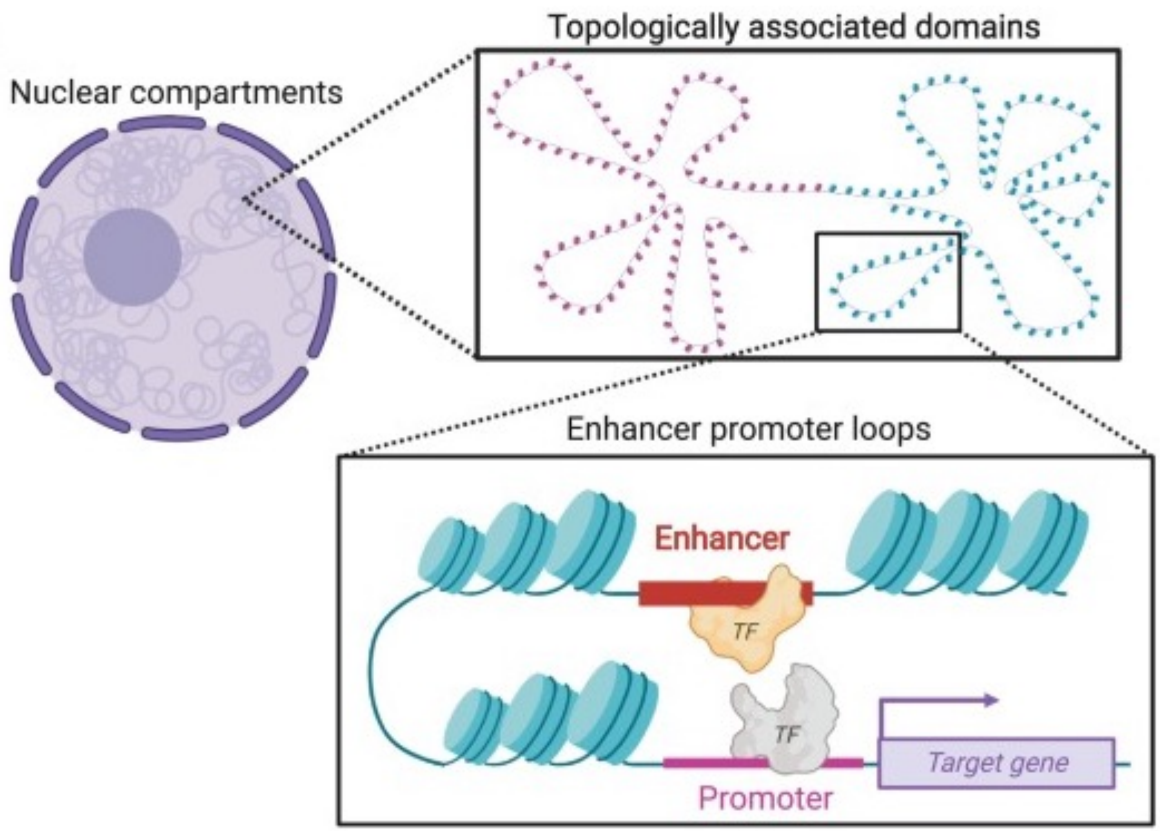
Joint tensor modeling of single cell 3D genome and epigenetic data with Muscle

Kwangmoon Park, Sündüz Keleş
Department of Statistics
University of Wisconsin-Madison

BIRS 2023, Jul 3, 2023

scHi-C data

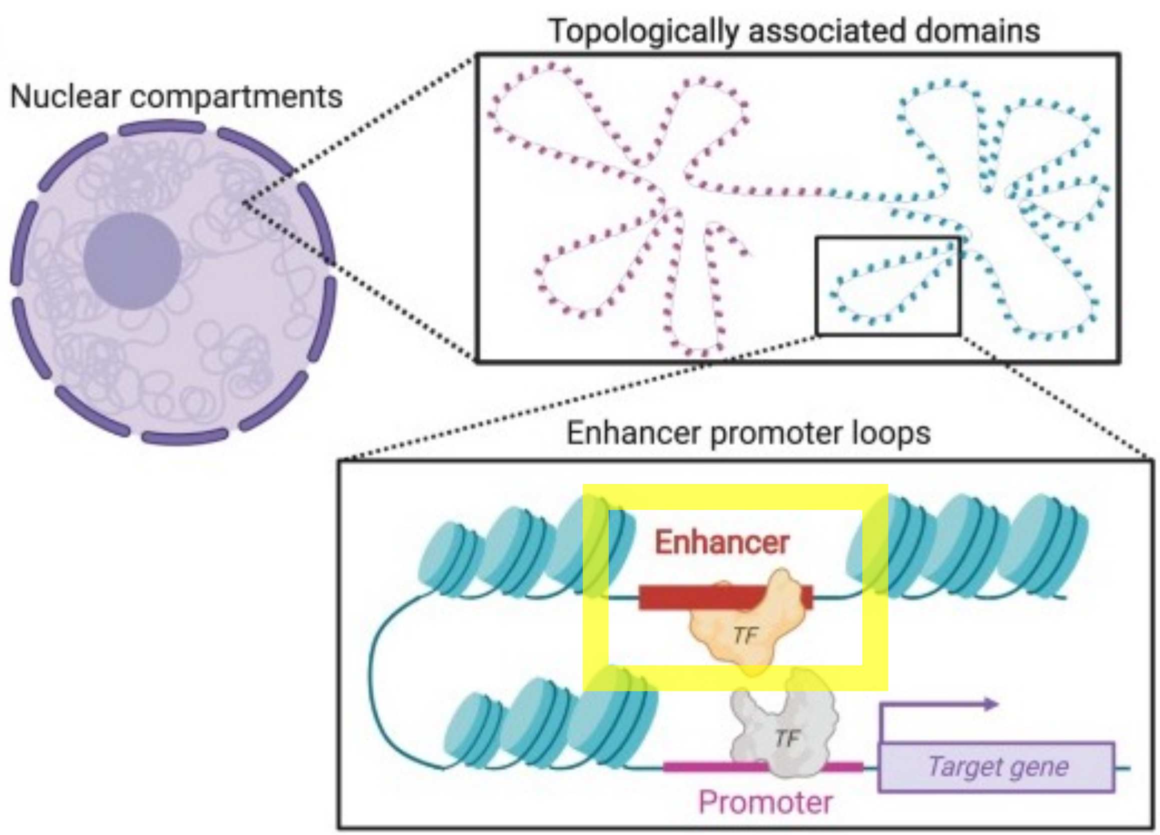
- Non-coding regions (99%) → Genes (1%)



Claringbould et al.

scHi-C data

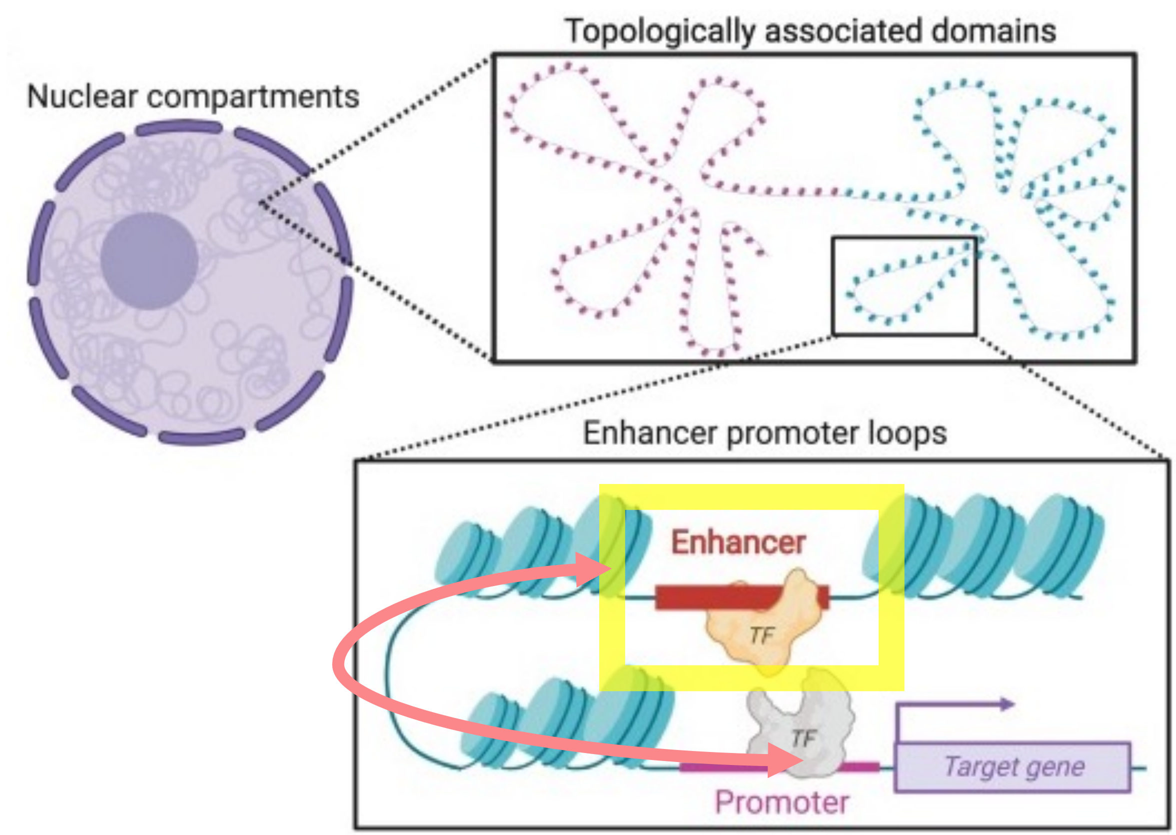
- Non-coding regions (99%) → Genes (1%)
- Far apart in **genomic** distance, but **physically** close



Claringbould et al.

