

Model-specific tests on variance heterogeneity for detection of potentially interacting genetic loci

Ludwig A. Hothorn^{*1,2}, Ondrej Libiger², Daniel Gerhard¹

¹Institute of Biostatistics, Leibniz University Hannover, D-30419 Hannover, Germany

²Scripps Genomic Medicine, The Scripps Research Institute, La Jolla, CA-92037, USA

Email: Ludwig A. Hothorn^{*} - hothorn@biostat.uni-hannover.de; olibiger@scripps.edu; gerhard@biostat.uni-hannover.de;

^{*}Corresponding author

Abstract

Background: Trait variances among genotype groups at a locus are expected to differ in the presence of an interaction between this locus and another locus or environment. A simple maximum test on variance heterogeneity can thus be used to identify potentially interacting single nucleotide polymorphisms (SNPs).

Results: We propose a multiple contrast test for variance heterogeneity that compares the mean of Levene residuals for each genotype group with their average as an alternative to a global Levene test. We applied this test to a Bogalusa Heart Study dataset to screen for potentially interacting SNPs across the whole genome that influence a number of quantitative traits. A user-friendly implementation of this method is available in the R statistical software package multcomp.

Conclusions: We show that the proposed multiple contrast test of model-specific variance heterogeneity can be used to test potentially interacting SNPs with an unknown allele, loci or covariates and provide valuable additional information compared with traditional tests. Care is needed in the interpretation, particularly in the commonly occurring unbalanced designs.

Author's summary: XXXX

Background

Statistical association between a biallelic marker and a quantitative trait is usually tested using either a two degree of freedom F-test in the one-way analysis of variance (ANOVA) [1], or a one degree of freedom F-test in a linear regression [2]. These approaches compare the means of quantitative trait values at genotype categories at a SNP locus (i.e., homozygous for major allele, heterozygous and homozygous for minor allele). The ANOVA is sensitive against any global heterogeneity, whereas the linear regression test is sensitive against an additive mode of inheritance. Less attention has been given to comparing the variances in the quantitative trait values associated with different genotype categories. Recently, [3] proposed using a standard Levene test [4] to identify variance heterogeneity due to potential interaction between a given locus and another allele at the same locus, alleles at different loci or the environment. Differences among the variances of quantitative trait values associated with each genotype category (denoted σ_j^2 with $j = 0, 1, 2$ interacting alleles) may reflect an interaction [5]. In contrast to approaches that explicitly test specific gene-gene or gene-environment interactions [6], methods that assess variance heterogeneity can be used to uncover loci that are not known to interact.

Levene test [7] tests a global null hypothesis $H_0 : \sigma_0^2 = \sigma_1^2 = \sigma_2^2$ against the alternative that a difference exists between any pair of variances: $H_1 : \exists j, j' : \sigma_j^2 \neq \sigma_{j'}^2, j \neq j'$. The test statistic consists of a quadratic form

$$T_{Levene}^2 = \frac{(N - J) \sum_{j=0}^J n_j (Z_{j.} - Z_{..})^2}{(J - 1) \sum_{j=0}^J (Z_{ji} - Z_{j.})^2}, \quad N = \sum_{j=0}^J n_j$$

(with $J = 3$) using the robust Levene residuals

$$Z_{ji} = \text{abs}(Y_{ji} - \text{Median}(Y_{j.})),$$

with n_j quantitative trait observations Y_{ij} per genotype j . The T_{Levene}^2 is F -distributed with $df_1 = J$ and $df_2 = N - J$.

This test is known to be relatively robust when data are not normally distributed. However, the main disadvantage associated with using Levene's test is that it can only be used to determine whether the group-specific variances differ among each other. However, in order to obtain a biologically or clinically relevant interpretation of the results, it is often valuable to additionally determine which pairs of genotype categories exhibit statistically significant variance heterogeneity.

To this end, [3] considered using three two-sample $df - 1$ tests for the three comparisons σ_0^2 vs. σ_{12}^2 , σ_1^2 vs. σ_{02}^2 , and σ_2^2 vs. σ_{01}^2 , where $\sigma_{jj'}^2$ denotes the variance estimator for the pooled groups jj' . However, these multiple tests do not control the family-wise type I error rate α .

In this paper, we propose a Levene-type multiple contrast test, a novel approach comprised of a global test on variance heterogeneity as well as the three specific tests on pairwise variance heterogeneity using a maximum test of linear forms. We apply this test in a genome-wide fashion using a Bogalusa Heart Study dataset [8].

