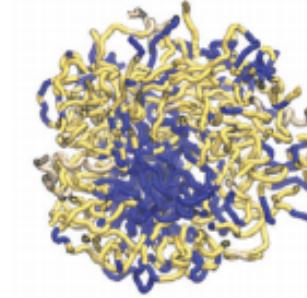


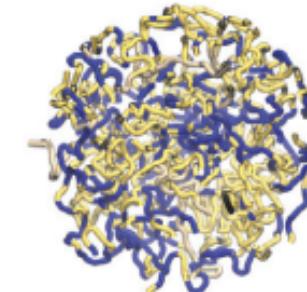
# **Large-scale topological changes restrict malignant progression in colorectal cancer**



**Alejandro Reyes,  
Ph.D.**

EuroBioc2020

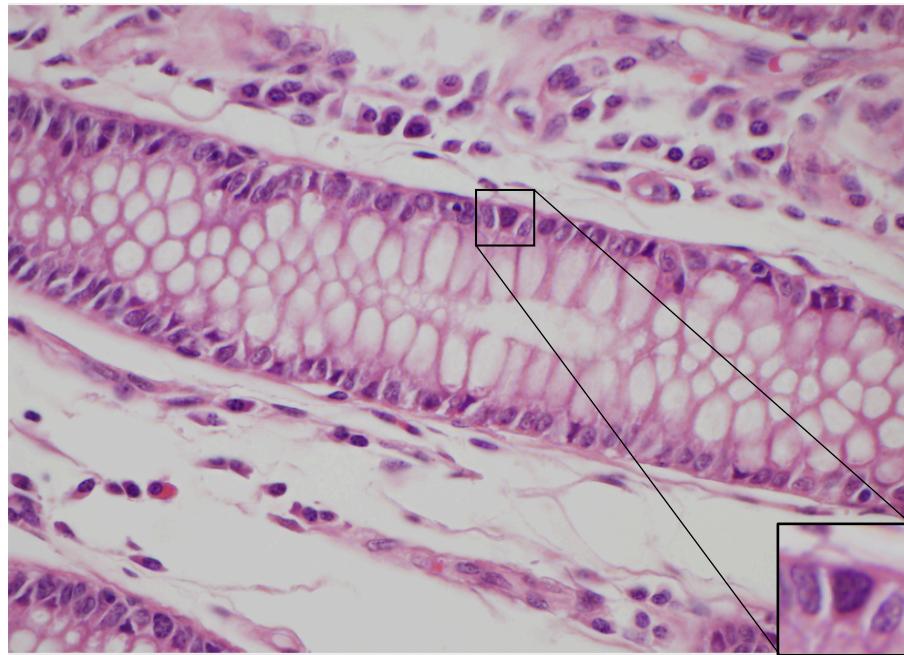
December 16th, 2020



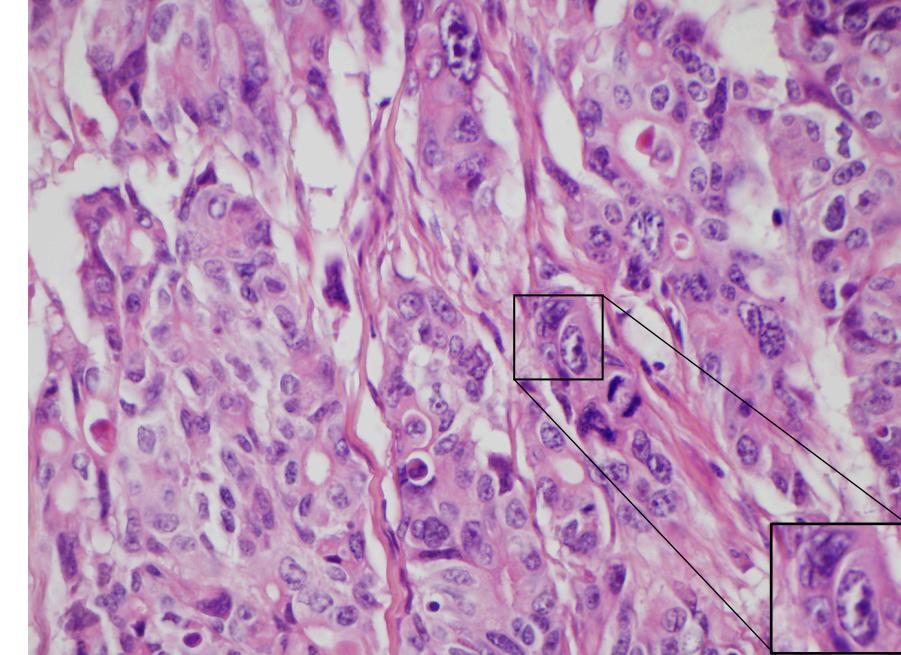
# Size and shape of the nucleus are hallmarks of cancer cells

---

Normal colon tissue

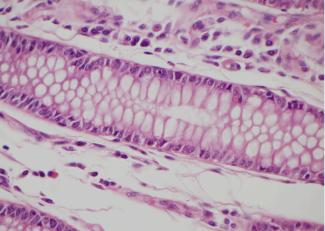
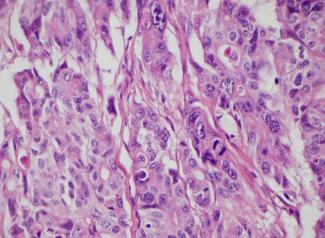
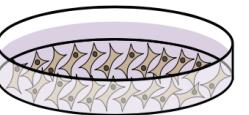


Colorectal cancer



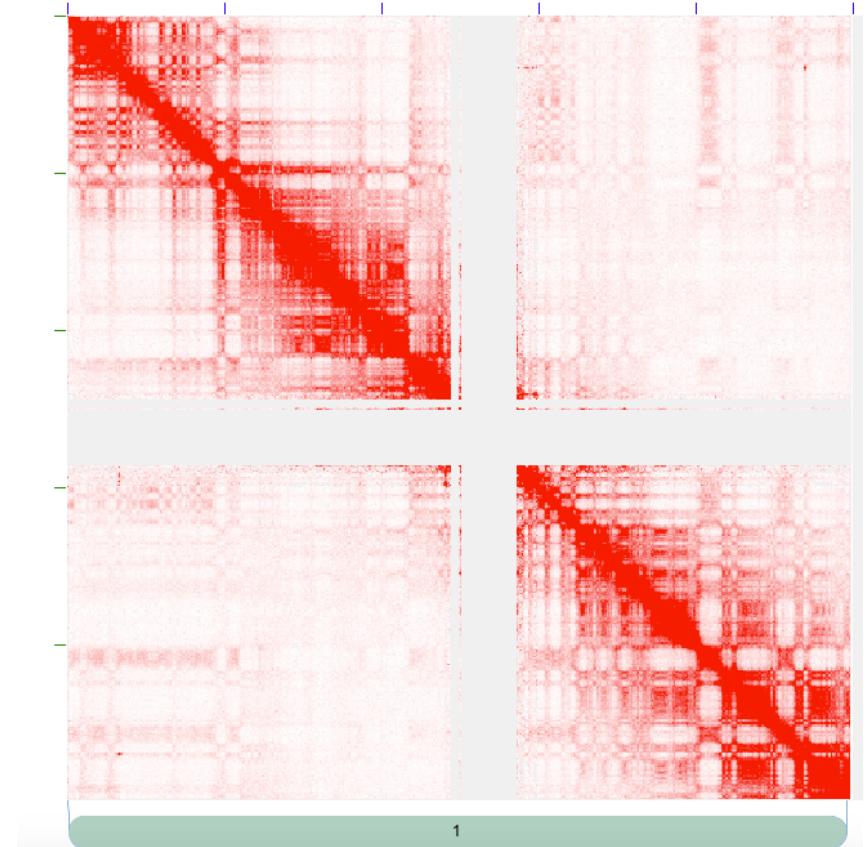
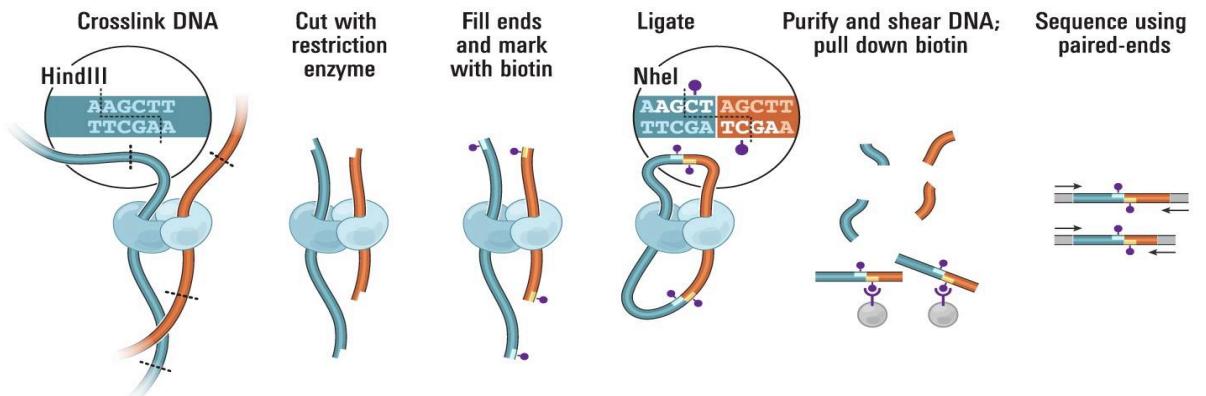
What are the molecular changes that occur in the 3D genome of cancer genomes?

