# ChromNet: Learning the human chromatin network from all ENCODE ChIP-seq data

## http://chromnet.cs.washington.edu
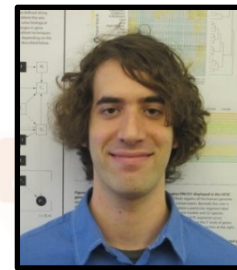
*ENCODE users meeting 2016*
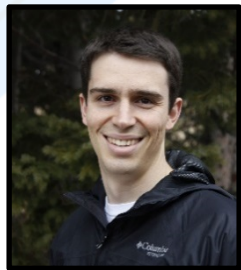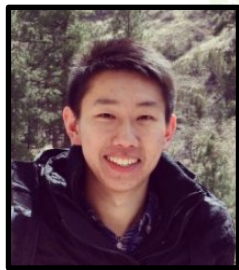
Scott Lundberg  William Tu  Brian Raught  Linda Penn  Michael Hoffman  Su-In Lee

# The *chromatin network**
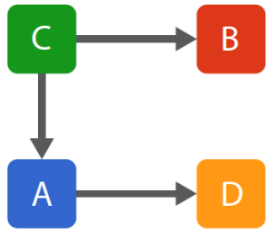## What we call the network of factors that interact to regulate the genome



*Lundberg et al. *Genome Biology* 2016

# Motivating example
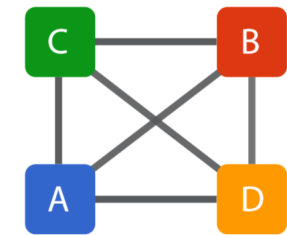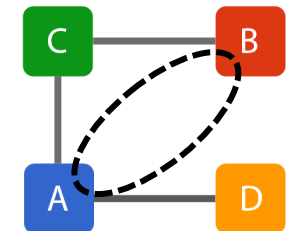## Illustrating the value of conditional dependence

# Measuring many factors is important
## Conditional dependence get better with a more complete model



True interactions

factors are variables

A

B

D

genomic positions are samples

Co-occurrences

Conditional dependence
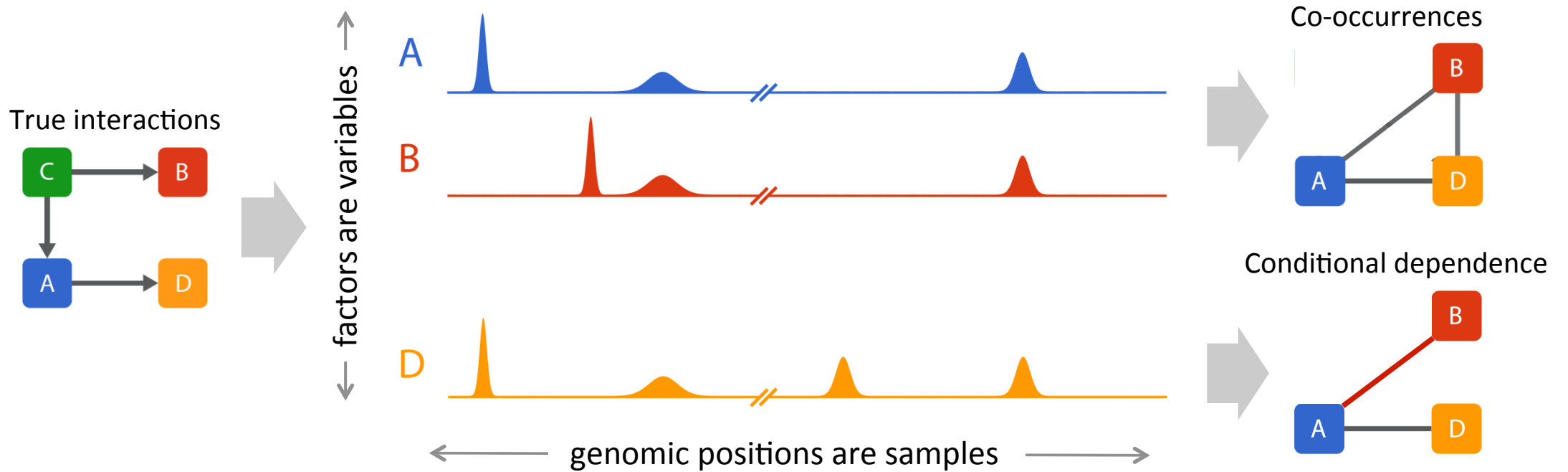
What if we didn't measure C? Then A and B get connected.

