

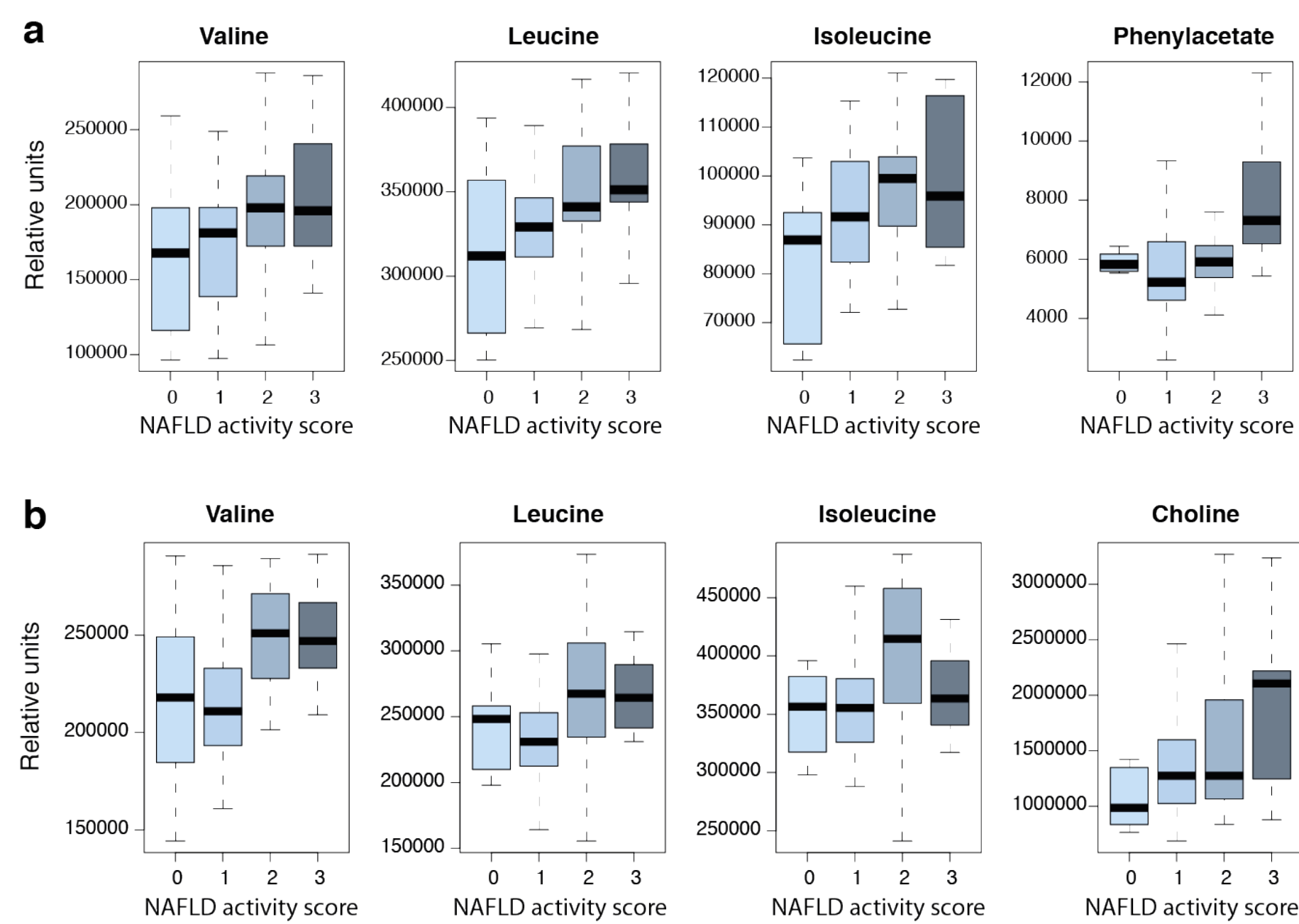
Integrated systems biology to study non-alcoholic fatty liver disease in obese women

