MCPerm: A Monte Carlo Permutation Method for Accurately Correcting the Multiple Testing in a Meta-Analysis of Genetic Association Studies

Yongshuai Jiang[1]*[¶], Lanying Zhang[1][¶], Fanwu Kong[2][¶], Mingming Zhang[1][¶], Hongchao Lv[1][¶], Guiyou Liu[3], Mingzhi Liao[4], Rennan Feng[5], Jin Li[1], Ruijie Zhang[1]*

1 College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China, 2 Department of Nephrology, The Second Affiliated Hospital, Harbin Medical University, Harbin, China, 3 Genome Analysis Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China, 4 College of Life Science, Northwest A&F University, Yangling, Shaanxi, China, 5 Department of Nutrition and Food Hygiene, School of Public Health, Harbin Medical University, Harbin, China

Abstract

Traditional permutation (TradPerm) tests are usually considered the gold standard for multiple testing corrections. However, they can be difficult to complete for the meta-analyses of genetic association studies based on multiple single nucleotide polymorphism loci as they depend on individual-level genotype and phenotype data to perform random shuffles, which are not easy to obtain. Most meta-analyses have therefore been performed using summary statistics from previously published studies. To carry out a permutation using only genotype counts without changing the size of the TradPerm P-value, we developed a Monte Carlo permutation (MCPerm) method. First, for each study included in the meta-analysis, we used a two-step hypergeometric distribution to generate a random number of genotypes in cases and controls. We then carried out a meta-analysis using these random genotype data. Finally, we obtained the corrected permutation P-value of the meta-analysis by repeating the entire process N times. We used five real datasets and five simulation datasets to evaluate the MCPerm method and our results showed the following: (1) MCPerm requires only the summary statistics of the genotype, without the need for individual-level data; (2) Genotype counts generated by our two-step hypergeometric distributions had the same distributions as genotype counts generated by shuffling; (3) MCPerm had almost exactly the same permutation P-values as TradPerm ($r = 0.999$; $P<2.2e-16$); (4) The calculation speed of MCPerm is much faster than that of TradPerm. In summary, MCPerm appears to be a viable alternative to TradPerm, and we have developed it as a freely available R package at CRAN: http://cran.r-project.org/web/packages/MCPerm/index.html.

Introduction

Meta-analysis is an important method that can improve statistical power by combining the results of multiple previously published studies [1,2]. Every year, thousands of meta-analyses of genetic association studies are published. Some focus on only one single nucleotide polymorphic (SNP) locus [3–5], while others consider multiple SNP loci [6–9]. For those meta-analyses based on multiple loci, multiple testing is a persistent problem. Many techniques have been devised to correct multiple hypotheses, such as the Bonferroni and Sidak corrections [10], false discovery rate [11], and permutation [12]. Although these methods are highly successful, the traditional permutation (TradPerm) method is still widely considered the gold standard for accurately correcting for multiple testing [12,13].

TradPerm is a type of non-parametric method in which the null distribution of the test statistic is estimated by shuffling uniform phenotype labels of cases and controls [13,14]. However, it relies on the original SNP genotype and phenotype data (individual-level data) to perform a large number of random shuffles [15]. For the meta-analyses of genetic association studies, individual-level data are difficult to obtain for several reasons including research project privacy. Compared with raw SNP genotype data, count-based summary statistics of genotype data are more readily available, so most meta-analyses of genetic associations have been performed using summary statistics from previously published studies. However, for researchers of meta-analyses, it can be challenging to perform the permutation test using only the summary statistics of the genotype when similarly high standards to TradPerm are expected.

Here, we developed a Monte Carlo permutation (MCPerm) method for use in the permutation test. MCPerm employs a two-step hypergeometric distribution to generate random genotype counts in cases and controls. We used these random genotype counts to construct the background distribution of meta-analysis P-values and
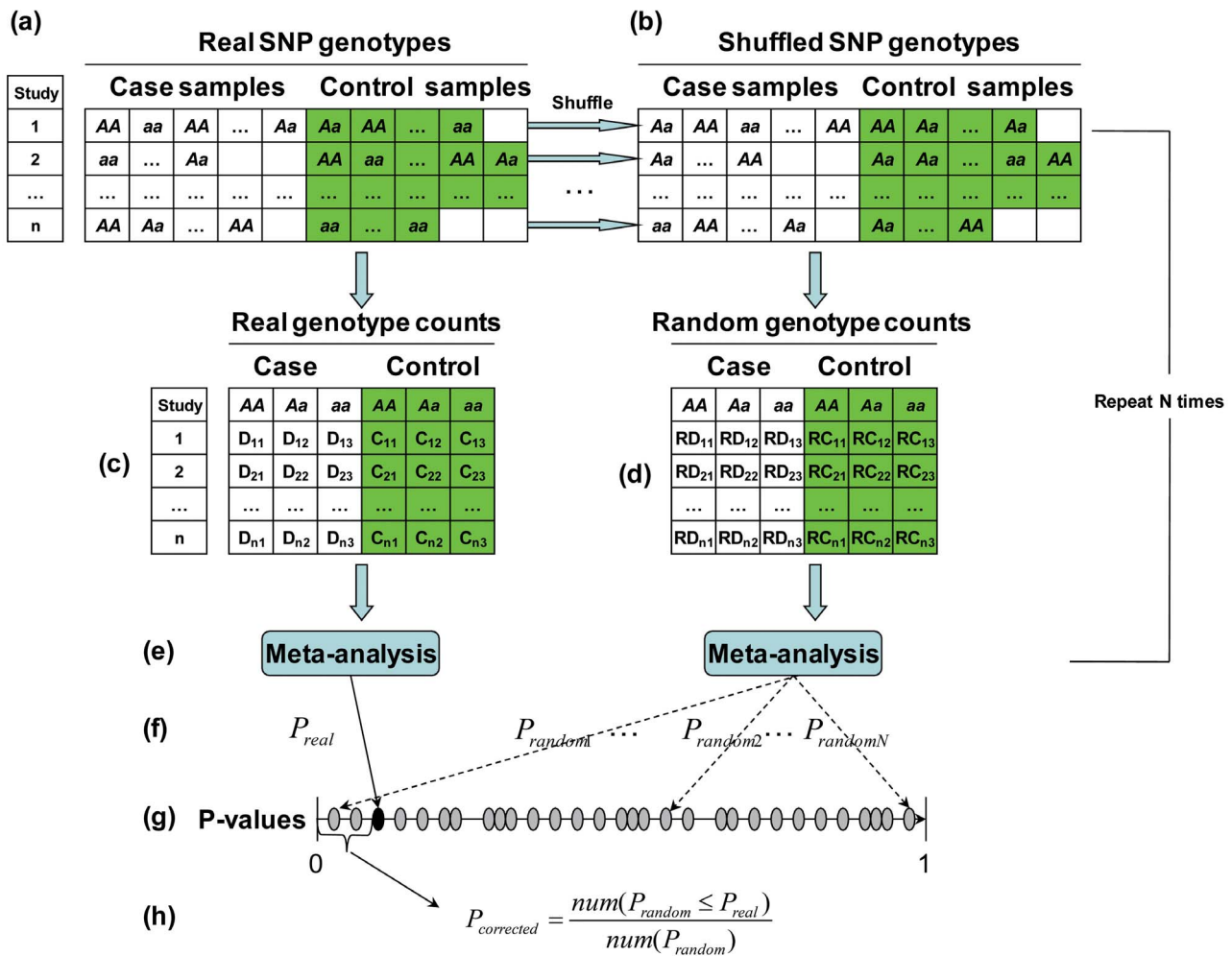
Figure 1. TradPerm method. (a) Real individual-level genotype. (b) Shuffled individual-level genotype. (c) Genotype count from real SNP genotype. (d) Genotype count from random SNP genotype. (e) Completing meta-analysis. (f) Calculating *P*-values of meta-analysis (real *P*-values and *N* random *P*-values). (g) Ranking all *P*-values. (h) Calculating the corrected *P*-value.
doi:10.1371/journal.pone.0089212.g001

