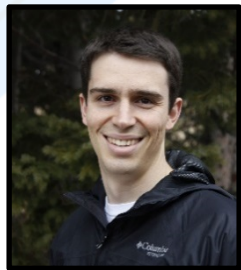


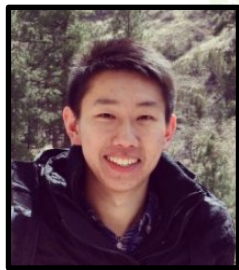
ChromNet: Learning the human chromatin network from all ENCODE ChIP-seq data

<http://chromnet.cs.washington.edu>

ENCODE users meeting 2016



Scott Lundberg



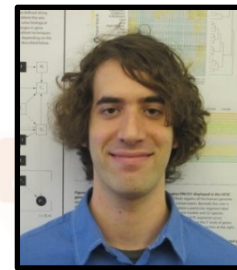
William Tu



Brian Raught



Linda Penn



Michael Hoffman



Su-In Lee



CAREER AWARD



The *chromatin network**

What we call the network of factors that interact to regulate the genome

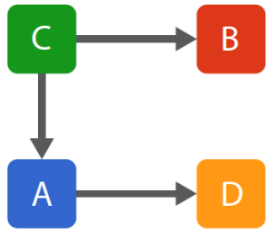


*Lundberg et al. *Genome Biology* 2016

Motivating example

Illustrating the value of conditional dependence

True interactions



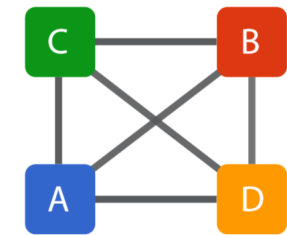
factors are variables

C

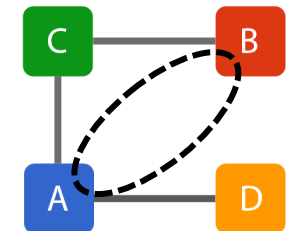


genomic positions are samples

Co-occurrences



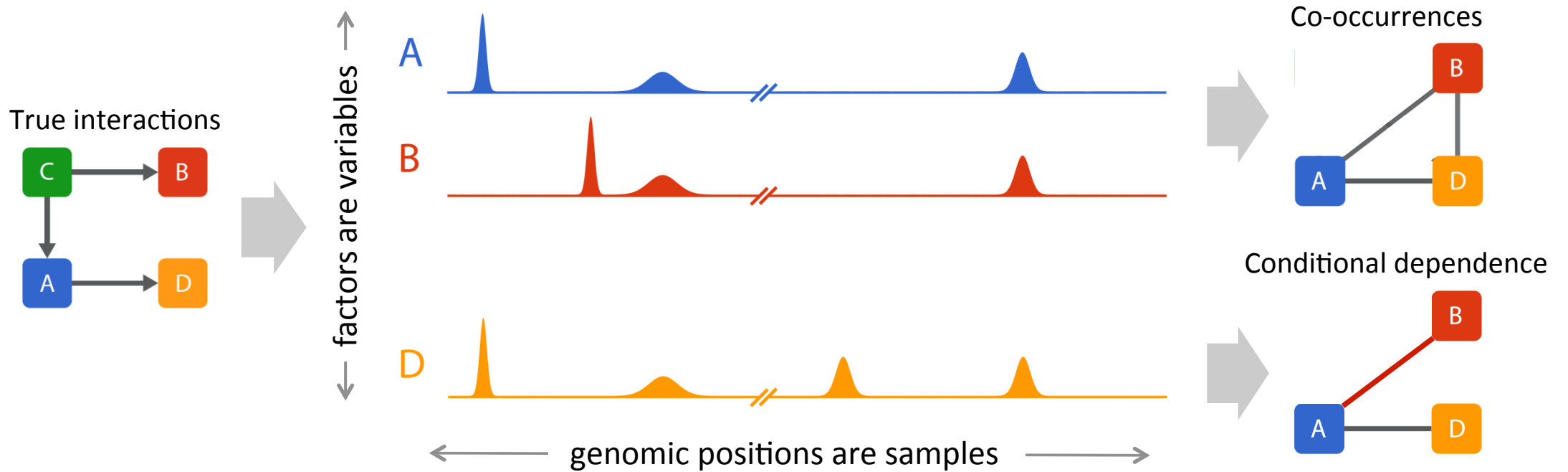
Conditional dependence



$$A = \beta_B B + \beta_C C + \beta_D D$$

Measuring many factors is important

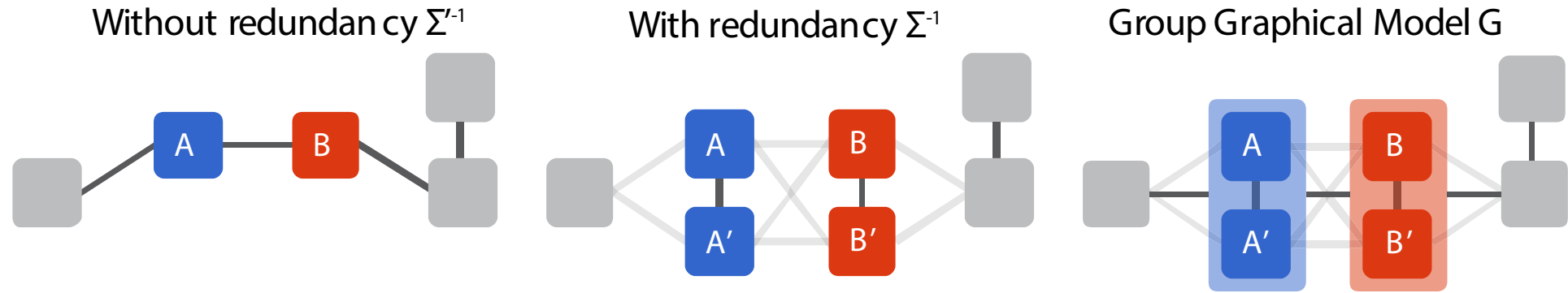
Conditional dependence get better with a more complete model



