

# Unravelling higher order chromatin organisation through statistical analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Ben Moore

19<sup>th</sup> November 2015

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

## Introduction

# Unravelling higher order chromatin organisation through statistical analysis



INSTITUTE OF GENETICS

Broad aim: investigate the relationship between structure and function of the genome

## Some specific questions:

1. How does higher order chromatin structure compare across human cell types?
  2. Can we predict higher order chromatin structure from locus-level features?
  3. How do the characteristics of boundaries demarcating higher order domains vary between cell types and domain classes?

# What's known about genome structure

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



Introduction

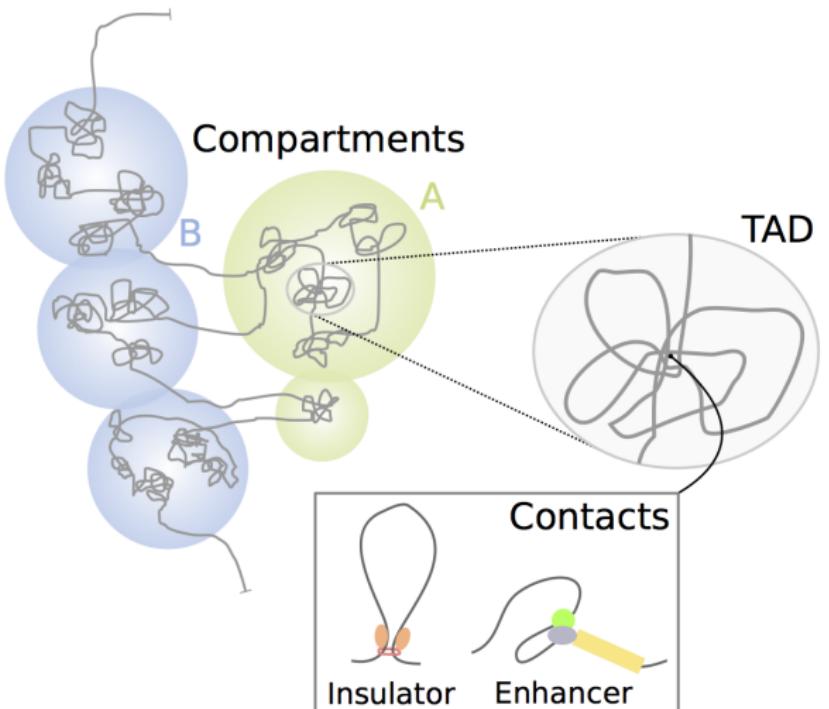
Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements



# Hi-C: a genome-wide chromosome conformation assay

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

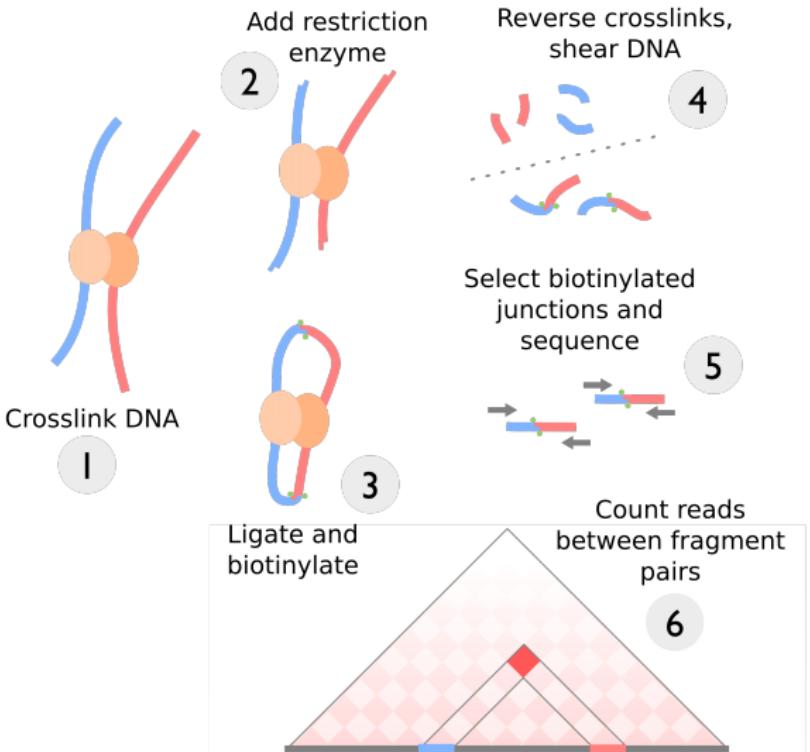
Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements



# Chromosome compartments from Hi-C data

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

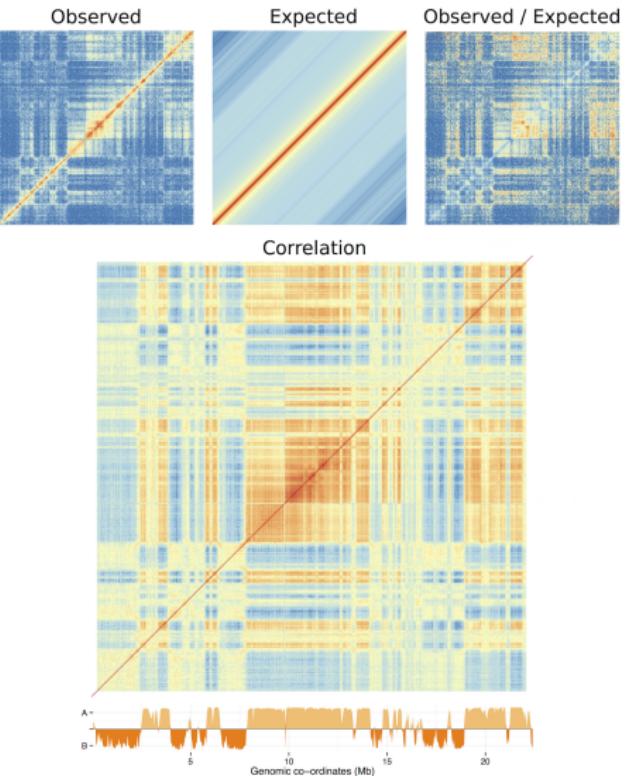
Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements



Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

# Strategy

- ▶ Integrate publicly available Hi-C data
- ▶ Uniformly reprocess each dataset
- ▶ Call compartments, TADs from reprocessed data
- ▶ Integrate cell-matched ENCODE epigenomic data

Then:

1. Compare/contrast cell types after reprocessing
2. Attempt predictive models of compartments and TADs from epigenomic features
3. Analyse boundary composition in terms of epigenomic features

Related collaborative work:

4. Investigate concept of "metaTADs"
5. Analyse conformation changes at specific locus of interest

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

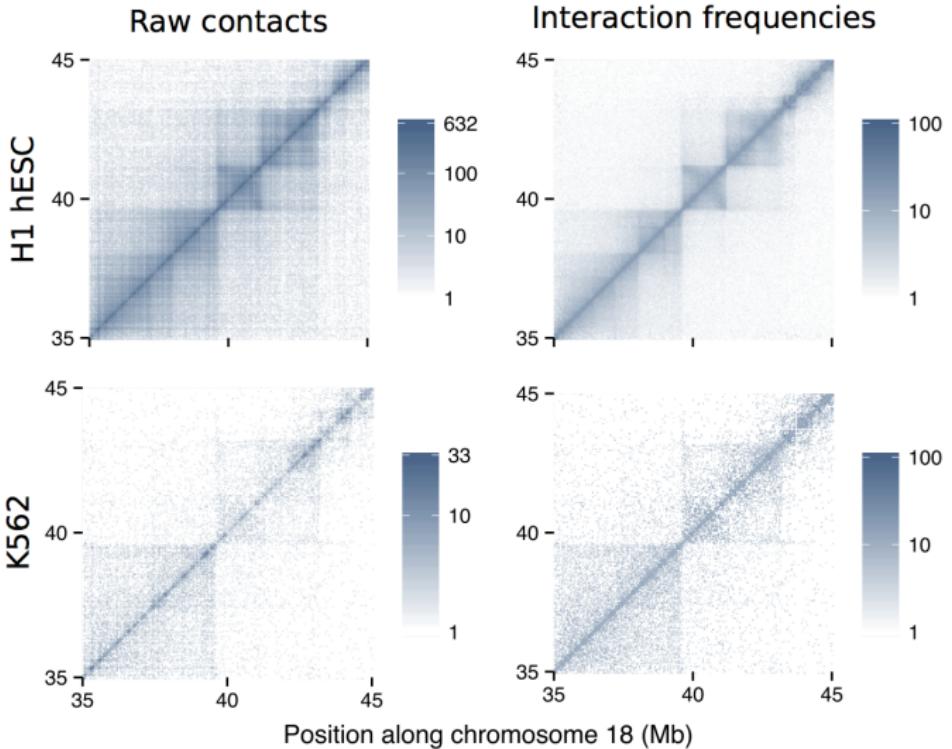
Next project

Thesis plan

Acknowledgements

# Reanalysis of Hi-C datasets

Very different sequencing depths between the input datasets:



# Reanalysis of Hi-C datasets

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

Reanalysis of Hi-C  
datasets

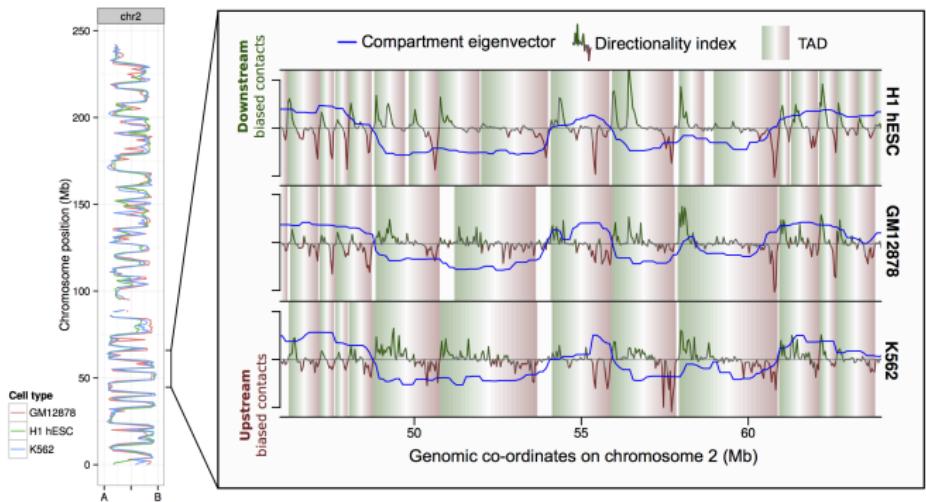
Collaborations

Next project

Thesis plan

Acknowledgements

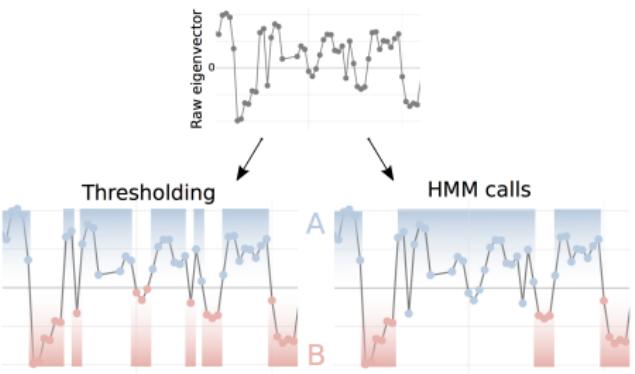
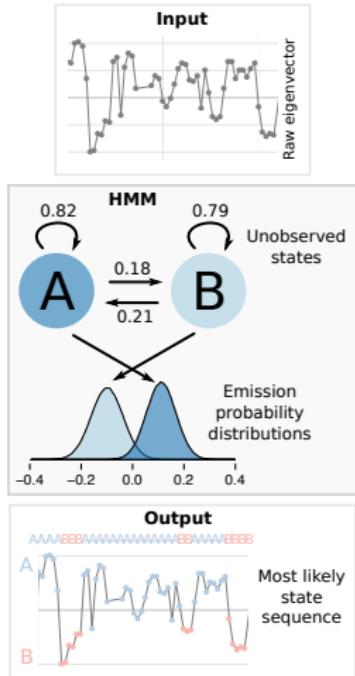
Despite this, reprocessed Hi-C data is well-correlated:



Justifies going forward with between cell-line analysis

# Improved compartment calling algorithm

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



igmm

INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

## Regions of variable compartment structure

# Unravelling higher order chromatin organisation through statistical analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