to complete the permutation correction. Finally, MCPerm was evaluated using five sets of real data and five sets of simulation data.

Methods

Data

The five real data sets used to evaluate the MCPerm method were the AlzGene database (Field Synopsis of Genetic Association Studies in Alzheimer disease, www.alzgene.org), SZGene database (Field Synopsis of Genetic Association Studies in Schizophrenia, www.szgene.org), PDGene database (Field Synopsis of Genetic Association Studies in Parkinson's Disease, www.pdgene.org), MSGene database (Field Synopsis of Genetic Association Studies in Multiple Sclerosis, www.msgene.org), and AlsGene database (Field Synopsis of Genetic Association Studies in Amyotrophic Lateral Sclerosis, www.alsgene.org) [6–9]. We used PERLscript to extract all records from the five databases. The following criteria were used to filter the data: (1) all records must have a case-control design; (2) all records must have SNP genotype counts for both cases and controls; (3) the minor allele frequency must be greater than 0.01 for single records; (4) the *P*-value of the Hardy-Weinberg equilibrium test in the control group must be greater than 0.001 for single records; (5) each SNP must be reported by at least four studies (which will be used in the meta-analysis). We also excluded family-based studies.

Each record included the following information: study ID, the first author's name, the year of publication, ethnicity, SNP names, SNP alleles, SNP genotypes, and genotype counts for cases and controls. A total of 850 SNP loci were obtained for the meta-analysis, and each locus had at least four records. These included 287 SNP loci from the AlzGene database (a total of 2,408 records),

Table 1. The 2×3 genotypic table.

	AA	Aa	Aa	Total
Case	D_{AA}	D_{Aa}	D_{aa}	N_{case}
Control	C_{AA}	C_{Aa}	C_{aa}	$N_{control}$
Total	N_{AA}	N_{Aa}	N_{aa}	N_{total}

doi:10.1371/journal.pone.0089212.t001

Table 2. Summary of simulated methods of six genotype counts.

genotypes	counts	generation methods
AA in case	D_{AA}	hypergeometric random number generated in step1
Aa in case	D_{Aa}	hypergeometric random number generated in step2
aa in case	D_{aa}	$N_{case} - D_{AA} - D_{Aa}$
AA in control	C_{AA}	$N_{AA} - D_{AA}$
Aa in control	C_{Aa}	$N_{Aa} - D_{Aa}$
aa in control	C_{aa}	$N_{control} - C_{AA} - C_{Aa}$
total	N_{total}	

doi:10.1371/journal.pone.0089212.t002

273 from the SZGene database (a total of 2,179 records), 182 from the PDGene database (a total of 1,625 records), 88 from the MSGene database (a total of 604 records), and 20 from the AlsGene database (a total of 161 records).

As TradPerm relies on individual-level data, we also constructed five simulation data sets. We generated the individual's genotype (AA, Aa, or aa) for each record based on the genotype count in cases and controls. By shuffling these simulated data, we constructed the random background and completed the TradPerm correction.

Meta-analysis

For ease of understanding, we will first describe the meta-analysis method. Suppose there are two alleles, A (A is the risk allele) and a at SNP locus A. Three possible genotypes are denoted by AA, Aa, and aa. Suppose also that a meta-analysis includes n previously published studies. Without loss of generality, we mainly considered the allele model in the present study (the effect of the A allele vs. the a allele); however, we also provided some findings about the dominant and recessive models, and used $\ln OR$ to measure the effect size. To carry out a meta-analysis, we first tested for heterogeneity between studies. The commonly used indicators are Cochran's Q-statistics and I^2 [16,17]. In this study, we used Cochran's Q-statistics to evaluate heterogeneity. The null hypothesis was that the n studies have the same effect, and the significance level was α_{het}. It should be noted that the multiple testing problem ought to be considered for the heterogeneity test in the meta-analyses of multiple loci, although few researchers do so. We provided a permutation correction method for the heterogeneity test, and, after testing for heterogeneity, we combined the effect sizes and evaluated the statistical association between genotype and phenotype. If heterogeneity existed ($P < \alpha_{het}$), we used a random effects model (assuming that each study has a specific effect size) to combine the effect sizes. If heterogeneity did not exist ($P \geq \alpha_{het}$), we used a fixed effects model (assuming that all studies share a common effect size) to combine the effect sizes [18].

TradPerm

For multiple SNP loci, the problem of multiple testing correction should be considered, so we describe below the TradPerm method for multiple testing corrections.

Suppose that we have obtained the original SNP genotype and phenotype data (individual-level data) from n previously published studies (Figure 1a). Each study has a different sample size. For each study, we can count the number of each of the three genotypes (AA, Aa, and aa) in the cases and controls (Figure 1c).

Using these genotype counts, we can perform a meta-analysis and obtain a meta-analysis P-value (P_{real}), then carry out a TradPerm procedure to correct the P_{real}.

