Published in final edited form as:

Ann Hum Genet. 2012 September ; 76(5): 402–409. doi:10.1111/j.1469-1809.2012.00718.x.

Statistical tests for detecting rare variants using variance-stabilizing transformations

Kai Wang and

Department of Biostatistics, College of Public Health, The University of Iowa, Iowa City, Iowa

John H. Fingert

Department of Ophthalmology and Visual Sciences, Carver College of Medicine, The University of Iowa, Iowa City, Iowa

Abstract

Next generation sequencing holds great promise for detecting rare variants underlying complex human traits. Due to their extremely low allele frequencies, the normality approximation for a proportion no longer works well. The Fisher's exact method appears to be suitable but it is conservative. We investigate the utility of various variance-stabilizing transformations in single marker association analysis on rare variants. Unlike a proportion itself, the variance of the transformed proportions no longer depends on the proportion, making application of such transformations to rare variant association analysis extremely appealing. Simulation studies demonstrate that tests based on such transformations are more powerful than the Fisher's exact test while controlling for type I error rate. Based on theoretical considerations and results from simulation studies, we recommend the test based on the Anscombe transformation over tests with other transformations.

Keywords

rare variants; sequencing; variance-stabilizing transformation; association

INTRODUCTION

With the recent advances in high-throughput sequencing technologies, the time is ripe to unveil the genetic variants that are missing from conventional genome-wide association studies that target common variants. These variants are believed to be extremely rare in the general population with disease allele frequencies between 1% ~ 5% or even lower [Schork et al., 2009]. Such a low frequency makes it challenging to conduct valid and powerful statistical inference. The normal approximation to a proportion no longer works well [Zar, 1999, Section 13.3] and the Fisher's exact test is conservative [Altham, 1969; Howard, 1998] (see also our simulation results presented later).

The prevailing methods for association analysis of variants are gene- or set-based. Typically single-nucleotide polymorphisms (SNPs) in a gene or a set are collapsed in various ways [Li & Leal, 2008; Sun et al., 2011] in order to enrich the genetic information content such that conventional association techniques are applicable. The kernel-based association test [Kwee et al., 2008; Wu et al., 2010] provides an alternative to combining multiple rare variants.

While testing for gene- or set-based association is of great interest, it is of even greater interest to pin-point the variants with high precision, for example, through single-marker analysis. However, single-marker methods specially designed for rare variants are surprisingly rare. The Uniform-test and the Beta-test [Li et al., 2010] might be the only ones available. These two tests adopt a sequential Bayesian approach. The prior distribution of the parameter is trained using the control sample and then fed into the case sample. One disadvantage of this approach is that switching the case sample and the control sample will lead to a different result.

The goal of this report is to explore an alternative approach to association testing of rare variants in case-control studies. This approach builds upon the rich literature on variance-stabilizing transformations of a proportion such that the distribution of the transformed proportion is much better approximated by a normal distribution than that of the proportion itself, especially when the proportion is close to 0 or 1.

METHOD

Let r be the probability that an individual carries at least one copy of a variant. If there are X such carriers out of n individuals, then a natural estimate of r is $\hat{r} = X/n$. An estimate of the variance of $\hat{r}$ is $\hat{r}(1 - \hat{r})/n$. Note that this quantity depends on the value of $\hat{r}$. If the variant is rare, r would be very close to 0 such that $\hat{r} = 0$ with a non-negligible probability. So the distribution of $\hat{r}$ is skewed and the normality approximation fails. To get around of this problem, one approach is to use the angular transformation defined as [Zar, 1999; Mosteller & Tukey, 1949]:

$$\begin{aligned}\theta(X, n) &= \arcsin \sqrt{\frac{X}{n}} \quad (\text{in radians}) \\ &= \frac{1}{2} \arcsin \left(\frac{2X}{n} - 1 \right) + \frac{\pi}{4}.\end{aligned}$$

This transformation is very natural [Mosteller & Tukey, 1949] considering that, for an angle θ , there is $\sin^2 \theta + \cos^2 \theta = 1$. The appealing feature of this transformation is that the variance of $\theta(X, n)$ can be estimated by $1/4n$. That is, it no longer depends on $\hat{r}$.