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

Group graphical model

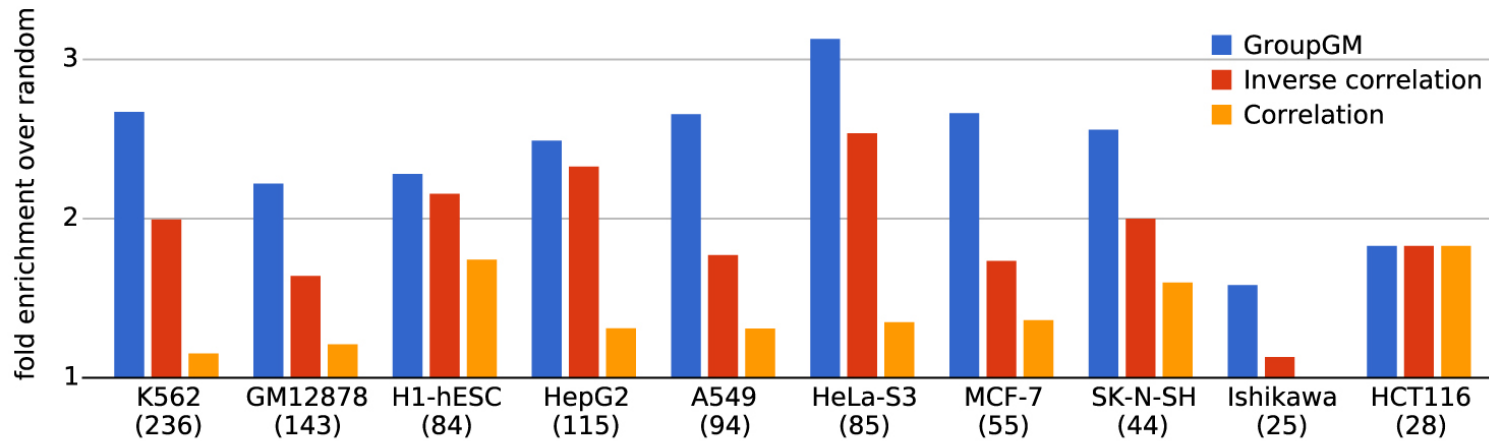
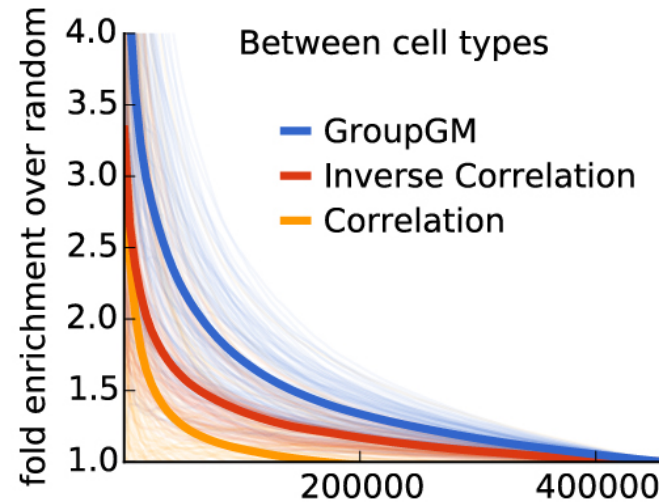
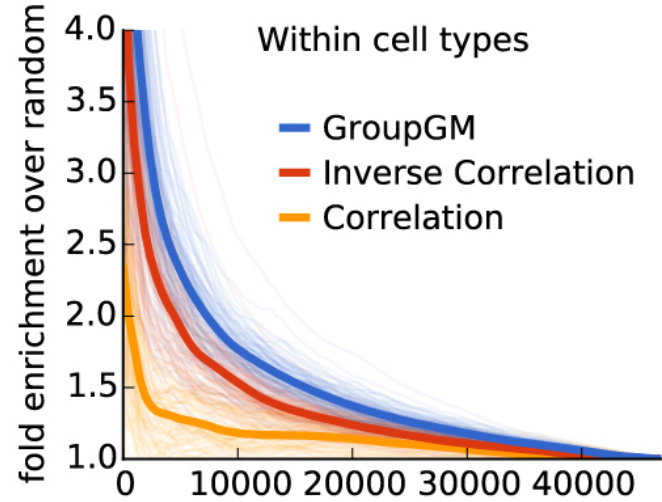
Making conditional dependence robust to redundancy



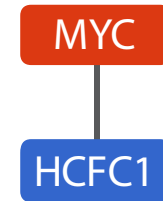
↑
Redundant variables cause lost or unstable connections.