The principle of TradPerm is if there is no association between genotypes and phenotypes, then any of the individual phenotype labels may be associated with any one of the genotypes. Therefore, the background distribution of the meta-analysis P-value can be generated by randomly shuffling the phenotype labels of cases and controls. Specifically, for each study included in the meta-analysis, we shuffled the phenotype labels (Figure 1b) and re-counted the number of the three genotypes in cases and controls (Figure 1d). Using these random genotype counts, we performed a meta-analysis (Figure 1e) and obtained a random meta-analysis P-value (P_{random}). The entire process was repeated N times to obtain $P_{random1}, P_{random2}, \cdots, P_{randomN}$ (Figure 1f). If there is a significant association between genotypes and phenotypes, the P_{real} calculated using the real data will appear small relative to the distribution of the P-values obtained under permutation (Figure 1g). In other words, if the association does not exist, the P_{real} for real data is unlikely to be obtained. The TradPerm P-value ($P_{corrected}$) under the null hypothesis can be calculated as the proportion of P-values under permutation that are less than or equal to P_{real} (Figure 1h). $P_{corrected}$ is defined as follows: [19].

$$P_{corrected} = \frac{(no.\ of\ P_{randomi} \leq P_{real})}{(total\ no.\ of\ P_{randomi})}$$

where P_{real} is the real meta-analysis P-value, and $P_{randomi}$ is the P-value generated by permutation. Similarly, the P-value of the heterogeneity test can be corrected by shuffling the phenotype labels of cases and controls.
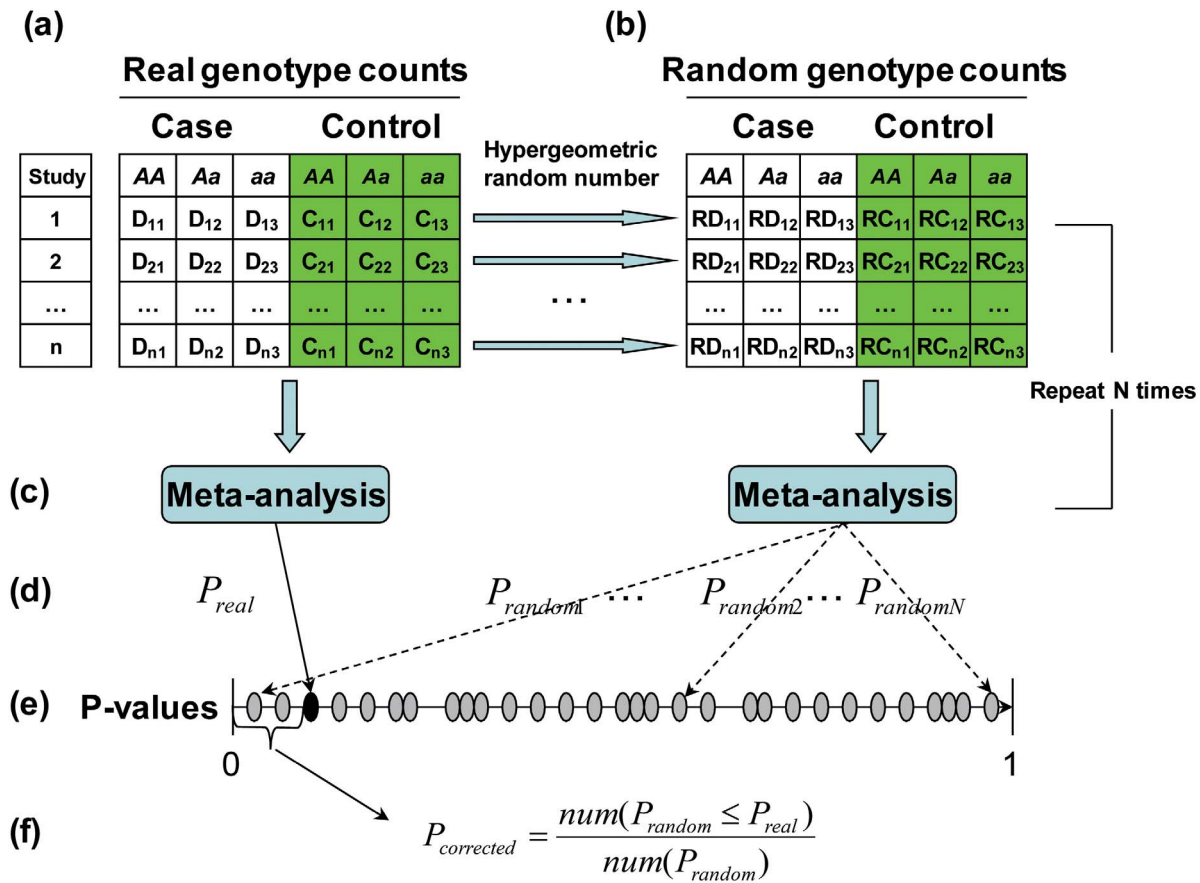
MCPerm

As described above, TradPerm relies on individual-level data which are very difficult to obtain. Therefore, our MCPerm method only uses count-based summary genotype data. MCPerm employs a two-step hypergeometric distribution to generate the random genotype counts in cases and controls.

Step 1: Simulate genotype count of AA in cases. First, we will consider one study in a meta-analysis. We suppose that there are a total of N_{total} samples, including N_{case} cases and $N_{control}$ controls in this study (Table 1). Consider locus A, where the counts of the three genotypes in cases and controls were denoted as: D_{AA} for AA in cases, D_{Aa} for Aa in cases, D_{aa} for aa in cases, C_{AA} for AA in controls, C_{Aa} for Aa in controls, and C_{aa} for aa in controls. The total number of counts of the three genotypes in all samples was denoted as: N_{AA} for AA, N_{Aa} for Aa, and N_{aa} for aa. Next, we generated the genotype counts, D_{AA}, by stochastic simulation.

First, we will introduce the hypergeometric distribution in statistics. Suppose an urn contains M balls of which K are black and $M-K$ are white. Consider an experiment in which m balls are drawn without replacement from the run. The number of black balls in the sample of m obeys a hypergeometric distribution: $X \sim H(m, K, M)$.

