Genome-wide association study of inflammatory proteins in ALSPAC Mothers and offspring

Neil Goulding

University of Bristol, UK

26th September 2024

University of BRISTOL

JOHN TEMPLETON FOUNDATION

Inspiring Awe & Wonder

CHILDREN OF THE 90S

WCE

WORLD CONGRESS OF EPIDEMIOLOGY 2024

Background

- pQTL - protein Quantitative Trait Loci, which are robust connections between a gene variant and the levels of a protein
- 2 main proteomic platforms – Olink and Somalogic

somalogic

Olink

- UKBB – 3,000 proteins; 54,219 participants (Olink)

Plasma proteomic associations with genetics and health in the UK Biobank

https://doi.org/10.1038/s41586-023-06592-6
Received: 17 June 2022
Accepted: 31 August 2023
Published online: 4 October 2023
Open access
Check for updates

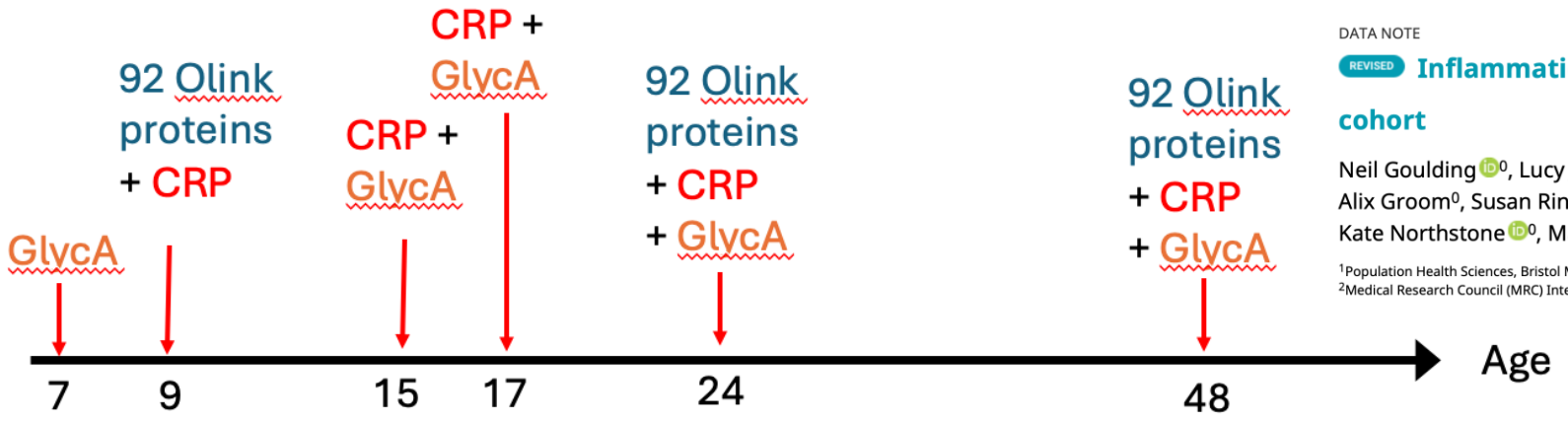
Benjamin B. Sun[1], Joshua Chiou[2,26], Matthew Traylor[3,26], Christian Benner[4,26], Yi-Hsiang Hsu[5,26], Tom G. Richardson[3,6,26], Praveen Surendran[6,26], Anubha Mahajan[4,26], Chloe Robins[7,26], Steven G. Vasquez-Grinnell[8,26], Liping Hou[9,26], Erika M. Kvikstad[8,26], Oliver S. Burren[10], Jonathan Davitte[7], Kyle L. Ferber[11], Christopher E. Gillies[12], Åsa K. Hedman[13], Sile Hu[3], Tinchi Lin[14], Rajesh Mikkilineni[15], Rion K. Pendergrass[4], Corran Pickering[16], Bram Prins[10], Denis Baird[1], Chia-Yen Chen[1], Lucas D. Ward[17], Aimee M. Deaton[17], Samantha Welsh[18], Carissa M. Willis[17], Nick Lehner[18], Matthias Arnold[18,19], Maria A. Wörheide[19], Karsten Suhre[20], Gabi Kastenmüller[18], Anurag Sethi[21], Madeleine Cule[21], Anil Raj[21], Alnylam Human Genetics*, AstraZeneca Genomics Initiative*, Biogen Biobank Team*, Bristol Myers Squibb*, Genentech Human Genetics*, GlaxoSmithKline Genomic Sciences*, Pfizer Integrative Biology*, Population Analytics of Janssen Data Sciences*, Regeneron Genetics Center*, Lucy Burkitt-Gray[16], Eugene Melamud[21], Mary Helen Black[9], Eric B. Fauman[2], Joanna M. M. Howson[3], Hyun Min Kang[12], Mark I. McCarthy[4], Paul Nioi[17], Slavé Petrovski[10,22], Robert A. Scott[8], Erin N. Smith[23], Sándor Szalma[23], Dawn M. Waterworth[24], Lyndon J. Mitnaul[12], Joseph D. Szustakowski[8,27], Bradford W. Gibson[5,27], Melissa R. Miller[2,27] & Christopher D. Whelan[1,25,27]