Lesley Hoyles^{1*}, José-Manuel Fernández-Real^{2*}, Massimo Federici^{3*}, Matteo Serino^{4,5}, Vincent Azalbert^{4,5}, Vincent Blasco^{4,5}, James Abbott¹, Richard H. Barton¹, Josep Puig², Gemma Xifra², Wifredo Ricart², Mark Woodbridge¹, Christopher Tomlinson¹, Marina Cardellini³, Francesca Davato³, Iris Cardolini³, Ottavia Porzio⁶, Paolo Gentilieschi⁶, Frédéric Lopez^{4,5}, Fabienne Fougelle⁷, Catherine Postic⁸, Sarah A. Butcher¹, Elaine Holmes¹, Jeremy K. Nicholson¹, Rémy Burcelin^{4,5}, Marc-Emmanuel Dumas¹

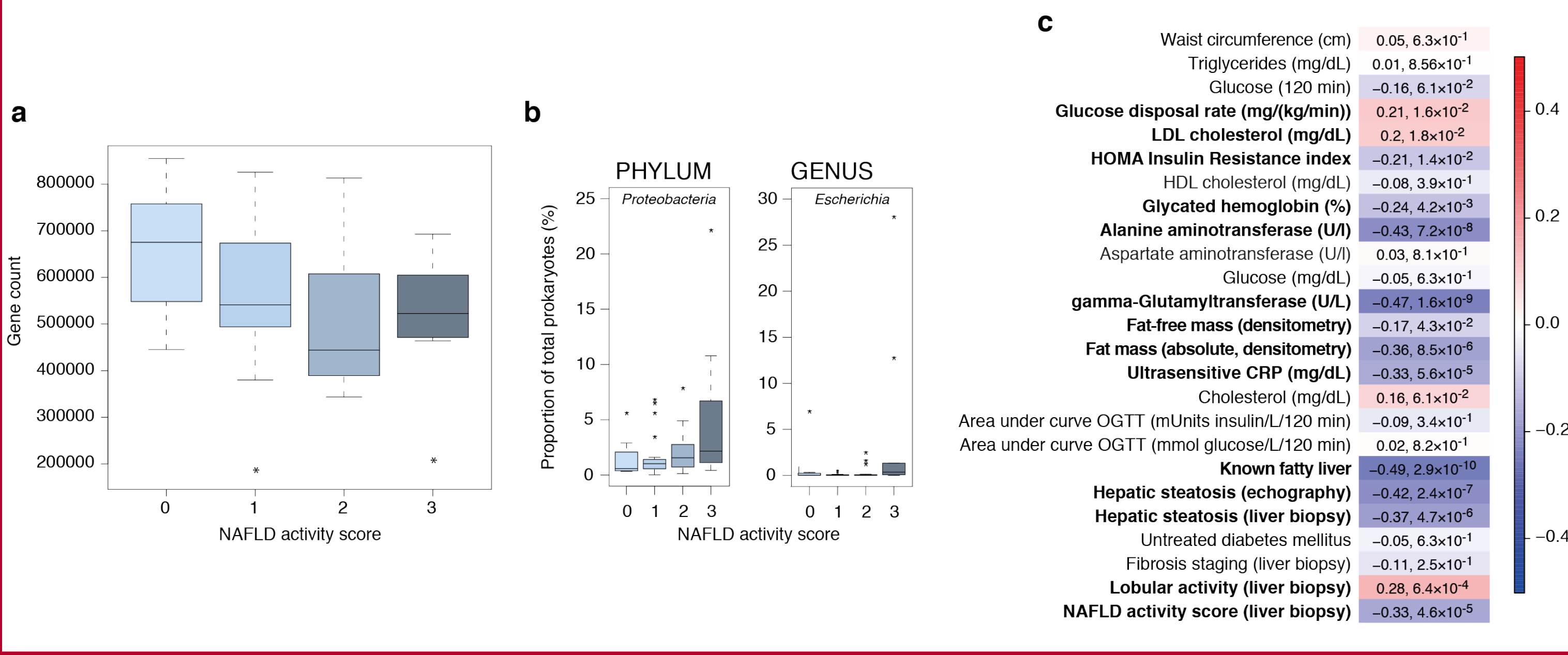
¹Surgery and Cancer, MRC-PHE Centre for Environment and Health, Imperial College London; ²Department of Endocrinology, Diabetes and Nutrition, Universitat de Girona, Spain; ³Department of Systems Medicine, University of Rome Tor Vergata, Italy; ⁴Institut National de la Santé et de la Recherche Médicale (INSERM), Toulouse, France; ⁵Université Paul Sabatier (UPS), Institut des Maladies Métaboliques et Cardiovasculaires (I2MC), France; ⁶Department of Experimental Medicine and Surgery, University of Rome Tor Vergata, Rome; ⁷Sorbonne Universités, UPMC Univ Paris 06, UMRs 1138, Centre de Recherche des Cordeliers, Paris, France