Measuring accuracy improvements

Recovery of known interactions and relationships

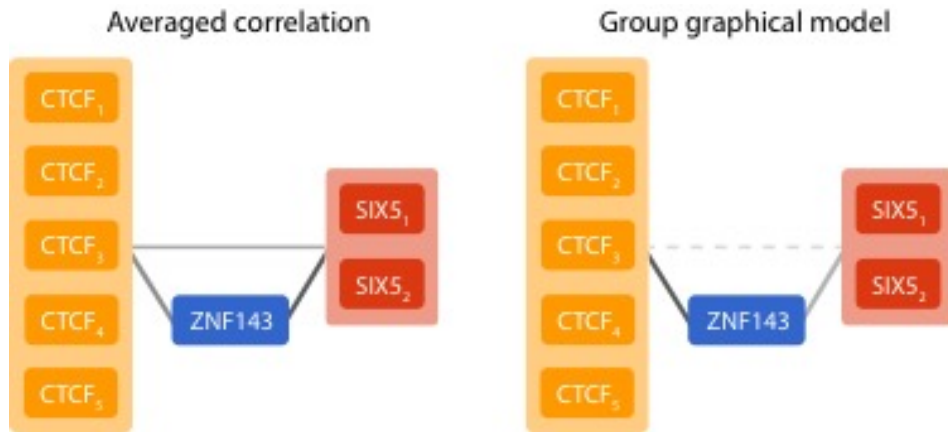


Validated novel interaction



Context specificity

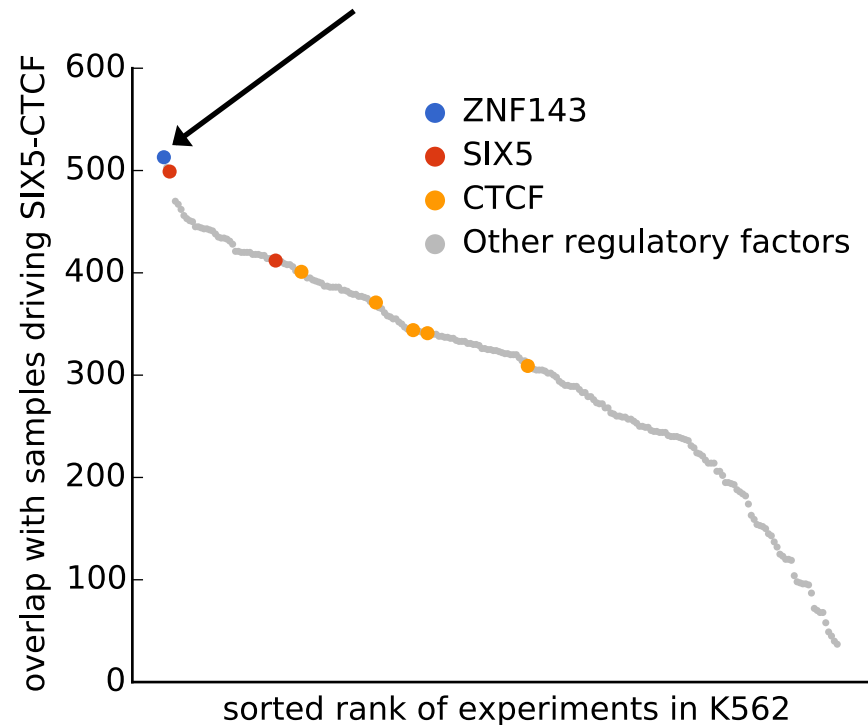