## Introduction

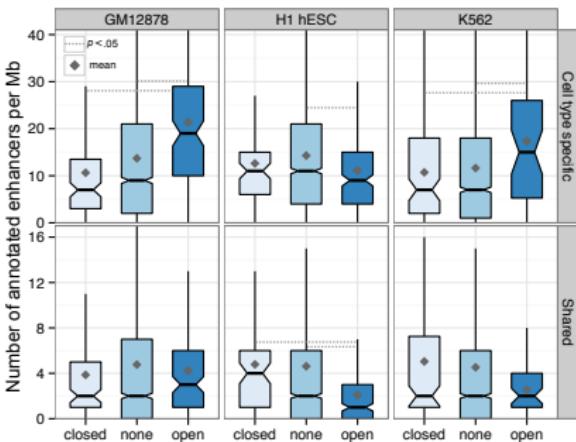
## Reanalysis of Hi-C datasets

## Collaborations

Next project

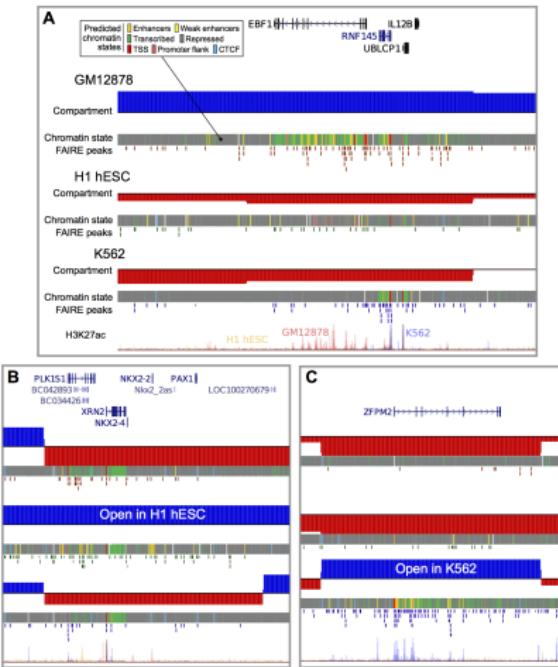
## Thesis plan

#### Acknowledgement



# Regions of variable compartment structure

Some harbour genes with known cell type specific function



igmm

INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

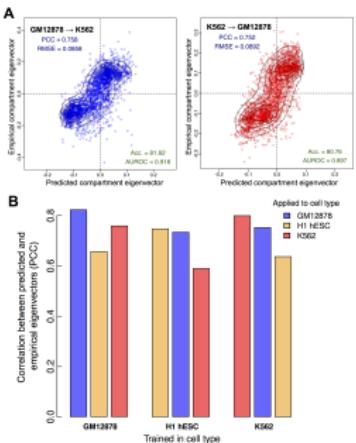
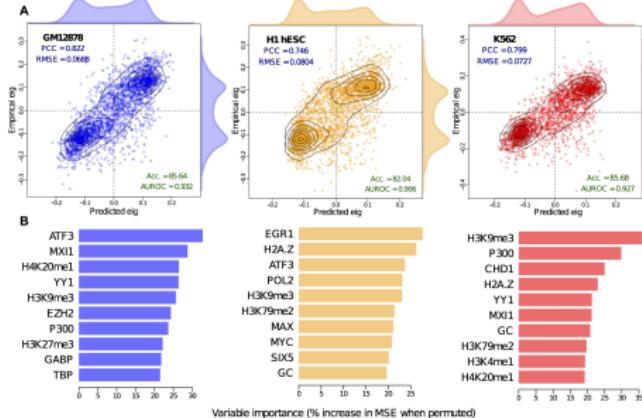
# Main project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Can model compartments, cross-apply between cell types:



Introduction

Reanalysis of Hi-C datasets

Collaborations

Next project

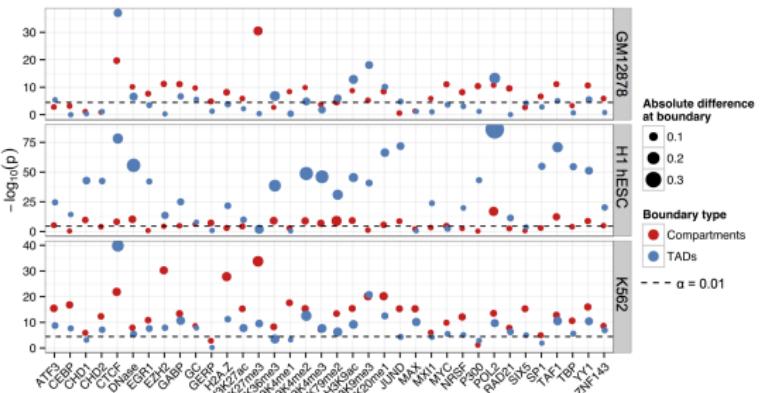
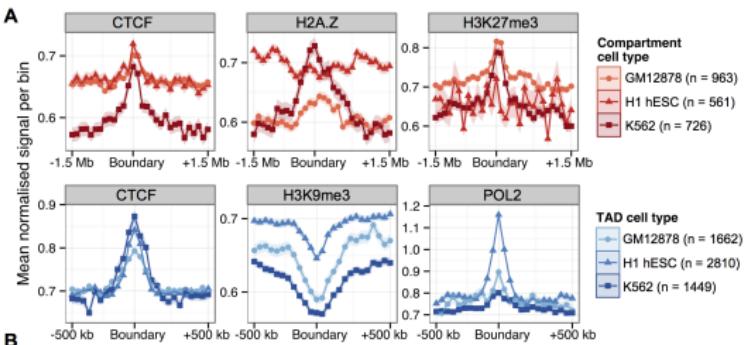
Thesis plan

Acknowledgements

## Main project

Lots of new boundary enrichments, quantitatively tested:

# Unravelling higher order chromatin organisation through statistical analysis



## Main project

## Unravelling higher order chromatin organisation through statistical analysis



## Currently in review:

RESEARCH

# Integrative modelling reveals the principles of multi-scale chromatin boundary formation in human nuclear organization

Benjamin L Moore, Stuart Aitken and Colin A Semple\*

## Introduction

## Reanalysis of Hi-C datasets

### **Collaborations**

[Next project](#)