# Topological atlas of colorectal cancer genomes

# of NGS libraries	Hi-C	Hi-ChIP H3K27ac	DNAme	ChIP-seq (5 marks)	RNA-seq
Normal (n = 7) 	11	2	3	10	1
Tumor (n = 26) 	20	7	26	21	7
Cell lines (n = 6) 	20	0	19	1	14

175 libraries, > 28 billion reads

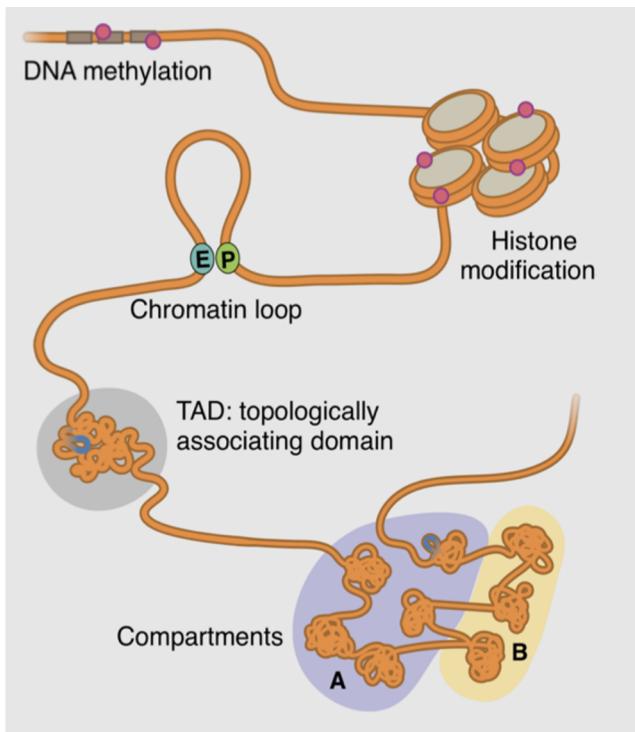
# Hi-C data measures 3D interactions genome-wide



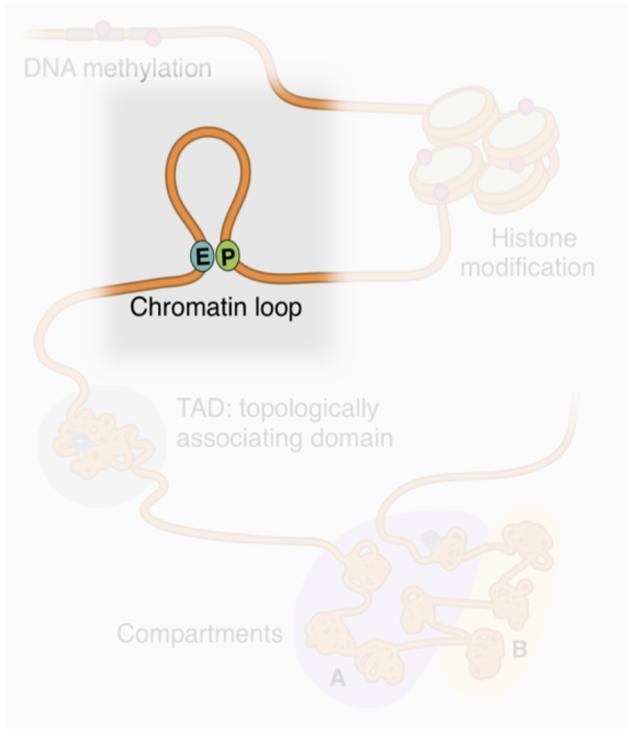
Hi-C: unbiased all-to-all interactions

HiChIP: interactions mediated by a DNA binding protein

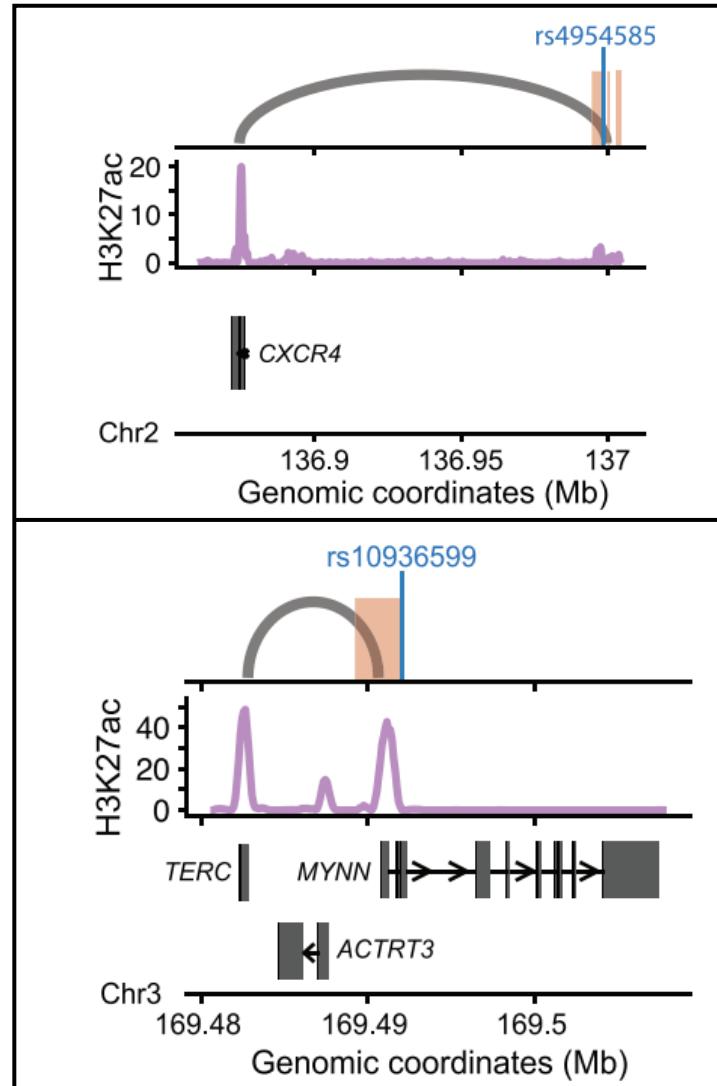
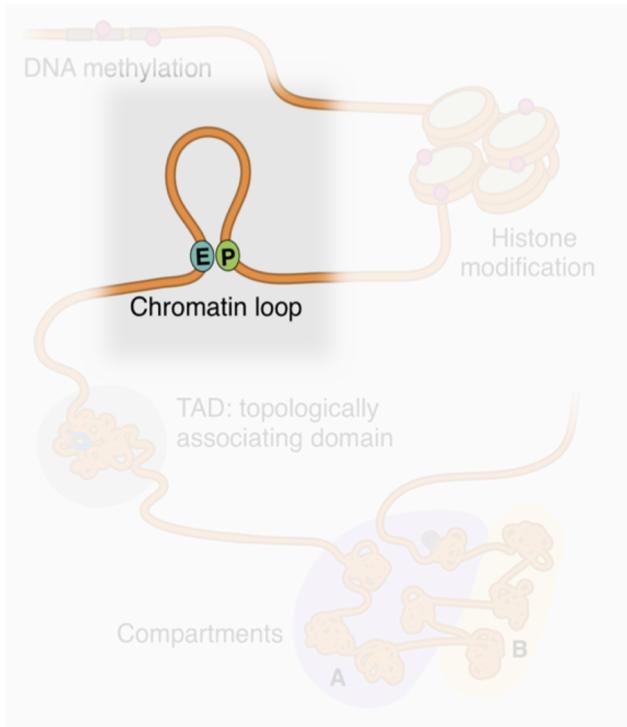
# Chromatin is folded in a hierarchical structure



# Chromatin loops are 2D distant genomic regions that are closeby in the 3D space

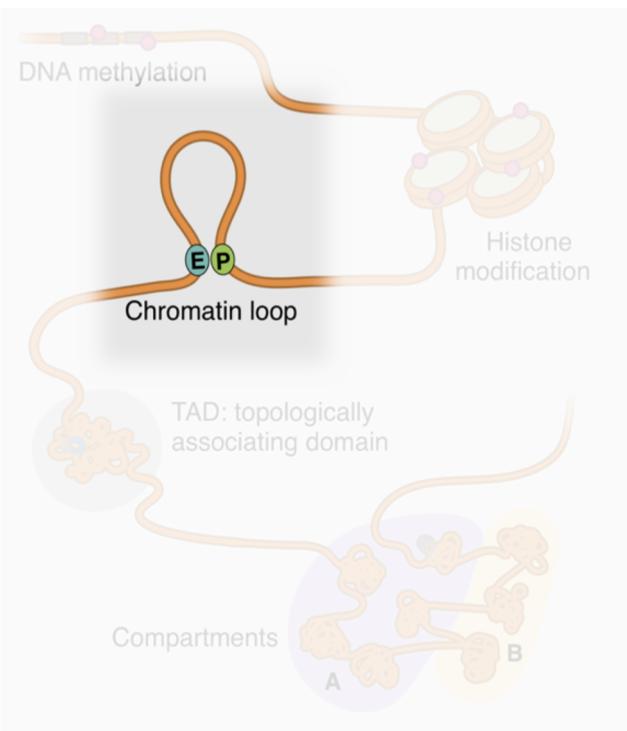


# Chromatin loops can link SNPs to distal genes



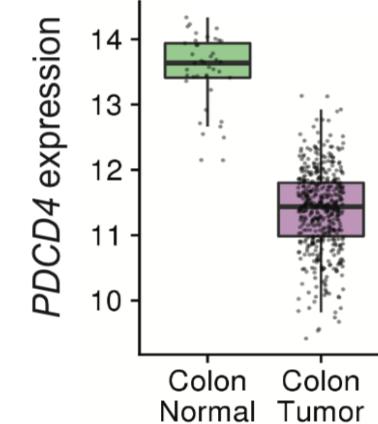
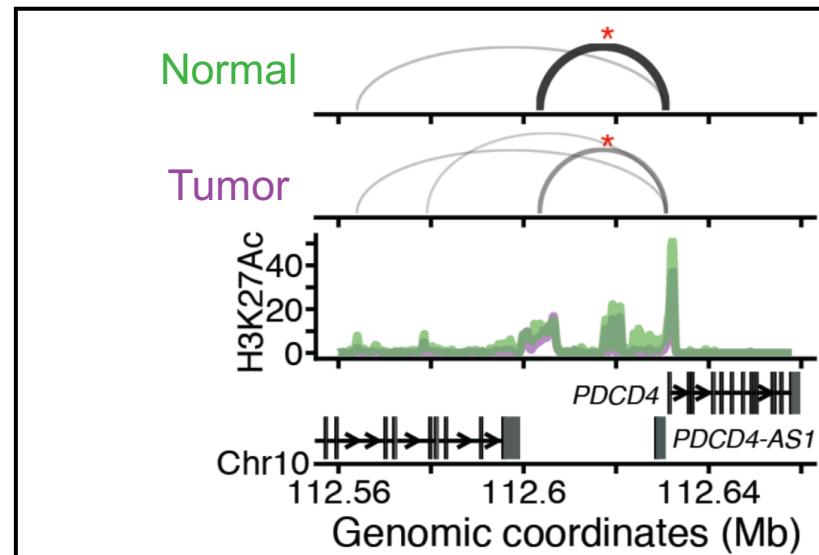
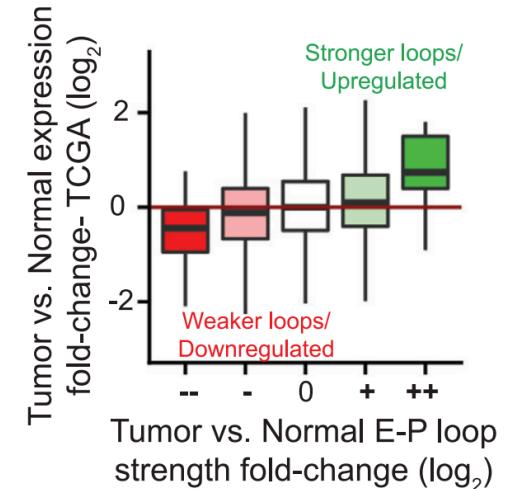
- Colorectal cancer GWAS SNPs are enriched among chromatin loop anchors ( $p < 10^{-5}$ ).
- SNPs in close gene 3D proximity to dozens of genes.

