Letter to the editor: expression of concern, reaffirmed

Andrew D. Paterson

Received: 29 May 2012 / Accepted: 22 July 2013 / Published online: 24 August 2013
© American Aging Association 2013

Dear Dr. Ingram,

I would like to outline a series of limitations in the analysis presented in two articles published recently in *Age*.

Kulminski et al. (2013) Inter-chromosomal level of genome organization and longevity-related phenotypes in humans. *Age* 35(2):501–518

In Kulminski et al. (2013), the authors fail to present a QQ plot describing the observed vs. expected distribution of SNP p values for the association with cardiovascular disease (CVD). The authors stated that “We screened all qualified SNPs for their potential associations with CVD (note, these analyses were not intended to ascertain if these associations were true)...”. It is unusual to perform an analysis without intending that the results of the analysis should be interpreted. This point is important since the authors used the results of these association tests for the rest of the analysis in the paper, and inclusion of SNPs with high error rates may mean that any conclusions regarding inter-chromosomal

association are spurious. Inspection of Supplementary Table 1 from the paper lists 69 SNPs with $p \leq 5 \times 10^{-7}$, many with extremely highly significant p values—for example, rs5491 (minor allele frequency, MAF=7 %) has odds ratio (OR)=6.81, with $p=2.7 \times 10^{-87}$. This is just one of many extreme results: rs11568688 (MAF=5.5 %) has an unbelievable OR=14.19 with $p=5.5 \times 10^{-60}$. Standard epidemiologic training teaches students to be sceptical about massive effects, since they may indicate confounding. We have plotted the minor allele frequency against the $-\log_{10}p$ value for the SNPs in this list (Fig. 1), and it is clear that SNPs with the smallest p values also tend to have lower MAF—a red flag that there may be problems with genotype calling for these SNPs. The Wellcome Trust Case Control Consortium (2007) had to develop a novel method for genotype calling to deal with technical problems with genotype calling on Affymetrix chips similar to the chip used to generate the data that Kulminski used. Notably, even this new method preferentially resulted in the exclusion of SNPs with low minor allele frequencies (see Supplementary Fig. 1 from WTCCC 2007). A focus on strict quality control of genotype data has received attention from a number of authors (Clayton et al. 2005; Laurie et al. 2010; Turner et al. 2011). Unfortunately, the authors did not provide conventional quality control metrics for these SNPs, such as genotyping call rate, deviation from Hardy–Weinberg Equilibrium, and crucially whether visual examination of the cluster plots of two allele intensity values to determine whether the genotype calling by the software used was appropriate. As previously stated (Paterson 2012), a study on the sample size

A. D. Paterson (✉)
Program in Genetics and Genome Biology,
The Hospital for Sick Children,
TMDT East Tower, Room 15-707, 101 College Street,
Toronto, ON M5G 1L7, Canada
e-mail: andrew.paterson@utoronto.ca

A. D. Paterson
Biostatistics Division, Dalla Lana School of Public Health,
University of Toronto,
Toronto, ON, Canada

Fig. 1 Plot of minor allele frequency (*X*-axis) vs. $-\log(10)p$ value for 69 SNPs reported to be associated with cardiovascular disease in Supplementary Table 1 from Kulminski et al., 2013

analysed by Kulminski et al. (2013) would be unlikely to have good power to detect any loci apart from the chromosome 9p21 locus for cardiovascular disease at $\alpha < 5 \times 10^{-8}$.

For certain SNPs that are the focus of much of the paper of Kulminski et al. (2013), there are additional red flags that strongly suggest major problems with the genotype calling. Specifically, in Online Resource 9, both the text and Table S2 provide some clues to the problems with genotype calling. The authors stated that “minor alleles at SNPs were found predominantly in phases 3 and 4.” What was not stated is that there were highly significant differences in genotype frequency for these SNPs between these phases. How the genotypes were called is not described in the methods (e.g. plate-by-plate or by phase); however, the results presented in Table S2 are consistent with major genotyping error. Our Table 1 shows the genotype counts for the three SNPs that are the focus of

analysis in much of the paper of Kulminski et al. (2013), separately by phase. First, it is clear that call rates separately by phase are low—so low that for all three SNPs in phase 4, these SNPs would be excluded from any standard analysis. Second, rs1889019 deviates significantly from Hardy–Weinberg equilibrium in both phases 3 and 4, as does rs9330200 in phase 4, and rs139069 in phase 3, although the authors make no mention of this. As expected, there are highly significant differences in the genotype frequency at all three SNPs between phases 3 and 4, for rs9330200 and rs139069, with Fisher’s exact test at $p < 100-100!$ Since phases of genotyping are confounded with generation (i.e. cohorts) in the study, this would be expected to demonstrate association with any phenotype that differs by generation, and/or age, such as CVD. Finally, since phases are related to the cohorts, one would imagine that drastic differences in genotypes between family members would be expected to result in Mendelian errors.

Table 1 Genotype counts for three SNPs extracted from Table S2 of Kulminski et al. (2013), separately by phase of genotyping

SNP	Phase 3				Phase 4			
	Call rate (%)	AA	Aa	aa	Call rate (%)	AA	Aa	aa
rs9330200	93	306	0	0	79	15	333	10
rs139069	100	0	328	0	76	343	0	0
rs1889019	58	19	167	6	48	92	121	3

The authors stated that “No apparent clustering of minor alleles or potentially problematic DNA samples on genotyping plates was identified (Pluzhnikov et al., 2010).” Exactly what the authors did is not clear—for example, were cluster plots visually inspected? They were not included as supplementary figures to the paper. The cited reference describes plate effects, whereas it appears that part of the technical problem in analysis performed by Kulminski et al. (2013) is related to phase effects, which may be missed by analysis of plates.