## Thesis plan

#### Acknowledgements

# Side project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis

igmm

INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

Reanalysis of Hi-C  
datasets

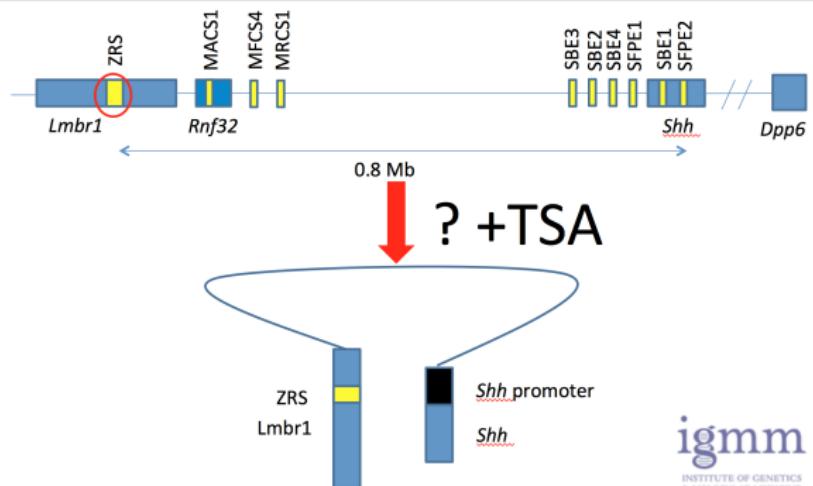
Collaborations

Next project

Thesis plan

Acknowledgements

## Does TSA Treatment Alter Chromatin Dynamics?



igmm  
INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

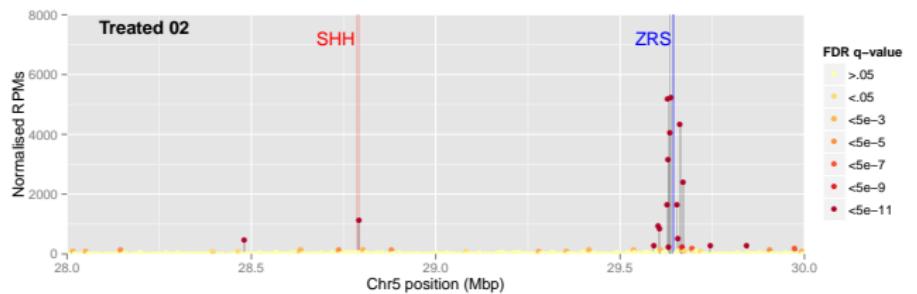
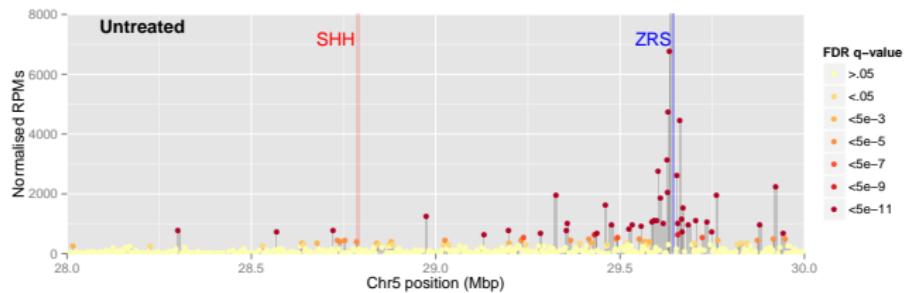
# Side project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis

igmm

INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

## Initial results:



Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

# Side project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

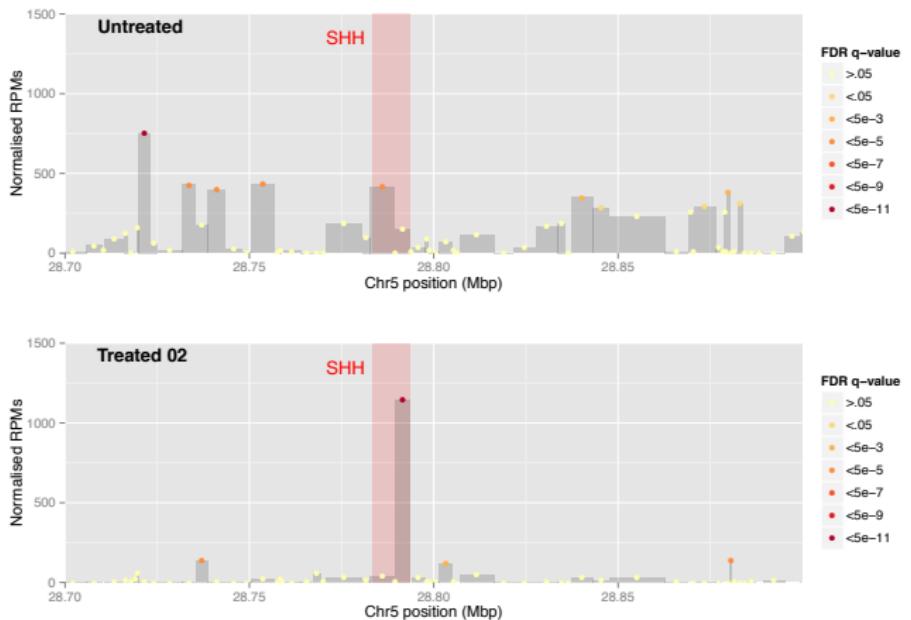
Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

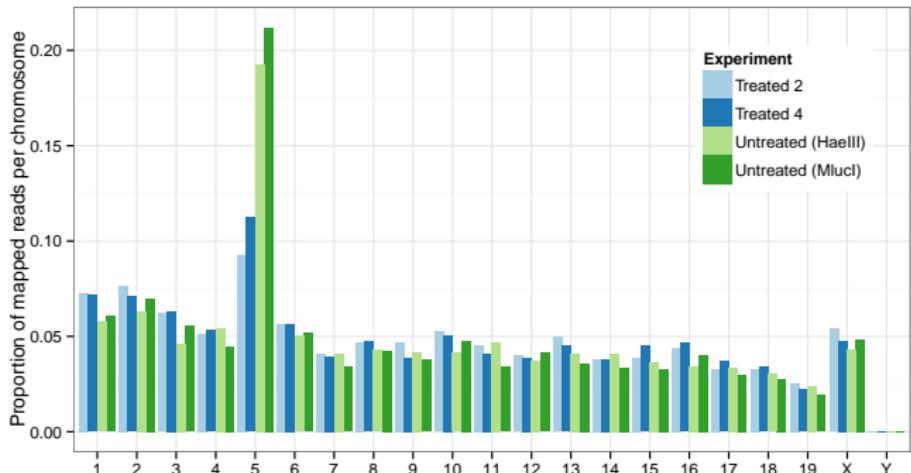


Incoming: repeats, 5C / Capture-C (?), FAIRE-seq ...