# Enhancer-promoter loops contribute to oncogenic programs

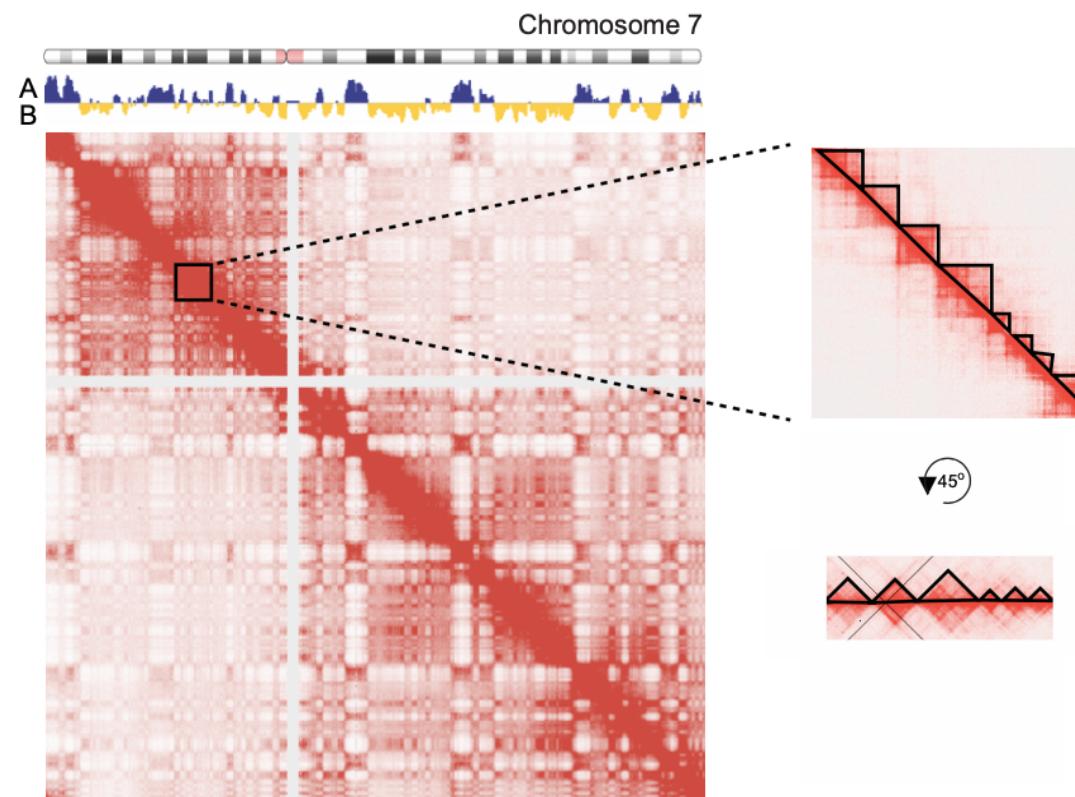
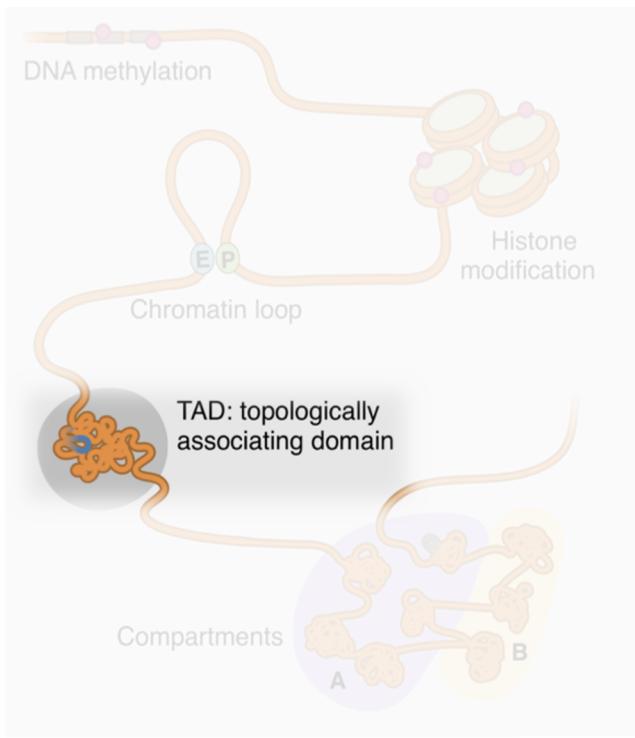


Identified 25,125 loops in HiChIP:

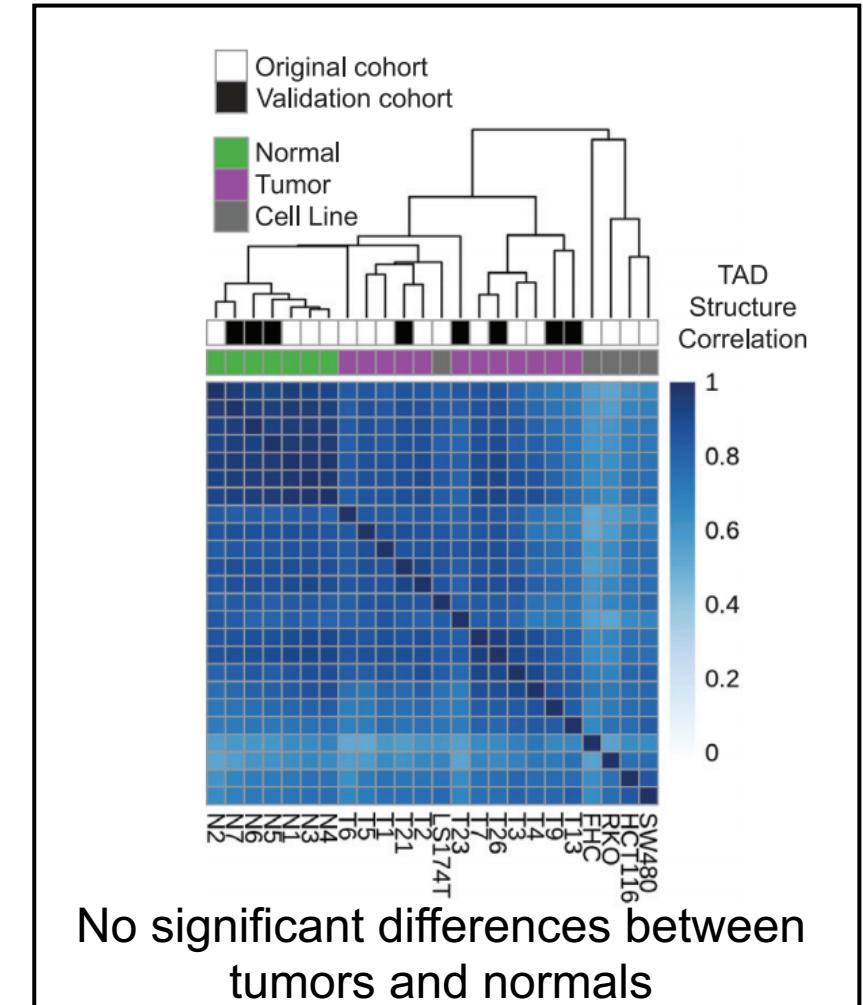
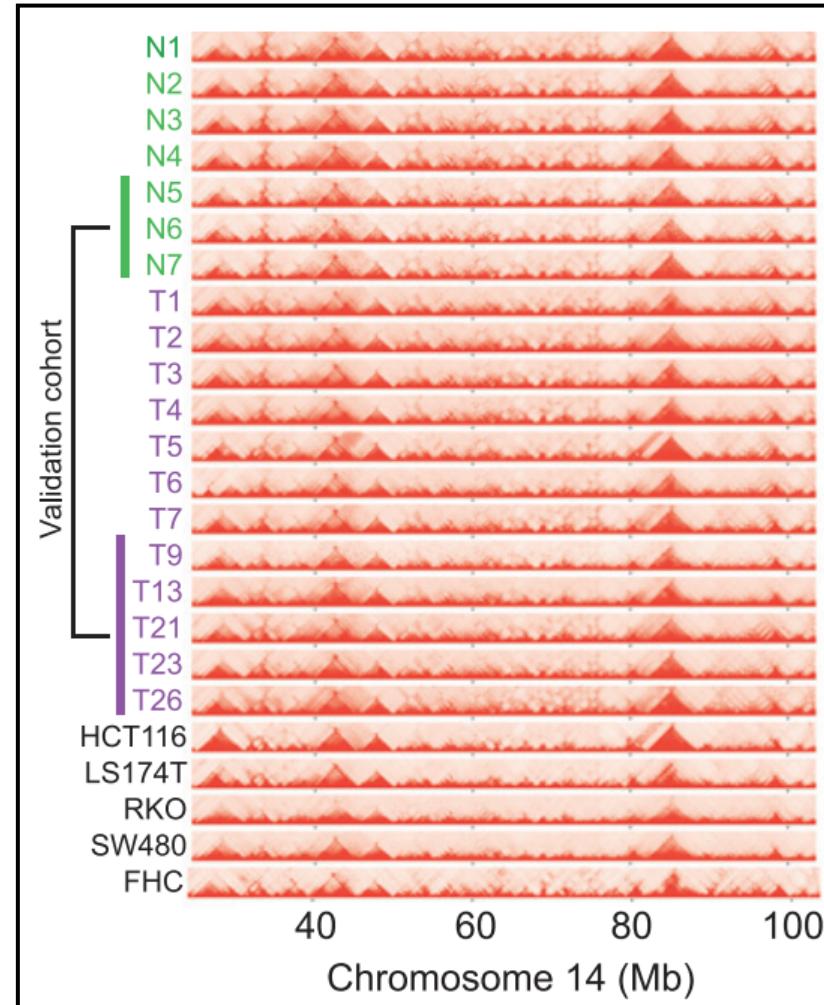
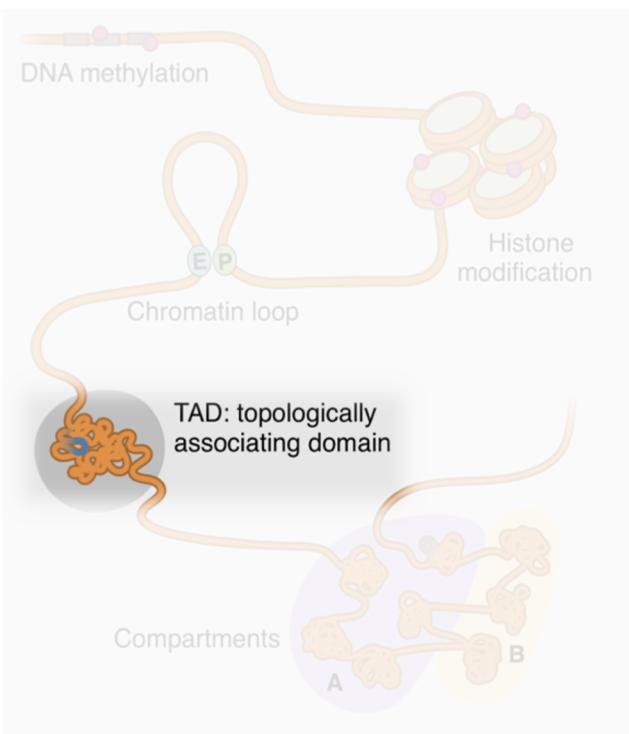
- 14,121 are E-P loops
- 571 are stronger in tumors
- 248 are weaker in tumors