scHi-C data

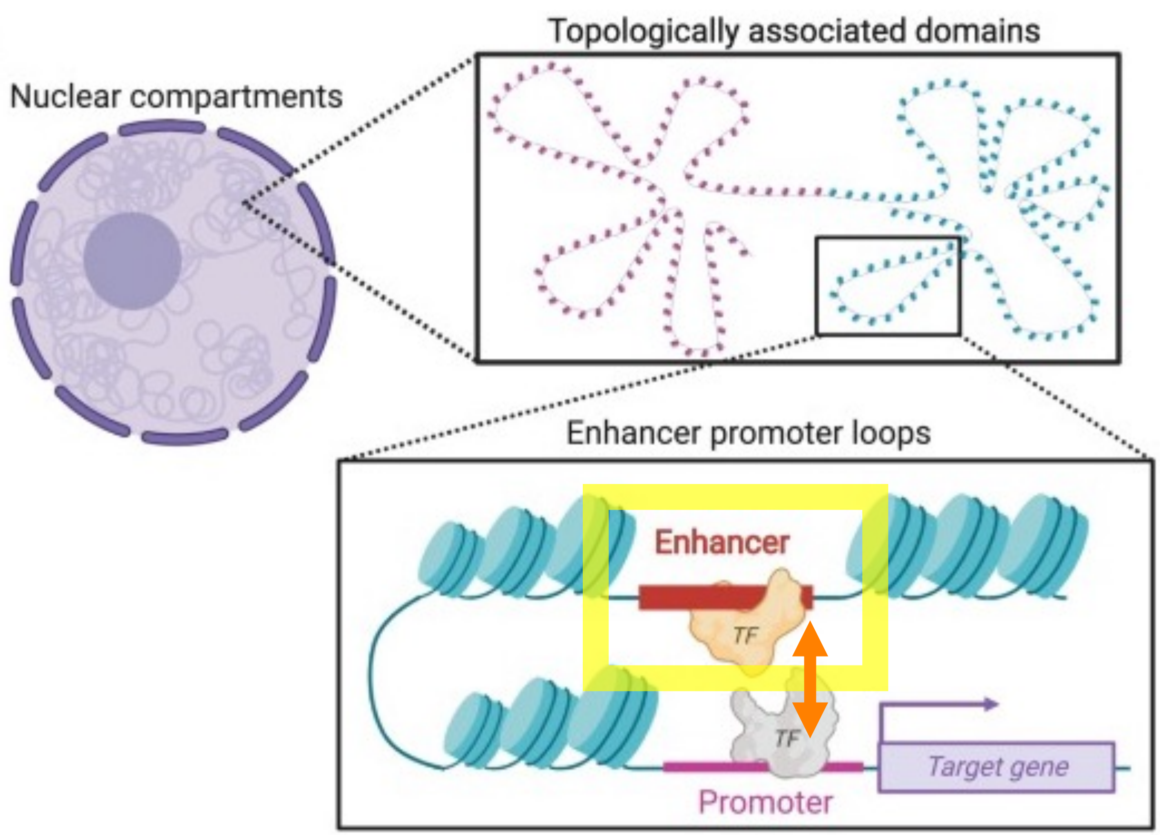
- Non-coding regions (99%) → Genes (1%)
- Far apart in **genomic** distance, but **physically** close



Claringbould et al.

scHi-C data

- Non-coding regions (99%) → Genes (1%)
- Far apart in **genomic** distance, but **physically** close



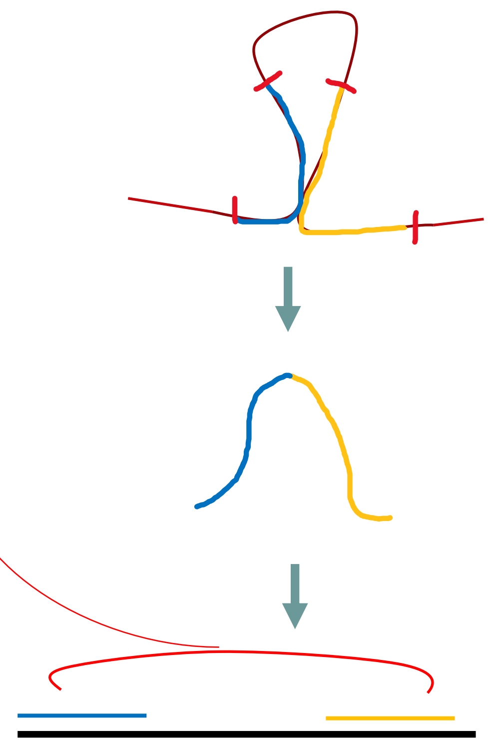
Claringbould et al.

scHi-C data

- Hi-C : Genome-wide physical 3D contact level
- Contact map ~ adjacency matrix (weighted graph)

Loci

Loci					
	5	1	0	...	0
	1	3	1	...	0
	0	1	⋮		⋮
	⋮	⋮	⋮		⋮
	0	0	0	...	1



scHi-C data

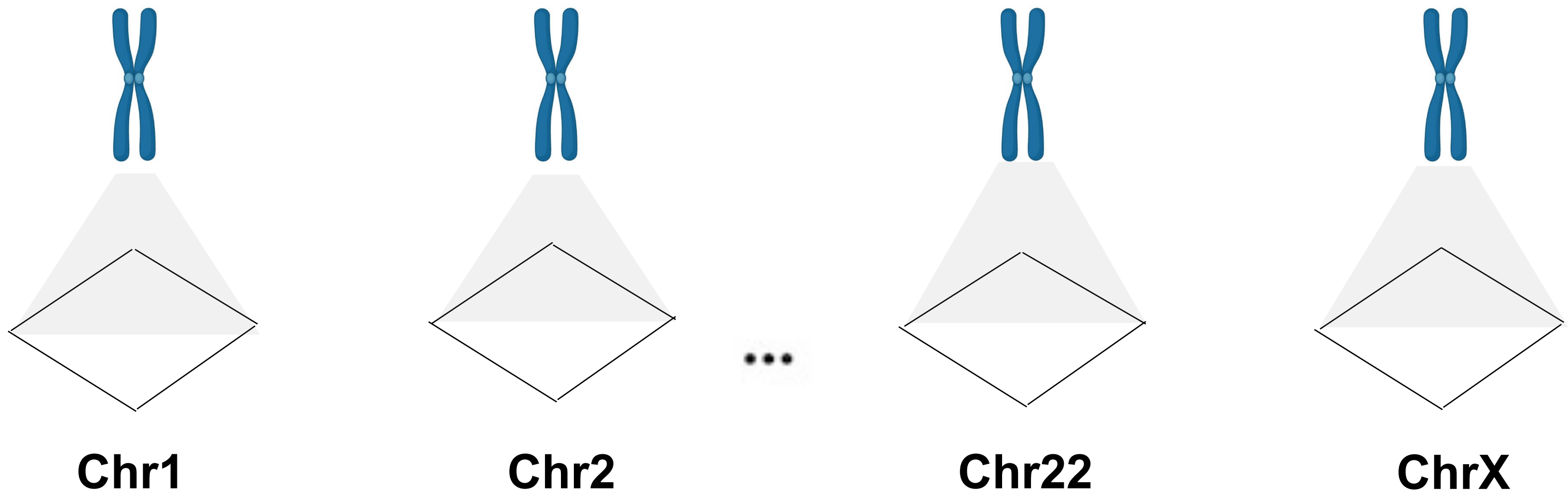
- Bulk Hi-C : 23 (or 20) matrices

...

