Supplemental material: An omnibus test for the global null hypothesis

Andreas Futschik[1], Thomas Taus[2], and Sonja Zehetmayer[3]

[1]Department of Applied Statistics, Johannes Kepler University Linz, Austria
[2]Institute of Population Genetics, University of Veterinary Medicine Vienna, Austria
[3]Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Austria

January 31, 2018

1 Influence of the chosen transformation on the omnibus method

We present additional power curves for the omnibus test for $m_1 \in \{1,3,5,10\}$ and $\Delta/\sigma = 0.3/\sqrt{m_1}$ for $m = 5$ (Fig. 1) as well as $m = 20$ (Fig. 2) for the following transformations:

- $h(p) = 1 - p$ (omnibus p)
- $h(p) = -\log p$ (omnibus *log p*)
- $h(p) = z_{1-p}$ (omnibus z)
- $h(p) = p^{-\alpha}$ with $\alpha = 0.5$ (omnibus *power*)

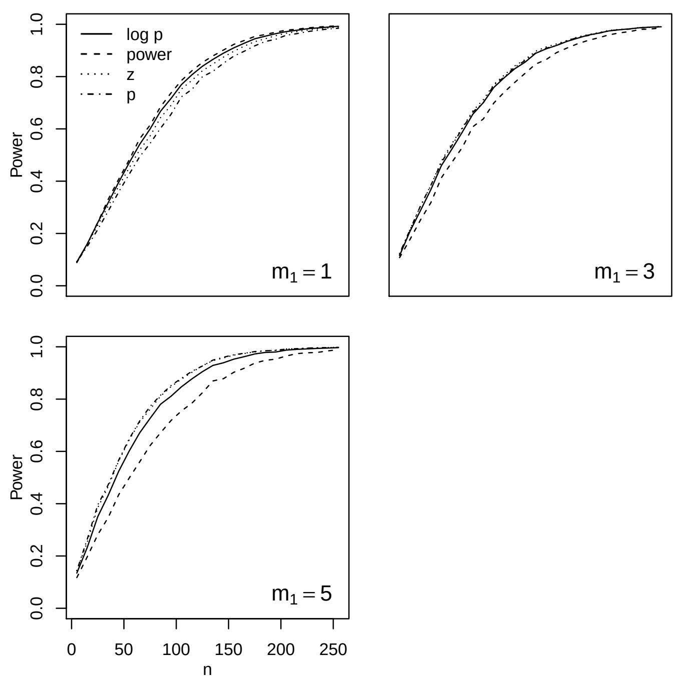

Figure 1: Power values for omnibus log p, power, z, and p are given for increasing n, $m = 5$, $m_1 \in \{1,3,5\}$, $\Delta/\sigma = 0.3/\sqrt{m_1}$.

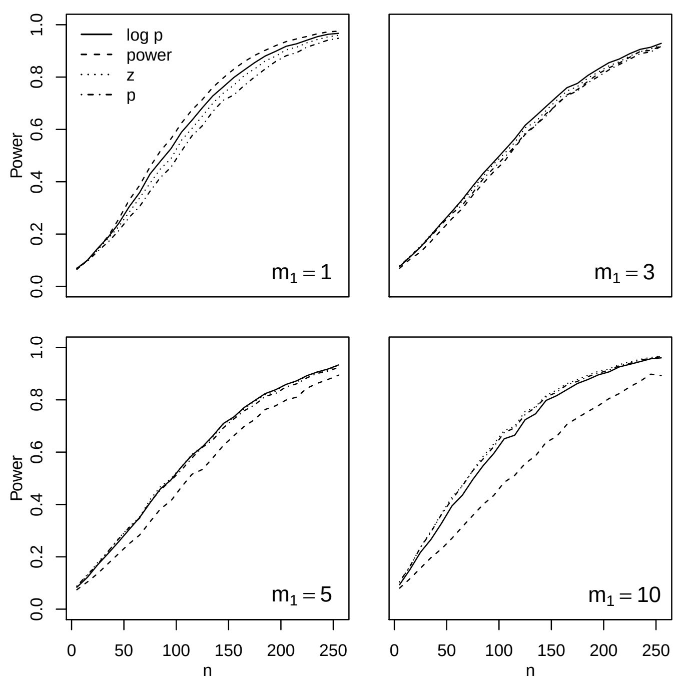

Figure 2: Power values for omnibus log p, power, z, and p are given for increasing n, $m = 20$, $m_1 \in \{1,3,5,10\}$, $\Delta/\sigma = 0.3/\sqrt{m}_1$.

