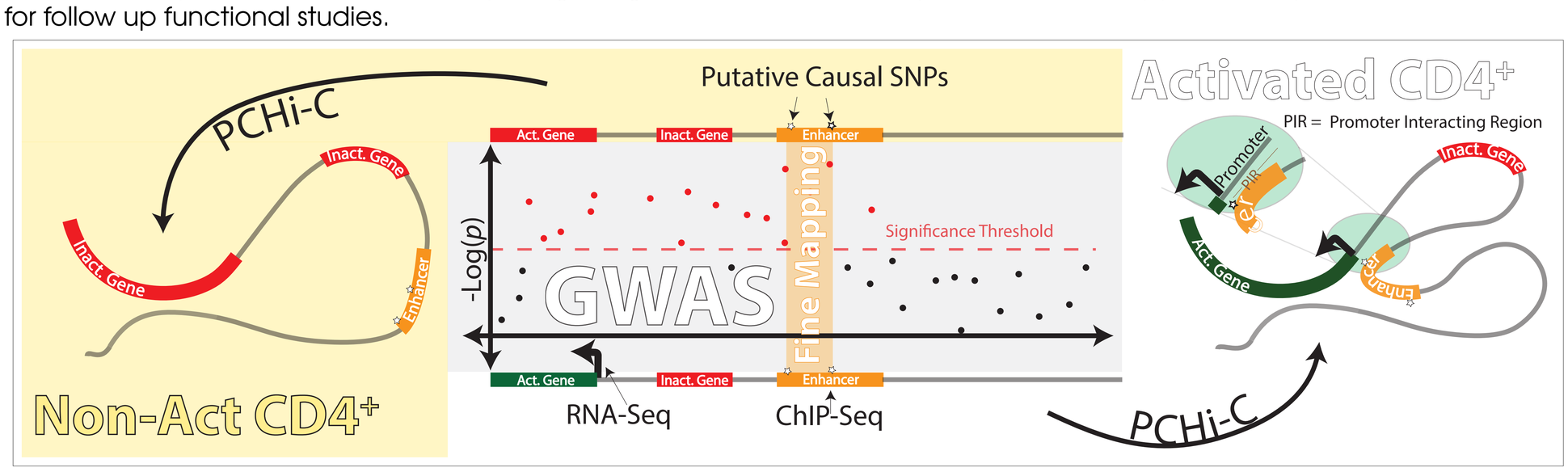
Chromosome contacts in activated T cells identify autoimmune disease-candidate genes.

Oliver S Burren[1,8], Arcadio Rubio García[1], Biola-Maria Javierre[2], Daniel B Rainbow[1], Jonathan Cairns[2], Nicholas J Cooper[1], John J Lambourne[3,4], Ellen Schofield[1], Xaquin Castro Dopico[1], Ricardo C Ferreira[1], Richard Coulson[1], Frances Burden[3,4], Sophia P Rowlston[3,4], Kate Downes[3,4] Steven W Wingett[2], Mattia Frontini[3,4,5], Willem H Ouwehand[3,4,5,6], Peter Fraser[2], Mikhail Spivakov[2], John A Todd[1], Linda S Wicker[1], Antony J Cutler[1], Chris Wallace[1,7,8].

[1]JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, University of Cambridge, UK; [2]The Babraham Institute, Cambridge, UK; [3]Department of Haematology, University of Cambridge, Cambridge, UK; [4]NHS Blood and Transplant, NIHR Cambridge Biomedical Campus, Cambridge, UK; [5]British Heart Foundation Centre of Excellence, University of Cambridge, UK; [6]Wellcome Trust Sanger Institute, Hinxton, UK; [7]MRC Biostatistics Unit, Cambridge, UK; [8]Department of Medicine, University of Cambridge, UK. **Contact: ob219@cam.ac.uk**

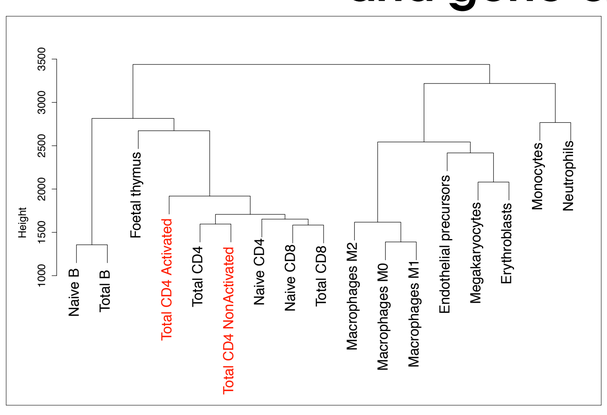
Promoter capture Hi-C (PCHi-C) can link putative causal variants to genes

Genome-wide association studies tell us about variation and phenotype. PCHi-C can indicate which genomic regions physically interact with gene promoters in specific tissue contexts. By integrating these we can develop a data-driven approach to prioritise causal candidate genes and tissue contexts for follow up functional studies.

