeck. MASTOR: Mixed-model association mapping of quantitative traits in samples with related individuals. *The American Journal of Human Genetics*, 92(5):652–666, 2013.
- [2] M.S. McPeck, X. Wu, and C. Ober. Best linear unbiased allele-frequency estimation in complex pedigrees. *Biometrics*, 60(2):359–367, 2004.
- [3] T. Thornton and M.S. McPeck. Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *The American Journal of Human Genetics*, 81(2):321–337, 2007.
- [4] Timothy Thornton, Qian Zhang, Xiaochen Cai, Carole Ober, and Mary Sara McPeck. Xm: Association testing on the x-chromosome in case-control samples with related individuals. *Genetic epidemiology*, 36(5):438–450, 2012.
- [5] Sheng Zhong, Duo Jiang, and Mary Sara McPeck. Ceramic: Case-control association testing in samples with related individuals, based on retrospective mixed model analysis with adjustment for covariates. *PLoS Genet*, 12(10):e1006329, 2016.