# Group graphical model
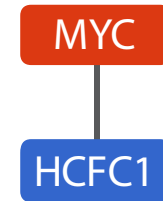## Making conditional dependence robust to redundancy



Without redundancy $\Sigma'^{-1}$     With redundancy $\Sigma^{-1}$     Group Graphical Model G
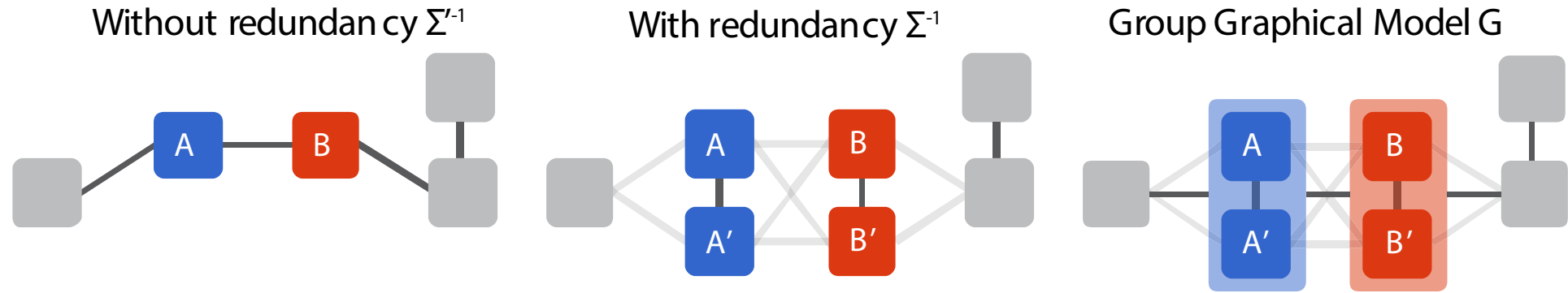
Redundant variables cause lost or unstable connections.