# Topologically associating domains



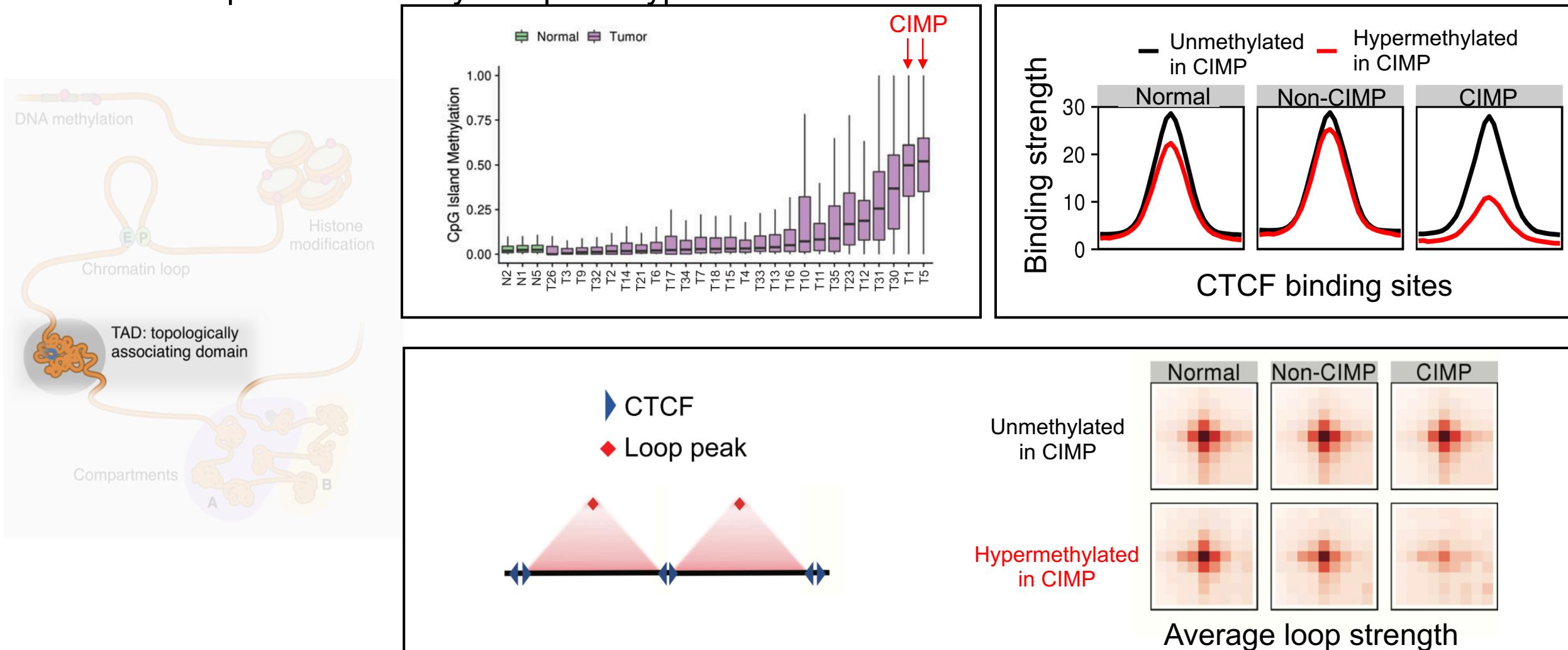
# Topologically associating domains are largely stable in tumors



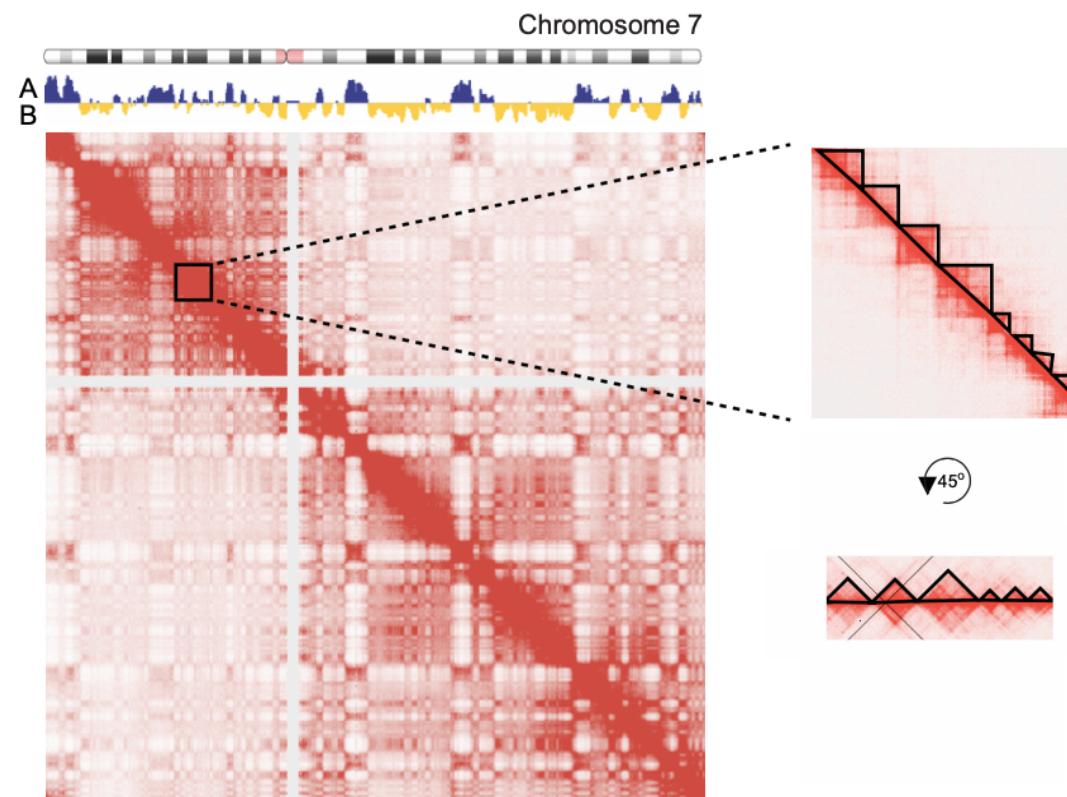
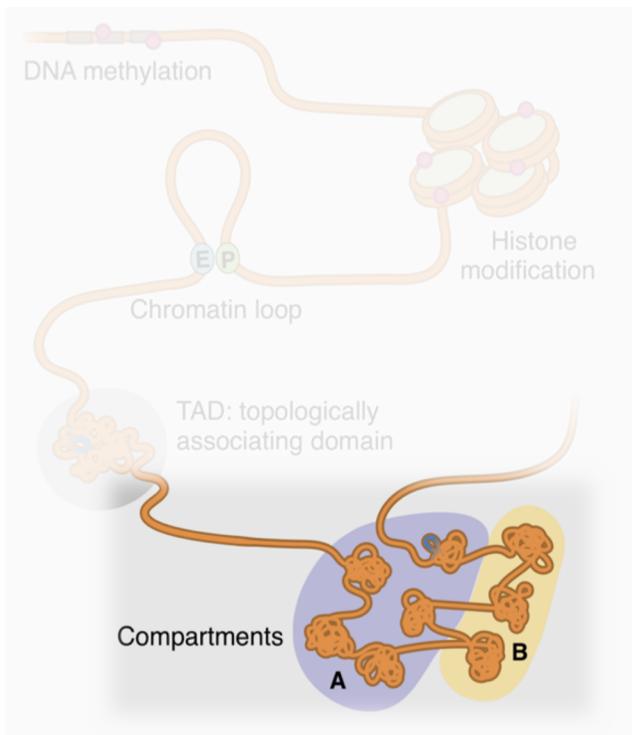
No significant differences between tumors and normals