Now we will consider the entire TradPerm process. There are a total of N_{total} samples of which N_{case} are cases and $N_{control}$ are controls (where $N_{control} = N_{total} - N_{case}$). When we randomly shuffle the phenotype labels of cases and controls, the count of genotype AA in all samples is unchanged (that is, N_{AA} is a constant). This is equivalent to randomly drawing N_{AA} samples without replacement. Therefore, the count of genotype AA in case samples obeys a hypergeometric distribution $D_{AA} \sim H(N_{AA}, N_{case}, N_{total})$; i.e.,

(a)

Real genotype counts

(b)

Random genotype counts

Figure 2. MCPerm method. (a) Genotype count collected from previous studies. (b) Genotype count generated by two-step hypergeometric distribution. (c) Completing meta-analysis. (d) Calculating *P*-values of meta-analysis (real *P*-values and *N* random *P*-values). (e) Ranking all *P*-values. (f) Calculating the corrected *P*-value.

$$P(X=D_{AA})=\frac{\binom{N_{case}}{D_{AA}}\binom{N_{total}-N_{case}}{N_{AA}-D_{AA}}}{\binom{N_{total}}{N_{AA}}}$$

$$P(X=D_{Aa}|D_{AA})=$$

$$\frac{\binom{N_{case}-D_{AA}}{D_{Aa}}\binom{(N_{total}-N_{AA})-(N_{case}-D_{AA})}{N_{Aa}-D_{Aa}}}{\binom{N_{total}-N_{AA}}{N_{Aa}}}$$

Finally, we used the hypergeometric random number to simulate the count of genotype *AA* in case samples (that is, D_{AA}). Kachitvichyanukul's method [20] was used to generate the hypergeometric random number. This can be achieved by the function "rhyper" in R language. Because $N_{AA}=D_{AA}+C_{AA}$ is a constant when we randomly shuffle, we can deduce that $C_{AA}=N_{AA}-D_{AA}$ (where D_{AA} is a hypergeometric random number generated by Kachitvichyanukul's method).

Step 2: Simulate genotype count of Aa in cases. At this stage, we have fixed the counts of genotype *AA* in case and control samples. As randomly shuffling is equivalent to randomly sampling without replacement, the remaining count of samples is $N_{total}-N_{AA}$. When we randomly shuffle the phenotype labels, the count of genotype *Aa* remains unchanged (that is, N_{Aa} is a constant). This is equivalent to randomly drawing N_{Aa} samples from the $N_{total}-N_{AA}$ remaining samples without replacement. Therefore, the count of genotype *Aa* in case samples also obeys the hypergeometric distribution $D_{Aa} \sim H(N_{Aa}, N_{case}-D_{AA}, N_{total}-N_{AA})$; i.e.,

Therefore, we still employ Kachitvichyanukul's method [20] to generate the hypergeometric random number. Because $N_{Aa}=D_{Aa}+C_{Aa}$ is a constant, we can deduce $C_{Aa}=N_{Aa}-D_{Aa}$ (where D_{Aa} is a hypergeometric random number generated by Kachitvichyanukul's method).

Finally, we can deduce $D_{aa}=N_{case}-D_{AA}-D_{Aa}$ and $C_{aa}=N_{control}-C_{AA}-C_{Aa}$. Table 2 summarizes the methods used to simulate the six genotype counts.

MCPerm procedure. The MCPerm *P*-value was calculated in the same way as for TradPerm. Specifically, for each study, we randomly generated genotype counts in cases and controls (Figure 2b). We then carried out a meta-analysis (Figure 2c) and obtained a random meta-analysis *P*-value (P_{random}). The entire process was repeated *N* times, resulting in *N* random P-values (Figures 2d and 2e). The MCPerm *P*-value ($P_{corrected}$) can also be calculated as the proportion of *P*-values under simulation that are less than or equal to P_{real}(Figure 2f). Similarly, the *P*-value of the

rs778294 (Jonsson, 2009) hist plot
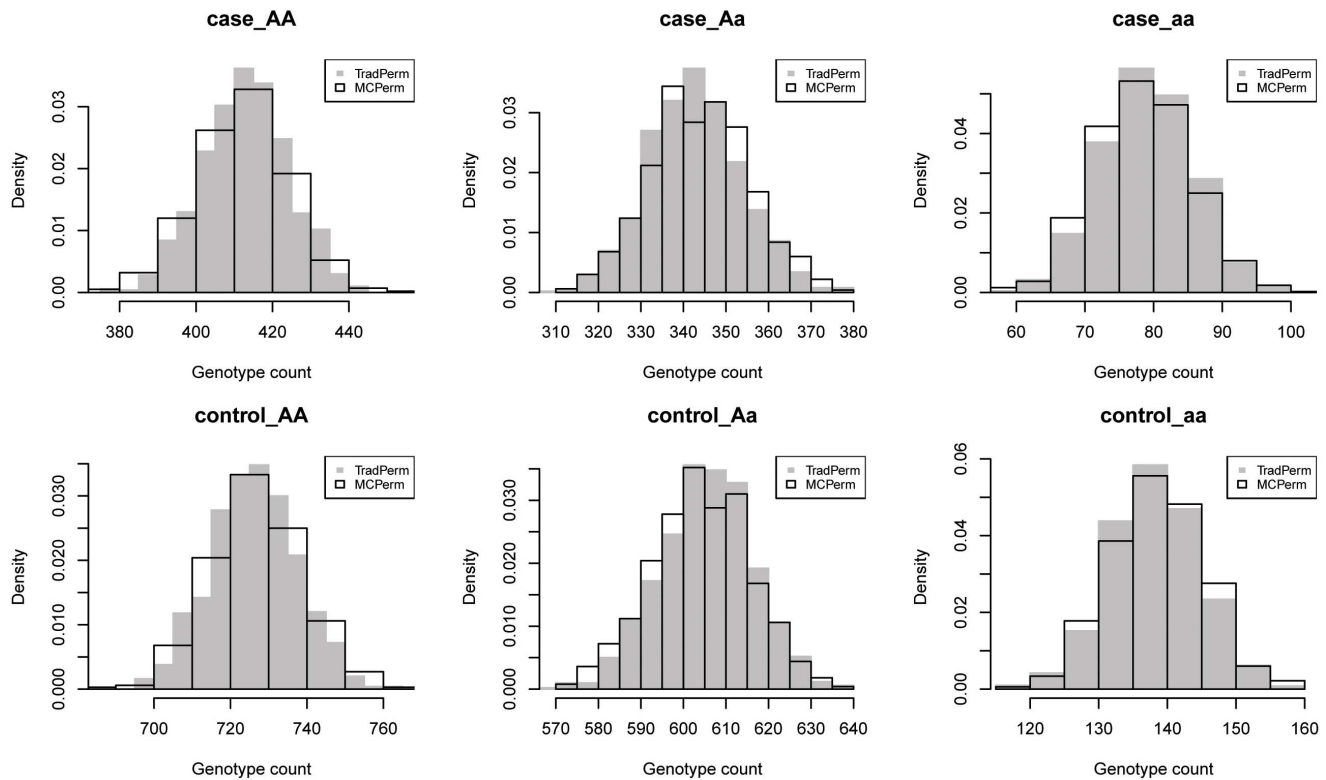
Figure 3. Distributions of the six genotype counts. *AA*, *Aa*, *aa* in cases, and *AA*, *Aa*, and *aa* in controls. Gray solid bars represent the distributions obtained under TradPerm, and black hollow bars represent the distributions generated by MCPerm. The distributions are the same between MCPerm and TradPerm.
doi:10.1371/journal.pone.0089212.g003

heterogeneity test can be corrected by simulating the genotype counts of cases and controls.