Most studies have just looked at one timepoint in adults

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

ALSPAC

- Avon Longitudinal Study of Parents and Children – also known as Children of the 90s
- Genetics - dataset contains > 8 million SNPs
- Proteomics - 3000 mothers, 3000 9-year-olds and 3000 24-year-olds
- 92 inflammatory proteins - Olink target 96 inflammatory panel



Wellcome Open Research
Wellcome Open Research 2024, 7:277 Last updated: 06 FEB 2024

Check for updates

DATA NOTE

REVISED **Inflammation proteomics datasets in the ALSPAC cohort**

Neil Goulding[0], Lucy J. Goudswaard[0], David A. Hughes[0], Laura J. Corbin[0], Alix Groom[0], Susan Ring[0], Nicholas J. Timpson[0], Abigail Fraser[0], Kate Northstone[0], Matthew Suderman[0]

[1]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[2]Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, Bristol, UK

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Aims

1. Identify pQTL across the lifecourse (ages 9, 24 and 48)
2. Which pQTL are robust across the lifecourse?
3. What proportion of UK Biobank pQTLs replicate at each of the three ALSPAC timepoints?
4. Investigate the relationship of pQTL with downstream phenotypes (e.g. asthma) at each timepoint

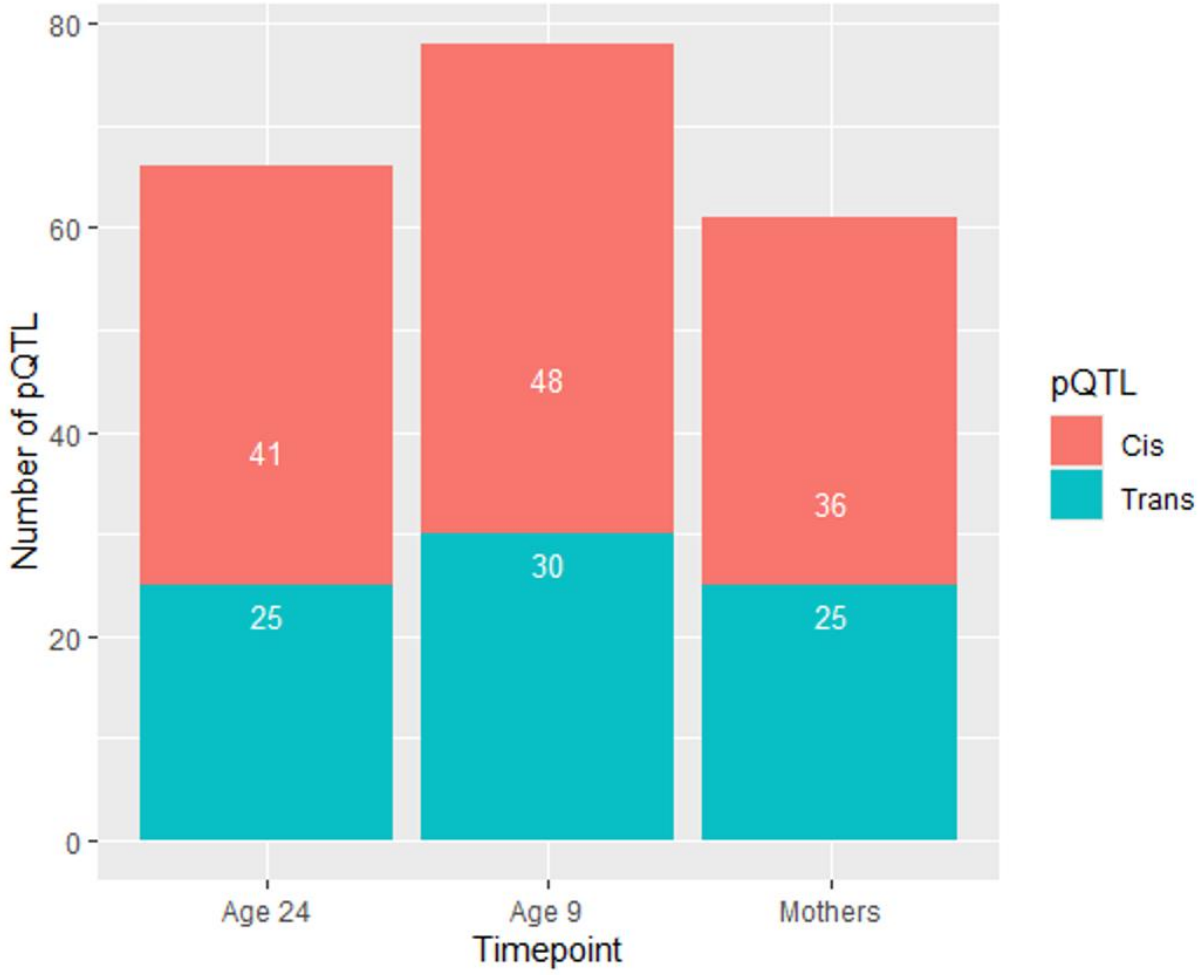
- **This talk will focus on aims 1 and 2**