Chr1 **Chr2** **Chr22** **ChrX**

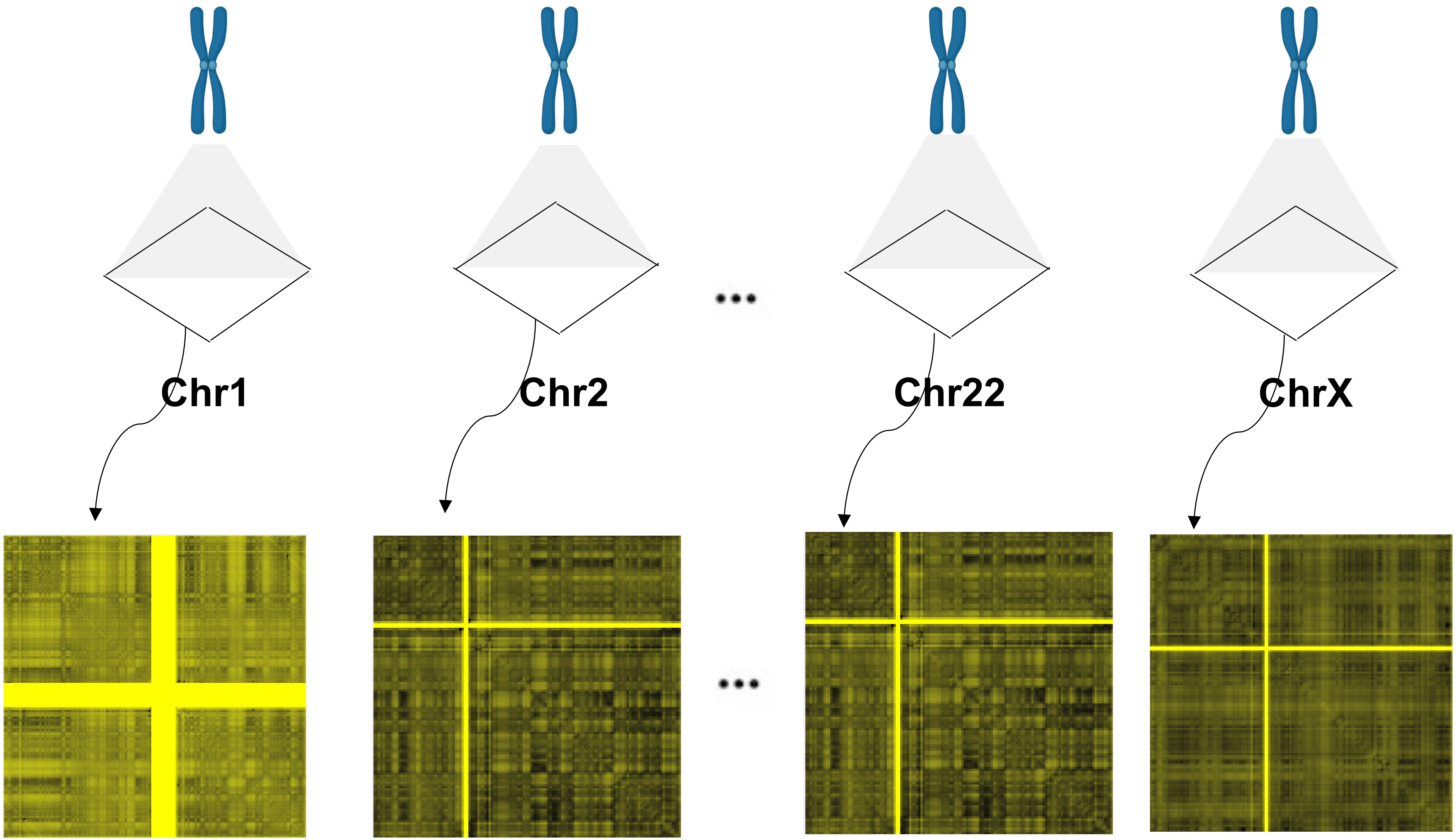
scHi-C data

- Bulk Hi-C : 23 (or 20) matrices



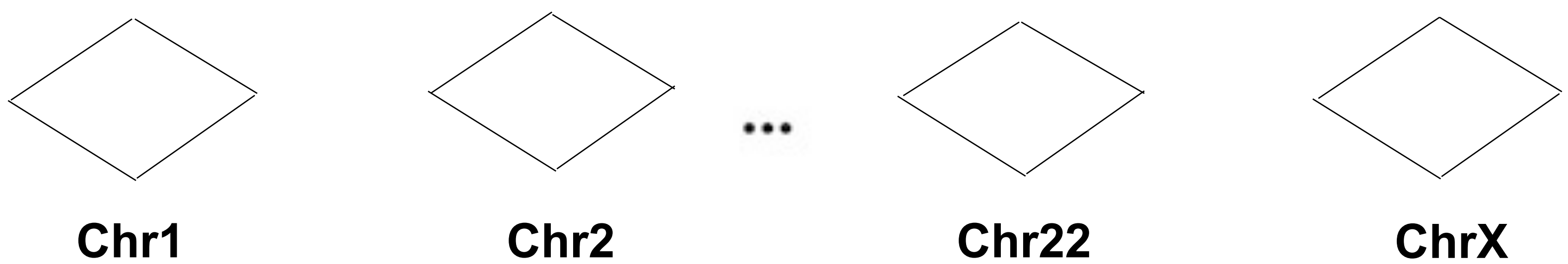
scHi-C data

- Bulk Hi-C : 23 (or 20) matrices



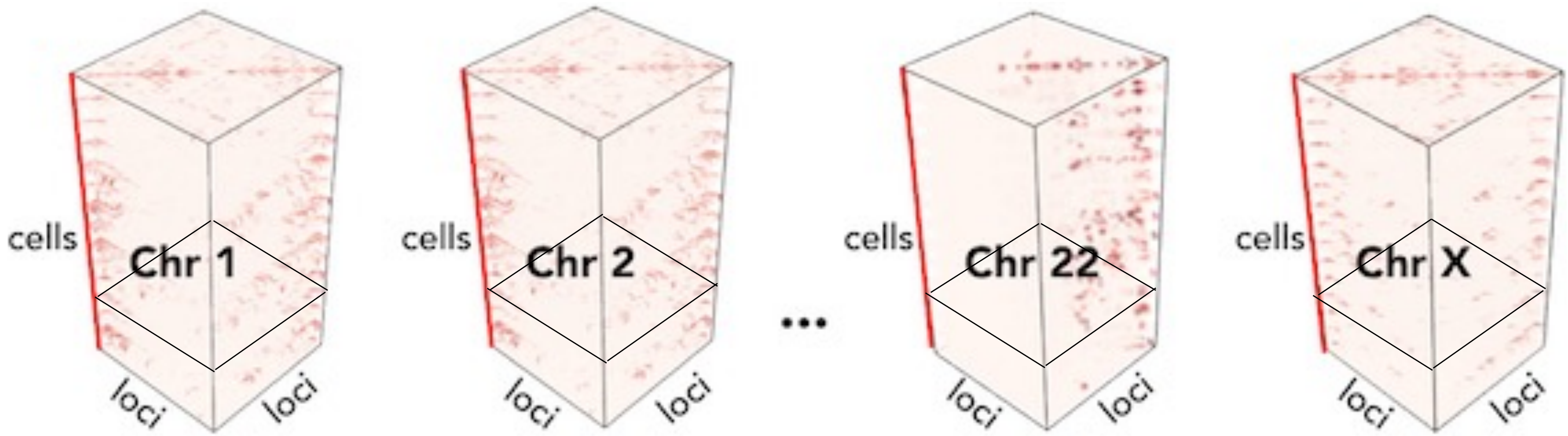
scHi-C data

- Bulk Hi-C : 23 (or 20) matrices



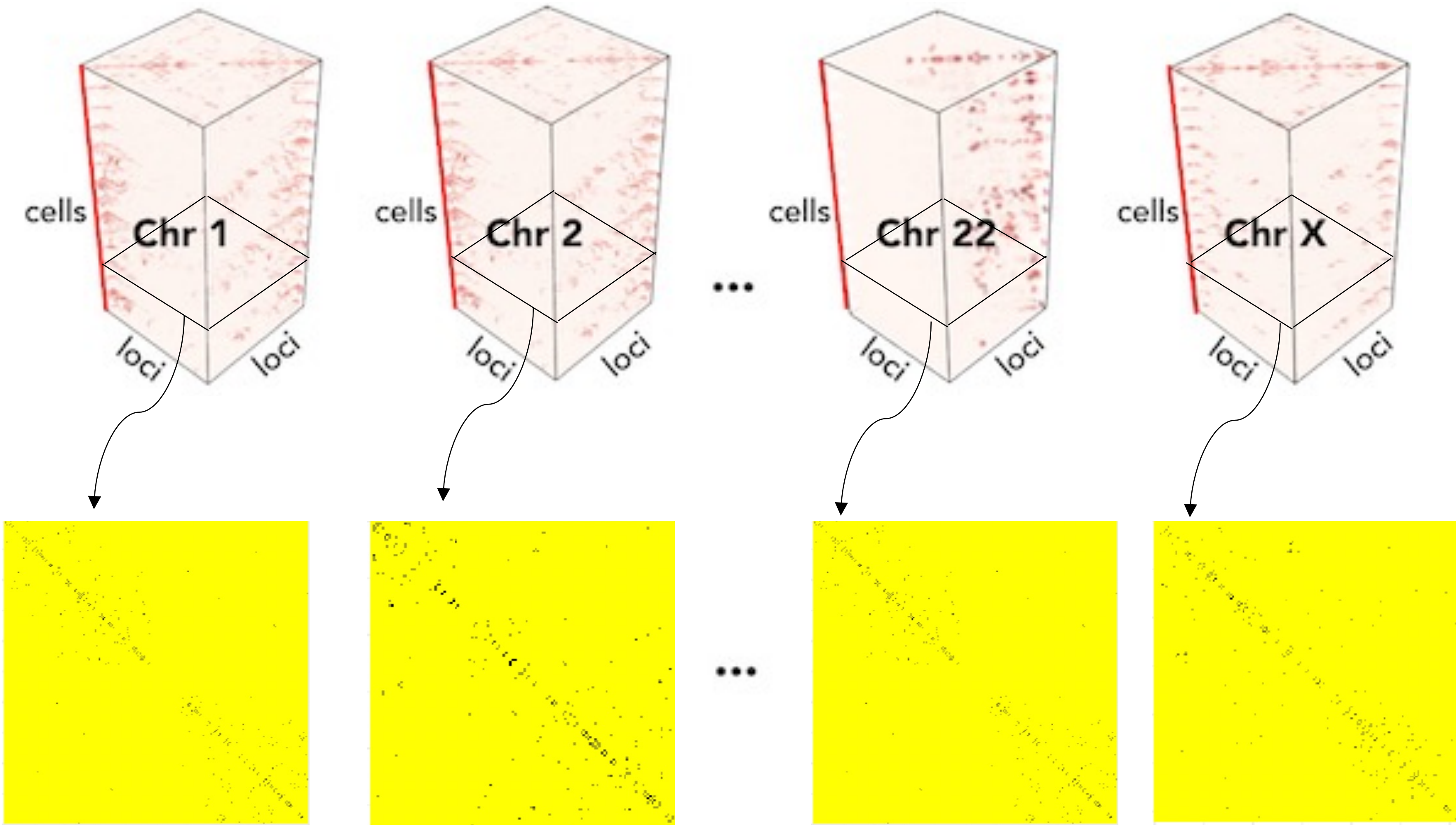
scHi-C data

- scHi-C : 23 (or 20) tensors



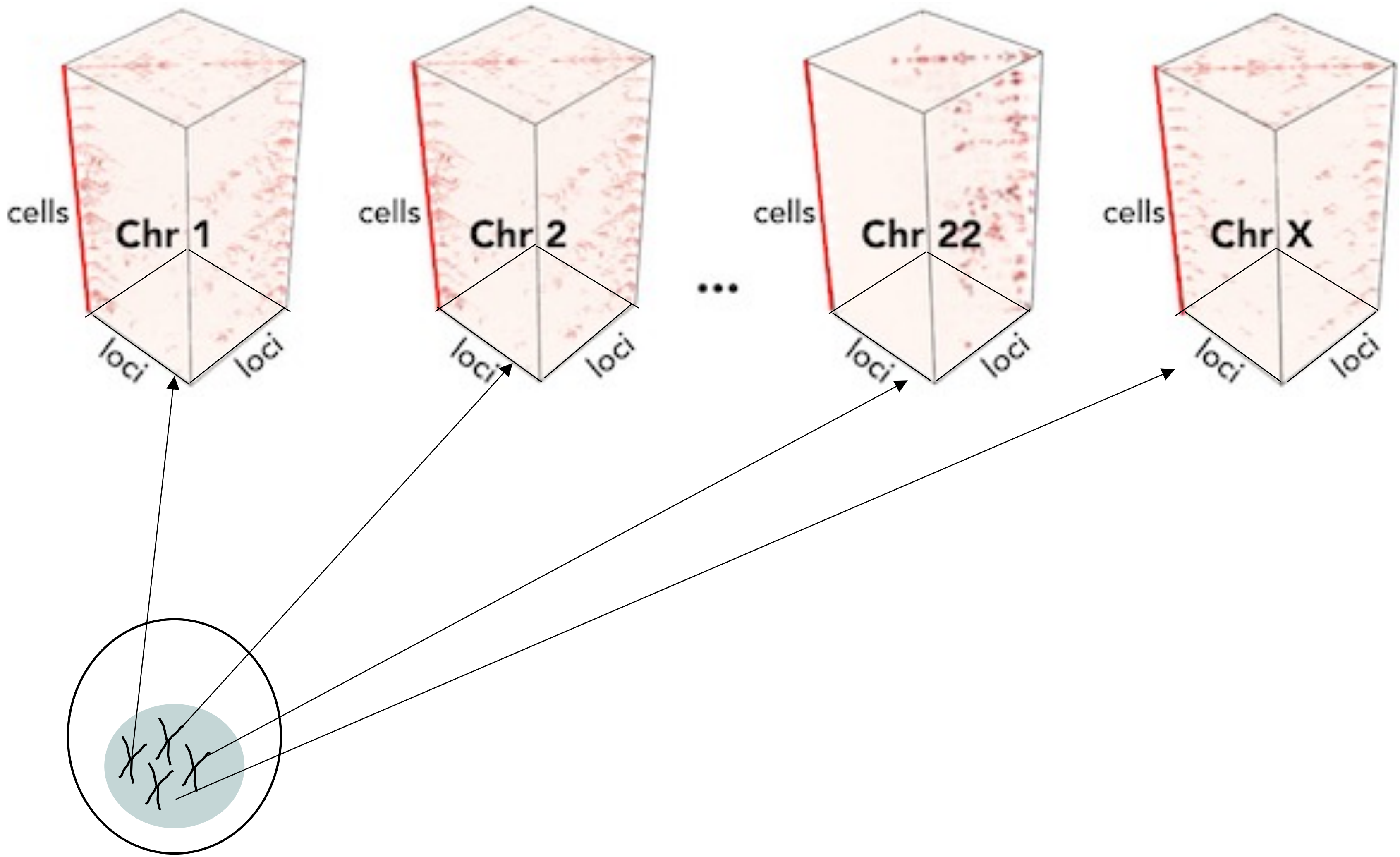
scHi-C data

- scHi-C : 23 (or 20) tensors



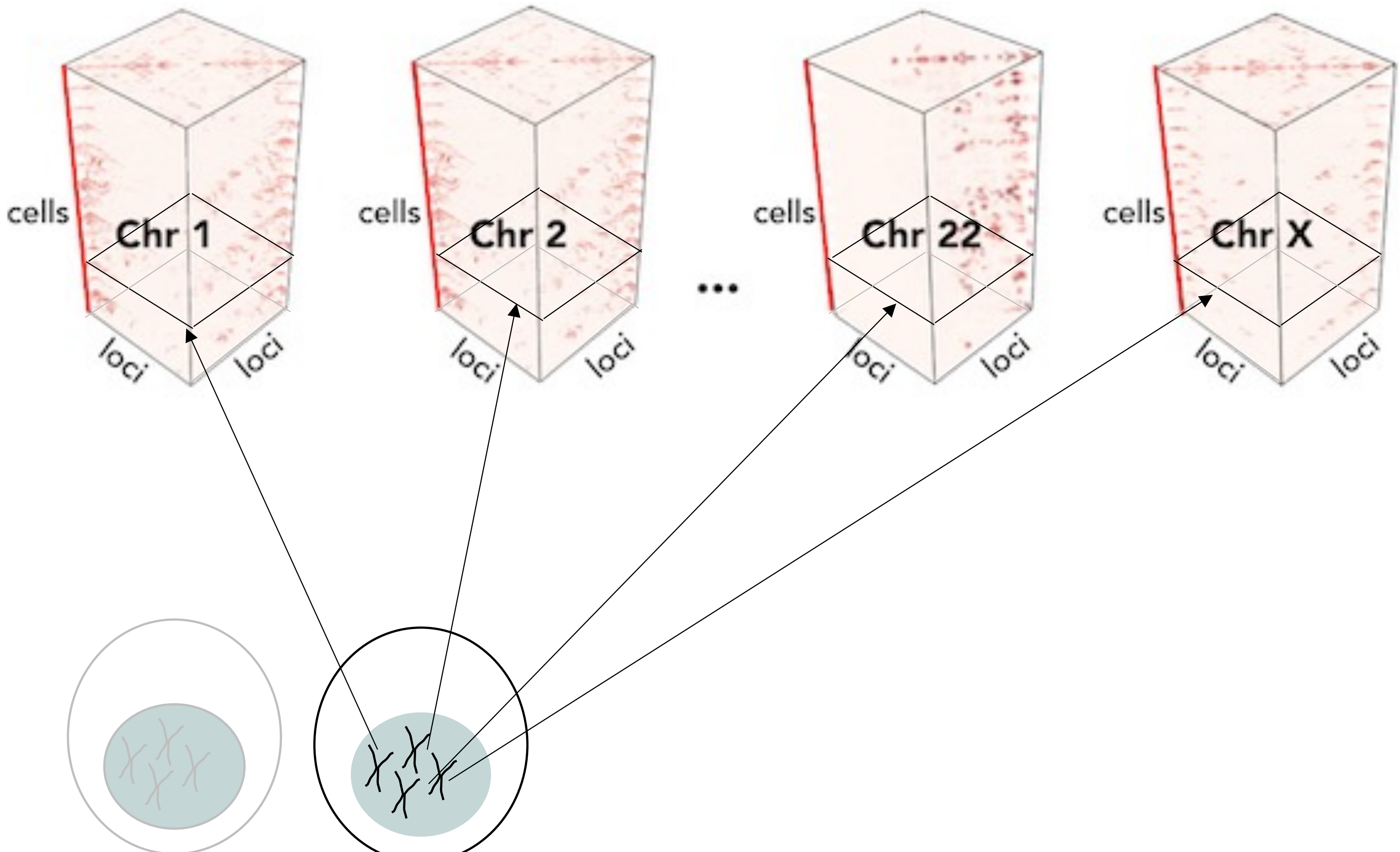
scHi-C data

- scHi-C : 23 (or 20) tensors



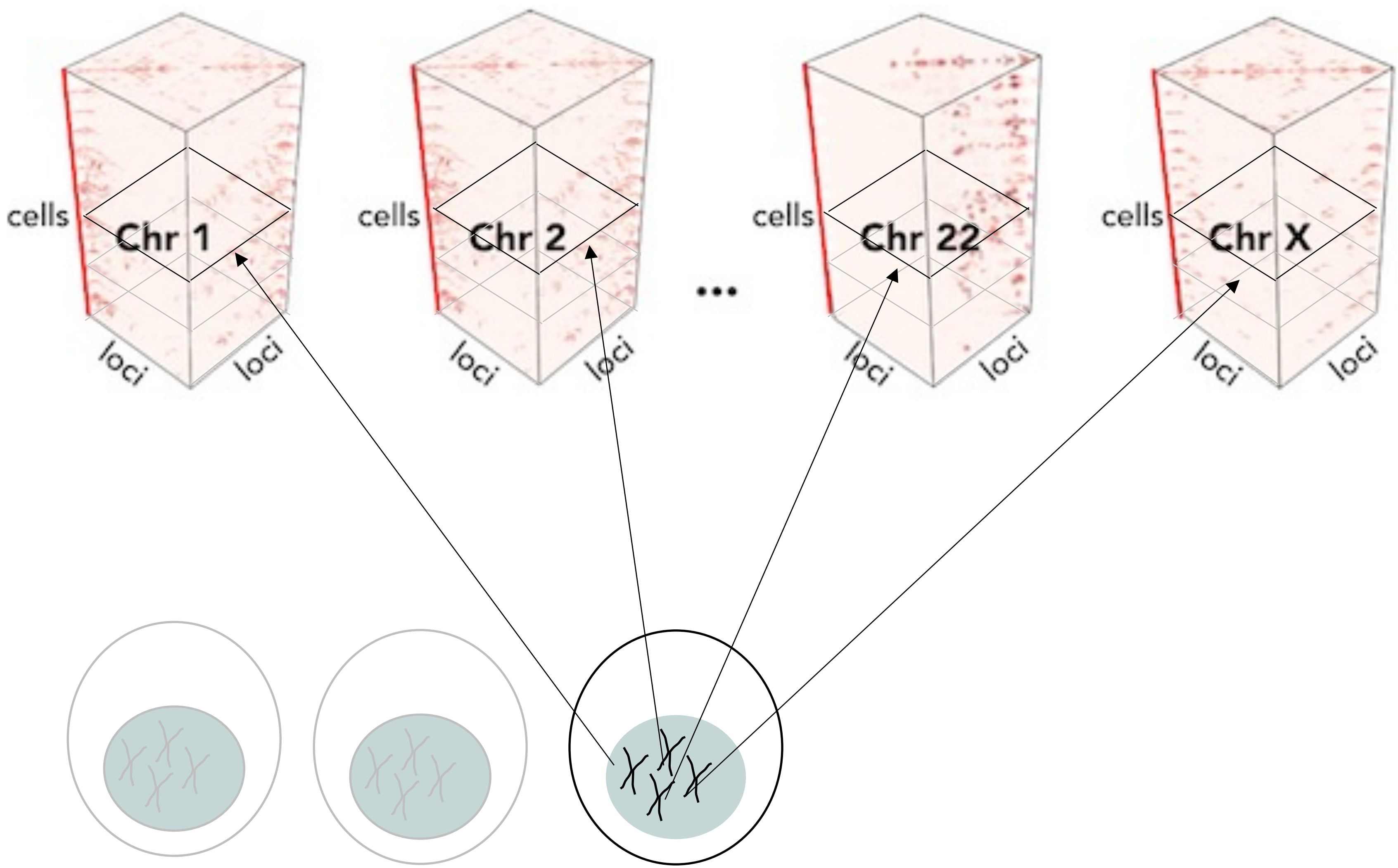
scHi-C data

- scHi-C : 23 (or 20) tensors



scHi-C data

- scHi-C : 23 (or 20) tensors



Main objectives