A better transformation that has the same feature is the Anscombe transformation [Anscombe, 1948]:

$$\begin{aligned}\theta'(X, n) &= \arcsin \sqrt{\frac{X+3/8}{n+3/4}} \\ &= \theta(X+3/8, n+3/4).\end{aligned}$$

Anscombe transformation is motivated by optimizing the variance of the following transformation over d_1 , d_2 , and c :

$$\sqrt{n+d_2} \arcsin \sqrt{\frac{X+c}{n+d_1}}.$$

When $d_1 = 3/4$, $d_2 = 1/2$, and $c = 3/8$, its variance is $1/4 + O(n^{-2})$, which is of the lowest order. The variance of $\theta'(X, n)$ is estimated by $1/4(n + 1/2)$. Another variation is the Freeman and Tukey transformation [Freeman & Tukey, 1950]:

$$\begin{aligned}\theta''(X, n) &= \arcsin \sqrt{\frac{X}{n+1}} + \arcsin \sqrt{\frac{X+1}{n+1}} \\ &= \theta(X, n+1) + \theta(X+1, n+1).\end{aligned}$$

It has been found that its variance is within 6% of $1/(n+1/2)$ [Freeman & Tukey, 1950].

Although much better approximated by normality, especially for low frequency x , than the angular transformation, the Anscombe transformation and the Freeman-Tukey transformation are not invariant to proportional change in X and n . For instance, the transformed value for $X=2$ and $n=100$ is different from $X=20$ and 1000 although the ratio X/n remains the same. The following transformation by Chanter [1975] avoids this phenomenon:

$$\theta'''(X, n) = \arcsin \sqrt{d + (1-2d)\frac{X}{n}},$$

where $d = 1/4n$. Its variance is estimated by $1/4n$.

All these transformations share the feature that their variances do not depend on the value of x . Their distributions are better approximated by normal distributions than the distribution of $\hat{x}$. Application of these transformations to case-control association testing is straightforward. Let A denote the allele of minor frequency and a the other allele. Using subscript 1 for cases and subscript 2 for controls, the data from a case-control study can be summarized in Table 1:

Since the cases are independent of the controls, the variance of the difference of the transformed values between cases and controls is the sum of their variance in each sample. Four tests can be defined as follows, one corresponding to an aforementioned transformation:

$$\begin{aligned}\text{angular} &= \frac{4[\theta(X_1, n_1) - \theta(X_2, n_2)]^2}{1/n_1 + 1/n_2}, \\ \text{Anscombe} &= \frac{4[\theta'(X_1, n_1) - \theta'(X_2, n_2)]^2}{1/(n_1 + 1/2) + 1/(n_2 + 1/2)}, \\ \text{FT} &= \frac{[\theta''(X_1, n_1) - \theta''(X_2, n_2)]^2}{1/(n_1 + 1/2) + 1/(n_2 + 1/2)},\end{aligned}$$

and

$$\text{Chanter} = \frac{4[\theta'''(X_1, n_1) - \theta'''(X_2, n_2)]^2}{1/n_1 + 1/n_2}.$$

Under the null hypothesis that there is no association between the variant and affection status, all four tests asymptotically follow a chi-square distribution with 1 degree of freedom.

Substituting X_1 and X_2 by $n_1 - X_1$ and $n_2 - X_2$ (the number of non-carriers) in these four tests, respectively, does not change their values as it changes only the sign of the difference in the numerators. We show this below by using the identity $\arcsin \sqrt{1-X/n} + \arcsin \sqrt{X/n} = \pi/2$. For the angular test,

$$\begin{aligned}\theta(n_1 - X_1, n_1) - \theta(n_2 - X_2, n_2) &= \left(\frac{\pi}{2} - \arcsin \sqrt{\frac{X_1}{n_1}}\right) - \left(\frac{\pi}{2} - \arcsin \sqrt{\frac{X_2}{n_2}}\right) \\ &= -[\theta(X_1, n_1) - \theta(X_2, n_2)].\end{aligned}$$