Results

As TradPerm is often treated as the gold standard for multiple testing corrections, we did not compare our MCPerm with other multiple testing corrections methods. Instead, we compared the consistency of genotype distributions and permutation *P*-values between TradPerm and MCPerm. Our aim was to illustrate that MCPerm does not change the size of TradPerm *P*-values, so is a useful alternative to TradPerm.

Comparing Six Genotype Distributions between MCPerm and TradPerm

As an example, we randomly selected one SNP, rs778294, from the SZGene database. There were a total of 12 studies about rs778294 in the SZ database, and we selected one ([21]) to compare six genotype distributions between MCPerm and TradPerm. We generated the individual's genotype for this study based on the genotype counts in cases and controls. First, we used TradPerm to shuffle the sample labels 1,000 times and calculated the counts of six genotypes for each shuffle. Then we used MCPerm to generate the counts of the six genotypes (again 1,000 times). Figure 3 shows the distributions of the genotype counts, and we can see that the six genotype counts under MCPerm had similar distributions to genotype counts under TradPerm. This can also be seen from the cumulative distribution curve (Figure S1).

We next used the Kolmogorov-Smirnov test (KS-test) to compare genotype distributions between MCPerm and TradPerm. The null hypothesis was that the genotype counts had the same distribution between MCPerm and TradPerm, and this null hypothesis was not rejected after performing the KS-test in all six tests. The six KS-test *P*-values were $P = 0.069$ for *AA* genotypes in cases, $P = 0.148$ for *Aa* genotypes in cases, $P = 0.828$ for *aa* genotypes in cases, $P = 0.069$ for *AA* genotypes in controls, $P = 0.148$ for *Aa* genotypes in controls, and $P = 0.828$ for *aa* genotypes in controls. Because $D_{AA} = N_{AA} - C_{AA}$, the *AA* genotypes in the cases had the same KS-test *P*-value as the *AA* genotypes in the controls ($P = 0.999$). Similarly, *Aa* and *aa* genotypes in the cases had the same KS-test *P*-values as the controls. This indicated that the six genotype counts under MCPerm had the same distributions as genotype counts under TradPerm.

We also used the quantile-quantile plot (QQ-plot) to evaluate the consistency of the genotype distributions between MCPerm and TradPerm. This plotted the quantiles of the genotype counts generated by MCPerm vs. the quantiles of the genotype counts calculated by TradPerm (from 1,000 permutations). Figure 4 shows the QQ-plots of the six genotypes which were all linear, indicating that the genotype distributions are the same between MCPerm and TradPerm.

In addition to genotype distributions, we also evaluated the distributions of the allele model (*A* allele vs. *a* allele, Figure S2), the dominant model (*AA*+*Aa* vs. *aa* genotypes, Figure S3), and the

rs778294 (Jonsson, 2009) QQ plot
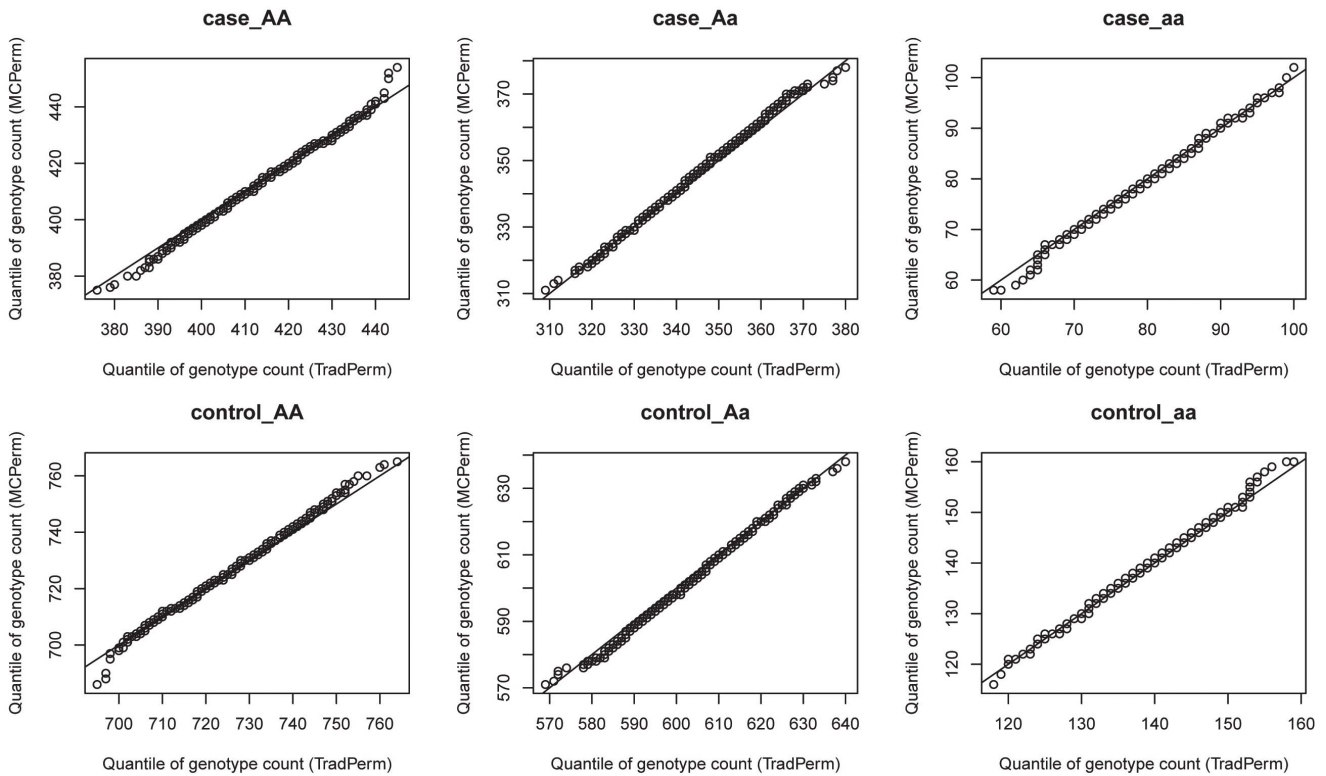
Figure 4. QQ-plots of six genotype counts. *AA*, *Aa*, and *aa* in cases, *AA*, *Aa*, and *aa* in controls.
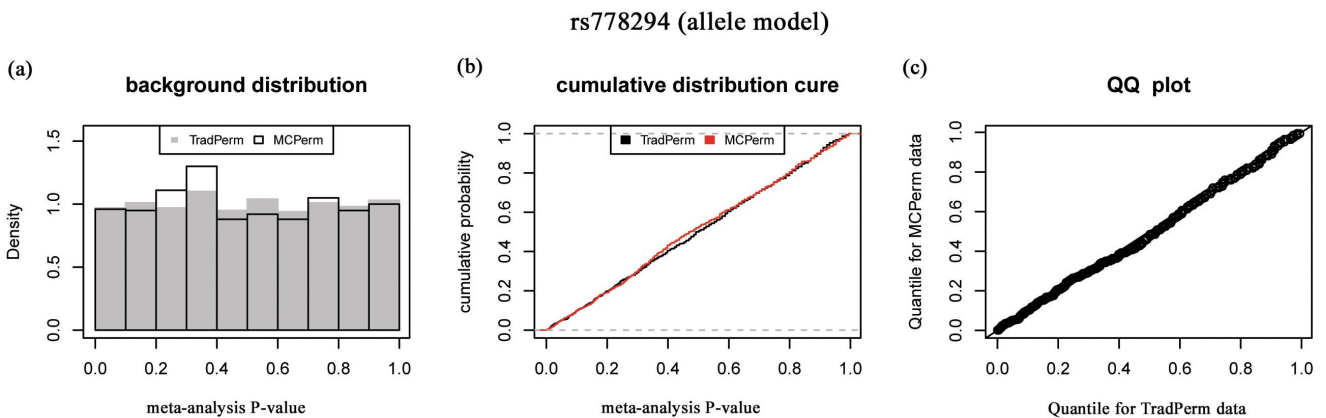doi:10.1371/journal.pone.0089212.g004