# Side project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis

All looks good except . . .



Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

# Next project

So far haven't looked at any expression data, but CAGE available for each cell line.

Initial idea:

- ▶ Investigate TADs as "regulons"—new paper reports 20% domains act as "discrete regulatory units" with relatively homogenous epigenetic states.<sup>1</sup>

Can I find evidence for this with my data?

---

<sup>1</sup> Le Dily *et al.* (2014) Distinct structural transitions of chromatin correlate with coordinated hormone-induced gene regulation. *Genes and Development*, **28**:2151-62

# Next project

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

Introduction

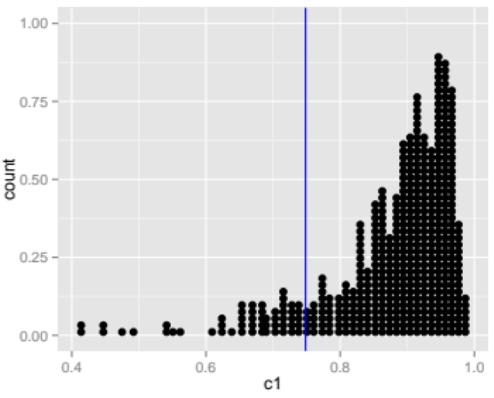
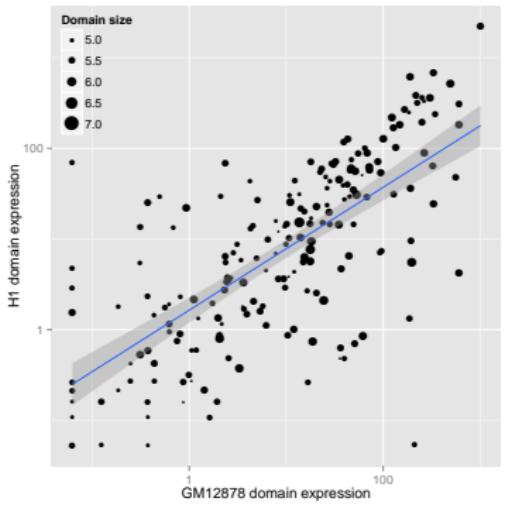
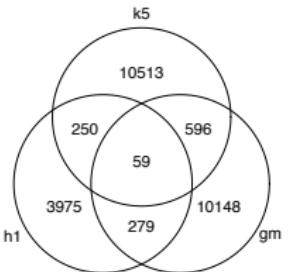
Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements



Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

# Next project

- ▶ Higher resolution Hi-C now available, better for identifying domains
- ▶ Improved methods of calling domains, where / why are some well-conserved between cell types?
- ▶ Mouse data also available

## Thesis plan p1

# Unravelling higher order chromatin organisation through statistical analysis



INSTITUTE OF GENETICS  
& MOLECULAR MEDICINE

1. Introduction
  2. Methods
  3. **Modelling transcription and chromatin** Replicating and extending ENCODE project predicting txn output; adapt techniques to genome organisation
  4. **Model dissection** How the chromatin structure models cross-apply; regularised models; variable importance
  5. **Boundaries** Comparing TADs/compartments across boundaries; “super bounds”; Giemsa bands
  6. ...

# Thesis plan p2

Unravelling higher  
order chromatin  
organisation  
through statistical  
analysis



6. **Investigating the function of self-interacting domains** Work from future project (very early stages); possibly two chapters from the next year
7. **C-methods collaborations** Write-up Hill lab collaborations, possibly include other minor collabs
8. Discussion
9. End materials, code

Introduction

Reanalysis of Hi-C datasets

Collaborations

Next project

Thesis plan

Acknowledgements

Introduction

Reanalysis of Hi-C  
datasets

Collaborations

Next project

Thesis plan

Acknowledgements

Thanks to supervisors:  
Colin Semple and Stuart Aitken