Methods

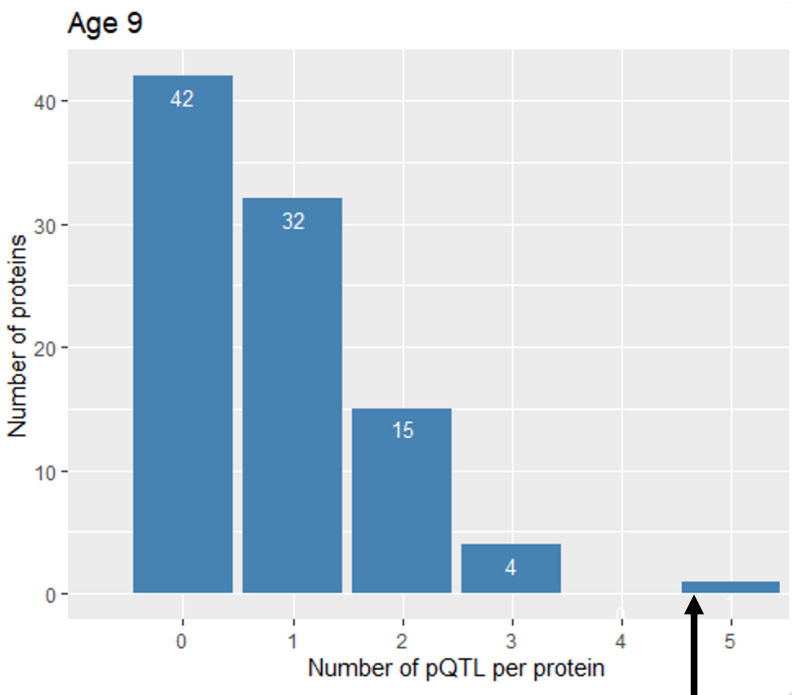
- Genome-wide testing performed using PLINK 2.3, fitting linear regression models between protein levels and 1000 Genomes imputed SNPs.
- Adjusted for age, sex and the top 10 genetic principal components.
- Extracted pQTLs with $p < 5.3 \times 10^{-10}$, which is the threshold of genome-wide significance (5×10^{-8}) adjusted for multiple testing (94 proteins).
- For each of the proteins there is a protein-coding gene. We defined associations with independent SNPs within 1 million base pairs (1 Mb) window of the gene boundaries of the protein-coding gene as *cis*-signals, and otherwise in *trans* (***except GlycA***)
- Clumped $\pm$ 1MB around leading variants using PLINK