The derivation for the Anscombe test and the Chanter test is similar since

$$\frac{(n-X)+3/8}{n+3/4} = 1 - \frac{X+3/8}{n+3/4}$$

and

$$\frac{1}{4n} + \left(1 - 2 \cdot \frac{1}{4n}\right) \frac{(n-X)}{n} = 1 - \left[\frac{1}{4n} + \left(1 - 2 \cdot \frac{1}{4n}\right) \frac{X}{n}\right].$$

For the FT test, note that

$$\begin{aligned}\arcsin \sqrt{\frac{(n-X)}{n+1}} + \arcsin \sqrt{\frac{(n-X)+1}{n+1}} &= \arcsin \sqrt{1 - \frac{X+1}{n+1}} + \arcsin \sqrt{1 - \frac{X}{n+1}} \\ &= \pi - \left(\arcsin \sqrt{\frac{X}{n+1}} + \arcsin \sqrt{\frac{X+1}{n+1}}\right).\end{aligned}$$

SIMULATION AND APPLICATION

The simulation setup is identical to that used by Li et al. [2010]. The Hardy-Weinberg equilibrium is assumed for the general population. Let p denote the population frequency of minor allele A. Let a denote the major allele. The population genotype frequency of genotypes aa , aA , and AA are $(1-p)^2$, $2p(1-p)$, and p^2 , respectively. Since allele A is assumed to be rare, genotypes for the controls are simulated using their population frequencies. The genotype frequencies in cases are $(1-p)^2/K$, $2p(1-p)\gamma_1/K$, and $p^2\gamma_2/K$, respectively. Here γ_1 is the relative risk of genotype aA versus genotype aa , γ_2 is the relative risk of genotype AA versus genotype aa , and $K = (1-p)^2 + 2p(1-p)\gamma_1 + p^2\gamma_2$. The values of p used in the simulation are 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, and 0.05. Given a significance level, the rejection rate is computed as the proportion of rejections over 10,000 simulation replicates.

To study type I error rate, we set $\gamma_1 = \gamma_2 = 1$. The rejection rates of the angular test, the FT test, the Anscombe test, the Chanter test, and the Fisher's exact test are reported in Table 2 for 250 cases and 250 controls and in Table 3 for 500 cases and 500 controls. It is clear that 1) Fisher's exact test is too conservative, especially when the minor allele frequency is low; 2) The angular test tends to be anti-conservative, and 3) the FT test, the Chanter test, and the Anscombe test have similar performance. All have size close to the nominal significance level. However, the FT test and the Chanter test are sometimes a little bit anti-conservative when the significance level is low (i.e., 0.001, 0.005, and 0.01) and the allele frequency is at the lower end of the range.

To study the power, a dominant model ($\gamma_1 = \gamma_2 = 2$) and a multiplicative model ($\gamma_1 = 2$, $\gamma_2 = 4$) are considered. The rejection rates at significance level 0.05 are graphed in Figure 1 for

the dominant model and in Figure 2 for the multiplicative model. The rejection rates at significance level 0.01 are also graphed for these two models, respectively (Figures 3 and 4). Since the angular test has inflated type I error rate, it is not included in the power analysis. For both models, the FT test, the Chanter test, and the Anscombe test have very similar power performance. All are more powerful than the Fisher's exact test.

Like Li et al. [2010], we also used the data from two empirical studies. In a study of association between copy number variation and schizophrenia [Need et al., 2009], 14 copy number variants are observed in 1013 cases and 3 in 1084 controls. The p -values are 0.003831 for the FT test, 0.003621 for the Anscombe test, 0.003358 for the Chanter test, and 0.005915 for the Fisher's exact test. All the FT test, the Anscombe test, and the Chanter test lead to a smaller p -value than the Fisher's exact test. The other example is on the association between microduplications and schizophrenia [McCarthy et al., 2009]. There are 9 objects with microduplications in 2645 cases and 1 object in 2420 controls. The p -values are 0.01183 for the FT test, 0.01107 for the Anscombe test, 0.009577 for the Chanter test, and 0.02258 for the Fisher's exact test. Again, the former three tests have smaller p -values.

