European Journal of Human Genetics (2020) 28:1629–1630
https://doi.org/10.1038/s41431-020-00731-7

COMMENT

A powerful new method for rare-variant analysis of quantitative traits in families

Joan E. Bailey-Wilson[1]

Received: 8 September 2020 / Accepted: 16 September 2020 / Published online: 21 September 2020

The ability to use whole exome and whole-genome sequencing to detect rare germline genetic variants across the genome in a cost-effective manner has led to the revitalization of family studies and linkage analysis [1, 2]. While association studies are powerful to detect the effects of common variants on the variation of quantitative traits, these variants often have only small individual effects on the trait, particularly for clinically important quantitative traits, which may have been affected by selection pressure over evolutionary time. Family-based linkage studies, in contrast, are powerful to detect the effects of rare variants (RVs) with large to moderate effects on variation of the trait. Traditional two-point linkage of each RV with the trait is theoretically straightforward for dominant or additive rare risk alleles in large pedigrees although more difficult for rare recessive-acting variants. However, linkage can be complicated in the presence of allelic heterogeneity (many different rare causal variants in the same gene exist in the population) and/or locus heterogeneity (some families in a population are segregating a variant in a certain gene that contributes to the variation of the quantitative trait but other families have no causal variants in that gene). This can cause loss of power in a collection of small pedigrees when each family is segregating a different causal variant in the same or different genes. Allelic heterogeneity is particularly problematic in the case of recessive action of RVs since an affected individual may be carrying two different RVs in the same gene. In this issue, Zhao et al. present a powerful new method, the RV-QNPL, that extends this group's previous work on RV linkage methods [3, 4] to quantitative trait nonparametric linkage analysis (QNPL). This new approach is based on the well-known Haseman–Elston nonparametric linkage method for quantitative traits in general pedigrees (H-Eg) [5, 6]. The RV-QNPL is more powerful than traditional QNPL methods designed for common variants. Importantly, because there is rarely strong linkage disequilibrium between RVs across a large genomic region, the RV-QNPL can narrow in on small linked regions, often to a single linked causal gene or regulatory element. Such fine-mapping is quite unlikely when utilizing traditional QNPL methods on common variants that usually can only narrow down the location of the causal variant(s) to a much larger genomic region (often >30–40 Mb).

The RV-QNPL's novelty is due to two advances over traditional QNPL methods. First, it combines the RVs in a gene or region into "condensed haplotype pattern" variants [3] so that each gene, regulatory element or intergenic region has one multi-allelic "marker locus" where the alleles are different haplotypes of RVs plus one haplotype that contains no RVs (if such exists in the data). The RV-QNPL then further modifies the H-Eg such that allele sharing between relatives is only estimated and compared to the expected distribution for the rare-variant containing haplotypes and not for the haplotype that does not contain any RVs. The new RV-QNPL is more powerful than other H-Eg methods when the causal variants are rare and it loses less power than these other H-Eg methods in the presence of allelic and locus heterogeneity, inclusion of non-causal RVs in the analysis and missing phenotype and genotype data on pedigree founders. Particularly encouraging was the small observed loss of power due to locus heterogeneity in the simulation studies. Due to minor increases in false positive rates in both the RV-QNPL and the other evaluated H-Eg methods when applied to pedigrees, the authors recommend using permutation to derive empiric significance levels and to control Type-1 error when using this method. The RV-QNPL is also appropriate for combining families of different ethnic groups, as long as population-specific allele frequencies are used to select which variants are rare in the population from which each family is derived. This new

✉ Joan E. Bailey-Wilson
jebw@mail.nih.gov

[1] Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD, USA

approach is a powerful and exciting addition to the tools available to analyze RVs in family data. Its improvement on fine-mapping over traditional QNPL methods should lead to the identification of RVs that make moderate to large contributions to the variation of clinically important quantitative traits.

Funding JEB-W is supported solely by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health, USA.

Compliance with ethical standards

Conflict of interest The author declares no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Bailey-Wilson JE, Wilson AF. Linkage analysis in the next-generation sequencing era. Hum Hered. 2011;72:228–36.
2. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. Nat Rev Genet. 2015;16:275–84.
3. Wang GT, Zhang D, Li B, Dai H, Leal SM. Collapsed haplotype pattern method for linkage analysis of next-generation sequence data. Eur J Hum Genet. 2015;23:1739–43.
4. Zhao L, He Z, Zhang D, Wang GT, Renton AE, Vardarajan BN, et al. A rare variant nonparametric linkage method for nuclear and extended pedigrees with application to late-onset alzheimer disease via WGS data. Am J Hum Genet. 2019;105:822–35.
5. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. Behav Genet. 1972; 2:3–19.
6. Sham PC, Purcell S, Cherny SS, Abecasis GR. Powerful regression-based quantitative-trait linkage analysis of general pedigrees. Am J Hum Genet. 2002;71:238–53.