Methods

A Levene-type multiple contrast test

For the Levene-type transformed variable Z_{ij} and the factor *genotype* with the levels $j = 0, 1, 2$ the following contrast test for the one-way layout is used:

$$T_k^{Levene} = \frac{\sum_{j=1}^J c_{kj} \bar{Z}_j}{S \sqrt{\sum_{j=1}^J c_{kj}^2 / n_j}}$$

where S is the root of the common mean square error estimate of the linear model on the transformed data Z_{ji} (absolute Levene residuals). Contrast coefficients c_{kj} are used for the $k = 3$ comparisons between the means of Levene residuals for each genotype with the average of the means. These comparisons coding for σ_0^2 vs. σ_{12}^2 , σ_1^2 vs. σ_{02}^2 , and σ_2^2 vs. σ_{01}^2 [9], are formulated in the 3×3 contrast matrix:

$$(c_{kj}) = \begin{pmatrix} -1 & 0.5 & 0.5 \\ 0.5 & -1 & 0.5 \\ 0.5 & 0.5 & -1 \end{pmatrix}.$$

A priori it is unknown which elementary test is mostly under the alternative. Therefore, the maximum of the test statistics $\max(T_j)$ is used, implying a family-wise type-I-error rate α for all of the three comparisons.

Under the approximate assumption of multivariate normal distributed errors with a homogeneous global variance of the transformed variable Z_{ij} , the vector $(T_j)'$ follows jointly a tri-variate t -distribution with $\nu = \sum_{j=1}^J (n_j - 1)$ degrees of freedom and a correlation matrix $\mathbf{R} = (\rho_{kk'})$ given by its elements

$$\rho_{kk'} = \frac{\sum_{j=1}^J c_{kj} c_{k'j} / n_j}{\sqrt{\left(\sum_{j=1}^J c_{kj}^2 / n_j\right) \left(\sum_{j=1}^J c_{k'j}^2 / n_j\right)}}. \quad (1)$$

Under the above assumptions the correlations $\rho_{kk'}$ depend only on the contrast coefficients c_{kj} and the sample sizes n_j . This approach controls the familywise error rate α and reveals a reasonable power for unbalanced designs as long the above assumptions hold true. It provides a global decision whenever any of the contrasts is under the alternative *and* additionally the elementary decisions by multiplicity-adjusted p-values for the three specific comparisons. Although simultaneous confidence intervals for both differences

and ratios to the average are available as well [10], they will not be recommended since a genetically interpretation for the transformed variable Z_{ji} seems to be not appropriate. Recently simultaneous confidence intervals for the pairwise ratios of variances was proposed using a maximum test on jackknifed $\log(s_j^2)$ [11]. This approach would be an alternative when modified for arbitrarily unbalanced designs. The question arises whether the quadratic form of the Levene test or the maximum contrast version of linear forms is more powerful. A general answer will be not available. For the comparison of normal distributed means, the least favorable configuration approach for $\mu_j - \bar{\mu}$. even conclude higher power for the maximum of linear contrasts relative to the quadratic form of the F-test of the one-way analysis of variance [12].

Alternatively, a multiple contrast test based on pairwise comparisons of the genotype-specific variances MCT^{pair} can be used by means of the following 3×3 contrast matrix:

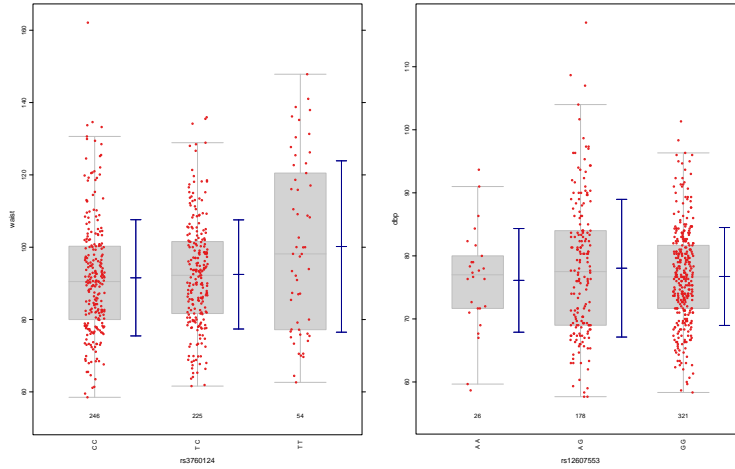
$$(c_{kj}) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