# Measuring accuracy improvements
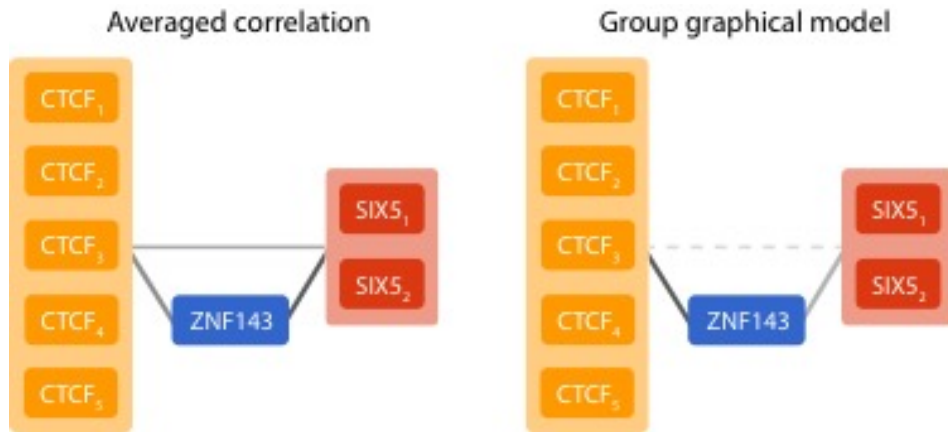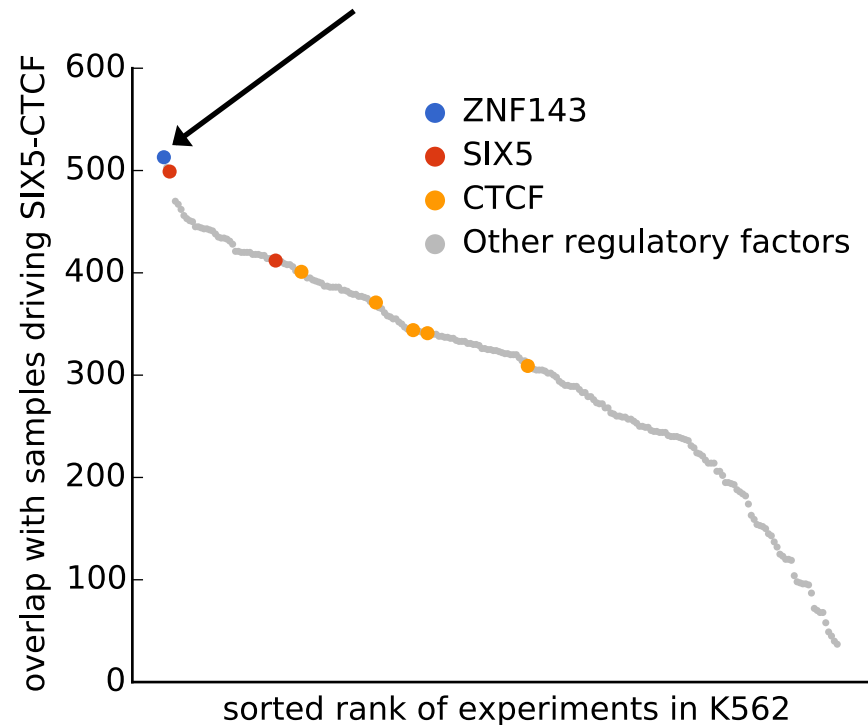## Recovery of known interactions and relationships
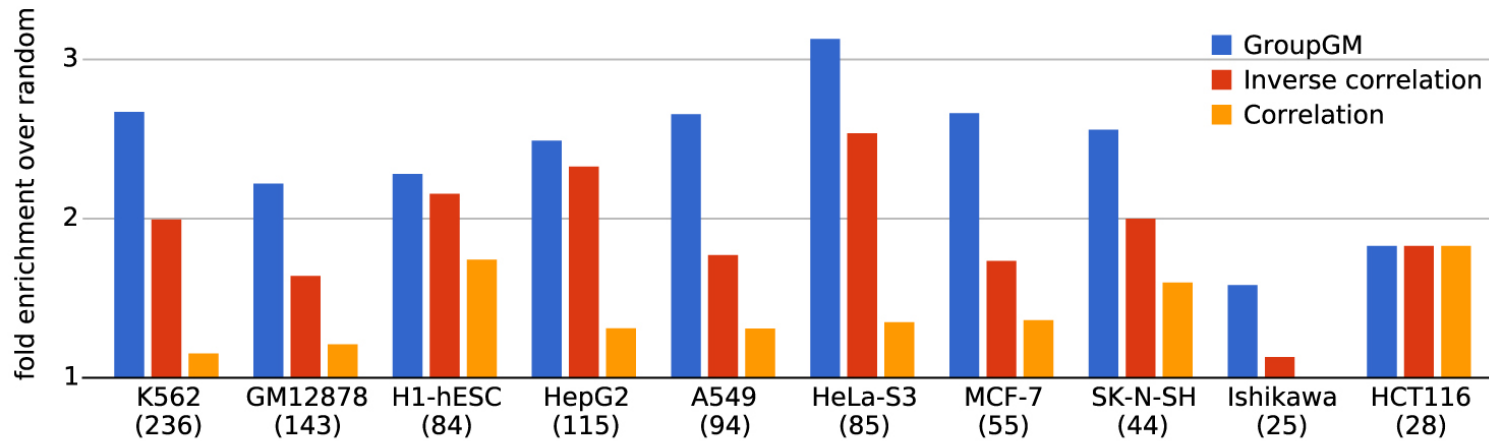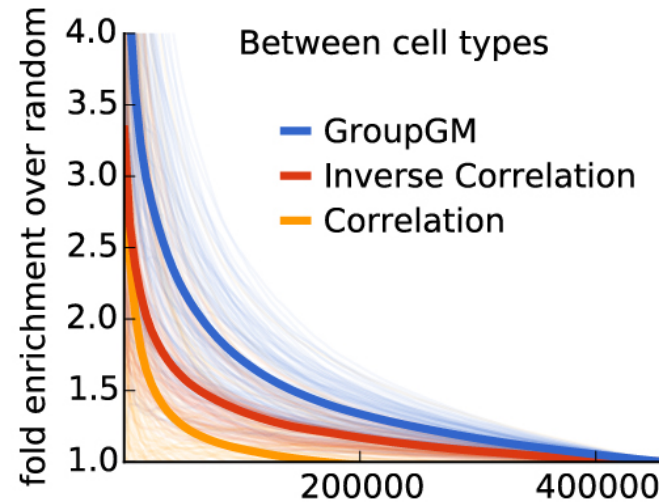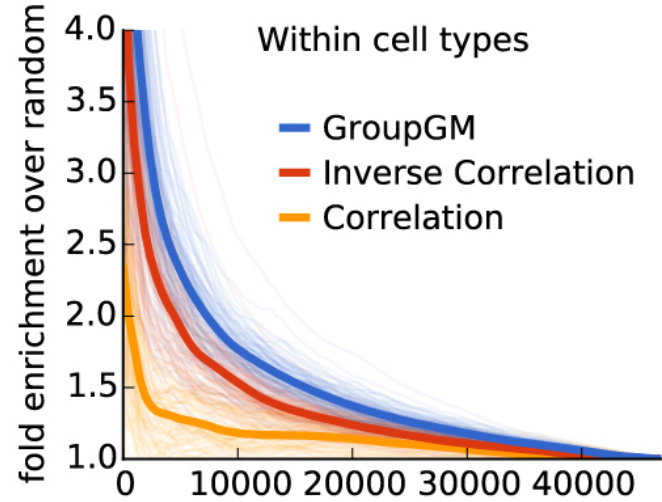
# Context specificity
## Identifying which genomic regions drive an interaction