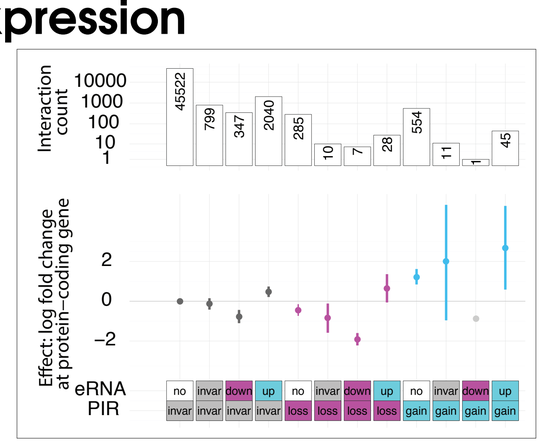


As part of a larger study encompassing 17 primary human haematopoietic tissues, we used a combination of PCHi-C, total RNA-Seq and CHiP-Seq to examine the effect of activation on human CD4+ T cells at a 4 hour time point. We prioritised genes and tissue contexts for functional validation by integrating these data with summary GWAS statisitcs for 11 autoimmune traits.

Chromosome interactions influence lineage fate and gene expression

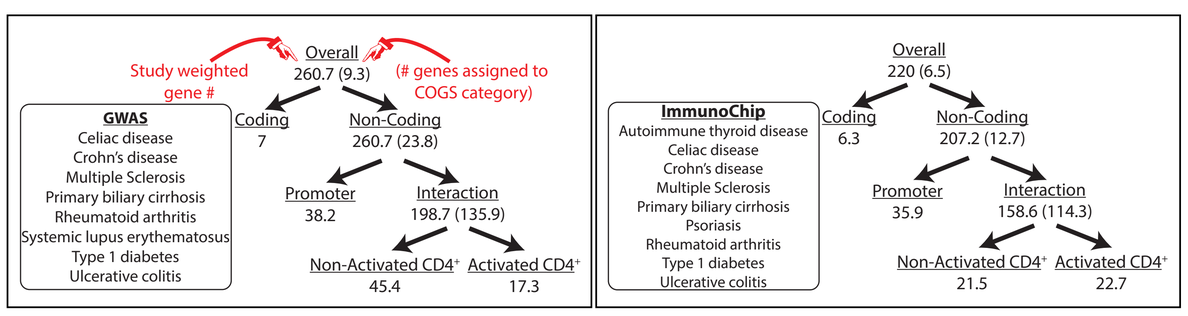


Clustering of chromatin interactions for 17 primary human tissues identified by PCHi-C reacapitulates the haematopoietic lineage tree.



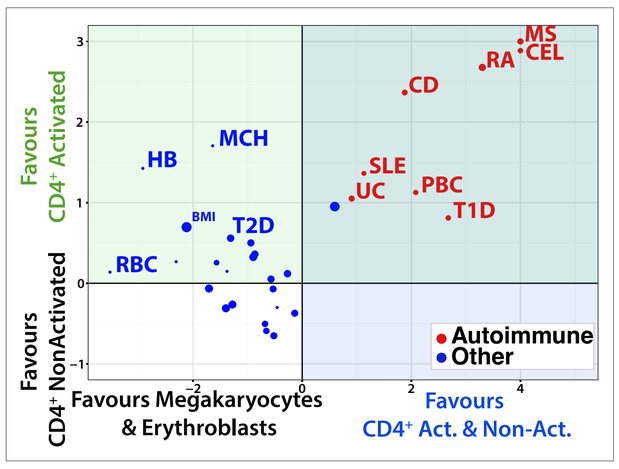
PCHi-C target gene expression correlates synergistically with promoter interaction region (PIR) gain and enhancer RNA (eRNA) expression on activation.

PCHi-C assisted gene prioritisation across 11 autoimmune traits



Functional gene prioritisation across 11 autoimmune diseases using genome wide (GWAS) or targeted genotyping array (ImmunoChip) data for autoimmune disease loci curated in `http://www.immunobase.org`