recessive model (*AA* vs. *Aa*+*aa* genotypes, Figure S4), and again observed the same distributions between MCPerm and TradPerm.

Comparing Meta-analysis *P*-values between MCPerm and TradPerm

In this section, we will describe the consistency of meta-analysis *P*-values between TradPerm and MCPerm. For each of the 850 SNPs, we carried out 1,000 permutations and calculated the permutation *P*-values of the meta-analysis (allele model) using TradPerm and MCPerm. For SNP rs778294, the probability density (Figure 5A), cumulative distribution curve (Figure 5B), and

QQ-plot (Figure 5C) showed that permutation *P*-values had the same distributions between MCPerm and TradPerm. The same conclusions were obtained for both the dominant model and recessive model (Figures S5 and S6), and for the other 849 SNP loci. The comparison results for all 850 loci are shown in Supporting Information S1, which can be downloaded at http://www.bioapp.org/research/MCPerm/index.html.

We next plotted a scatter plot of 850 MCPerm meta-analysis *P*-values against 850 TradPerm meta-analysis *P*-values (Figure 6). This revealed a highly linear relationship between MCPerm *P*-values and TradPerm *P*-values. We also calculated the Pearson's

rs778294 (allele model)

Figure 5. Comparison of meta-analysis *P*-values between MCPerm and TradPerm. (a) Probability density. (b) Cumulative distribution curve. (c) QQ-plot.
doi:10.1371/journal.pone.0089212.g005
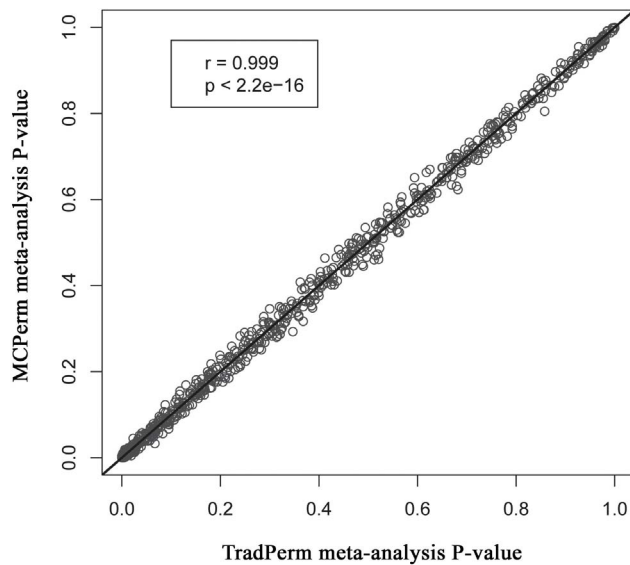
Figure 6. Scatter plots of 850 MCPerm vs. 850 TradPerm meta-analysis *P*-values. The MCPerm *P*-values are highly consistent with TradPerm *P*-values (*r* = 0.999; *P*<2.2e-16).
doi:10.1371/journal.pone.0089212.g006

correlation coefficient of the correlation between MCPerm *P*-values and TradPerm *P*-values (*r* = 0.999), and the correlation test *P*-value was <2.2e-16. These findings indicate that the MCPerm *P*-values are highly consistent with TradPerm *P*-values. The same conclusions were obtained for both the dominant model and recessive model.

Comparing Heterogeneity Test *P*-values between MCPerm and TradPerm

For the 850 SNPs, we compared the consistency of heterogeneity test *P*-values. For SNP rs778294, the probability density, cumulative distribution curve, and QQ-plot results revealed the same distributions of heterogeneity test *P*-values between MCPerm and TradPerm (Figures S7, S8, and S9). The comparison results for all 850 loci are shown in Supporting Information S1. The correlation test (*r* = 0.999, *P*-value <2.2e-16) also showed a high consistency of heterogeneity test *P*-values between MCPerm and TradPerm (Figure S10).

Estimating the Variance of lnOR and a Permutation Box Plot

In general, for a single study, the variance of $\ln OR$ (in an allele model, dominant model, or recessive model) is estimated using the following formula: $\mathrm{var}(\ln OR) = 1/a + 1/b + 1/c + 1/d$, where a, b, c, and d represent the observed counts in fourfold tables [22]. The variance of $\ln OR$ is used as the weight of a single study when multiple studies are combined in a meta-analysis. Obviously, different observed counts cause different estimations of the variance of $\ln OR$, and over-reliance on the original observed counts will affect the stability of the results. An advantage of MCPerm is that we can simulate the background of $\ln OR$ by a two-step hypergeometric random number. This enables us to directly calculate the variance of $\ln OR$ using the simulated $\ln OR$ values, thus increasing the reliability of the meta-analysis results.