- **Cell type** specific **3D genome structure**!

  - Cell type clustering
  - Contact map of each cell type (TAD / AB comp)
  - Multi modal data integration including Hi-C

- Tool : **tensor/matrix** decomposition

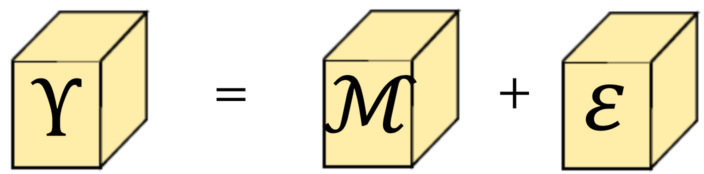
Contributions

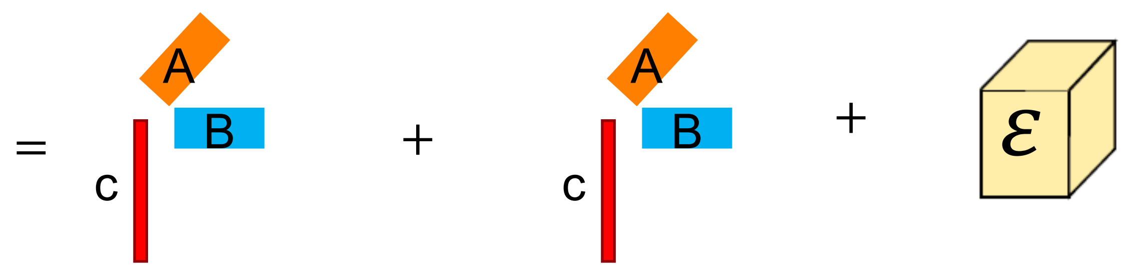
- **Methodology**
  - Joint decomposition of multiple tensor objects
  (**Common** parameter / Semi-**nonnegative** Tensor / Data balancing)
  - Optimality properties for the **A**lternating **L**east **S**quares algorithm
- **Interpretation**
  - **Unification** of estimation target and parameter of interests
    (e.g., Mean contact pattern, cell type information)
  - **Direct** / does not require complex modification on parameters