Faecal, liver biopsy, blood and urine samples and data for 28 clinical variables were collected for 56 obese (BMI >35 kg/m²) women from Italy ($n = 31$) and Spain ($n = 25$) who elected for bariatric surgery. Confounder analyses of clinical data were done using linear modeling. Histological examination of liver biopsies was used to grade NAFLD (NAFLD activity score: 0, 1, 2, 3). Faecal metagenomes were generated and analysed using the Scalable Automated Metagenomics Pipeline (ScAMP)¹. Differentially expressed genes were identified in hepatic transcriptomes using limma², and analysed using Enrichr³, network analyses and Signaling Pathway Impact Analysis (SPIA)⁴. ¹H-NMR data were generated for plasma and urinary metabolomes⁵. Clinical, metagenomic, transcriptomic and metabolomic data were integrated using partial Spearman's correlation, taking confounders (age, BMI and cohort) into account. Interactions between the faecal microbiome and the clinical/molecular phenome were quantified, with the resulting Receiver Operating Characteristic (ROC) curve confirming the diagnostic power of molecular phenomic and metagenomic indices in NAFLD.

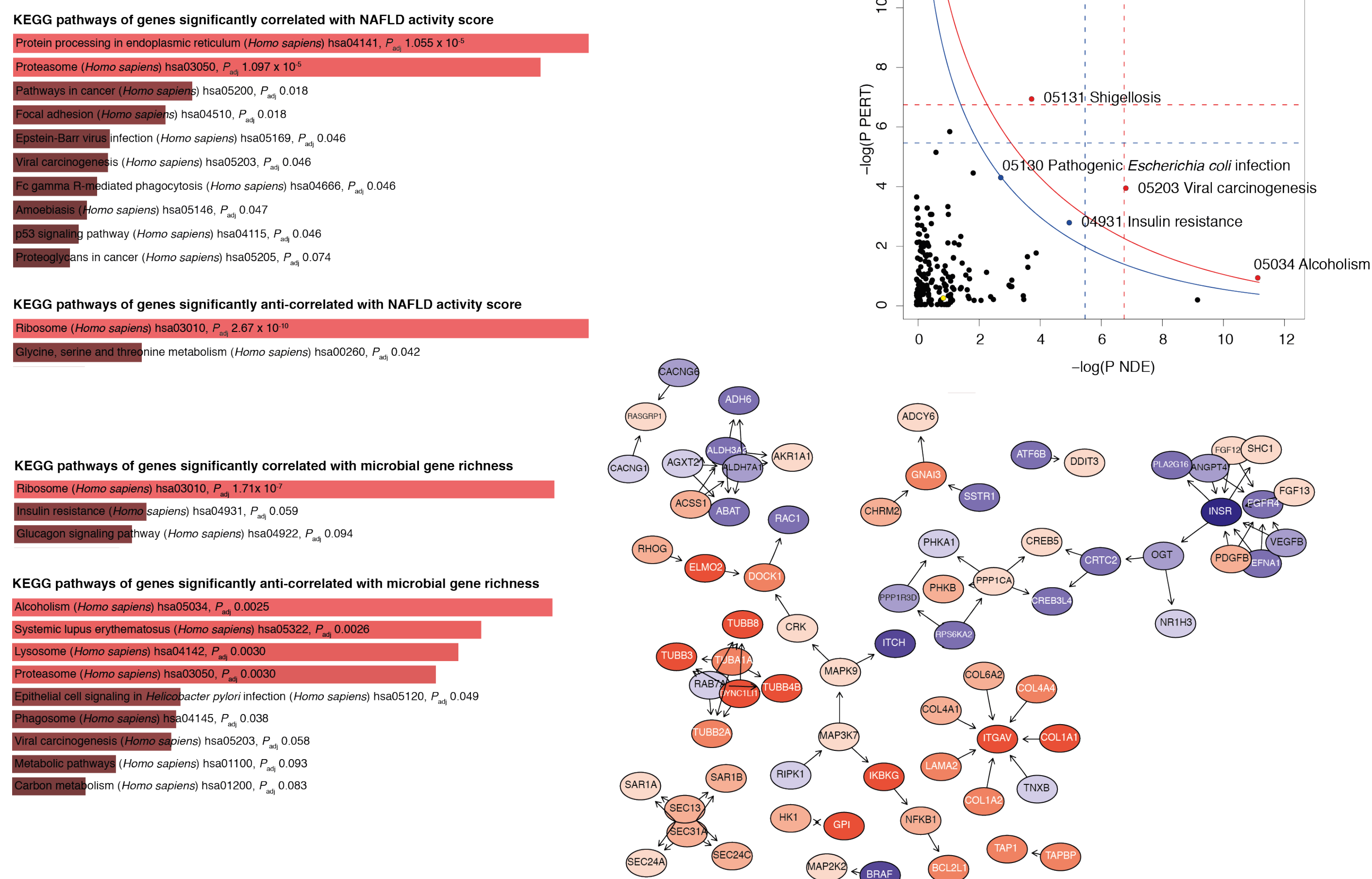
(a) Plasma and (b) urine metabolomic profiles highlighted imbalances in branched-chain amino acid metabolism, gut-derived microbial metabolites resulting from metabolism of amino acids and choline metabolism. There was no significant correlation of TMAO and TMA with NAFLD. Instead our data suggest choline is excreted rather than used by the gut microbiota.