Averaged correlation

Group graphical model

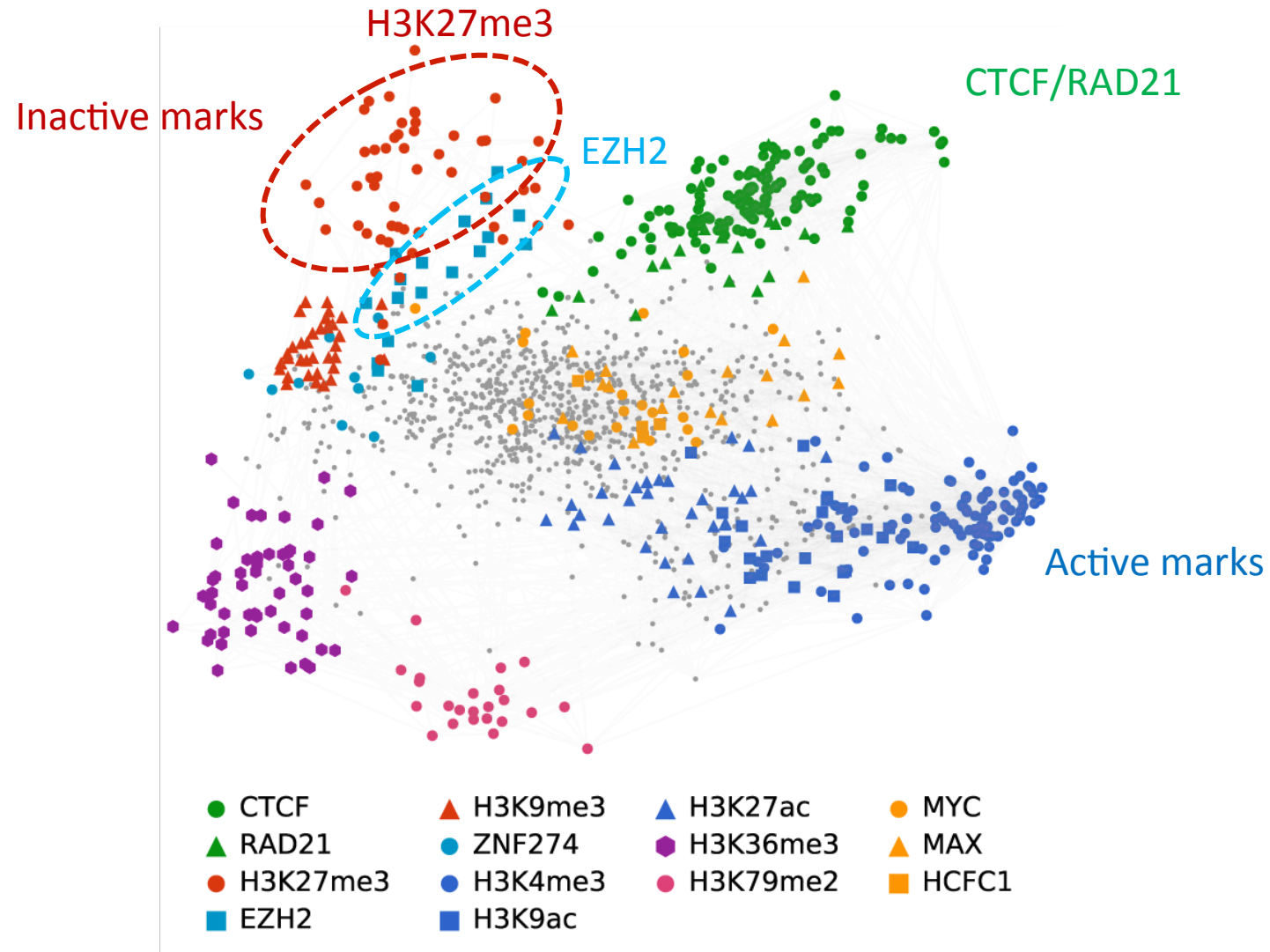CTCF and SIX5 associate in the presence of ZNF143.