DISCUSSION

In this report, we have investigated the utility of variance-stabilizing transformations in detecting rare variants in case-control genetic association studies. Due the extremely low frequency of rare variants, the regular normal distribution approximation in deriving the limiting distribution of a test statistic no longer works well. Even the Fisher's exact test deviates from the nominal significance level (by being conservative) as demonstrated here and elsewhere [Altham, 1969; Howard, 1998]. By using a variance-stabilizing transformation, the variance of the effect size does not depend on the sample proportion any more and the precision of normality approximation is dramatically improved.

The Fisher's exact test reports p -values only for the outcomes that are as extreme or more extreme than the observed outcome. When the observed outcome is extreme itself, the number of as extreme or more extreme outcomes is very limited. The p -values is thus highly "discretized" and becomes conservative. In contrast, the proposed methods interpolates p -values between two possible outcomes. The ordinary normal approximation that is alternative to the Fisher's exact test also does such interpolation, but performs worse than the proposed tests.

Current methods for rare variant association studies avoid the issue of rarity by using various collapsing methods within a gene or an SNP set. The only single marker analysis method we know of that specifically designed for association studies on rare variants is the Bayesian method [Li et al., 2010]. However, it is not symmetric about the case sample and the control sample – switching the samples will result in a different test result. In comparison, the tests based on variance-stabilizing transformations are computationally very efficient due to their explicit expressions.

Variance-stabilizing transformations for a proportion have been long studied and have been used extensively in fields such as ecology. Their application to association studies of rare variants is not only conceptually appealing but also, as demonstrated by our research, very promising. The performance of the proposed tests is much better than the Fisher's exact test in terms of type I error rate and power. We note that their type I error rate is less satisfactory when the allele frequency is in the lower end of the range considered in the simulation. It has been demonstrated that variance-stabilizing transformations does not work well for proportions close to 0 [Chanter, 1975]. This observation is in support of the importance of variance stabilization in conducting valid inference.

Our simulation findings confirm previous reports [Chanter, 1975] that the Anscombe transformation achieves variance stabilization starting at a lower proportion. We recommend the Anscombe test over the angular test, the FT test, the Chanter test, and the Fisher's exact test for association studies on rare variants.

It should be noted that although variance-stabilizing transformations leads to statistics that have higher power than the Fisher's exact method, the power increase is moderate and diminishes as the minor allele frequency or the sample size increases. When the variant is rare, the power of these variance-stabilizing transformation based methods are also low, just like any other methods. This is a limitation of single marker analysis on rare variants. There have been a large number of "collapsing" or "burden" methods by considering a larger genomic unit. A recent review of these methods is provided in Dering et al. [2011]. Another review is Bansal et al. [2010].

We focused on single marker analyses as the main goal of this report is to promote the use of variance-stabilizing transformations in association studies of rare variants. In addition, the single marker analysis is by itself a very important case but its usefulness in rare variant association studies is under-estimated because previous studies are typically based on untransformed allele frequencies and are under powered. However, the application of the variance-stabilizing transformations to gene-based analysis is straightforward if the rare variants are simply collapsed using either indicator coding or dosage coding [Dering et al., 2011]. In collapsing, a variant is often weighted inversely to its variance $1/\sqrt{\hat{p}(1-\hat{p})}$ where $\hat{p}$ is its estimated frequency either in the controls or in the combined sample of cases and controls. It is interesting to note that the first-order derivative of the arcsin function is

$$\frac{d}{dp} \arcsin \sqrt{p} = \frac{1}{2\sqrt{p(1-p)}}.$$

Applying Taylor expansion, $\arcsin \sqrt{p}$ can be approximated by $0.5p^{-1/2}(1-p)^{-1/2} \cdot p$ when p is close to 0. If the variance-stabilizing transformations are applied to each rare variant before collapsing, an weighting is automatically achieved.

