Narrow-sense heritability estimation of complex traits using identity-by-descent information.

Luke M. Evans[1,8], Rasool Tahmasbi[1], Matthew Jones[2], Scott I. Vrieze[3], Gonçalo R. Abecasis[4], Sayantan Das[4], Doug W. Bjelland[1], Teresa R. deCandia[1], Haplotype Reference Consortium, Jian Yang[5], Michael E. Goddard[6,7], Peter M. Visscher[5], Matthew C. Keller[1,2,8]

[1]Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80309

[2]Department of Psychology and Neuroscience, University of Colorado, Boulder, CO, 80309

[3]Department of Psychology, University of Minnesota, Minneapolis, MN 55455

[4]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109

[5]Institute for Molecular Bioscience and the Queensland Brain Institute, University of Queensland, Brisbane, 4072, Queensland, Australia

[6]Faculty of Veterinary and Agricultural Science, University of Melbourne, Parkville, Victoria, Australia

[7] Biosciences Research, Department of Economic Development, Jobs, Transport and Resources, Victoria, Australia

[8]Corresponding authors luke.m.evans@colorado.edu and matthew.c.keller@gmail.com

Running title: IBD-based heritability estimation
Word count: 7,304 for main text excluding references, tables, and figures

ABSTRACT

Heritability is a fundamental parameter in genetics. Traditional estimates based on family or twin studies can be biased due to shared environmental or non-additive genetic variance. Alternatively, those based on genotyped or imputed variants typically underestimate narrow-sense heritability contributed by rare or otherwise poorly-tagged causal variants. Identical-by-descent (IBD) segments of the genome share all variants between pairs of chromosomes except new mutations that have arisen since the last common ancestor. Therefore, relating phenotypic similarity to degree of IBD sharing among classically unrelated individuals is an appealing approach to estimating the near full additive genetic variance while possibly avoiding biases that can occur when modeling close relatives. We applied an IBD-based approach (GREML-IBD) to estimate heritability in unrelated individuals using phenotypic simulation with thousands of whole genome sequences across a range of stratification, polygenicity levels, and the minor allele frequencies of causal variants (CVs). In the absence of population stratification, the IBD-based approach produced heritability estimates, even when CVs were extremely rare, although precision was low. However, population stratification and non-genetic familial environmental effects passed vertically across generations led to strong biases in IBD-based heritability. We used data on two traits in ~120,000 people from the UK Biobank to demonstrate that, depending on the trait and possible confounding environmental effects, GREML-IBD can be applied to very large genetic datasets to infer the contribution of very rare variants lost using other methods. However, we observed apparent biases in these real data, suggesting that more work may be

46 required to understand and mitigate factors that influence IBD-based heritability
47 estimates.
48
49 **KEYWORDS**
50 Identical-by-descent, IBD, heritability, missing heritability, GCTA, shared haplotype
51