Block Term tensor Decomposition (BTD)

- Estimation of signal tensor($\mathcal{M}$) with the existence of noise

(Chr=1, R=2 case)



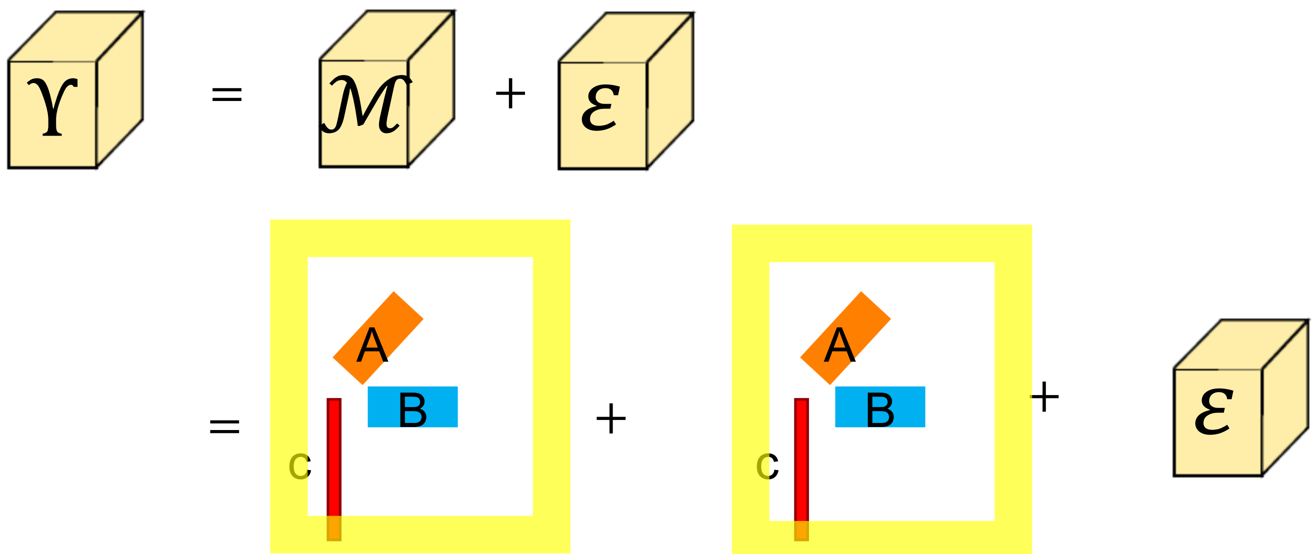


$$\begin{aligned}\mathcal{Y} &= \mathcal{M} + \mathcal{E} \\ &= \sum_{r=1}^{R} (A_r B_r^T) \circ c_r + \mathcal{E},\end{aligned}$$

Block Term tensor Decomposition (BTD)

- Estimation of signal tensor($\mathcal{M}$) with the existence of noise

(Chr=1, R=2 case)

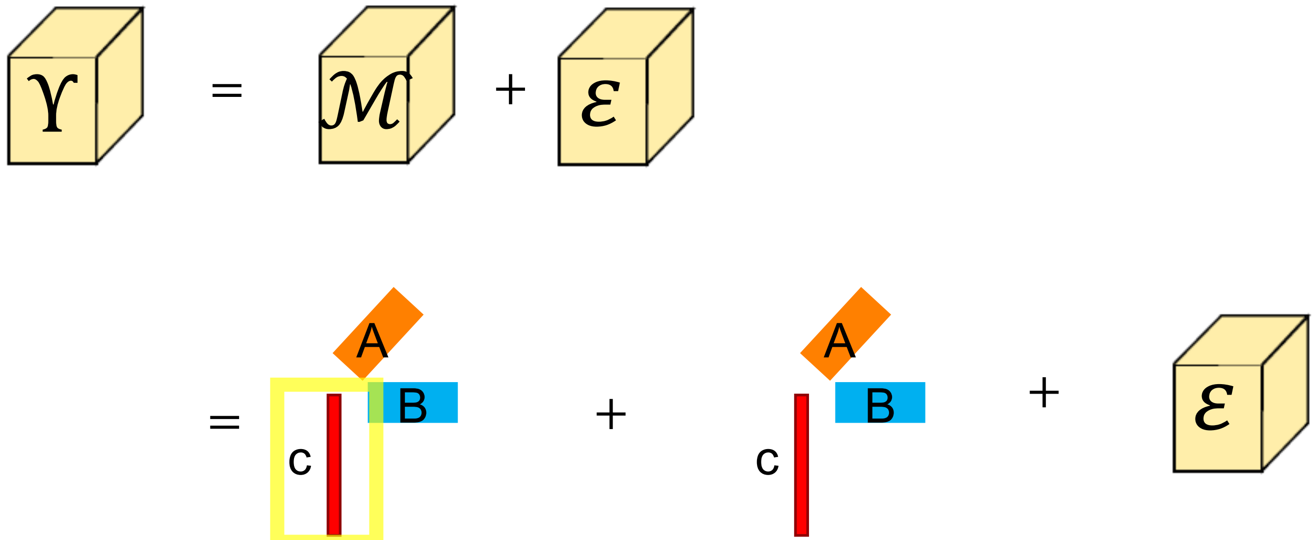


$$\begin{aligned}\mathcal{Y} &= \mathcal{M} + \mathcal{E} \\ &= \sum_{r=1}^{R} (A_r B_r^T) \circ c_r + \mathcal{E},\end{aligned}$$

Block Term tensor Decomposition (BTD)

- Estimation of signal tensor($\mathcal{M}$) with the existence of noise

(Chr=1, R=2 case)

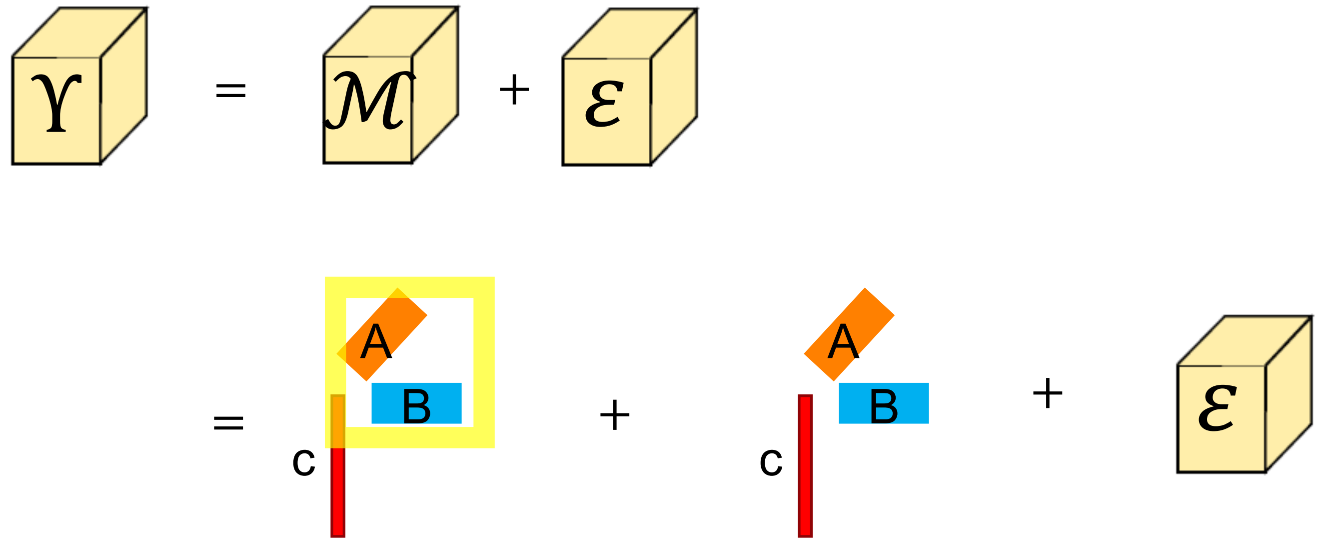


$$
\begin{aligned}
\mathcal{Y} &= \mathcal{M} + \mathcal{E} \\
&= \sum_{r=1}^{R} (A_r B_r^T) \circ c_r + \mathcal{E},
\end{aligned}
$$

Block Term tensor Decomposition (BTD)

- Estimation of signal tensor($\mathcal{M}$) with the existence of noise

  (Chr=1, R=2 case)

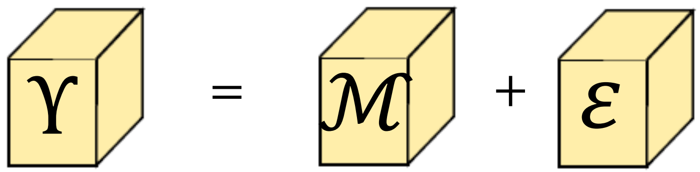


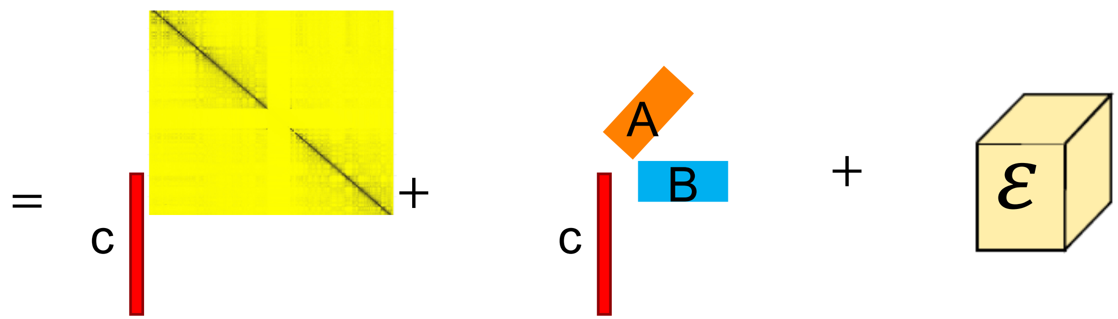
$$\mathcal{Y} = \mathcal{M} + \mathcal{E}$$
$$= \sum_{r=1}^{R} (A_r B_r^T) \circ c_r + \mathcal{E},$$

Block Term tensor Decomposition (BTD)

- Estimation of signal tensor($\mathcal{M}$) with the existence of noise

(Chr=1, R=2 case)