All the tests based on the variance-stabilizing transformations discussed in this report have been implemented in the iGasso package which is available at the R repository (URL: <http://cran.r-project.org/web/packages/iGasso/index.html>).

Acknowledgments

We thank two anonymous reviewers and a handling editor for their helpful comments. This research was supported in part by the National Institutes of Health (NEI R01EY018825).

References

- Altham PME. Exact Bayesian analysis of a 2×2 contingency table, and Fisher's "Exact" significance test. *Journal of the Royal Statistical Society Series B (Methodological)*. 1969; 31(2):261–269.
- Anscombe FJ. The transformation of Poisson, binomial and negative-binomial data. *Biometrika*. 1948; 35(3/4):246–254.
- Bansal V, Libiger O, Torkamani A, Schork N. Statistical Analysis strategies for association studies involving rare variants. *Nature Reviews Genetics*. 2010; 11:773–785.
- Chanter DO. Modifications of the angular transformation. *Journal of the Royal Statistical Society Series B (Applied Statistics)*. 1975; 24(3):354–359.

- Dering C, Hemmelmann C, Pugh E, Ziegler A. Statistical Analysis of Rare Sequence Variants: An Overview of Collapsing Methods. *Genet Epidemiol.* 2011; 35:S12–S17. [PubMed: 22128052]
- Freeman MF, Tukey JW. Transformations related to the angular and the square root. *The Annals of Mathematical Statistics.* 1950; 21(4):607–611.
- Howard JV. The 2×2 table: a discussion from a bayesian viewpoint. *Statistical Science.* 1998; 13(4): 351–367.
- Kwee LC, Liu D, Lin X, Ghosh D, Epstein MP. A powerful and flexible multilocus association test for quantitative traits. *Am J Hum Genet.* 2008; 82(2):386–397. [PubMed: 18252219]
- Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet.* 2008; 83:311–321. [PubMed: 18691683]
- Li Q, Zhang H, Yu K. Approaches for evaluating rare polymorphisms in genetic association studies. *Hum Hered.* 2010; 69:219–228. [PubMed: 20332646]
- McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, Perkins DO, Dickel DE, Kusenda M, Krastoshevsky O, Krause V, Kumar RA, Grozeva D, Malhotra D, Walsh T, Zackai EH, Kaplan P, Ganesh J, Krantz ID, Spinner NB, Roccanova P, Bhandari A, Pavon K, Lakshmi B, Leotta A, Kendall J, Lee YH, Vacic V, Gary S, Iakoucheva LM, Crow TJ, Christian SL, Lieberman JA, Stroup TS, Lehtimäki T, Puura K, Haldeman-Englert C, Pearl J, Goodell M, Willour VL, DeRosse P, Steele J, Kassem L, Wolff J, Chitkara N, McMahon FJ, Malhotra AK, Potash JB, Schulze TG, Nöthen MM, Cichon S, Rietschel M, Leibenluft E, Kustanovich V, Lajonchere CM, Sutcliffe JS, Skuse D, Gill M, Gallagher L, Mendell NR, Craddock N, Owen MJ, O'Donovan MC, Shaikh TH, Susser E, DeLisi LE, Sullivan PF, Deutsch CK, Rapoport J, Levy DL, King M-C, Sebat J. Wellcome Trust Case Control Consortium. Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genetics.* 2009; 41:1223–1227. [PubMed: 19855392]
- Mosteller F, Tukey JW. The uses and usefulness of binomial probability paper. *Journal of the American Statistical Association.* 1949; 44(246):174–212. [PubMed: 18131512]
- Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, Shianna KV, Yoon W, Kasperaviciute D, Gennarelli M, Strittmatter WJ, Bonvicini C, Rossi G, Jayathilake K, Cola PA, McEvoy JP, Keefe RS, Fisher EM, St Jean PL, Giegling I, Hartmann AM, Möller HJ, Ruppert A, Fraser G, Crombie C, Middleton LT, St Clair D, Roses AD, Muglia P, Francks C, Rujescu D, Meltzer HY, Goldstein DB. A genome-wide investigation of snps and cnvs in schizophrenia. *PLoS Genet.* 2009; 5(2):e1000373. [PubMed: 19197363]
- Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev.* 2009; 19(3):212–9. [PubMed: 19481926]
- Sun YV, Sung YJ, Tintle N, Ziegler A. Identification of genetic association of multiple rare variants using collapsing methods. *Genet Epidemiol.* 2011; 35:S101–S106. [PubMed: 22128049]
- Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. *Am J Hum Genet.* 2010; 86(6):929–942. [PubMed: 20560208]
- Zar, JH. *Biostatistical Analysis.* 4. New Jersey: Prentice-Hall, Inc; 1999.