Identifying which genomic regions drive an interaction



CTCF and SIX5 associate in the presence of ZNF143.

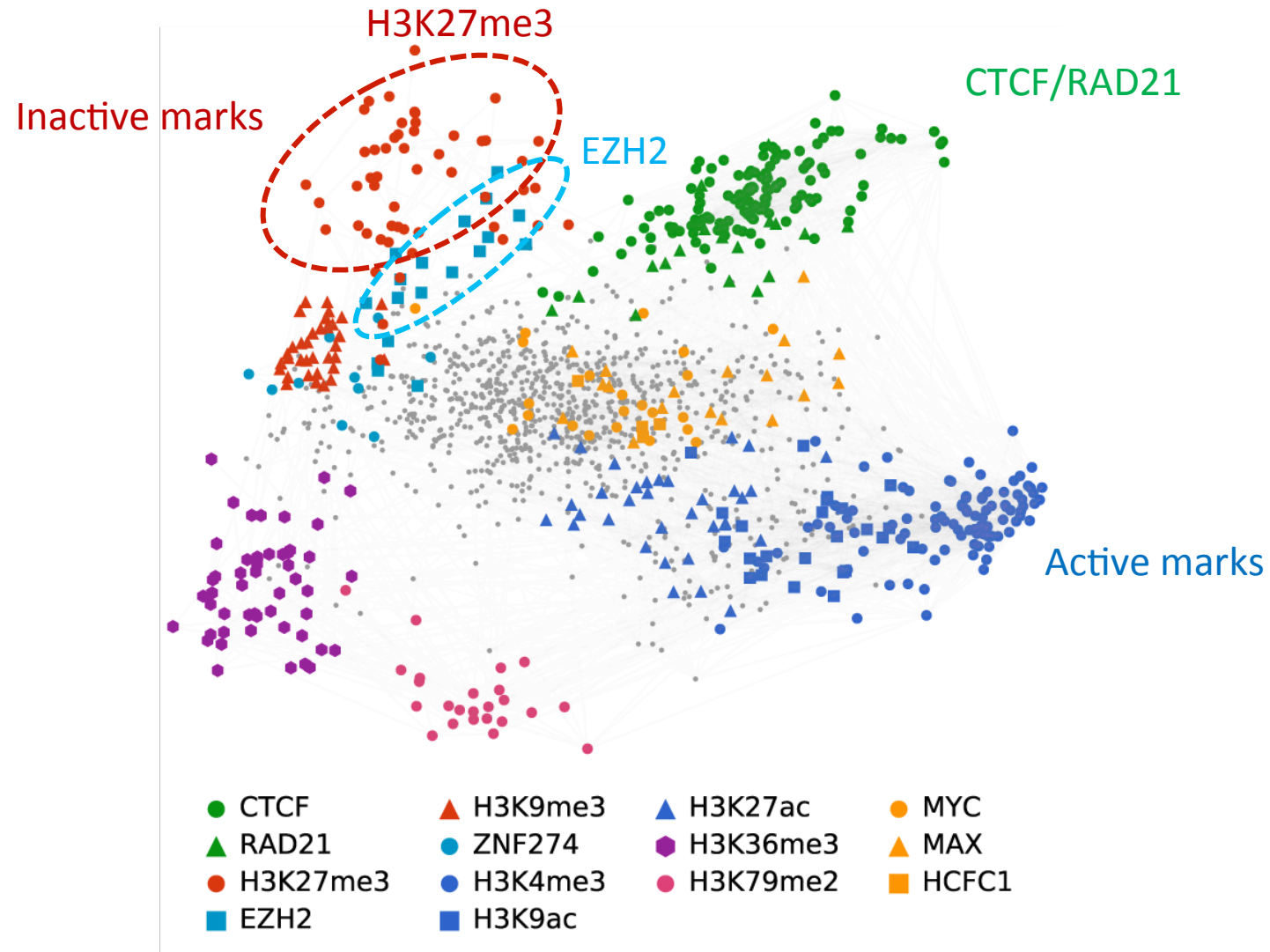
We removed ZNF143 from ChromNet then checked which genomic regions drove the CTCF-SIX5 edge.