# Topologically associating domains are largely stable, except for CIMP tumors

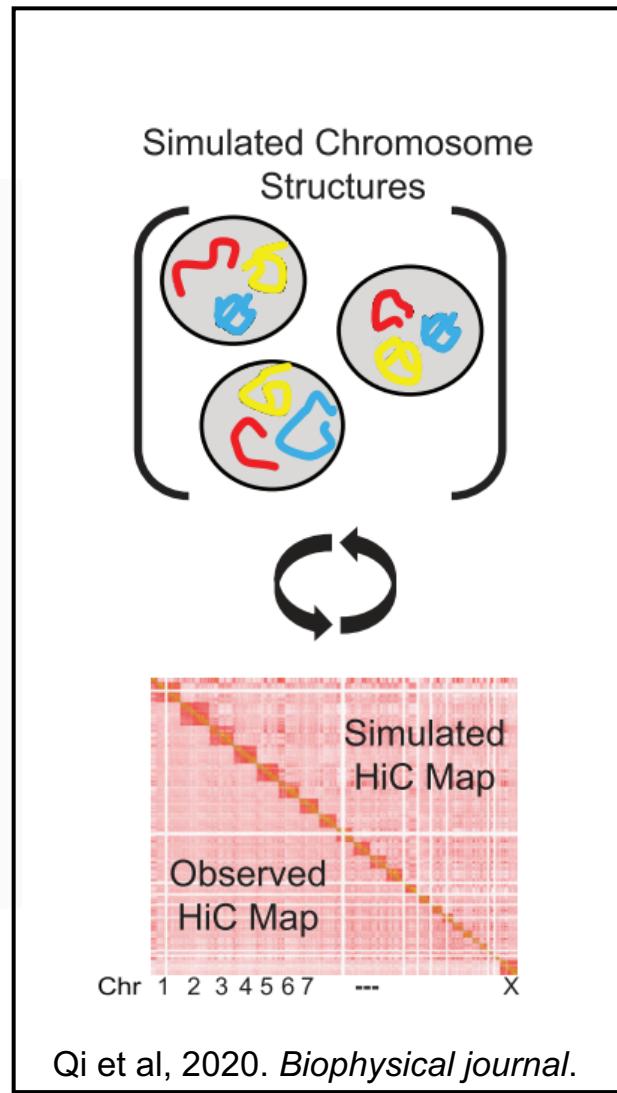
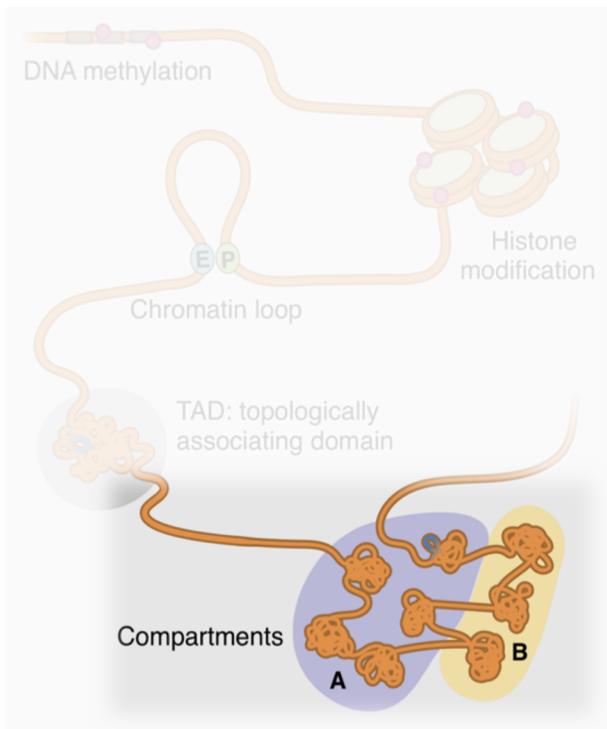
CIMP: CpG island methylator phenotype



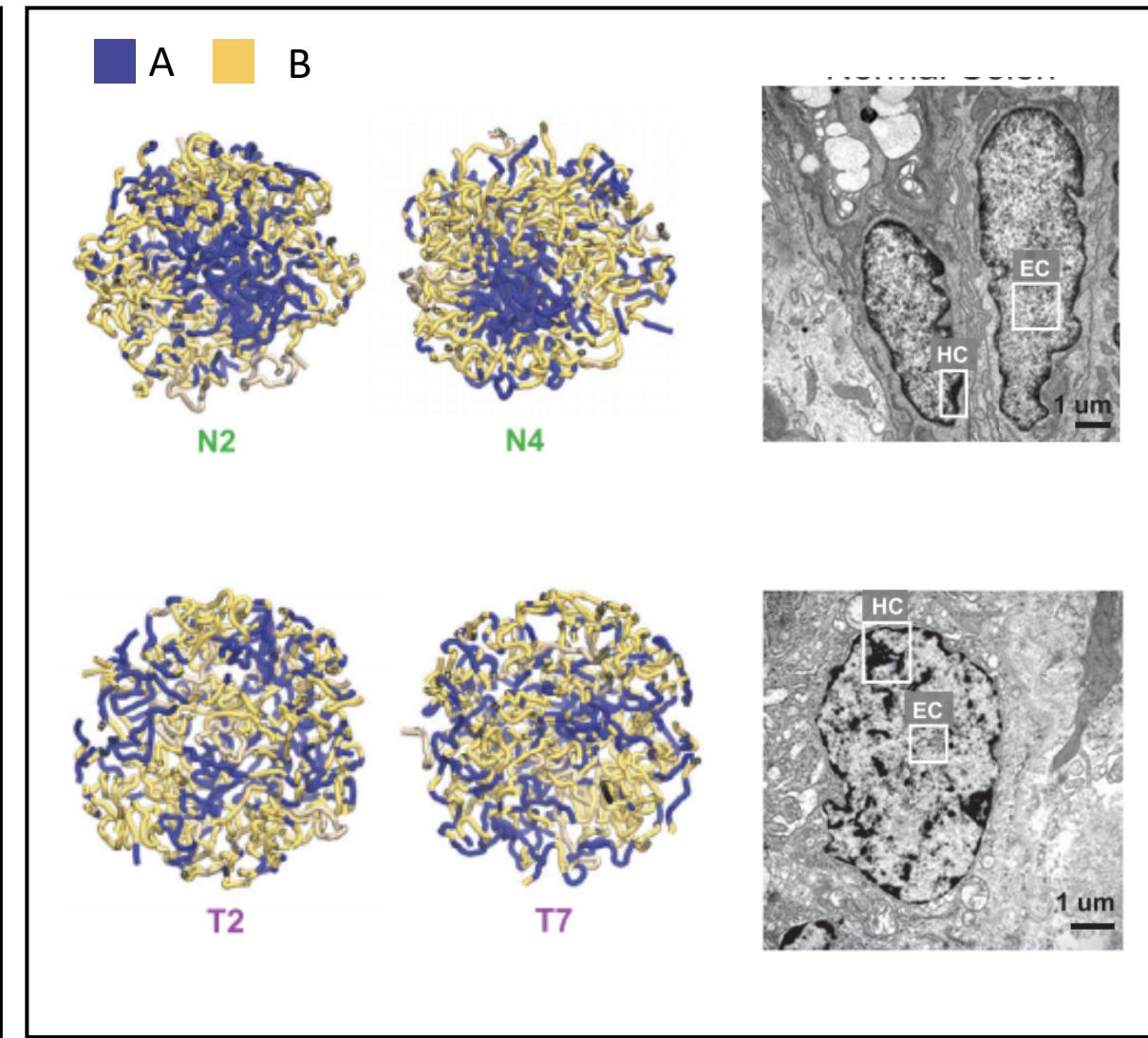
# Chromatin compartments



# Peripheral localization of compartments is disrupted in tumours

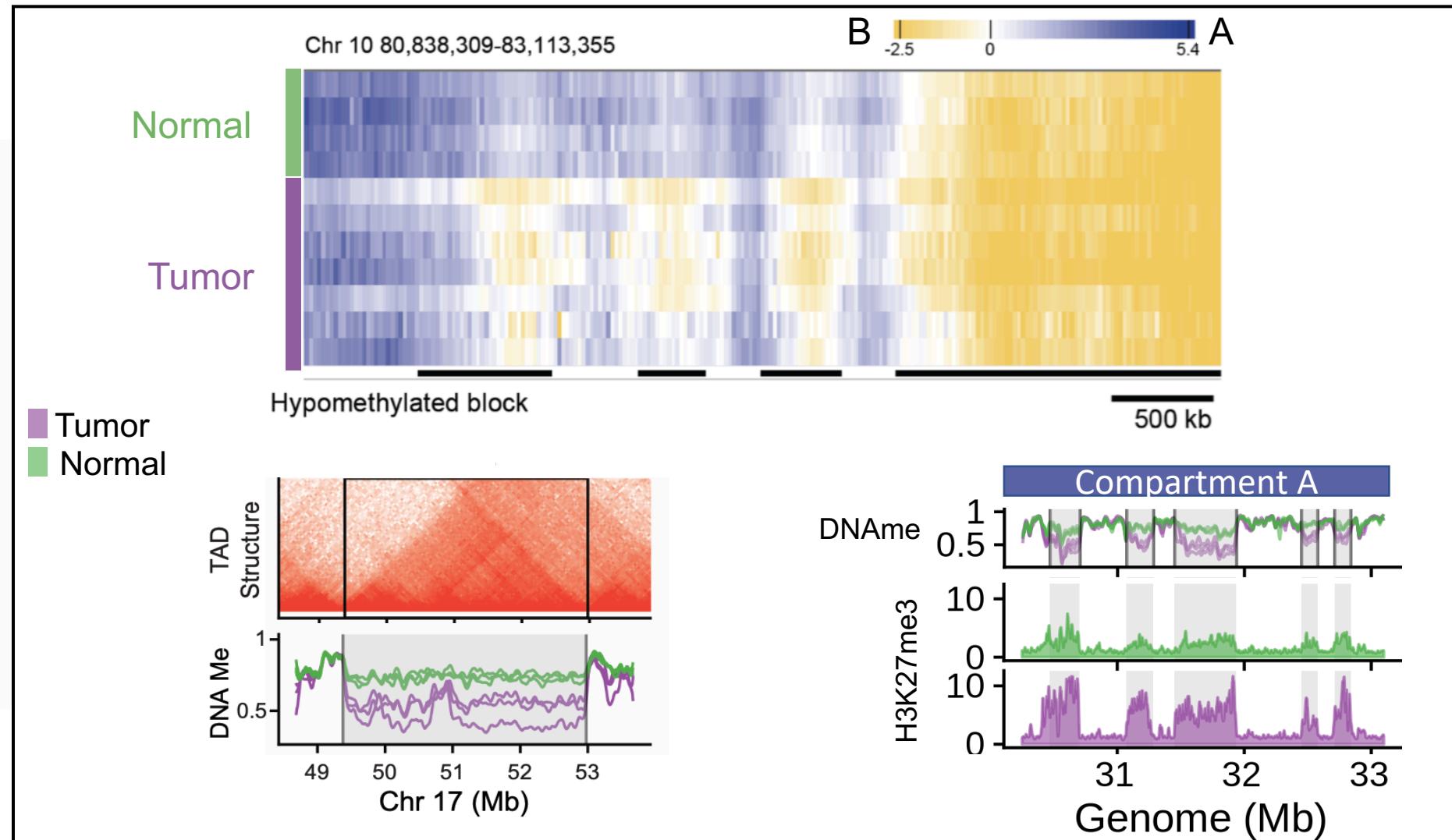
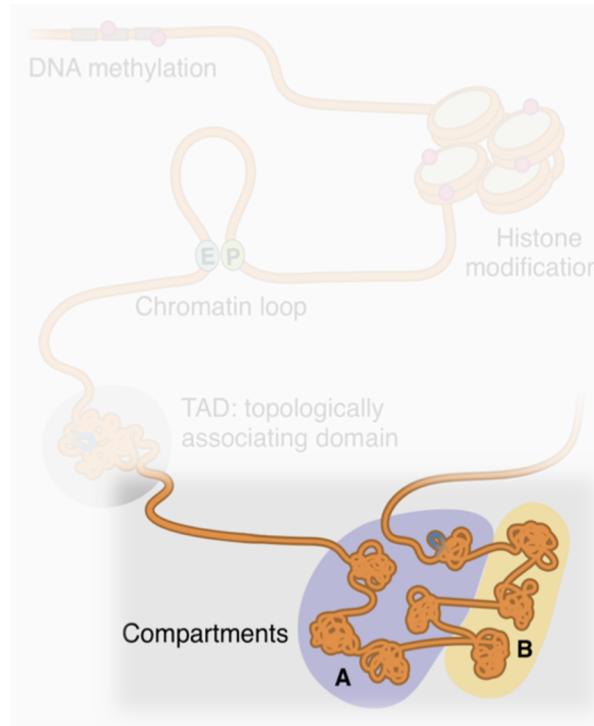


Qi et al, 2020. *Biophysical journal*.