Figure 1.
Power at significance level 0.05. Dominant model ($\gamma_1 = \gamma_2 = 2$)

Figure 2. Power at significance level 0.05. Multiplicative model ($\gamma_1 = \gamma_2 = 4$)

Figure 3. Power at significance level 0.01 for the dominant model ($\gamma_1 = \gamma_2 = 2$). There are 10,000 simulation replicates.

Figure 4. Power at significance level 0.01 for the multiplicative model ($\gamma_1 = \gamma_2$, $\gamma_2 = 4$).

Table 1

Summary of data from a case-control study

Disease Status	Genotype		Total
	aa	aA or AA	
Case	X_1	$n_1 - X_1$	n_1
Control	X_2	$n_2 - X_2$	n_2

Table 2

Type I error rate computed as the proportion of rejections over 10,000 simulation replicates. There are 250 cases and 250 controls.

Significance Level	Allele Frequency	Test Statistic				
		Angular	FT	Anscombe	Chatter	Fisher
0.001	0.005	0.0064	0.0009	0.0003	0.0009	0.0000
	0.006	0.0088	0.0012	0.0004	0.0012	0.0000
	0.007	0.0092	0.0012	0.0004	0.0012	0.0000
	0.008	0.0074	0.0016	0.0006	0.0016	0.0000
	0.009	0.0057	0.0017	0.0008	0.0017	0.0001
	0.01	0.0051	0.0019	0.0009	0.0019	0.0003
	0.02	0.0012	0.0008	0.0008	0.0008	0.0001
	0.03	0.0018	0.0017	0.0017	0.0018	0.0011
	0.04	0.0010	0.0009	0.0009	0.0010	0.0005
	0.05	0.0014	0.0011	0.0011	0.0012	0.0007
0.005	0.005	0.0388	0.0064	0.0018	0.0064	0.0003
	0.006	0.0371	0.0090	0.0037	0.0090	0.0004
	0.007	0.0305	0.0098	0.0048	0.0098	0.0004
	0.008	0.0235	0.0080	0.0044	0.0081	0.0006
	0.009	0.0190	0.0079	0.0056	0.0079	0.0008
	0.01	0.0144	0.0072	0.0061	0.0077	0.0009
	0.02	0.0076	0.0058	0.0073	0.0073	0.0027
	0.03	0.0060	0.0056	0.0057	0.0057	0.0028
	0.04	0.0055	0.0049	0.0053	0.0055	0.0027
	0.05	0.0061	0.0056	0.0059	0.0060	0.0035
0.01	0.005	0.0394	0.0176	0.0066	0.0066	0.0010
	0.006	0.0389	0.0206	0.0095	0.0095	0.0014
	0.007	0.0344	0.0184	0.0113	0.0113	0.0013
	0.008	0.0292	0.0172	0.0102	0.0102	0.0022
	0.009	0.0249	0.0155	0.0116	0.0116	0.0019
	0.01	0.0192	0.0130	0.0111	0.0111	0.0026
	0.02	0.0123	0.0110	0.0110	0.0110	0.0052