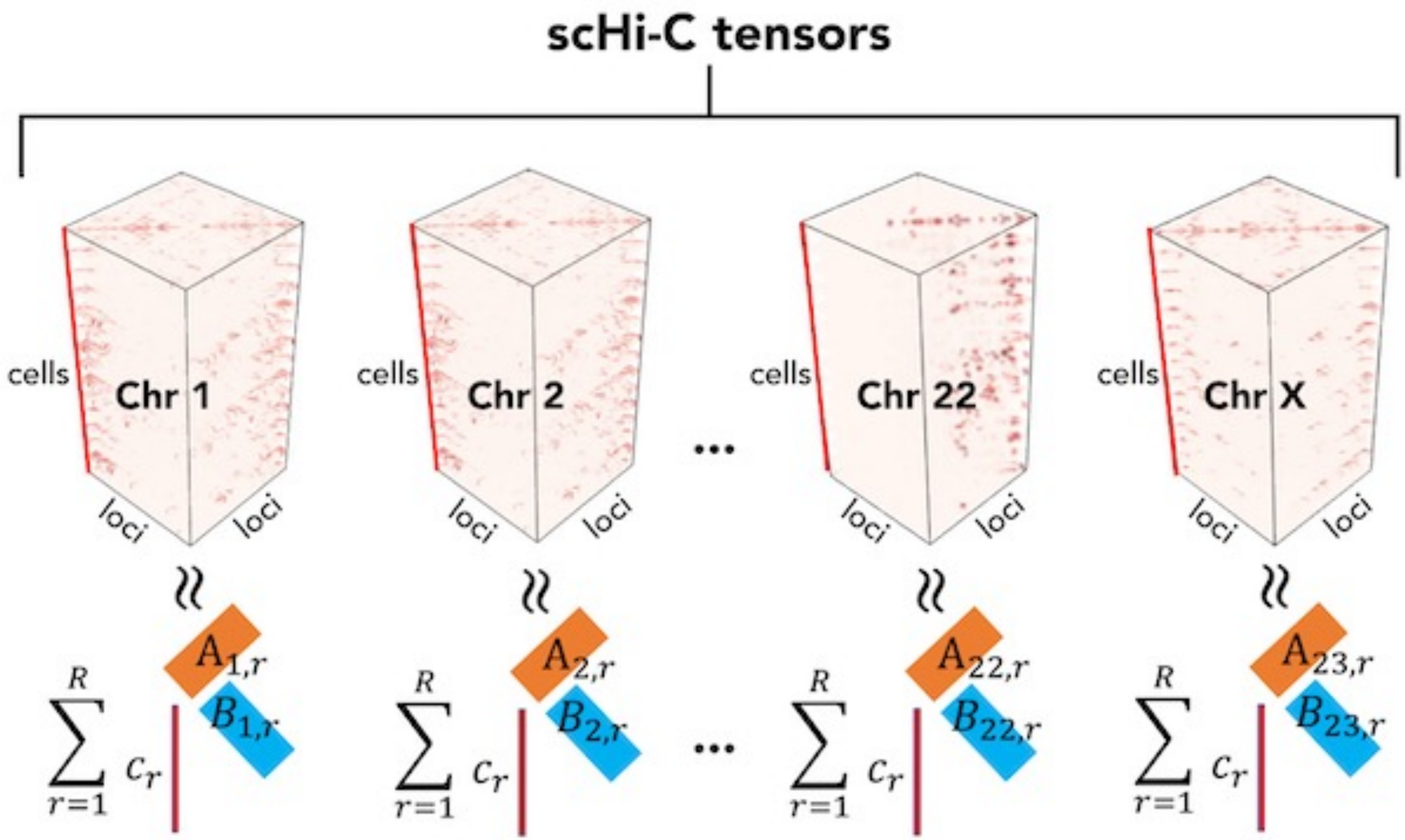




$$
\begin{aligned}
\mathcal{Y} &= \mathcal{M} + \mathcal{E} \\
&= \sum_{r=1}^{R} (A_r B_r^T) \circ c_r + \mathcal{E},
\end{aligned}
$$

Joint BTD

- Use all chromosome(tensor)'s cell information
  - Tensor decomposition with common cell loading (|)



Joint BTD

- Model

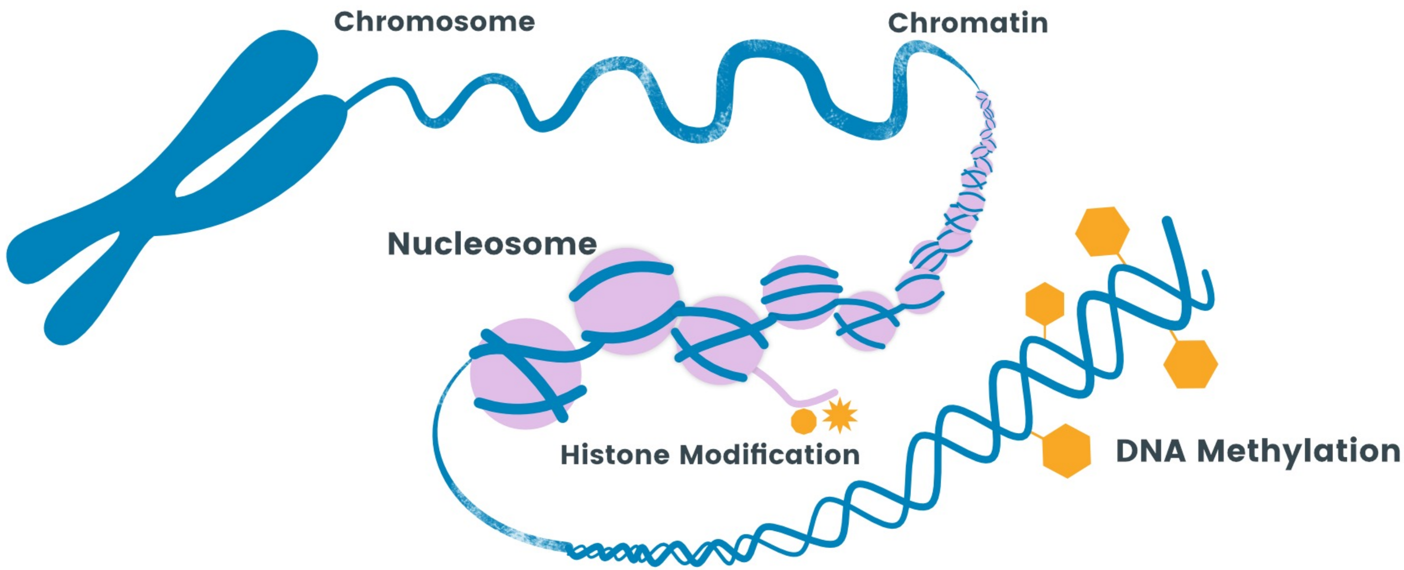
$$\min \sum_{chr=1}^{23} \|\mathcal{Y}_{chr} - \mathcal{M}_{chr}\|_F^2$$

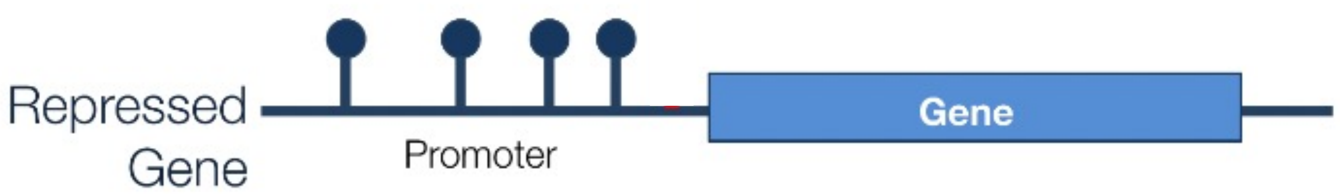
$$\text{s.t.} \quad \mathcal{M}_{chr} = \sum_{r=1}^{R} (A_{chr,r} B_{chr,r}^T) \circ c_r$$

$$A_{chr,r}, B_{chr,r} : \text{Loci loadings specific for each chr}$$

$$c_r \geq 0 : \text{Cell loading common for all chr (Non-negative)}$$

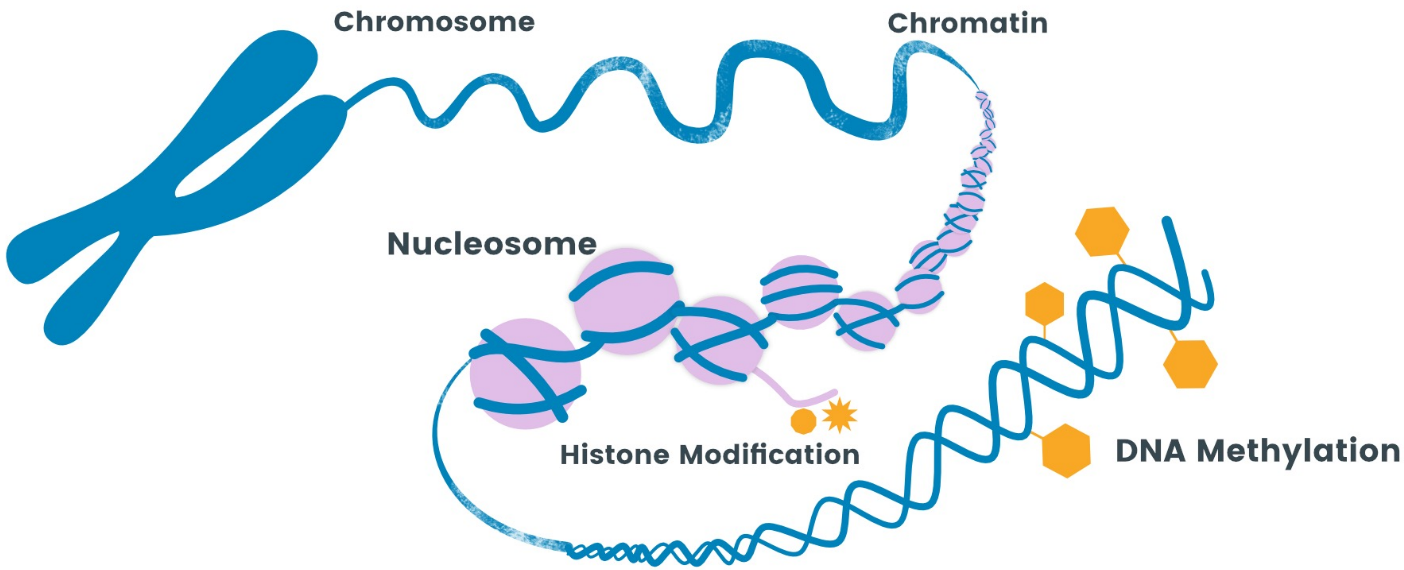
Multimodality and Muscle (DNA methylation)

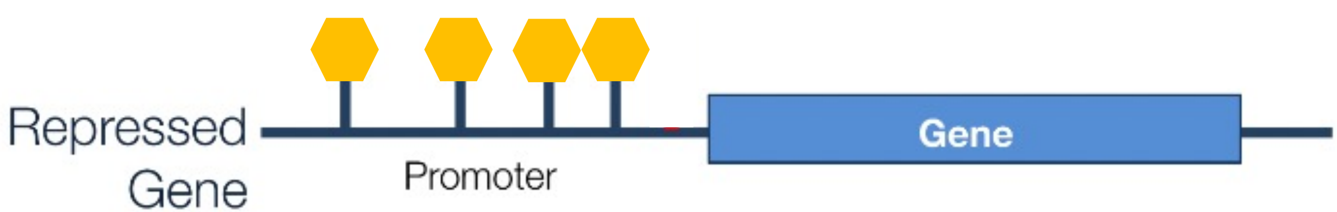




DNA methylation represses gene expression !