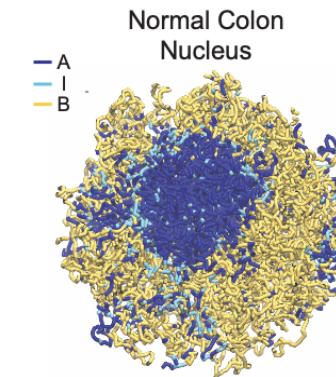
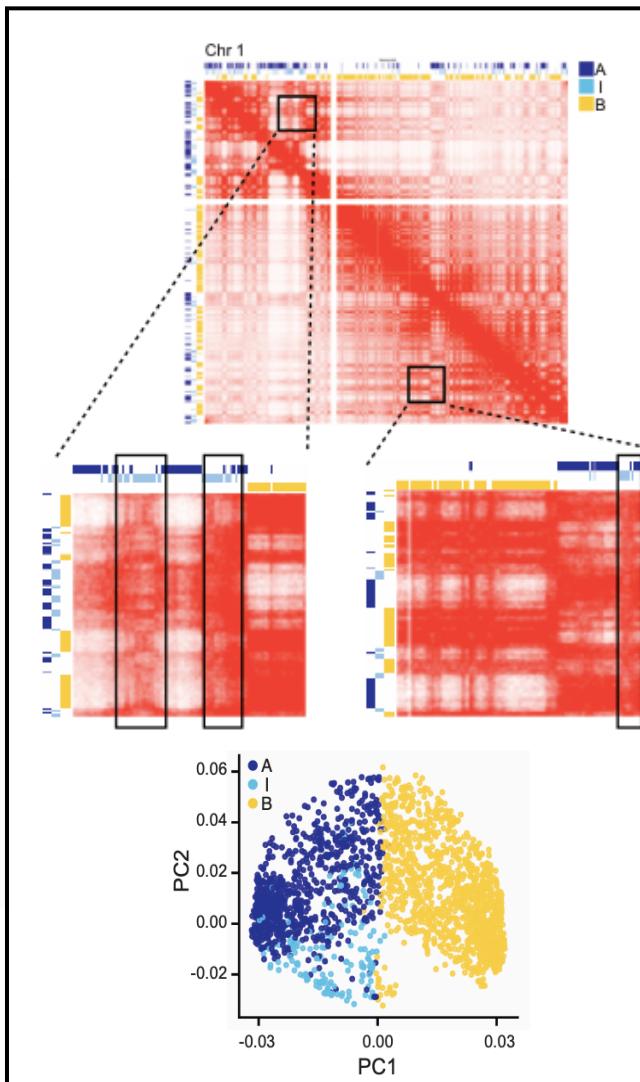
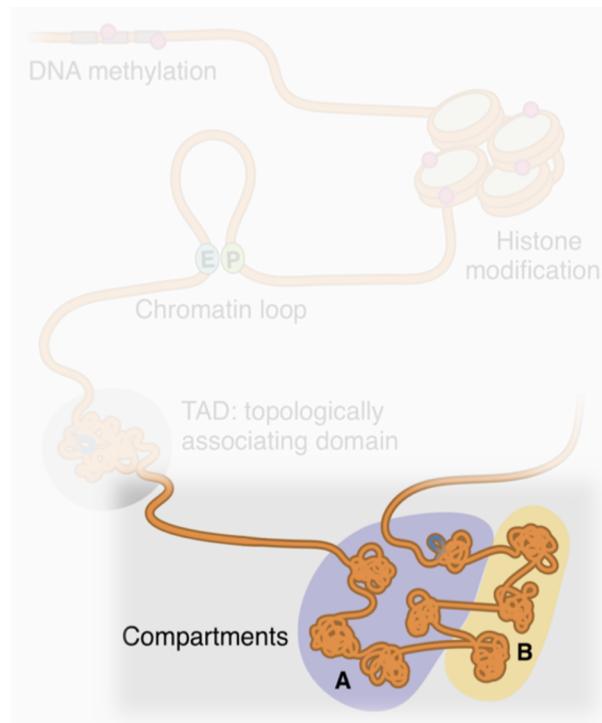


Figures from Johnstone, Reyes, et al. *Cell*, 2020

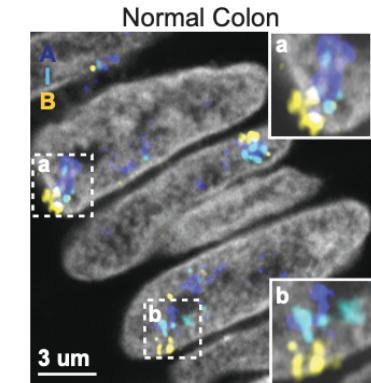
# Hypomethylated A TADs change their interaction preferences towards B



# HiC interaction revealed a novel compartment with intermediate properties



Polymer model

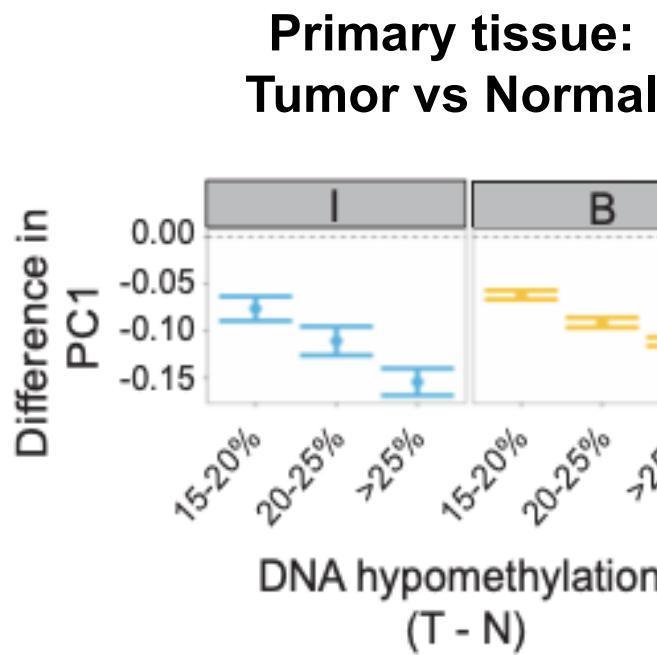


Oligopainting

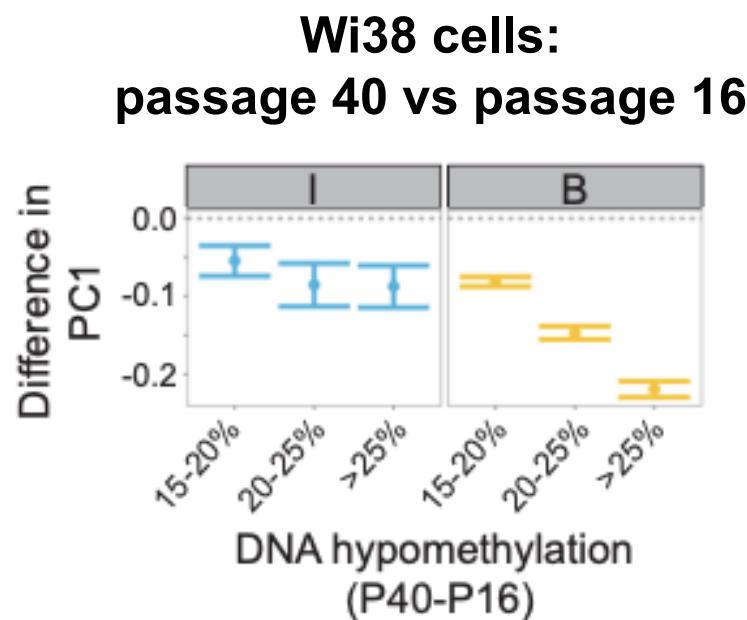
## Compartment I:

- Intermediate interactions with A and B,
- Intermediate gene activity and gene density
- Intermediate nuclear peripheral positioning
- Marked by H3K27me3
- Hypomethylated in tumors
- Changing towards B in tumors

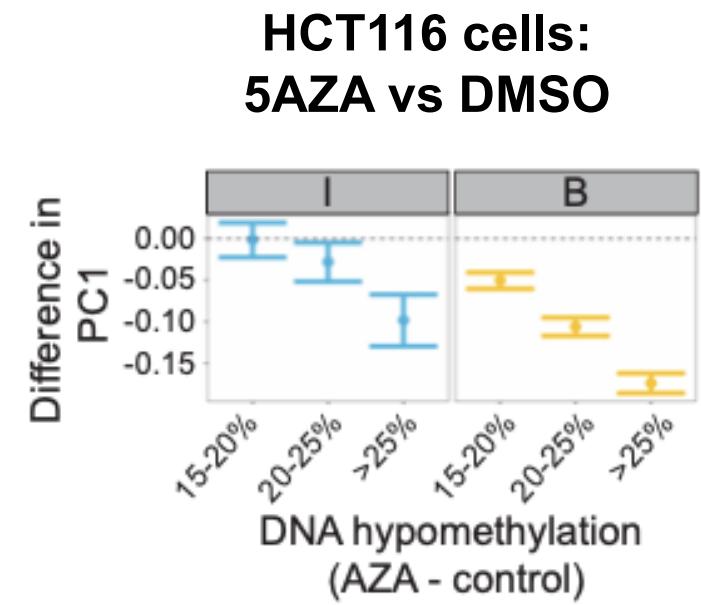
# Topological changes in tumour mirror features of accumulated cell divisions