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



Global network

All cells types integrated into a single network.

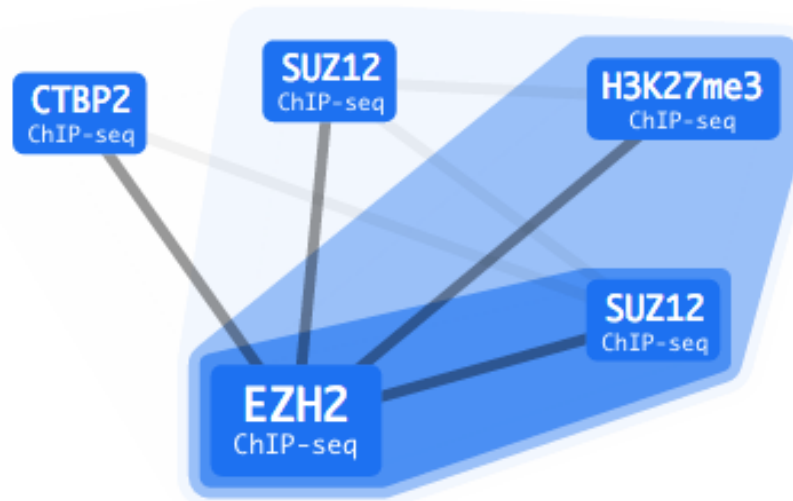


Explore, and even integrate your own data!

