A Genome‐wide Association Study of non‐alcoholic fatty liver disease in India: is there divergence in the genetic risk profile?

Niraj C Doshi[1,2], Gautam Mehta[1], Jane Chalmers[3], Kondarapassery B Leena[4,5], Stuart Astbury[3], Jane I Grove[3], Andrew McQuillin[2], Marsha Y Morgan[1], Kotacherry T Shenoy[4] and Guruprasad P Aithal[3], (1)UCL Institute for Liver & Digestive Health, Division of Medicine, Royal Free Campus, University College London, United Kingdom, (2)Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, United Kingdom, (3)National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and the University of Nottingham, United Kingdom, (4)Population Health and Research Institute (PHRI), Trivandrum, India, (5)Holistic Health and Research Institute, Trivandrum, India

Background: The global pooled prevalence of non-alcoholic fatty liver disease (NAFLD) is 25% (Youssoni 2016). Genetic factors have been shown to play a significant role in determining the risk for the development and progression of NAFLD in Caucasian/East Asian populations where robustly replicated associations have been found, many at genome-wide significance, with variants in *PNPLA3*, *TM6SF2*, *MBOAT7*, *GCKR*, and *HSD17B13*(Trépo 2020). South Asians develop NAFLD and at a lower body mass index (BMI) than their Caucasian counterparts, with reported prevalences of almost 50% among adults in some areas of India (Chalmers 2019). Nevertheless, there has been no systematic attempt, to date, to undertake a genome wide association study (GWAS) of NAFLD in the South Asian community. This study aimed to redress the balance. **Methods:** The Trivandrum population-based NAFLD cohort comprises of 2222 individuals from Kerala, South India. Genomic DNA was available for 908 participants. Cases were defined as participants who, based on ultrasound scanning, had evidence of fatty infiltration of the liver (n = 454). Controls were defined as participants with no evidence of fatty infiltration (n = 454). Samples were genotyped using the Global Screening Array-24 v3.0 BeadChip (Illumina, Inc), and were imputed against the Haplotype Reference Consortium panel and the population-specific GenomeAsia pilot panel. The analysis was conducted in GMMAT to control for a degree of cryptic relatedness, with the model adjusted for age, sex, and BMI as fixed effects and family structure as a random effect by a relationship matrix calculated in LDAK. Ethical approval for the study was granted by the Ethics Committee of Sree Gokulam Medical

College and Research Foundation, India, and written informed consent was obtained from each participant. **Results:** No associations were identified with NAFLD at genome-wide significance ($p < 5 \times 10^{-8}$). However, a number of genetic variants implicated in liver-related and metabolic pathology showed suggestive evidence of association (p value range 1.93×10^{-6} to 5.83×10^{-6}) including single-nucleotide polymorphisms (SNPs) in *chromosome 20 open reading frame 78* (*C20orf78*), *Transmembrane Protein 63C* (*TMEM63C*), *Bromodomain Adjacent To Zinc Finger Domain 1A* (*BAZ1A*), *Fibroblast Growth Factor 21* (*FGF21*), and *Insulin Receptor Substrate 1* (*IRS1*). Variants in previously associated loci viz *PNPLA3, MBOAT7,TM6SF2, GCKR* and *HSD17B13*, showed much lower significance than expected (p value range 1.20×10^{-3} to 0.730). **Conclusion:** These findings suggest that there may be divergence in the genetic risk profile for developing NAFLD in the South Asian population compared with European/East Asian populations. This is concordant with the observed divergence in NAFLD phenotype, but needs to be validated in further, larger cohorts.

ND, GM, JC share first authorship. MM, KS, GA share senior authorship.