Tumor changes in methylation and genome compartments



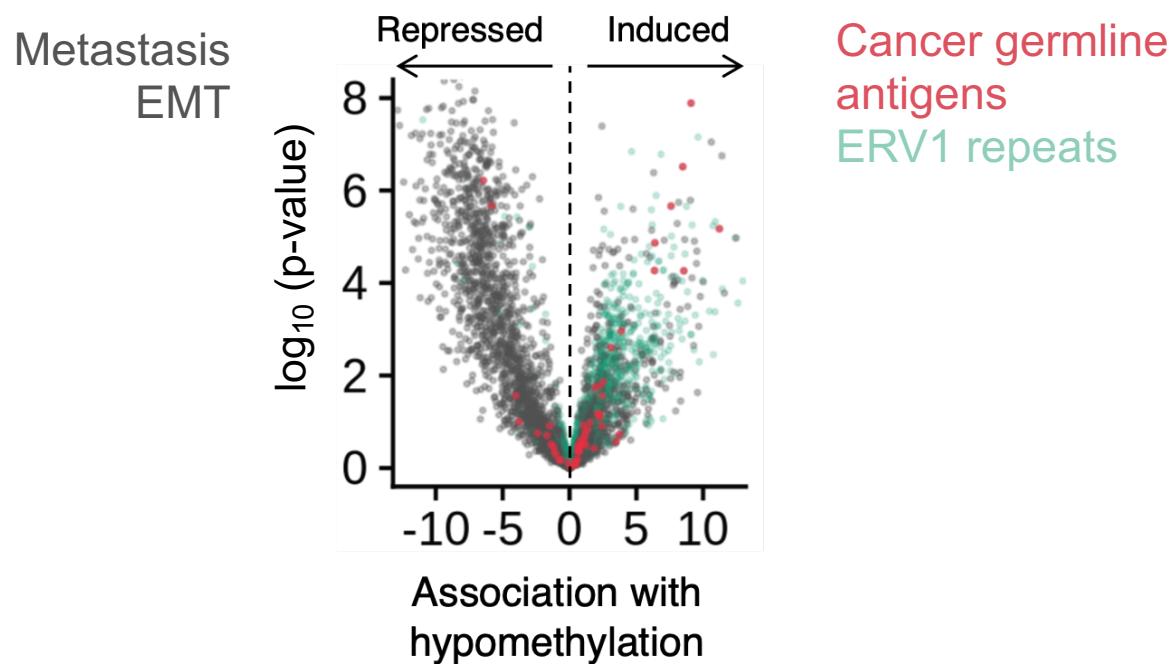
Same pattern in cells with accumulated cell divisions



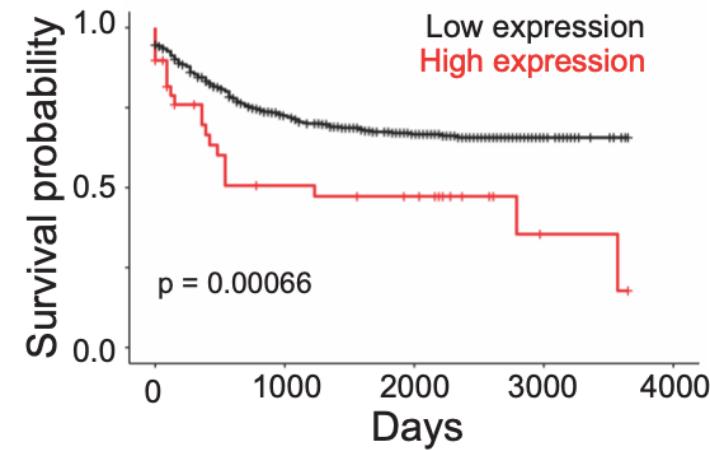
DNA methylation is likely causal of 3D genome changes

# Topological changes induce gene expression programs that restrict tumour progression

TCGA data: hypomethylation as surrogate of topological changes



Score built on genes downregulated with hypomethylation



Reproduced in two independent survival datasets.

# Conclusions

- Enhancer-promoter loops contribute to rewiring of oncogenic expression program.
- TADs are largely stable except for boundaries that are hypermethylated in CIMP tumors.
- Compartments structure changes in association with hypomethylation to induce tumor-supressing gene expression programs.

# Authors of R/BioC packages used in this project

@javrucebo (data.table)  
@marc-outins (data.table)  
Aaron Lun (csw)  
Aaron Lun (DelayedArray)  
Aaron Lun (limma)  
Aaron Lun (SingleCellExperiment)  
Aaron Statham (RepiTools)  
Aaron TL Lun (edgeR)  
Adrian Dragulescu (xlsx)  
Alboukadel Kassambara (ggpubr)  
Alboukadel Kassambara (survminer)  
Alejandro Reyes (GenomeMatrix)  
Alexander Gosdschar (genomation)  
Alicia Oshlack (limma)  
Altuna Akalin (genomation)  
Andrea Riebler (RepiTools)  
Andrew E. Jaffe (minfi)  
Andrew J. Bass (biobrook)  
Andrey Riabushenko (data.table)  
Andrzej Oleś (EBImage)  
Angad P. Singh (PureCN)  
Anthony Damico (data.table)  
Anton Antonov (gridExtra)  
Antonio Colaprico (TCGAbiolinks)  
Brief Gusnanto (CNAnorm)  
Arun Srinivasan (data.table)  
Atkinson Elizabeth (survival)  
Ayappan Perumal (data.table)  
Baptiste Auguie (gridExtra)  
Belinda Phipson (edgeR)  
Belinda Phipson (limma)  
Ben Bolstad (preprocessCore)  
Bernd Klaus (fdrttool)  
Bioconductor (AnnotationHub)  
Bioconductor (BioCParallel)  
Bioconductor (liftOver)  
Bioconductor (TxDb.Hsapiens....)  
Biocore (geneplotter)  
Bozena Mika-Gospodz (genomation)  
Brian Montgomery (matrixStats)  
Caleb Lareau (diffloop)  
Caleb Lareau (sparseHiC)  
Carolyn de Graaf (limma)  
Catharina Olsen (TCGAbiolinks)  
Catherine Hurley (gclus)  
Celine Levy-Leduc (HiCseg)  
Cheng Lee (data.table)  
Christian Weichenberger. (ensemblDB)  
Claudia Cava (TCGAbiolinks)  
Christian Weichenberger. (ensemblDB)  
Claudia Cava (TCGAbiolinks)  
Claus O. Wilke (cowplot)  
Claus O. Wilke (grgrides)  
Cole Arendt (xlsx)  
Cole Miller (data.table)  
Constantin Ahlmann-Eltz (DESeq2)  
Constantin Ahlmann-Eltz (matrixStats)  
contributors worldwide (utils)  
Corporation (doParallel)  
Crownow Cynthia (survival)  
D. Sarkar (GenomicFeatures)  
Dan McGlinn (vegan)  
Dan Tenenbaum (AnnotationHub)  
Dan Tenenbaum (doParallel)  
Dana Seidel (scales)

Daniel Ho (MatchIt)  
Daniel Possenriede (data.table)  
Dario Strbenac (RepiTools)  
David Arenburg (data.table)  
David G. Robinson (biobrook)  
David Hunter (mixtools)  
David Simons (data.table)  
Davide Garolini (TCGAbiolinks)  
Davide Risso (SingleCellExperiment)  
Davis J McCarthy (edgeR)  
Davis McCarthy (limma)  
Davis Vaughan (data.table)  
Declan Groves (data.table)  
Deeqnes Toth (data.table)  
Derek Young (mixtools)  
Di Wu (limma)  
Dianne Cook (biovizBase)  
Dianne Cook (plyranges)  
Didier Chauveau (mixtools)  
Dirk Eddelbuettel (data.table)  
Dmitry Selivanov (data.table)  
Dongcan Jiang (matrixStats)  
Dongseok Choi (limma)  
Douglas Bates (Matrix)  
E. Andres Houseman (minfi)  
Eduard Antonyan (data.table)  
Eduard Szoeccs (vegan)  
Egil Ferkingstad (limma)  
Elliott Sales de Andrade (data.table)  
Emily Nelson (biobrook)  
Erich Neuwirth (RColorBrewer)  
Ethan Smith (data.table)  
Martin Maechler (bitops)  
F. Guillaume Blanchet (vegan)  
F. Hahne (genefilter)  
Felipe Parages (data.table)  
Fengjuan Xuan (mixtools)  
Di Wu (limma)  
Dianne Cook (biovizBase)  
Dianne Cook (plyranges)  
Didier Chauveau (mixtools)  
Dirk Eddelbuettel (data.table)  
Dmitry Selivanov (data.table)  
Dongcan Jiang (matrixStats)  
Dongseok Choi (limma)  
Douglas Bates (Matrix)  
E. Andres Houseman (minfi)  
Eduard Antonyan (data.table)  
Eduard Szoeccs (vegan)  
Egil Ferkingstad (limma)  
Elliott Sales de Andrade (data.table)  
Emily Nelson (biobrook)  
Erich Neuwirth (RColorBrewer)  
Ethan Smith (data.table)  
Martin Maechler (bitops)  
F. Guillaume Blanchet (vegan)  
F. Hahne (genefilter)  
Felipe Parages (data.table)  
Fengjuan Xuan (mixtools)  
Francois Cocquemais (data.table)  
Francois Pepin (biomaRt)  
Francois Pepin (limma)  
Gavin L. Simpson (vegan)  
Gianluca Bontempi (TCGAbiolinks)  
Gordon K Smyth (Rsubread)