Results – total number of pQTLs

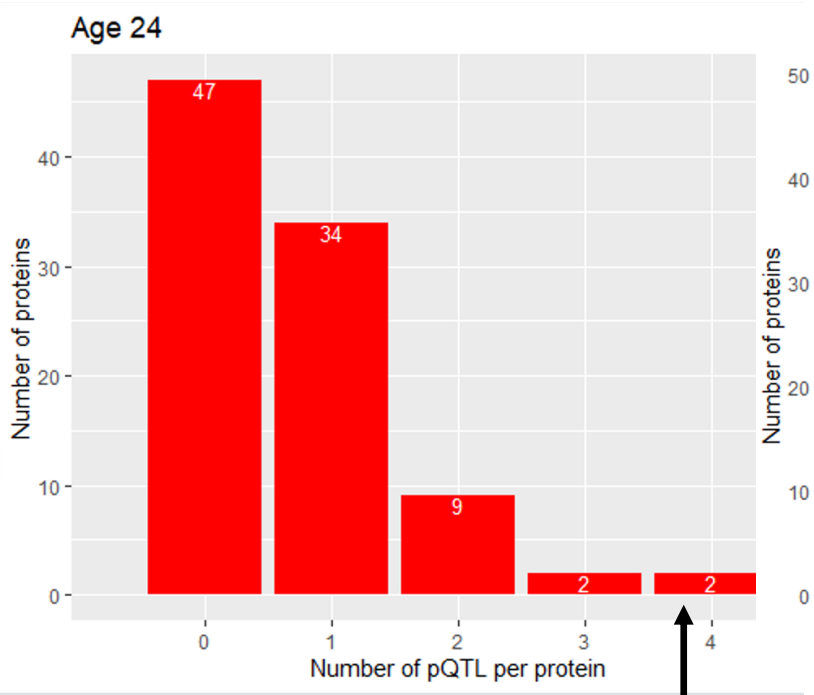
- After clumping – 205 pQTL in total across all 3 timepoints (61% cis)
- Most pQTLs at age 9
- 59%-61% cis-pQTL at each timepoint
- N = 2426 (age 9)
- N = 2170 (age 24)
- N = 2120 (mothers)