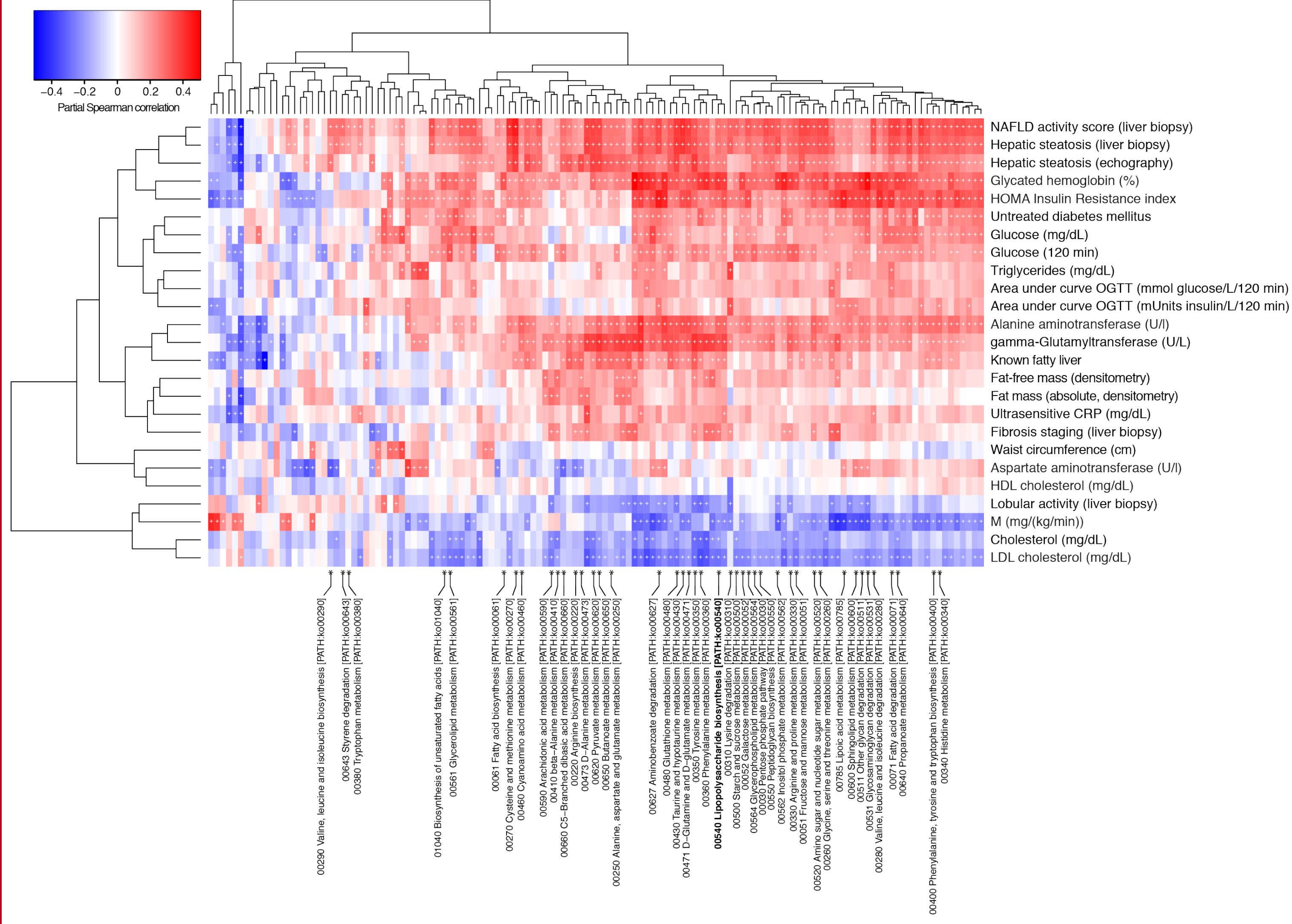
NAFLD activity score was (a) anti-correlated with microbial gene richness, and (b) correlated with abundance of *Proteobacteria*. (c) Microbial gene richness was correlated with clinical markers of NAFLD.