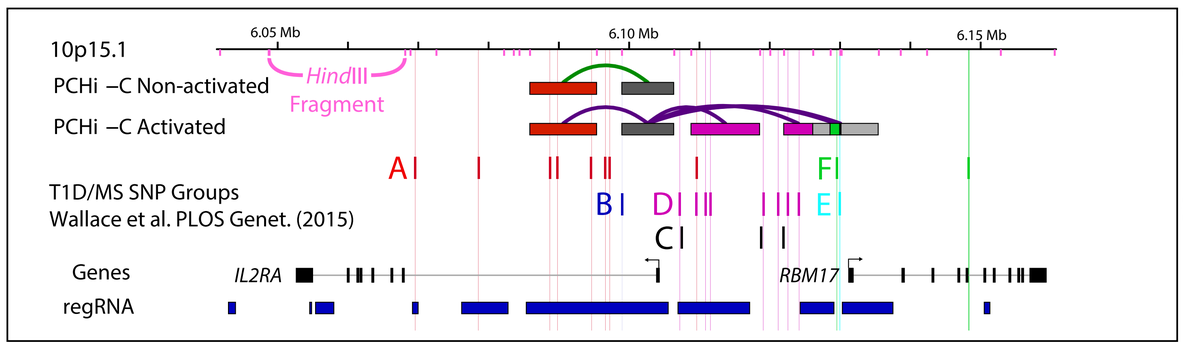
Autoimmune GWAS signals are enriched in activated CD4+ T Cell PIRs



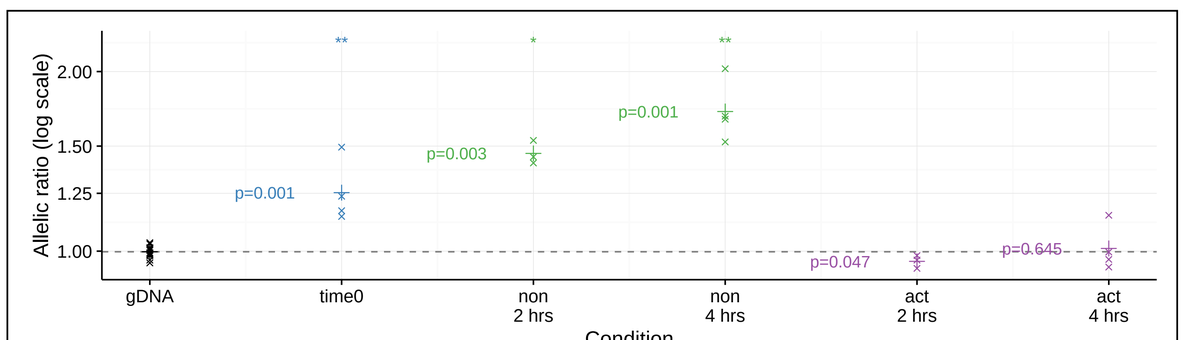
Inspired by GOSHIFTER(2), we developed *blockshifter*, a method to examine the enrichment of GWAS summary statistics between tissue specific PIRs in the presence of correlation. We examined summary statistics for 31 GWAS traits and found autoimmune traits are most strongly enriched in activated CD4+ T cell PIRs

Interaction-mediated regulation of *IL2RA* expression

COGS analysis prioritised *IL2RA* in multiple autoimmune diseases. We previously fine mapped five signals (A-F) in type 1 diabetes (T1D) and multiple sclerosis (MS)(2). PCHi-C data shows constitutive interactions between the promoter of *IL2RA* and SNP set 'A' regardless of tissue state. On activation other interactions between downstream regions overlapping SNP sets 'D-F' form.

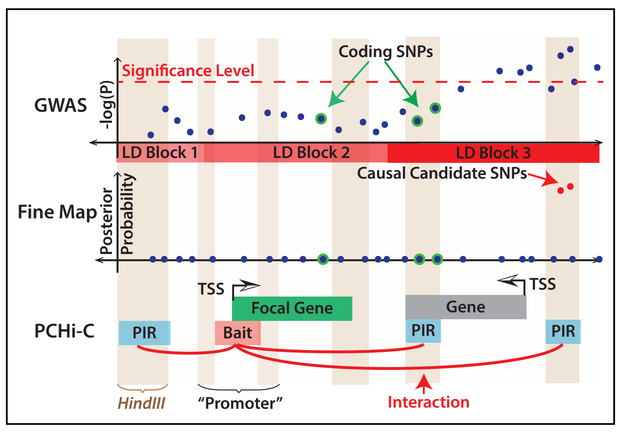


Allelically imbalanced transcription of *IL2RA* within CD4+ T cells of individuals heterozygous for PCHi- C linked autoimmne-associated variants is specific to non-activated cells.



Black indicates genomic DNA, which we expect to be balanced (1:1), blue indicates cDNA collected at time zero and green and purple indicate cDNA allele read counts collected under non-activated and activated conditions

Integrating PCHi-C with GWAS summary statistics



Capture Hi-C **O**mnibus **G**ene **S**core (COGS) is a Bayesian method for integrating PCHi-C data and GWAS to prioritise gene and tissue contexts according to SNP overlap with :

1. Coding Regions.
2. "Promoter" regions/short range interactions refractive to PCHi-C interrogation.
3. Constitutive and tissue specific PIRs.

References

(1) Trynka et al. (2015) *Am. J. Hum. Genet.*, **(2)** Wallace et al. (2015) *PLOS Genet.*.

To browse data visit http://www.chicp.org