For each study in a meta-analysis, researchers expect to see the position of a real $\ln OR$ in a random background. To represent this, we designed a permutation box plot as shown for SNP

rs778294 in Figure 7a. This allowed us to readily compare the real $\ln OR$ with the random background. In addition, for each study, we plotted the distribution of the random background and the position of the real $\ln OR$ (Figure 7b) to deepen our understanding of the meta-analysis.

Comparing Running Time between MCPerm and TradPerm

We carried out 1,000 permutations for all 850 SNP loci on a 3.2 GHz desktop PC with a 2G memory running the Windows XP system. TradPerm took about 6 h of computation time, while MCPerm only required 18 min of computation time. This suggests that, for the 850 SNPs examined, MCPerm is about 20 times faster than TradPerm.

We further simulated three genome-wide data sets (100,000, 500,000, and 1,000,000 SNPs) to compare the running time of the two methods. Each data set included 10 genome-wide association studies. We then carried out 1,000 permutations of the three data sets. For 100,000 SNPs, 500,000 SNPs, and 1,000,000 SNPs, MCPerm required approximately 1.5 days, 6–7 days, and 13.5 days of computation time, respectively. The running time was not tested for TradPerm as it would take too long (more than 40 days for 100,000 SNPs).

In addition, for MCPerm, the genotype counts were directly simulated rather than calculated from the shuffled data, so the computation time did not depend on the sample size of a single study. This is directly relevant for a meta-analysis, as the number of studies affects the computation time.

MCPerm R Package

To facilitate the use of MCPerm, we developed a freely available R package, named MCPerm. The package has detailed instructions and examples and has been uploaded to the Comprehensive R Archive Network (CRAN). Users can download the package at the following website: http://cran.r-project.org/web/packages/MCPerm/index.html. It can be run using both Linux and the Windows environment. The current version is v1.1.4. There are a total of 45 functions that facilitate the implementation of the meta-analysis, TradPerm, and MCPerm. A detailed list of functions is given in Table S1.

Discussion

Multiple testing is a challenging issue in SNP-based meta-analysis. Among the many multiple testing correction methods, TradPerm is usually considered to be the gold standard [4,5]. It estimates the background distribution of test *P*-values by shuffling phenotype labels, and thus is an accurate correcting method. However, the over-reliance on raw SNP genotype data and the large amount of computation are two obstacles that limit the scope of TradPerm. To overcome these limitations, we developed a MCPerm method that only uses the summary statistics of genotypes to perform the permutation.

MCPerm simulated the random genotype counts in cases and controls using a two-step hypergeometric distribution. We used five real data sets and five simulation data sets to prove that the genotype count distributions are the same between MCPerm and TradPerm. Furthermore, we showed that MCPerm can obtain the same size of permutation *P*-values of meta-analysis as TradPerm, and take less computation time than TradPerm to do so.

For the meta-analysis of a single SNP locus, we suggest that a MCPerm procedure is carried out as the background distribution of $\ln OR$ and the position of $\ln OR$ in the background (permutation box plot) will deepen our understanding of the analytical results. In
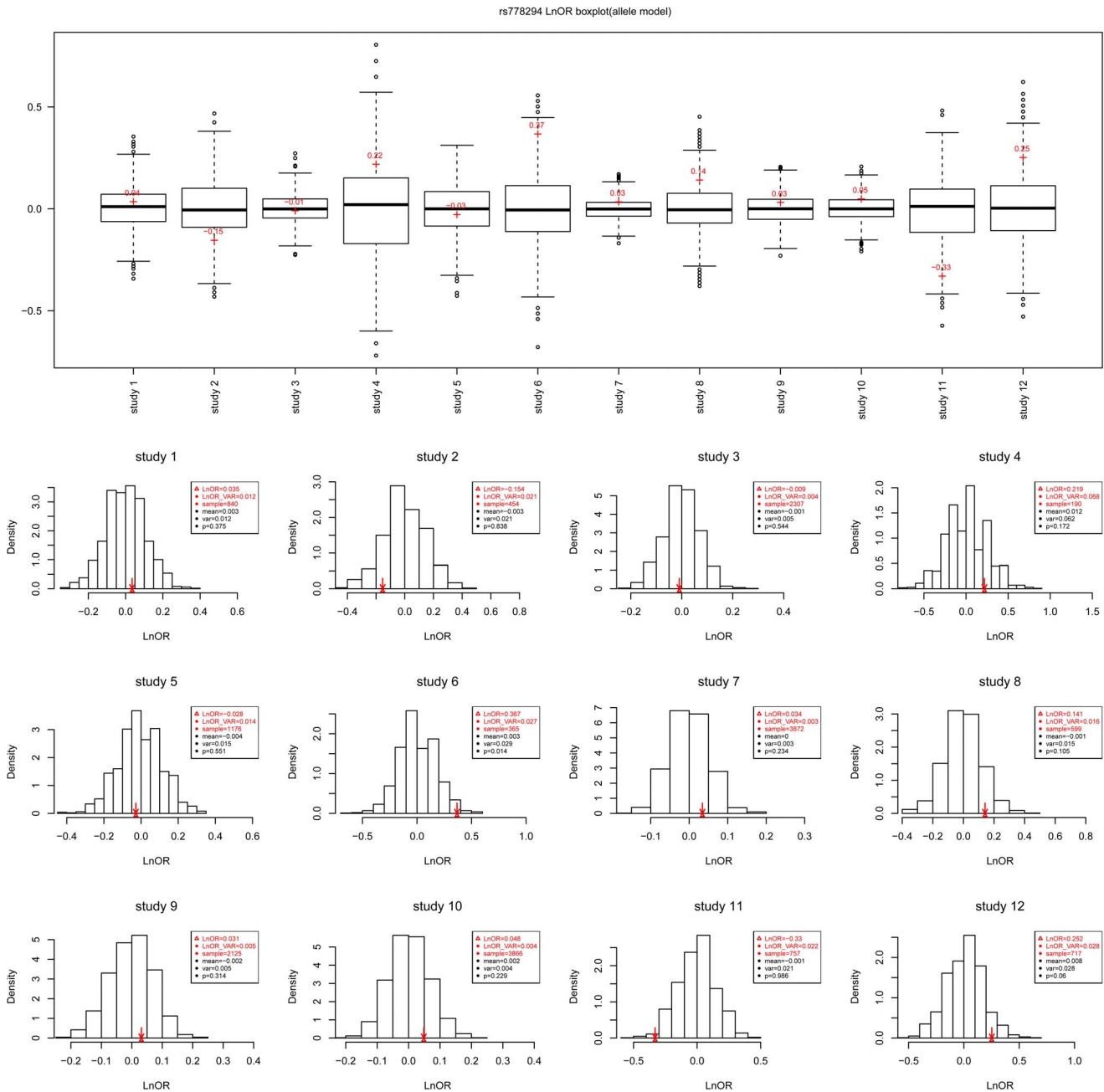
Figure 7. Real LnOR in a random background. (a) Permutation box plot. (b) Permutation probability density.
doi:10.1371/journal.pone.0089212.g007