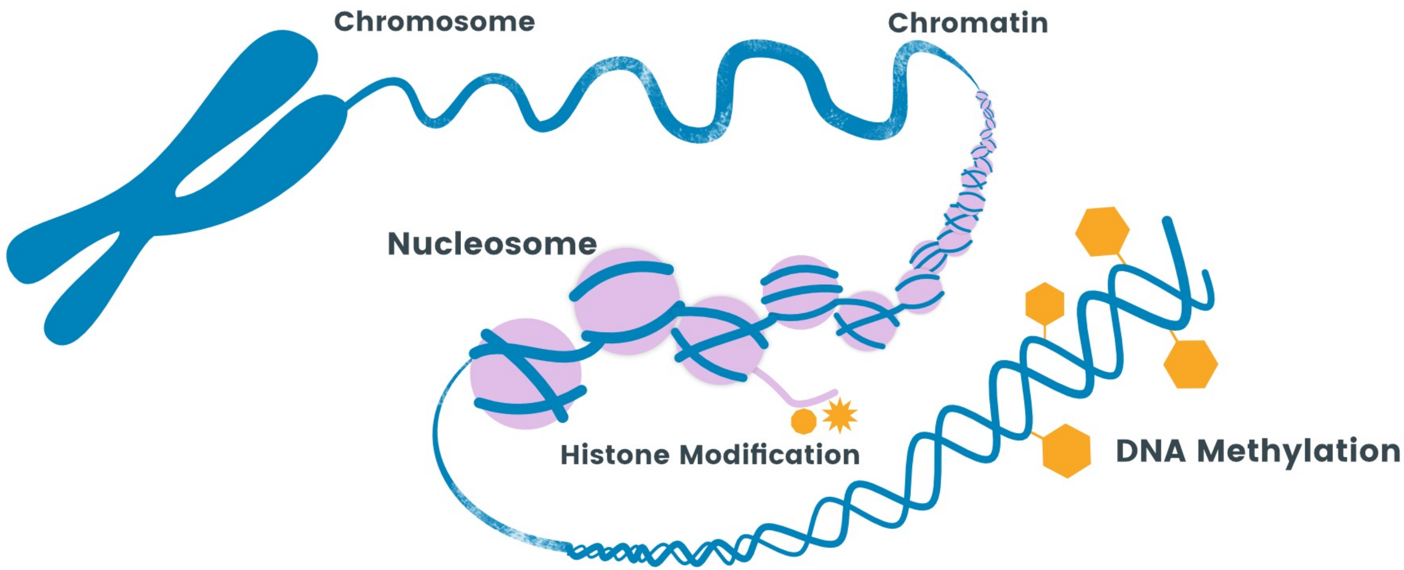
Multimodality and Muscle (DNA methylation)

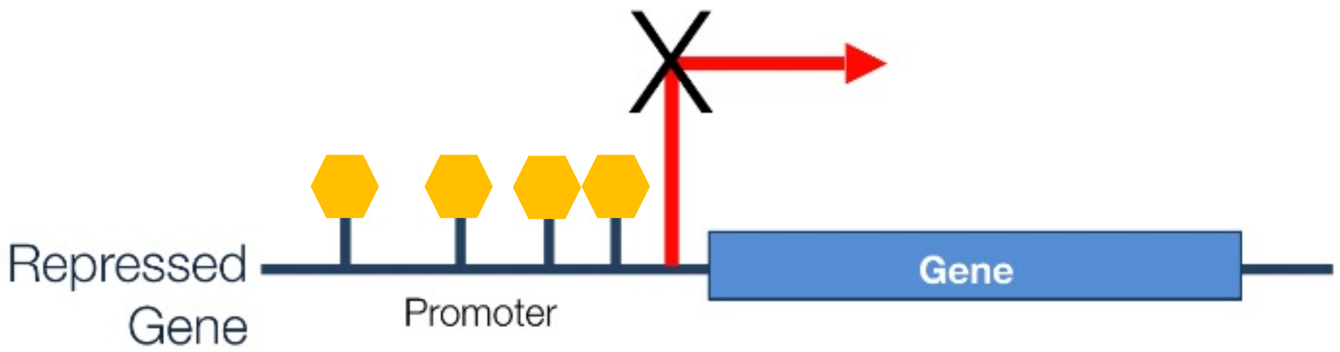




DNA methylation represses gene expression !

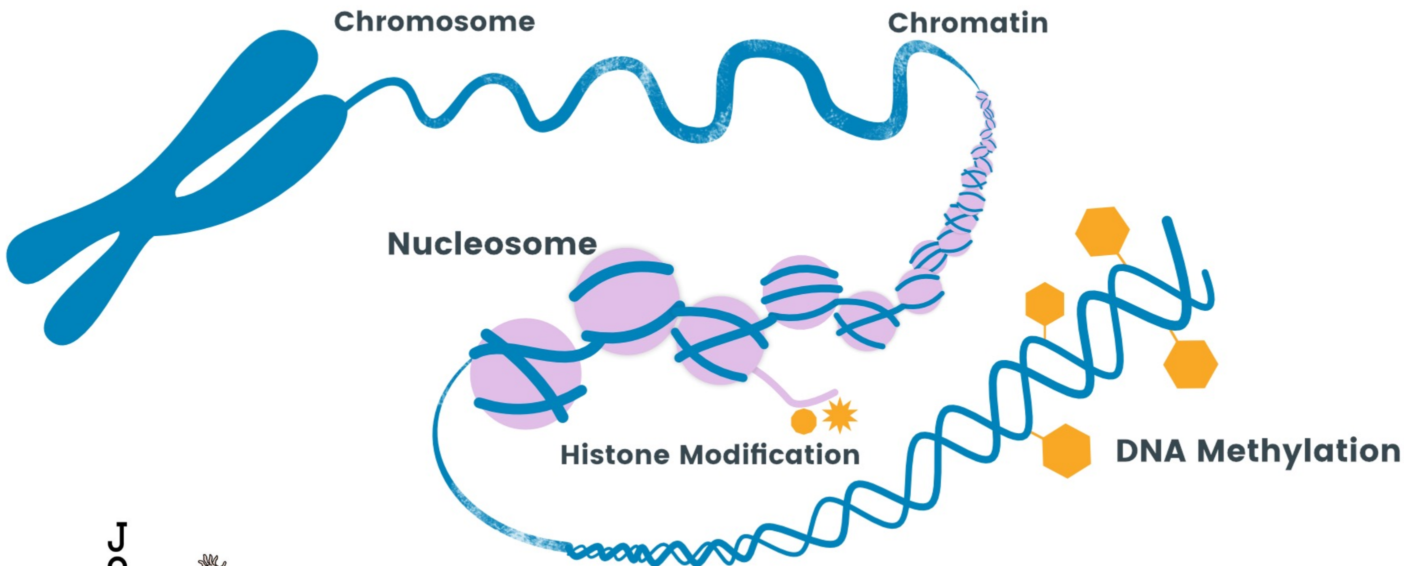
Multimodality and Muscle (DNA methylation)

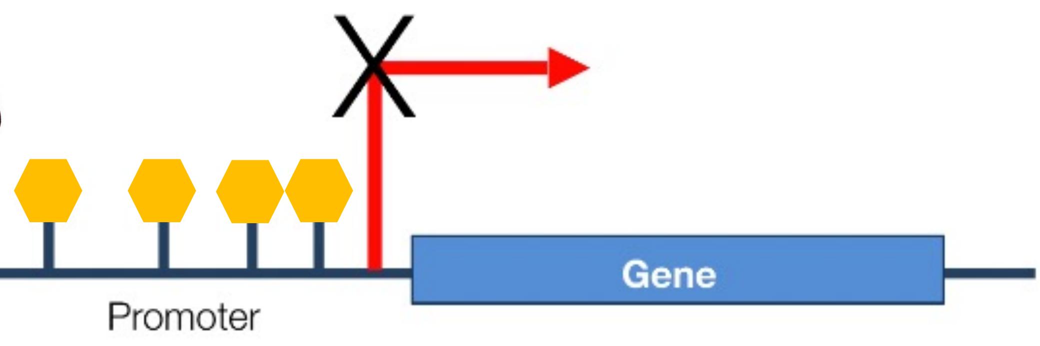




DNA methylation represses gene expression !