Results

Evaluation of a real data example

We evaluated the utility of the proposed test using data from Bogalusa Heart Study (BHS). The longitudinal study included genotype information on 525 unrelated individuals of European descent at 545,821 SNPs [12]. Twelve clinically-relevant quantitative traits were also measured for each study participant. We applied the multiple contrast test to each SNP throughout the genome, explicitly analyzing SNP by age interaction. While each test requires about X seconds of computing time, different genomic regions were analyzed on separate processor nodes in parallel, effectively decreasing the total computational time requirements for a genome-wide study by a factor proportional to the number of available computing nodes, in our case about 1,000. A number of SNP loci with a p-value smaller than ($< 10^{-8}$) were identified for several quantitative traits. In addition, the number of SNPs with p-values below a multiplicity threshold of $< 10^{-8}$ varied for different quantitative traits. For example, 23 SNP loci yielded p-values lower than that threshold when considering heart rate, while only 5 SNPs reached similar significance in the analysis involving body mass index. A thorough examination of the results showed that several significant tests occurred when the risk allele homozygote genotype group contained only a handful of data points. Although the approach described in this paper is statistically valid in these situations, we selected only tests in which each genotype group contained at least one percent of all data points for

further study. In the following text, we showcase two instances of variance heterogeneity in greater detail: rs3760124 and waist circumference, as well as rs12607553 and diastolic blood pressure.



The boxplots show relatively symmetric distributions of quantitative trait values in all genotype groups, thus ruling out the presence of outliers or extremely skewed distributions of trait values as sources of the observed variance heterogeneity. Furthermore, all three genotype categories contain a relatively large number of observed trait values resulting in reliable variance estimates.

Trait	SNP	Test	Comparison	p-value
waist circumference	rs3760124	Levene	global	$3.4 \cdot 10^{-07}$
		MCT^{ave}	σ_{CC}^2 vs. $\sigma_{CT,TT}^2$	$2.7 \cdot 10^{-01}$
			σ_{CT}^2 vs. $\sigma_{CC,TT}^2$	$1.5 \cdot 10^{-01}$
			σ_{TT}^2 vs. $\sigma_{CC,CT}^2$	$9.8 \cdot 10^{-08}$
		MCT^{pair}	σ_{CC}^2 vs. σ_{CT}^2	$9.6 \cdot 10^{-01}$
			σ_{CC}^2 vs. σ_{TT}^2	$5.3 \cdot 10^{-07}$
			σ_{CT}^2 vs. σ_{TT}^2	$4.4 \cdot 10^{-07}$
		diastolic blood pressure	rs12607553	Levene
MCT^{ave}	σ_{AA}^2 vs. $\sigma_{AG,GG}^2$			$5.9 \cdot 10^{-01}$
	σ_{AG}^2 vs. $\sigma_{AA,GG}^2$			$1.2 \cdot 10^{-07}$
	σ_{GG}^2 vs. $\sigma_{AA,AG}^2$			$1.8 \cdot 10^{-06}$
MCT^{pair}	σ_{AA}^2 vs. σ_{AG}^2			$3.2 \cdot 10^{-02}$
	σ_{AA}^2 vs. σ_{GG}^2			$9.9 \cdot 10^{-01}$
	σ_{AG}^2 vs. σ_{GG}^2			$1.9 \cdot 10^{-07}$

Table 1: P values for original and multiple contrast Levene-type tests

It can be seen from Table 1 that our proposed test (labeled as MCT^{ave}) yielded a more significant result compared with the Levene global test. More importantly however, it allowed us to interpret the effect of the SNP alleles on the variance heterogeneity. When we considered waist circumference, we found the

variance heterogeneity to be most significant between the risk allele homozygous genotype group (σ_{TT}^2), and the pooled variance for the two remaining genotype groups associated with the presence of at least one non-risk allele ($\sigma_{CC,CT}^2$). In addition, we can see that the variance associated with the risk allele homozygotes is significantly higher in this example. Conversely, in the second example involving diastolic blood pressure, the variance in the heterozygotes σ_{AG}^2 was significantly higher compared to homozygotes $\sigma_{AA,GG}^2$. This type of insights may prove to be valuable for forming hypotheses in further research, and can be obtained with the Levene-type multiple contrast test that we propose here. Furthermore, pairwise comparisons (labeled as MCT^{pair}) can also be performed (see Table 1).

An example R code for testing variance heterogeneity at a single SNP is provided in the Supplementary material. The multiplicity-adjusted p-values can be estimated by means of the R package `multcomp` [13]. Alternatively, a SAS procedure `GLIMMIX` can be used for a resampling based estimation of multiplicity-adjusted p-values [14].