We removed ZNF143 from ChromNet then checked which genomic regions drove the CTCF-SIX5 edge.

They matched the held-out ZNF143 better than even the CTCF and SIX5 tracks themselves!



- ● ZNF143
- ● SIX5
- ● CTCF
- ● Other regulatory factors

overlap with samples driving SIX5-CTCF

sorted rank of experiments in K562

# Global network

## All cells types integrated into a single network.

# Explore, and even integrate your own data!
## http://chromnet.cs.washington.edu

# Backup slides

# PLA results
## Following up on the MYC-HCFC1 network edge
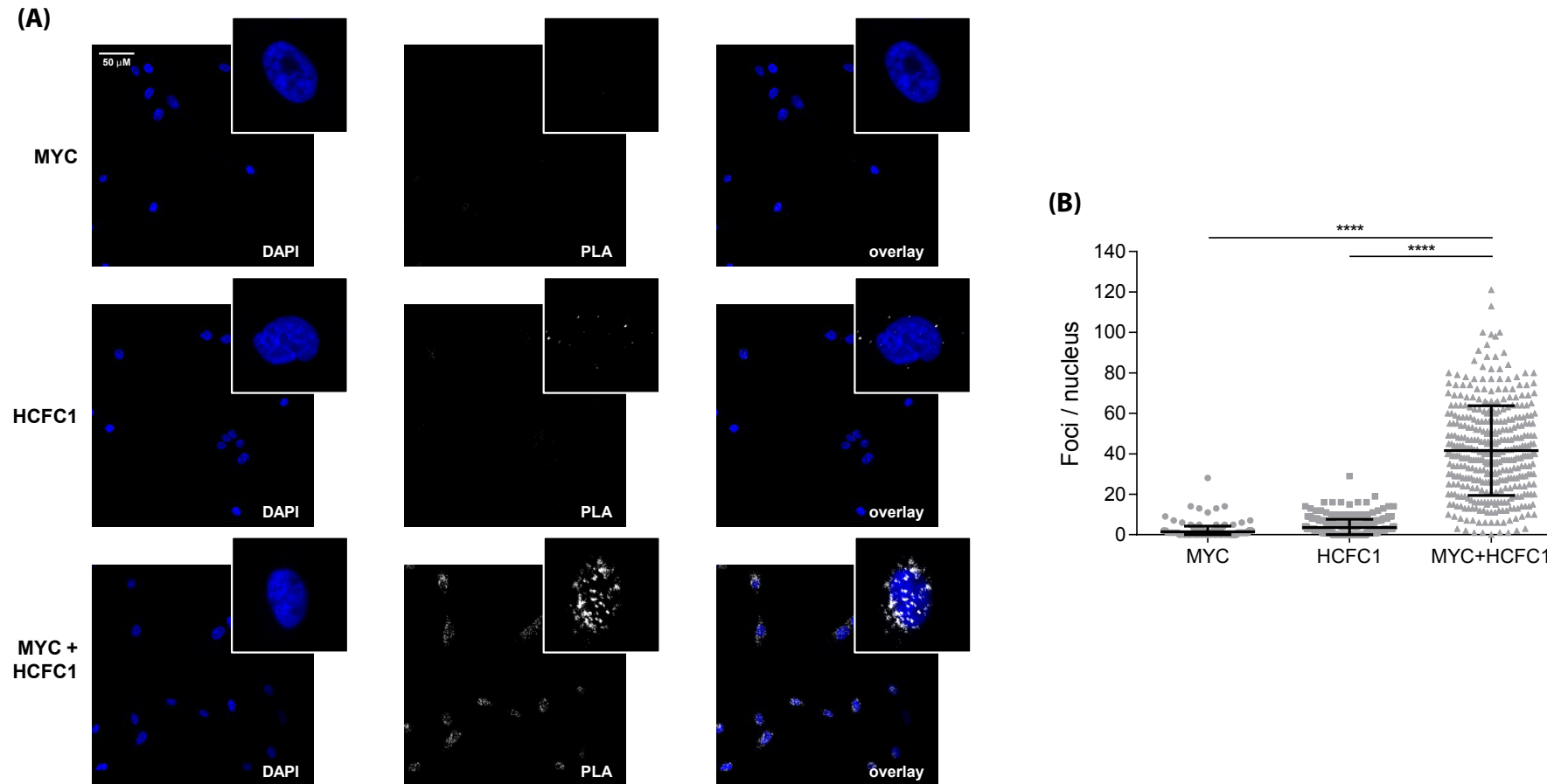
# Histone mark / writer connections