2 Power comparison between different testing methods

Figures 3 and 4 show power curves for omnibus log p, Bonferroni test, Simes test, Fisher combination test, Stouffer's z test, and higher critisism (HC) for $m_1 \in \{1,3,5,10\}$, and $\Delta/\sigma = 0.3/\sqrt{m}_1$ for $m = 5$ (Fig. 3) as well as $m = 20$ (Fig. 4).

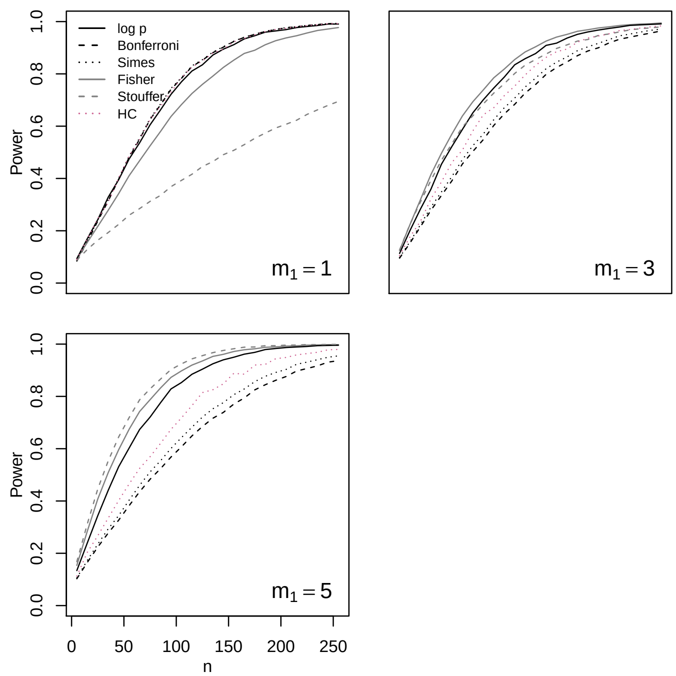

Figure 3: Power values for increasing n, $m = 10$, $m_1 \in \{1,3,5\}$, $\Delta/\sigma = 0.3/\sqrt{m}_1$ for omnibus log p, Bonferroni test, Simes test, Fisher combination test, and Stouffer's z test, and HC.

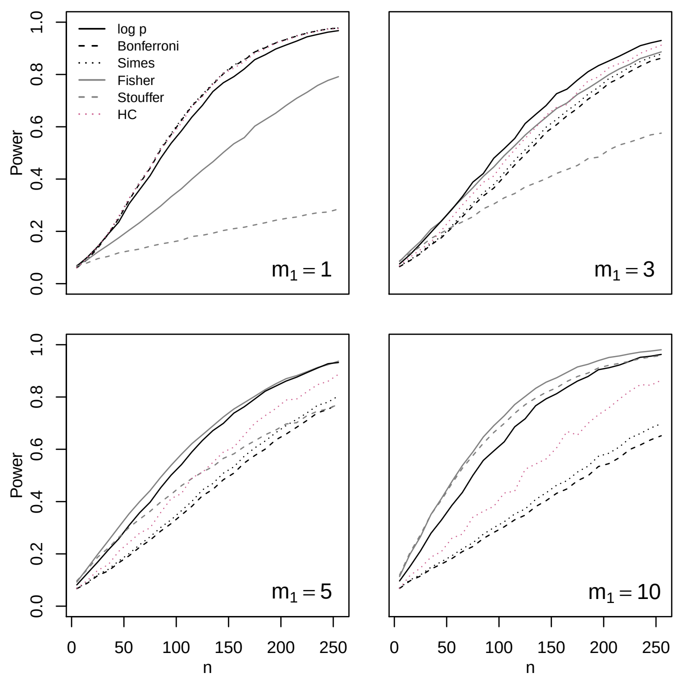

Figure 4: Power values for increasing n, $m = 10$, $m_1 \in \{1,3,5,10\}$, $\Delta/\sigma = 0.3/\sqrt{m_1}$ for omnibus log p, Bonferroni test, Simes test, Fisher combination test, and Stouffer's z test, and HC.

3 Type I error

Fig. 5 shows the type I error for the omnibus test for the four transformations (as supplement to Fig. 1 from the manuscript). Note that the upper limit of the y-axis is now 0.1. A horizontal line was drawn at 0.05.