Significance Level	Allele Frequency	Test Statistic			
		Angular	FT	Anscombe	Fisher
0.05	0.03	0.0111	0.0106	0.0106	0.0057
	0.04	0.0102	0.0100	0.0100	0.0061
	0.05	0.0122	0.0112	0.0112	0.0076
0.05	0.005	0.1366	0.0440	0.0440	0.0072
	0.006	0.1082	0.0477	0.0477	0.0113
	0.007	0.0933	0.0464	0.0464	0.0159
	0.008	0.0831	0.0477	0.0477	0.0174
	0.009	0.0724	0.0432	0.0432	0.0193
	0.01	0.0680	0.0436	0.0436	0.0180
	0.02	0.0578	0.0497	0.0497	0.0301
	0.03	0.0523	0.0476	0.0476	0.0318
	0.04	0.0537	0.0524	0.0524	0.0361
	0.05	0.0518	0.0493	0.0493	0.0358

Table 3

Type I error rate computed as the proportion of rejections over 10,000 simulation replicates. There are 500 cases and 500 controls.

Significance Level	Allele Frequency	Test Statistic				
		Angular	FT	Anscombe	Chanter	Fisher
0.001	0.005	0.0060	0.0022	0.0015	0.0022	0.0001
	0.006	0.0041	0.0023	0.0019	0.0023	0.0005
	0.007	0.0030	0.0014	0.0010	0.0014	0.0004
	0.008	0.0021	0.0014	0.0013	0.0014	0.0004
	0.009	0.0017	0.0009	0.0010	0.0011	0.0001
	0.01	0.0012	0.0006	0.0009	0.0009	0.0002
	0.02	0.0012	0.0010	0.0010	0.0010	0.0006
	0.03	0.0012	0.0011	0.0012	0.0012	0.0005
	0.04	0.0018	0.0017	0.0017	0.0017	0.0014
	0.05	0.0017	0.0016	0.0016	0.0016	0.0012
0.005	0.005	0.0150	0.0074	0.0063	0.0080	0.0016
	0.006	0.0108	0.0062	0.0064	0.0074	0.0022
	0.007	0.0076	0.0043	0.0049	0.0063	0.0015
	0.008	0.0084	0.0047	0.0059	0.0071	0.0018
	0.009	0.0070	0.0042	0.0048	0.0069	0.0018
	0.01	0.0062	0.0040	0.0046	0.0061	0.0021
	0.02	0.0055	0.0050	0.0052	0.0052	0.0025
	0.03	0.0050	0.0049	0.0050	0.0050	0.0030
	0.04	0.0070	0.0070	0.0070	0.0070	0.0048
	0.05	0.0059	0.0056	0.0056	0.0059	0.0036
0.01	0.005	0.0220	0.0148	0.0123	0.0123	0.0027
	0.006	0.0165	0.0120	0.0120	0.0120	0.0040
	0.007	0.0137	0.0104	0.0122	0.0122	0.0026
	0.008	0.0132	0.0110	0.0120	0.0120	0.0038
	0.009	0.0111	0.0098	0.0108	0.0108	0.0038
	0.01	0.0118	0.0102	0.0115	0.0115	0.0043
	0.02	0.0110	0.0099	0.0102	0.0102	0.0061

Significance Level	Allele Frequency	Test Statistic			
		Angular	FT	Anscombe	Fisher
0.03	0.03	0.0109	0.0097	0.0097	0.0066
	0.04	0.0124	0.0118	0.0122	0.0093
	0.05	0.0109	0.0109	0.0109	0.0084
	0.005	0.0689	0.0420	0.0461	0.0215
	0.006	0.0641	0.0413	0.0478	0.0222
0.04	0.007	0.0565	0.0431	0.0502	0.0243
	0.008	0.0581	0.0482	0.0537	0.0289
	0.009	0.0564	0.0512	0.0545	0.0316
	0.01	0.0511	0.0493	0.0503	0.0301
	0.02	0.0527	0.0490	0.0493	0.0344
0.05	0.03	0.0532	0.0494	0.0511	0.0354
	0.04	0.0509	0.0494	0.0494	0.0394
	0.05	0.0508	0.0503	0.0508	0.0400