addition, the procedure will enable the variance to be directly calculated from the simulated ln *OR* values, increasing the reliability of the meta-analysis.

With the arrival of the era of post-association studies, two-step hypergeometric random numbers in MCPerm may have broader application prospects than expected. In some gene-based and pathway-based association studies, researchers combine *P*-values for SNPs into an overall *P*-value for a gene or pathway [23]. In such studies, the correction of *P*-values for SNPs is essential. However, they are usually undertaken by bioinformatics researchers who are not always able to collect raw SNP genotype data because of privacy policies. The two-step hypergeometric random number in MCPerm will aid these researchers to complete their permutation corrections using only the summary statistics of the genotype data such as those provided by the Wellcome Trust Case Control Consortium [24]. In conclusion, we hope that MCPerm will be widely used in genome-wide association studies and genome-wide studies of meta-analysis.

Supporting Information

Figure S1 Cumulative distribution curve of six genotype counts.
(JPG)

Figure S2 Comparison of the distributions of the allele model (*A* allele vs. *a* allele).
(JPG)

Figure S3 Comparison of the distributions of the dominant model (*AA*+*Aa* vs. *aa* genotypes).
(JPG)

Figure S4 Comparison of the distributions of the recessive model (*AA* vs. *Aa*+*aa* genotypes).
(JPG)

Figure S5 Comparison of meta-analysis *P*-values of the dominant model.
(JPG)

Figure S6 Comparison of meta-analysis *P*-values of the recessive model.
(JPG)

Figure S7 Comparison of heterogeneity test *P*-values of the allele model.
(JPG)

Figure S8 Comparison of heterogeneity test *P*-values of the dominant model.
(JPG)

Figure S9 Comparison of heterogeneity test *P*-values of the recessive model.
(JPG)

Figure S10 Scatter plots of 850 MCPerm heterogeneity test *P*-values against 850 TradPerm heterogeneity test *P*-values. The MCPerm *P*-values are highly consistent with TradPerm *P*-values ($r = 0.999$; $P < 2.2e-16$)
(JPG)

Table S1 The functions in MCPerm R package (v 1.1.4).
(DOC)

Supporting Information S1
http://www.bioapp.org/research/MCPerm/index.html.

Author Contributions

Conceived and designed the experiments: YJ. Performed the experiments: YJ LZ. Analyzed the data: MZ HL. Contributed reagents/materials/analysis tools: ML RF. Wrote the paper: YJ RZ FK GL JL.

References

1. Green S (2005) Systematic reviews and meta-analysis. Singapore Med J 46: 270–273; quiz 274.
2. Engberg S (2008) Systematic reviews and meta-analysis: studies of studies. J Wound Ostomy Continence Nurs 35: 258–265.
3. Haerian BS, Haerian MS (2013) OPRM1 rs1799971 polymorphism and opioid dependence: evidence from a meta-analysis. Pharmacogenomics 14: 813–824.
4. Wu Z, Wang H, Chu X, Chen J, Fang S (2013) Association between FAS-1377 G/A polymorphism and susceptibility to gastric cancer: evidence from a meta-analysis. Tumour Biol.
5. Liu G, Zhang S, Cai Z, Li Y, Cui L, et al. (2013) BIN1 gene rs744373 polymorphism contributes to Alzheimer's disease in East Asian population. Neurosci Lett.
6. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, et al. (2008) Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. Nat Genet 40: 827–834.
7. Lill CM, Abel O, Bertram L, Al-Chalabi A (2011) Keeping up with genetic discoveries in amyotrophic lateral sclerosis: the ALSoD and ALSGene databases. Amyotroph Lateral Scler 12: 238–249.
8. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, et al. (2012) Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. PLoS Genet 8: e1002548.
9. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet 39: 17–23.
10. Sidak Z (1967) Rectangular confidence regions for the means of multivariate normal distributions. J Am Stat Assoc 62: 626–633.
11. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125: 279–284.
12. Westfall PH, Young SS (1993) Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment. New York: John Wiley & Sons, Inc.
13. Pahl R, Schafer H (2010) PERMORY: an LD-exploiting permutation test algorithm for powerful genome-wide association testing. Bioinformatics 26: 2093–2100.
14. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.
15. Gao X, Starmer J, Martin ER (2008) A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol 32: 361–369.
16. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J (2006) Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods 11: 193–206.
17. Ioannidis JP, Patsopoulos NA, Evangelou E (2007) Heterogeneity in meta-analyses of genome-wide association investigations. PLoS One 2: e841.
18. Lau J, Ioannidis JP, Schmid CH (1997) Quantitative synthesis in systematic reviews. Ann Intern Med 127: 820–826.
19. Anderson MJ (2001) Permutation tests for univariate or multivariate analysis of variance and regression. Canadian Journal of Fisheries & Aquatic Sciences 58: 626–639.
20. Kachitvichyanukul V, Schmeiser B (1985) Computer generation of hypergeometric random variates. Journal of Statistical Computation and Simulation 22: 127–145.
21. Bass NJ, Datta SR, McQuillin A, Puri V, Choudhury K, et al. (2009) Evidence for the association of the DAOA (G72) gene with schizophrenia and bipolar disorder but not for the association of the DAO gene with schizophrenia. Behav Brain Funct 5: 28.
22. Hauck WW (1979) The large-sample variance of the Mantel-Haenszel estimator of a common odds ratio. Biometrics 35: 817–819.
23. Peng G, Luo L, Siu H, Zhu Y, Hu P, et al. (2010) Gene and pathway-based second-wave analysis of genome-wide association studies. Eur J Hum Genet 18: 111–117.
24. WTCCC (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.