Gordon K Smyth (edgeR)  
Gordon Smyth (csw)  
Gordon Smyth (limma)  
Gregoire Pau (EBImage)  
H. Pagès (GenomicFeatures)  
H. Pagès (GenomicRanges)  
Hadley Wickham (dplyr)  
Hadley Wickham (grid)  
Hadley Wickham (magrittr)  
Hadley Wickham (plyr)  
Hadley Wickham (pryr)  
Hadley Wickham (readr)  
Hadley Wickham (reshape2)  
Hadley Wickham (rlang)  
Hadley Wickham (scales)  
Hadley Wickham (stringr)  
Hadley Wickham (tidy)  
Hajk-Georg Drost (philentropy)  
Harris Jaffee (matrixStats)  
Hector Corrada Bravo (matrixStats)  
Helena Wagner (vegan)  
Henrik Bengtsson (matrixStats)  
Henry M. Wood (CNAnorm)  
Herv'e Pag' es (Rsamttools)  
Hervé Pagès (DelayedArray)  
Hervé Pagès (GenomicAlignments)  
Hervé Pagès (HDF5Array)  
Hervé Pagès (SummarizedExperiment)  
Hoben Thomas (mixtools)  
Houtan Noushmehr (TCGAbiolinks)  
Hugh Parsonage (data.table)  
Idris Eckley (changePoint)  
Isabella Castiglioni (TCGAbiolinks)  
Jake Conway (UpSetR)  
James Sams (data.table)  
James Wettenhall (limma)  
Jamie Lee (changePoint)  
Jan Gorecki (data.table)  
Jari Oksanen (vegan)  
Jean-Philippe Fortin (minfi)  
Jens Peder Meldgaard (data.table)  
Jenny Dai (Rsubread)  
Jeremy Silver (limma)  
Jim Hester (data.table)  
Jim Hester (readr)  
Jim Hester (SummarizedExperiment)  
Johannes Rainer (biovizBase)  
Johannes Rainer (EnsDb.Hsapiens.v75)  
Johannes Rainer (ensemblDB)  
John D. Storey (biobrook)  
Jonathan Sidi (gedit)  
Joseph Barry (EBImage)  
Jovana Maksimovic (minfi)  
Kamil Slowikowski (grgrides)  
Karl Broman (data.table)  
Kasper Daniel Hansen (bsseq)  
Kasper Daniel Hansen (minfi)  
Kasper Daniel Hansen (minfiData)  
Katarzyna Wreczcka (genomation)  
Kayla Morrell (AnnotationHub)  
Kaylea Haynes (changePoint)  
Keegan Korthauer (SingleCellExperiment)  
Kevin Rue-Albrecht (SingleCellExperiment)  
Kevin Ushey (data.table)  
Korbinian Strümmer (fdrttool)  
Kun Ren (data.table)

Kurt Hornik (clue)  
Laurent Gatto (biobrook)  
Laurent Gatto (ensemblDB)  
Leonardo Collado-Torres (recount)  
Leonardo Silvestri (data.table)  
Lionel Henry (magrittr)  
Lionel Henry (rlang)  
Liz Ing-Simmons (genomation)  
Lori Shepherd (AnnotationHub)  
Luciano Garofano (TCGAbiolinks)  
M. Carlson (GenomicFeatures)  
M. Henry H. Stevens (vegan)  
M. Lawrence (GenomicFeatures)  
M. Lawrence (GenomicRanges)  
M. Morgan (GenomicFeatures)  
Manish Saraswat (data.table)  
Marc Carlson (AnnotationHub)  
Marc Carlson (org.Hs.eg.db)  
Marc Carlson  
(TxDb.Hsapiens.UCSC.hg19.knownGene)  
Marcel Ramos (RaggedExperiment)  
Marcin Kosinski (survminer)  
Marcus Davy (limma)  
Mark D. Robinson (edgeR)  
Mark Robinson (RepiTools)  
Markus Bonsch (data.table)  
Markus Riester (PureCN)  
Martin Aryee (diffloop)  
Martin Aryee (minfi)  
Martin Aryee (minfiData)  
Martin Aryee (sparseHiC)  
Martin Maechler (Matrix)  
Martin Morgan (AnnotationHub)  
Martin Morgan (BioCParallel)  
Martin Morgan (data.table)  
Martin Morgan (GenomicAlignments)  
Martin Morgan (RaggedExperiment)  
Martin Morgan (Rsamttools)  
Martin Morgan (SummarizedExperiment)  
Matt Dowd (data.table)  
Matthew E. Ritchie (edgeR)  
Matthew Ritchie (limma)  
Matthew Young (goseq)  
Matthieu Gomez (data.table)  
Mette Langaas (limma)  
Michael Chirico (data.table)  
Michael Friendly (vegan)  
Michael Lawrence (biovizBase)  
Michael Lawrence (plyranges)  
Michael Lawrence (tracklayer)  
Michael Love (DESeq2)  
Michael Quinn (data.table)  
Michael Schubmehl (data.table)  
Michel Berkelaar (ipSolve)  
Michel Lang (BioCParallel)  
Michel Lang (data.table)  
Michèle Ceccarelli (TCGAbiolinks)  
Michelle Wallig (doParallel)  
Michelle Wallig (foreach)  
Mike Smith (biomaRt)  
Mike Smith (EBImage)  
Mingxiang Teng (gcapc)  
Morgan Jacob (data.table)  
Mus Yaramaz-David (data.table)  
Natalie Thorne (limma)  
Nathaniel Hayden (Rsamttools)  
Nello Blaser (data.table)

Nikos Ignatiadis (IHW)  
Nils Gehlenborg (UpSetR)  
Nitesh Turaga (BioCParallel)  
Ola Hossjer (matrixStats)  
Oleg Sklyar (EBImage)  
others (ipSolve)  
Otto Seiskari (data.table)  
P. Aboyoun (GenomicFeatures)  
P. Aboyoun (GenomicRanges)  
Pasha Stetsenko (data.table)  
Paul Fearnhead (changePoint)  
Paul Murrell (grid)  
Peter F. Hickey (minfi)  
Peter Hickey (bsseq)  
Peter Hickey (DelayedArray)  
Peter Hickey (matrixStats)  
Peter Langfelder (matrixStats)  
Peter R. Minchin (vegan)  
Peter Solymos (vegan)  
Philip A. Marais (EBImage)  
Philippe Chataignon (data.table)  
Pierre Gramme (grgrides)  
Pierre Legendre (vegan)  
Przemyslaw Biecek (survminer)  
R Core team (pryr)  
R Core Team (readr)  
R Core Team (utils)  
R. B. O'Hara (vegan)  
R. Gentleman (genefilter)  
R. Gentleman (geneplotter)  
Rafael A. Iribarry (gcpc)  
Rafael A. Iribarry (minfi)  
Raivo Kolde (pheatmap)  
Rebecca Killick (changePoint)  
Richard Heiberger (abind)  
Rick Saporta (data.table)  
Robert Gentleman (matrixStats)  
Robert Gentleman (tracklayer)  
Roeland Kindt (vegan)  
Romain Francois (readr)  
Roy Storey (data.table)  
RStudio (dplyr)  
RStudio (ggplot2)  
RStudio (magrittr)  
RStudio (rlang)  
RStudio (scales)  
RStudio (stringr)  
RStudio (tidy)  
Ryan Elmore (mixtools)  
Ryan Thompson (BioCParallel)  
Steve Dutky (bitops)  
S. Falcon (GenomicFeatures)  
Scheipl Fabian (survminer)  
Scott Ritchie (data.table)  
Sean Davis (biomaRt)  
Sean Davis (GOQuery)  
Sebastian Freundt (data.table)  
Sebastian Gibb (ensemblDB)  
Seth Wenzel (data.table)  
Shan V. Andrews (minfi)  
Simon Anders (DESeq2)  
Simon Urbanek (fastmatch)  
Sonali Arora (AnnotationHub)  
Stefan Milton Baché (magrittr)  
Stefano Berri (CNAnorm)  
Stefano M. Pagnotta (TCGAbiolinks)  
Steffen Durinck (biomaRt)

Steve Dutky (bitops)  
Steve Lianoglou (biobrook)  
Steve Lianoglou (data.table)  
Steve Weston (doParallel)  
Steve Weston (foreach)  
Stuart Lee (plyranges)  
Tal Galili (dendextend)  
Tathiane Malta (TCGAbiolinks)  
Tatiana Benaglia (mixtools)  
Tengfei Yin (biovizBase)  
Terry M Therneau (survival)  
Thais Sabedot (TCGAbiolinks)  
The Bioconductor Dev Team (BSgenome...)  
Thomas Hettmansperger (mixtools)  
Thomas Lumley (survival)  
Tiago Chedraoui Silva (TCGAbiolinks)  
Tim Triche (ensemblDB)  
Tim Triche (minfi)  
Tobias Schmidt (data.table)  
Toby Hocking (data.table)  
Tom Short (data.table)  
Tony Plate (abind)

Tyson Barrett (data.table)  
V. Carey (genefilter)  
V. Obenchain (GenomicFeatures)  
Vaclav Tlapak (data.table)  
Valerie Obenchain (BioCParallel)  
Valerie Obenchain (GenomicAlignments)  
Valerie Obenchain (Rsamttools)  
Valerie Obenchain (SummarizedExperiment)  
Valerie Obenchain (AnnotationHub)  
Vedran Franke (genomation)  
Vince Carey (rtracklayer)  
Vince S Buffalo (biomaRt)  
VJ Carey (gwascat)  
W. Huber (genefilter)  
Walter Böhm (clue)  
Watal Iwasaki (data.table)  
Wei Shi (limma)  
Wei Shi (Rsubread)  
Willem Litjengberg (reactome.db)  
Winston Timp (minfiData)  
Wolfgang Huber (biomaRt)  
Wolfgang Huber (DESeq2)  
Wolfgang Huber (EBImage)  
Wolfgang Huber (IHW)  
Xianghui Dong (data.table)  
Xianying Tan (data.table)  
Xiaobei Zhou (edgeR)  
Yang Liu (Rsubread)  
Yifang Hu (edgeR)  
Yifang Hu (limma)  
Yunshun Chen (edgeR)  
Yunshun Chen (limma)

Yutani Hiroaki (grgrides)  
Zuguang Gu (circize)  
Zuguang Gu (ComplexHeatmap)  
Zuguang Gu (FastStringSimilarity)  
Zuguang Gu (GO)

# Acknowledgements

**Sarah E. Johnstone (MGH)**

Yifeng Qi (MIT)

Carmen Adriaens (MGH)

Esmat Hegazi (MGH)

Karin Pelka (MGH)

Jonathan Chen (MGH)

Luli Zou (HSPH)

Yotam Drier (MGH)

Vivian Hecht (Broad)

Noam Shoresh (Broad)

Martin K. Selig (MGH)

Caleb Lareau (Broad)

Sowmya Iyer (MGH)

Son C. Nguyen (UPenn)

**Bradley E. Bernstein (MGH)**

**Martin J. Aryee (MGH)**

**Rafael Irizarry (DFCI)**

Bin Zhang (MIT)

Nir Hacohen (MGH)

Eric F. Joyce (Upenn)

**Thank you**