Conclusions

The serious issue of missing heritability, which refers to the fact that common SNPs identified by genome-wide association studies as associated with a disease collectively explain only a small portion of the prevalence of this disease, may be due, in part, to the presence of unknown interactions among alleles at various SNP loci or environment, that have an effect on the disease. The identification of such interactions poses many challenges, primarily because of the large number of potentially interacting pairs, trios, etc. of alleles and environmental variables that need to be tested. A feasible alternative, as suggested by [5], is to test individual loci for the evidence of their involvement in an interaction with other alleles, loci or covariates. The idea of assessing variance heterogeneity between three genotype groups at a particular SNP locus as evidence of a potential interaction is appealing for its simplicity. The Levene-type maximum contrast test proposed in this paper allows one to not only test for global variance heterogeneity, but also perform groupwise test that allows one to elucidate the effect of the individual alleles on the quantitative trait variance. While this is an advantage over the standard Levene test, the price to pay is increased computing time. However, parallelization can be used to substantially alleviate this problem. Variance heterogeneity testing of the type proposed here can be easily embedded in a linear model which would allow us to test more elaborate models and account for the potential effects of covariates such as sex or age. R code implementing this test is available as part of the `multcomp` package. Care needs to be exercised in interpreting the results of this test in cases of low frequency variants or missing trait data, when one or

more of the genotype groups contains an extremely small number of observed trait values. The issues surrounding the sensitivity and specificity of this approach in these potentially common cases is an area in need of further work.

Authors contributions

LAH and DG derived the method. LAH and OL developed the software. OL performed the GWA analysis. LAH, OL and DG wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The work for the first author was partly supported by the German Science Foundation grand DfG-HO1687.

References

1. Liu YZ, Pei YF, Guo YF, Wang L, Liu XG, Yan H, Xiong DH, Zhang YP, Levy S, Li J, Haddock CK, Papasian CJ, Xu Q, Ma JZ, Payne TJ, Recker RR, Li MD, Deng HW: **Genome-wide association analyses suggested a novel mechanism for smoking behavior regulated by IL15**. *Molecular Psychiatry* 2009, **14**(7):668–680.
2. Zhao JH, Li MY, Bradfield JP, Zhang HT, Mentch FD, Wang K, Sleiman PM, Kim CE, Glessner JT, Hou CP, Keating BJ, Thomas KA, Garris ML, Deliard S, Frackelton EC, Otieno FG, Chiavacci RM, Berkowitz RI, Hakonarson H, Grant SFA: **The role of height-associated loci identified in genome wide association studies in the determination of pediatric stature**. *Bmc Medical Genetics* 2010, **11**:96.
3. Struchalin MV, Dehghan A, Wittelman JCM, van Duijn C, Aulchenko YS: **Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations**. *Bmc Genetics* 2010, **11**:92.
4. Levene H: *Robust tests for equality of variances*. Stanford University Press 1960.
5. Pare G, Cook NR, Ridker PM, Chasman DI: **On the Use of Variance per Genotype as a Tool to Identify Quantitative Trait Interaction Effects: A Report from the Women's Genome Health Study**. *Plos Genetics* 2010, **6**(6):e1000981.
6. Erhardt V, Bogdan M, Czado C: **Locating Multiple Interacting Quantitative Trait Loci with the Zero-Inflated Generalized Poisson Regression**. *Statistical Applications In Genetics and Molecular Biology* 2010, **9**:26.
7. Gastwirth JL, Gel YR, Miao WW: **The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice**. *Statistical Science* 2009, **24**(3):343–360.
8. Smith EN, Chen W, Kahonen M, Kettunen J, Lehtimaki T, Peltonen L, Raitakari OT, Salem RM, Schork NJ, Shaw M, Srinivasan SR, Topol EJ, Viikari JS, Berenson GS, Murray SS: **Longitudinal Genome-Wide Association of Cardiovascular Disease Risk Factors in the Bogalusa Heart Study**. *Plos Genetics* 2010, **6**(9):e1001094.
9. Hasler M, Hothorn LA: **Multiple Contrast Tests in the Presence of Heteroscedasticity**. *Biometrical Journal* 2008, **50**(5):793–800.
10. Djira GD, Hothorn LA: **Detecting Relative Changes in Multiple Comparisons with an Overall Mean**. *Journal Of Quality Technology* 2009, **41**:60–65.
11. Rublik F: **A Note on Simultaneous Confidence Intervals for Ratio of Variances**. *Communications In Statistics-theory and Methods* 2010, **39**(6):1038–1045.
12. Hayter AJ, Liu W: **The Power Function of the Studentised Range Test**. *Annals of Statistics* 1990, **18**:465–468.

13. Bretz F, Hothorn T, Westfall P: **On multiple comparisons in R**. *R News* 2002, **2**:14–17.
14. *SAS/STAT 9.2, Cary NC, USA (2010)*.

Figures

Additional Files

The R code

```
med26<-tapply(tgK$NCleanEpi, tgK$G26, median) # group-specific medians
med26V<-med26[as.integer(tgK$G26)] # median for each subject
tgK$NCE26<- abs(tgK$NCleanEpi-med26V) # Levene transformation

library(multcomp)
l26<-lm(NCE26~G26, data=tgK) # one-way ANOVA linear model fir
summary(glht(l26, linfct=mcp(G26 = "AVE"))) # multiple contrast test of "average type"
```