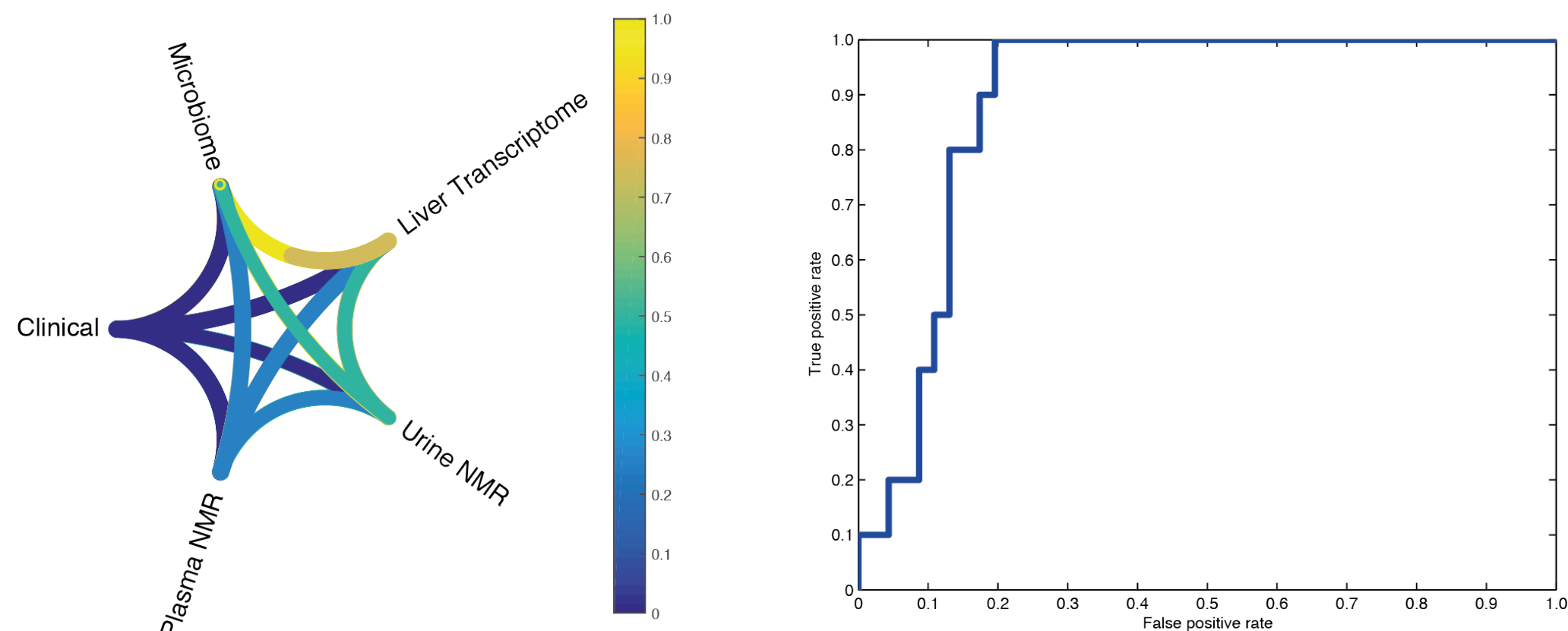
NAFLD-associated hepatic transcriptomes were associated with branched-chain amino acid metabolism, endoplasmic reticulum/phagosome. Hepatic genes significantly correlated with NAFLD activity score AND microbial gene richness were significantly associated with immune responses linked to non-specific microbial infections and insulin resistance, and the most connected gene was *INSR* (insulin receptor).



KEGG analyses of metagenomic data suggested increased microbial processing of dietary lipids and amino acids, as well as endotoxin-related (LPS) processes related to *Proteobacteria*.



There is a close association between the gut microbiome, host metabolome, hepatic steatosis, and clinical and molecular insulin resistance in morbid obesity; and interplay between the microbiome and host gene expression in inflammation and host metabolism.



*These authors made equal contributions to the work.

Acknowledgements

This work was supported by EU-FP7 FLORINASH (Health-F2-2009-241913) to RB, MF, JMFR, FF, CP, EH and JKN. This work used the computing resources of the UK MEDical BIOinformatics partnership – aggregation, integration, visualization and analysis of large, complex data (UK MED-BIO), which is supported by the Medical Research Council (grant number MR/L01632X/1). LH is in receipt of an MRC Intermediate Research Fellowship in Data Science (grant number MR/L01632X/1). This work was also supported by funding to RB (Region 2009-2014; Agence Nationale de la Recherche ANR Bactimmunodia & GAD), and to JMFR (CIBERobn Pathophysiology of Obesity and Nutrition and FEDER funds).



Contact
Lesley Hoyles
lesley.hoyles11@imperial.ac.uk

References

- [1] Abbott *et al.* (2017). In preparation.
- [2] Ritchie *et al.* (2015). *Nucleic Acids Res* **43**, e47.
- [3] Chen *et al.* (2013). *BMC Bioinformatics* **14**, 128.
- [4] Tarca *et al.* (2009). *Bioinformatics* **25**, 75–82.
- [5] Dumas *et al.* (2005). *Proc Natl Acad Sci USA* **103**, 12511–12516.