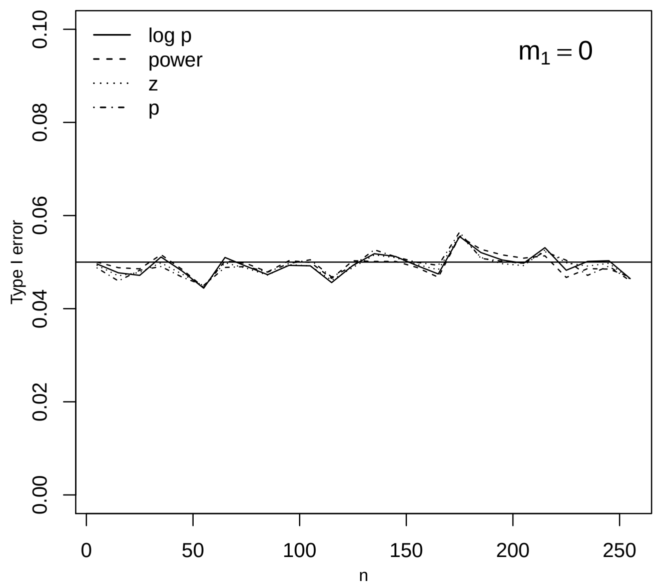

Figure 5: Type I error for omnibus log p, power, z, and p are given for increasing n, $m = 10$, $m_1 = 0$, $\alpha = 0.05$.

4 Two-sided tests

We additionally performed two-sided tests with approx. half of the alternatives with positive and the other half with negative effect sizes. Fig. 6 shows the two-sided equivalent to Fig. 1 from the manuscript. The power values for the two-sided case are lower than for the one-sided case due to the similar type I error rate of 0.05. However, the comparison of the transformations leads to analogous results. Again there are only small differences in power between the transformations, and the log p transform seems to lead to a particularly good trade off in power across many scenarios.

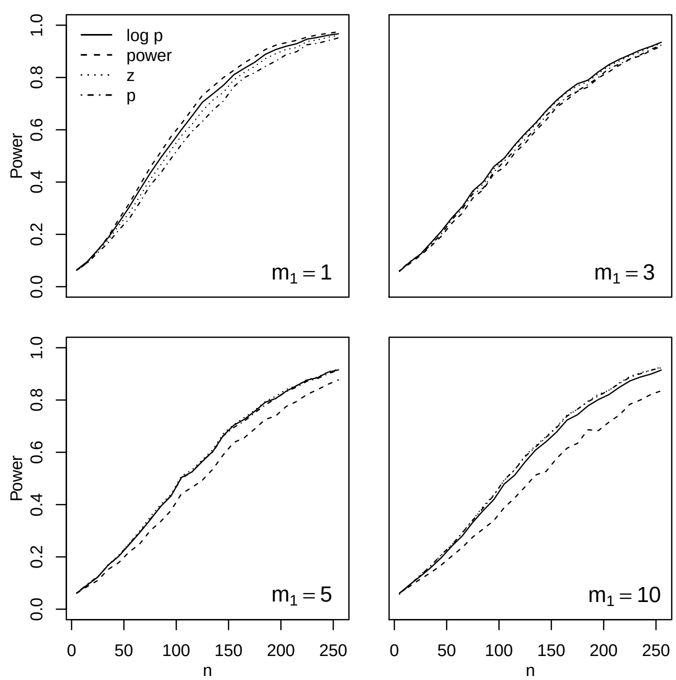

Figure 6: Power values for omnibus log p, power, z, and p are given for increasing n, $m = 10$, $m_1 \in \{1,3,5,10\}$, $\Delta/\sigma = 0.3/\sqrt{m_1}$, $\alpha = 0.05$ for the two-sided test.

5 Experimental Evolution

We provide a graphical summary of the approximately 2.6×10^6 p-values obtained by applying our omnibus test to five independent measurements of allele frequency change for each SNP considered in the experiment.

The Manhattan plots in Figure 7 provide p-values of separate tests for each of five replicate *Drosophila* populations. The corresponding overall p-values obtained using our omnibus test are displayed in Figure 8. For better visibility, we only plot each 50th SNP.