<http://chromnet.cs.washington.edu>

ChromNet 1 matches, 4 nearby dense sparse ? ⚙️

Cell Types and Treatments
H1-hESC

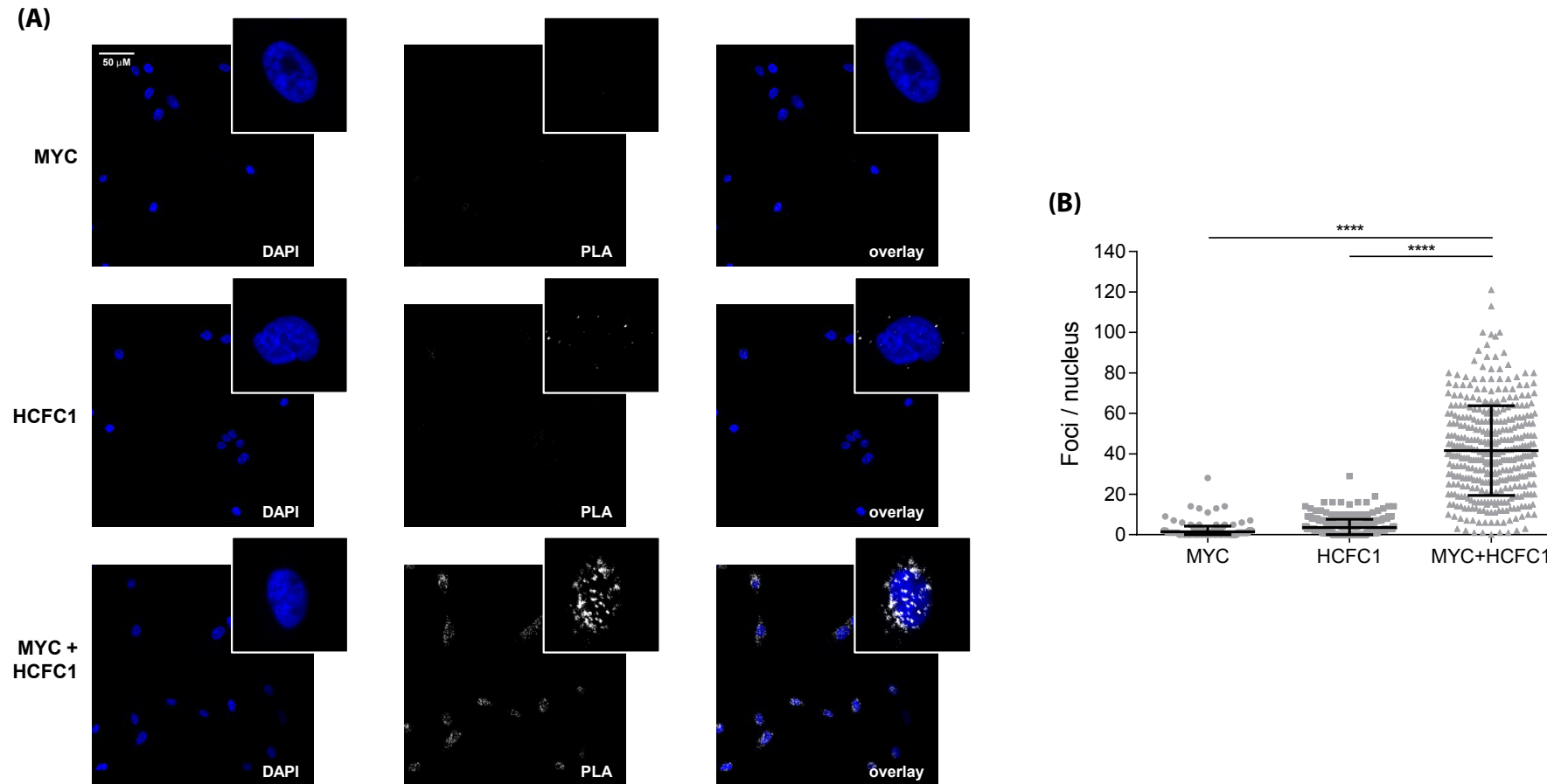


Backup slides

This work was supported by a National Science Foundation (NSF) Graduate Research Fellowship to SML (DGE-1256082), a NSF grant to SIL (DBI-1355899), a Natural Sciences and Engineering Research Council of Canada grant to MMH (RGPIN-2015-03948), a Canada Research Chair in Molecular Oncology to LZP, a Canadian Institute for Health Research grant to LZP and BR (MOP-275788), and a Canadian Breast Cancer Foundation Ontario Region Doctoral Fellowship to WBT. Cloud computing resources for this research were generously provided by Google.

PLA results

Following up on the MYC-HCFC1 network edge



Histone mark / writer connections