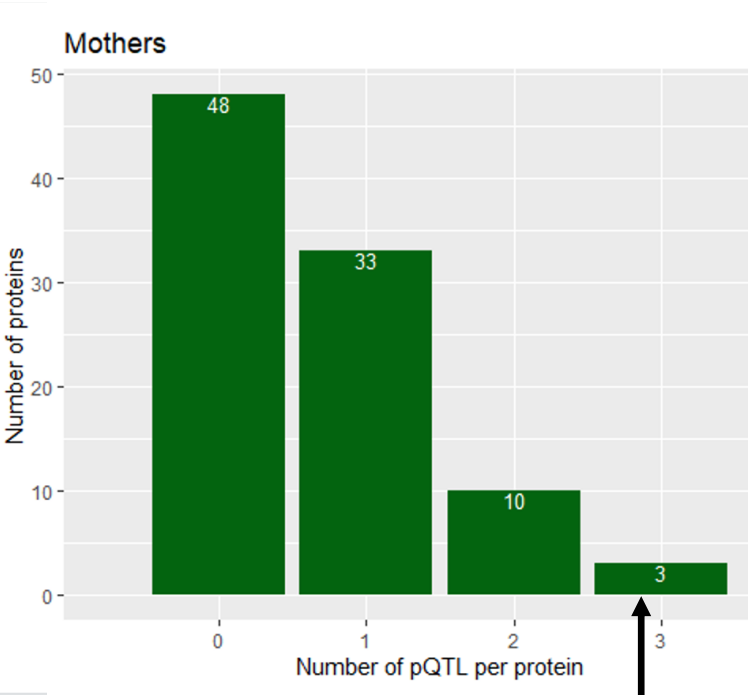
Number of pQTLs per protein



CCL25



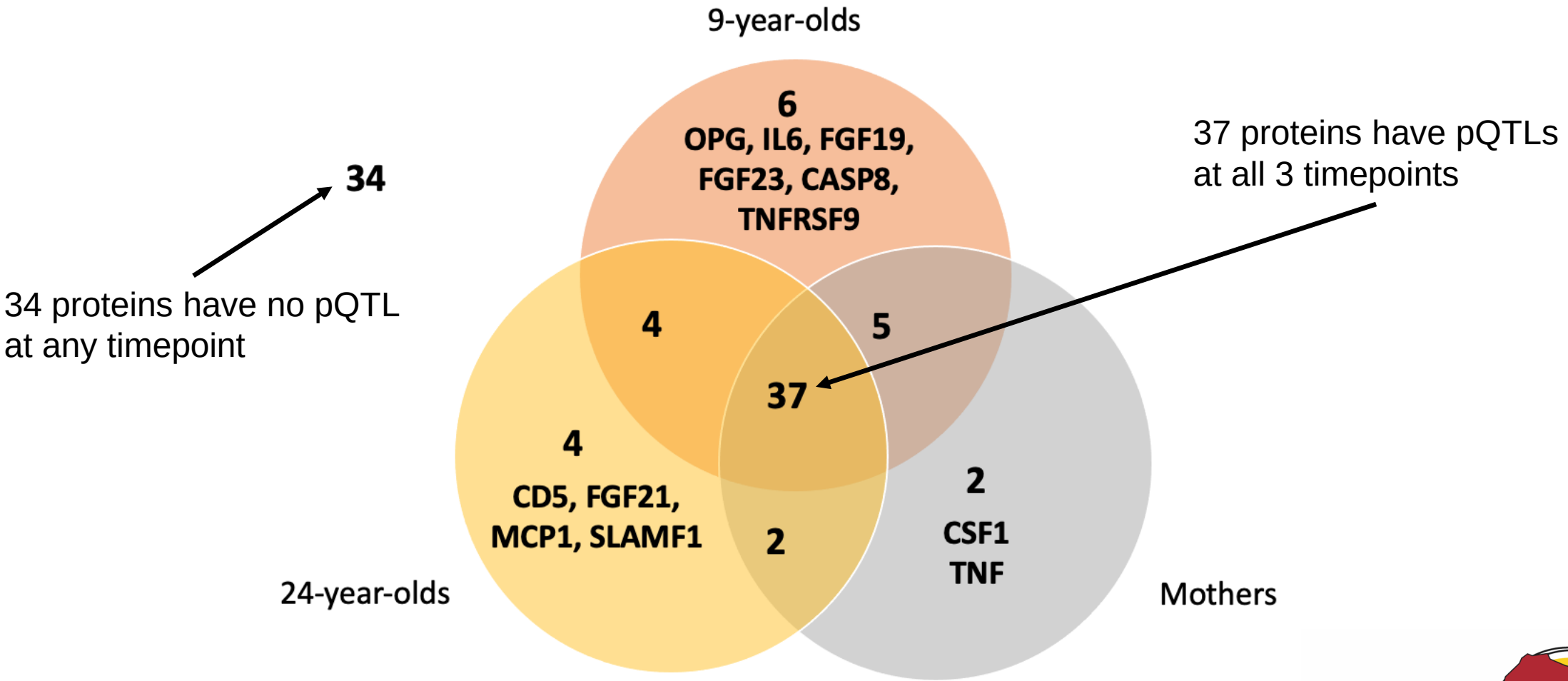
CCL25, MMP1



CCL25, CD8A, MCP4

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Results 3



Venn diagram of proteins that have pQTLs at each timepoint

Conclusions and Further Work

- Many of the proteins have pQTL that are robust across the lifecourse
- We have also found pQTL unique to each time point
- Is genetic architecture across development more complicated than initially thought?
- Next step - replication in UKBB
- Investigate the relationship of pQTL with downstream phenotypes (e.g. asthma) at each timepoint

Acknowledgements

- Collaborators: Matt Suderman (Bristol), Tom Richardson (Bristol and GSK) and Daisy Crick (Melbourne)

CHILDREN OF THE 90S

University of BRISTOL

JOHN TEMPLETON FOUNDATION

Inspiring Awe & Wonder

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024