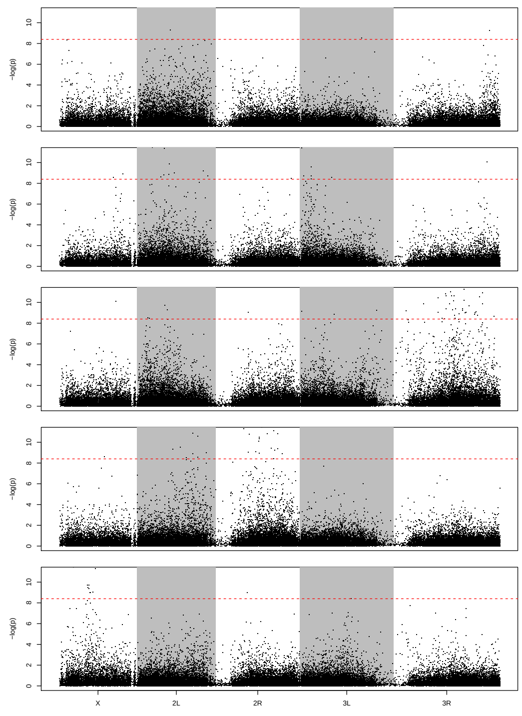

Figure 7: Manhattan plots of negative logarithm of p-values from a genome wide scan of five replicate populations. The red dashed line indicates the significance threshold when using the Bonferroni procedure to control the FWE at $\alpha = 0.01$. Data are taken from Griffin et al., (2017).

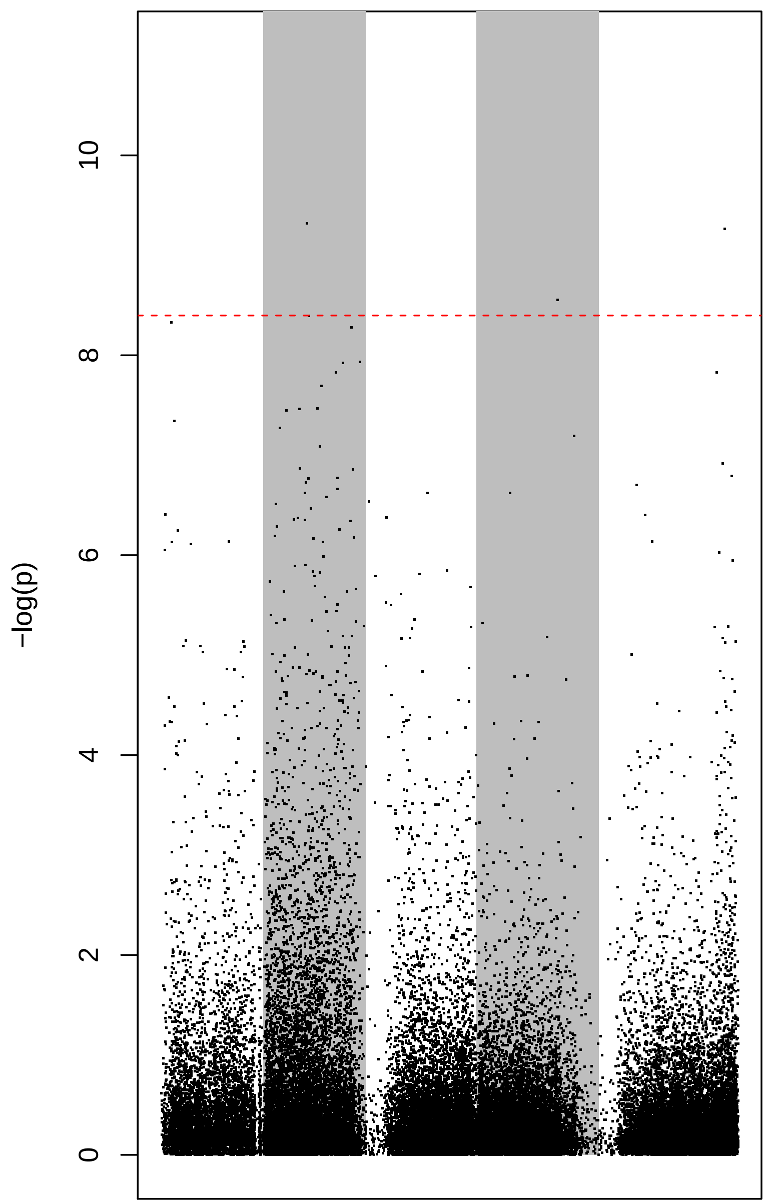

Figure 8: Plot of combined evidence across replicates. Manhattan plots of the negative logarithm of the p-values obtained with our omnibus *log p* test. The red dashed line indicates the significance threshold when using the Bonferroni procedure to control the FWE at $\alpha = 0.01$. Data are taken from Griffin et al., (2017).