Kulminski and Culminskaya (2013) Genomics of human health and aging. *Age* 35(2):455–469

Similar concerns also apply to another paper in *Age* (title above), which used much of the same data by the same authors (Kulminski and Culminskaya 2013). Specifically, in Supplementary Table 1, they list 63 SNPs with $p < 10^{-6}$ that are associated with either CVD, cancer, systolic blood pressure or total cholesterol—again no QQ plots are provided. In online resource no. 3, for selected SNPs, they compared the original SNP models (with no covariates) to those including age and sex as covariates. A general pattern, whereby the SNP associations are attenuated when age was included as a covariate, implies correlation of age (specifically phase) with genotype, a further indicator that there are phase-specific genotype differences, likely due to differential genotyping error between phases. These technical artefacts could easily produce spurious SNP association and confound any downstream analysis.

In a recent letter to the editor of the *Experimental Gerontology*, I brought up similar criticisms (Paterson 2012) of an earlier paper published by the same author (Kulminski 2011). However, in their reply (Kulminski 2012), none of the specific criticisms were addressed. Until the authors address the issues of genotyping quality, any conclusions that SNPs on different chromosomes are associated with each other is premature.

Response to comments by Kulminski in response to JAAA-D-12-00764 (received 31 Dec 2012)

In their response, the author provides theoretical arguments why it is unlikely that their results are due to confounding due to genotyping error. However, they do not provide any new data or analyses to support their claims. For example, they have elsewhere argued that

technical errors in SNP genotyping are unlikely to result in correlation between markers (their reference no. 3), but this argument assumes that technical errors are randomly distributed, and this may not be the case. In order to illustrate the potential problem, I quote from a recent paper by The Genome-Wide Association Working Group which is part of the MicroArray Quality Consortium:

One of the aims of the working group is to assess the variability of genotype calls within and between different genotype calling algorithms... Our results show that the choice of genotyping algorithm ... can introduce marked variability in the results of downstream case-control association analysis for the Affymetrix 500 K array. The amount of discordance between results is influenced by how samples are combined and processed through the respective genotype calling algorithm, indicating that systematic genotype errors due to computational batch effects are propagated to the list of single-nucleotide polymorphisms found to be significantly associated with the trait of interest (Miclaus et al. 2010).

My intention was not to contribute to the ongoing discussion in the literature regarding the relative proportion of variance for traits and diseases that are the result of common vs. rare genetic variation. I will just note that a recent review provided clear arguments that common variant GWAS has led to the identification of numerous genetic risk factors for complex diseases (Visscher et al. 2012).

Finally, when unexpectedly massive effects are observed, it behoves us to examine the raw data that contribute to such results. Numerous studies have documented the high-quality phenotype data in the Framingham Heart Study; the other variable in the model is the SNP genotype data—the quality of which can easily be examined—but for unstated reasons the authors prefer not to do this. Failing to do so runs the risk of misleading not only themselves, but others in the scientific community.

References

- Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD, Todd JA (2005) Population structure, differential bias and genomic control

- in a large-scale, case-control association study. *Nat Genet* 37(11):1243–1246
- Kulminski A (2011) Complex phenotypes and phenomenon of genome-wide inter-chromosomal linkage disequilibrium in the human genome. *Exp Gerontol*. doi:10.1016/j.exger.2011.08.010
- Kulminski A (2012) Have to or may? Re: expression of concern re: Kulminski AM (2011). Complex phenotypes and phenomenon of genome-wide inter-chromosomal linkage disequilibrium in the human genome. *Exp Gerontol*. 46, 979–986. *Exp Gerontol* 47(6):481–2
- Kulminski AM, Culminskaya I (2013) Genomics of human health and aging. *Age* 35(2): 455–469
- Kulminski AM, Culminskaya I, Yashin AI (2013) Inter-chromosomal level of genome organization and longevity-related phenotypes in humans. *Age* 35(2):501–518
- Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, Boehm F, Caporaso NE, Cornelis MC, Edenberg HJ, Gabriel SB, Harris EL, Hu FB, Jacobs KB, Kraft P, Landi MT, Lumley T, Manolio TA, McHugh C, Painter I, Paschall J, Rice JP, Rice KM, Zheng X, Weir BS, GENEVA Investigators (2010) Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* 34(6):591–602
- Miclaus K, Chierici M, Lambert C, Zhang L, Vega S, Hong H, Yin S, Furlanello C, Wolfinger R, Goodsaid F (2010) Variability in GWAS analysis: the impact of genotype calling algorithm inconsistencies. *Pharmacogenomics J* 10(4):324–335. doi:10.1038/tpj.2010.46
- Paterson AD (2012) Expression of concern re: Kulminski, A. 2011. Complex phenotypes and phenomenon of genome-wide inter-chromosomal linkage disequilibrium in the human genome. *Experimental Gerontology*. doi:10.1016/j.exger.2011.08.010. *Exp Gerontol* 47(6):479–80
- The Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661–678
- Turner S, Armstrong LL, Bradford Y, Carlson CS, Crawford DC, Crenshaw AT, de Andrade M, Doheny KF, Haines JL, Hayes G, Jarvik G, Jiang L, Kullo IJ, Li R, Ling H, Manolio TA, Matsumoto M, McCarty CA, McDavid AN, Mirel DB, Paschall JE, Pugh EW, Rasmussen LV, Wilke RA, Zuvich RL, Ritchie MD. (2011) Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet*. Chapter 1:Unit1.19
- Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90(1):7–24. doi:10.1016/j.ajhg.2011.11.029