Multimodality and Muscle (DNA methylation)

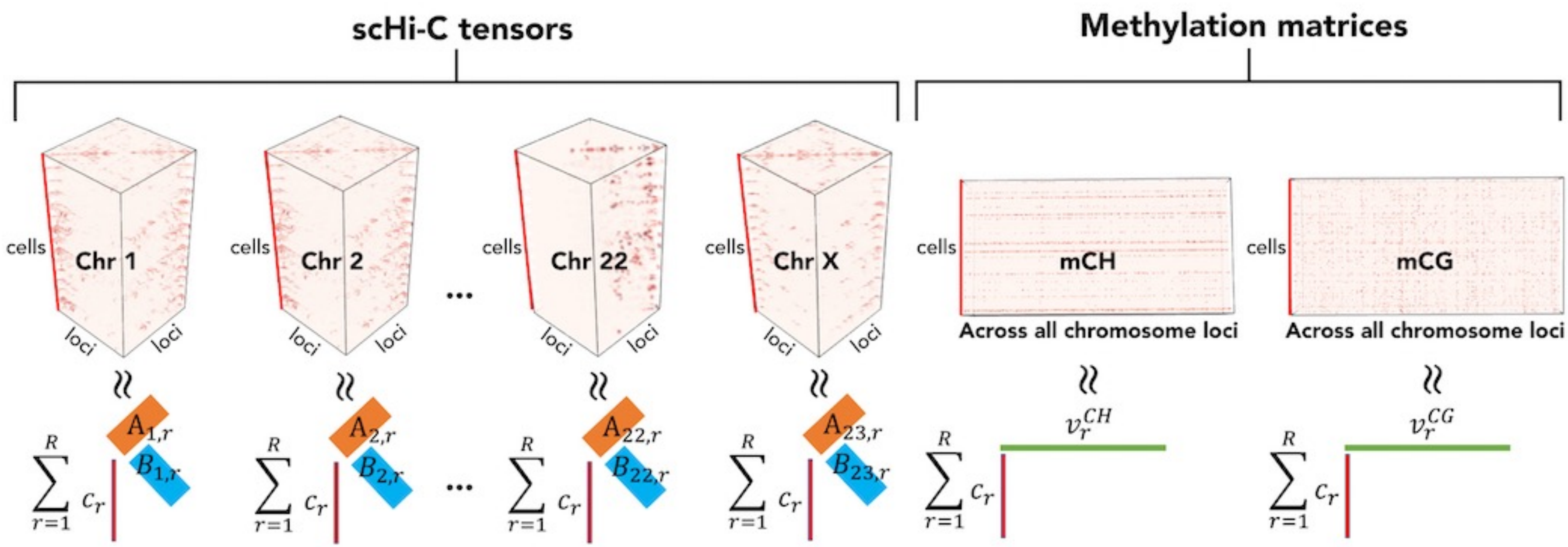




DNA methylation represses gene expression !

Multimodality and Muscle

- **DNA methylation** & **scHi-C** sequenced at the same cell level
- **MUSCLE :** A semi-non negative joint decomposition of **MU**ltiple **S**ingle **C**el**L** t**E**nsors



Muscle

- Model

$$
\begin{aligned}
\mathcal{Y}_{chr} &= \mathcal{M}_{chr} + \mathcal{E}_{chr}, \quad \epsilon_{i,j,c,chr} \overset{i.i.d}{\sim} N(0,\sigma_1^2), \quad \forall chr \in [Chr], \\
Y^k &= M^k + E^k, \quad \epsilon_{l,c}^k \overset{i.i.d}{\sim} N(0,\sigma_2^2), \quad \text{for } k \in \{CG, CH\}, \\
\text{s.t. } \mathcal{M}_{chr} &= \sum_{r=1}^{R} (A_{chr,r} B_{chr,r}^T) \circ c_r, \quad M^k = \sum_{r=1}^{R} v_r^k \circ c_r, \quad \text{for } k \in \{CG, CH\}, \quad \forall chr \in [Chr], \\
c_r &\geq 0, \quad \|c_r\| = 1, \quad \forall r \in [R], \text{ and } \frac{\sigma_1^2}{\sigma_2^2} = \frac{N_h}{N_m},
\end{aligned}
$$

- Variances are not ancillary to mean parameters
- Proportional variance model

Muscle

- Model

$$\mathcal{Y}_{chr} = \mathcal{M}_{chr} + \mathcal{E}_{chr}, \quad \epsilon_{i,j,c,chr} \overset{i.i.d}{\sim} N(0, \sigma_1^2), \quad \forall chr \in [Chr],$$

$$Y^k = M^k + E^k, \quad \epsilon_{l,c}^k \overset{i.i.d}{\sim} N(0, \sigma_2^2), \quad \text{for } k \in \{CG, CH\},$$

$$\text{s.t. } \mathcal{M}_{chr} = \sum_{r=1}^{R} (A_{chr,r} B_{chr,r}^T) \circ c_r, \quad M^k = \sum_{r=1}^{R} v_r^k \circ c_r, \quad \text{for } k \in \{CG, CH\}, \quad \forall chr \in [Chr],$$

$$c_r \geq 0, \quad \|c_r\| = 1, \quad \forall r \in [R], \text{ and } \frac{\sigma_1^2}{\sigma_2^2} = \frac{N_h}{N_m},$$

- Variances are not ancillary to mean parameters
- Proportional variance model

Muscle

- Model

$$
\begin{aligned}
\mathcal{Y}_{chr} &= \mathcal{M}_{chr} + \mathcal{E}_{chr}, \quad \epsilon_{i,j,c,chr} \overset{i.i.d}{\sim} N(0, \sigma_1^2), \quad \forall chr \in [Chr], \\
Y^k &= M^k + E^k, \quad \epsilon_{l,c}^k \overset{i.i.d}{\sim} N(0, \sigma_2^2), \quad \text{for } k \in \{CG, CH\}, \\
\text{s.t. } \mathcal{M}_{chr} &= \sum_{r=1}^{R} (A_{chr,r} B_{chr,r}^T) \circ c_r, \quad M^k = \sum_{r=1}^{R} v_r^k \circ c_r, \quad \text{for } k \in \{CG, CH\}, \quad \forall chr \in [Chr], \\
c_r &\geq 0, \quad \|c_r\| = 1, \quad \forall r \in [R], \text{ and } \frac{\sigma_1^2}{\sigma_2^2} = \frac{N_h}{N_m},
\end{aligned}
$$

- Variances are not ancillary to mean parameters
- Proportional variance model

Muscle

- MLE equivalent problem

$$\min_{\substack{A_{chr,r},B_{chr,r},c_r\geq 0\\ v_r^{CG},v_r^{CH},\|c_r\|_2=1}} \left\{ \underbrace{\frac{1}{N_h}\sum_{chr=1}^{23} \left\| \mathcal{Y}_{chr} - \sum_{r=1}^{R} (A_{chr,r}B_{chr,r}^{T}) \circ c_r \right\|_F^2}_{\text{scHi-C tensors}} + \underbrace{\frac{1}{N_m}\sum_{k\in\{CG,CH\}} \left\| Y^k - \sum_{r=1}^{R} v_r^k \circ c_r \right\|_F^2}_{\text{Methylation matrices}} \right\}$$

Muscle

- MLE equivalent problem

Data common cell loading c_r contains the shared information across modalities

$$\min_{\substack{A_{chr,r}, B_{chr,r}, c_r \geq 0 \\ v_r^{CG}, v_r^{CH}, \|c_r\|_2 = 1}} \left\{ \underbrace{\frac{1}{N_h} \sum_{chr=1}^{23} \left\| \mathcal{Y}_{chr} - \sum_{r=1}^{R} (A_{chr,r} B_{chr,r}^T) \circ c_r \right\|_F^2}_{\text{scHi-C tensors}} + \underbrace{\frac{1}{N_m} \sum_{k \in \{CG, CH\}} \left\| Y^k - \sum_{r=1}^{R} v_r^k \circ c_r \right\|_F^2}_{\text{Methylation matrices}} \right\}$$

Muscle – ALS algorithm

Algorithm 1 Muscle ALS algorithm

Input: scHi-C tensors $\mathcal{Y}_{chr} \in \mathbb{R}^{l_{chr} \times l_{chr} \times C}$, methylation matrices $Y^{CG}, Y^{CH} \in \mathbb{R}^{\sum_{chr} l_{chr} \times C}$, scHi-C loci loading rank K_{chr}, $\forall chr \in [Chr]$, data modality common rank R.

Output: scHi-C loci loadings $A_{chr,r}, B_{chr,r} \in \mathbb{R}^{l_{chr} \times K_{chr}}$, methylation loci loadings $v_r^{CG}, v_r^{CG} \in \mathbb{R}^{\sum_{chr} l_{chr}}$, and data modality common cell loading vector $c_r \in \mathbb{R}_+^C$, $\forall chr \in [Chr]$ and $\forall r \in [R]$.

1: Initialize the decomposition objects $\tilde{\mathcal{Y}}_{chr} \leftarrow \mathcal{Y}_{chr}$ $\forall chr \in [Chr]$, $\tilde{Y}^{CG} \leftarrow Y^{CG}$, $\tilde{Y}^{CH} \leftarrow Y^{CH}$.

2: **for** $r = 1$ to R **do**.

3: Initialize $\hat{A}_{chr,r}$, and $\hat{B}_{chr,r}$ by rank … here that the core tensor size is absorb…

Pools all the loci loading information across the data

4: Initialize $\hat{v}_r^k$ by first left singular vector of $\tilde{Y}^k$ (singular value is absorbed), $\forall k \in \{CG, CH\}$.

5: **while** the convergence criterion is not met **do**

6: Update $\hat{c}_r \leftarrow \frac{\left(\frac{1}{N_h}\sum_{chr} Y_{chr}^T X_{chr} + \frac{1}{N_m}(\tilde{Y}^{CG})^T v_r^{CG} + \frac{1}{N_m}(\tilde{Y}^{CH})^T v_r^{CH}\right)_+}{\left\|\left(\frac{1}{N_h}\sum_{chr} Y_{chr}^T X_{chr} + \frac{1}{N_m}(\tilde{Y}^{CG})^T v_r^{CG} + \frac{1}{N_m}(\tilde{Y}^{CH})^T v_r^{CH}\right)_+\right\|_2}$,

with $Y_{chr} = \text{unfold}_3(\tilde{\mathcal{Y}}_{chr}) \in \mathbb{R}^{(l_{chr} * l_{chr}) \times C}$ and $X_{chr} = [(\hat{A}_{chr,r} \odot \hat{B}_{chr,r})\mathbf{1}_{K_{chr}}] \in \mathbb{R}^{(l_{chr} * l_{chr})}$.

7: Update $(\hat{A}_{chr,r}, \hat{B}_{chr,r}) \leftarrow \text{Eigen}_{K_{chr}}(\tilde{\mathcal{Y}}_{chr} \times_3 \hat{c}_r)$, $\forall chr \in [Chr]$. Note here that the eigenvalues are absorbed into $\hat{A}_{chr,r}$.

8: Update methylation loci loadings $v_r^k \leftarrow \tilde{Y}^k \hat{c}_r$. $\forall k \in \{CG, CH\}$.

9: **end while**

10: Update $\tilde{\mathcal{Y}}_{chr} \leftarrow \tilde{\mathcal{Y}}_{chr} - (\hat{A}_{chr,r}\hat{B}_{chr,r}^T) \circ \hat{c}_r$ and $\tilde{Y}^k \leftarrow \tilde{Y}^k - \hat{v}_r^k \circ \hat{c}_c$ for $\forall k \in \{CG, CH\}$.

11: **end for**

Muscle – ALS algorithm

Optimality properties for ALS algorithm

Algorithm 1 Muscle ALS algorithm

Input: scHi-C tensors $\mathcal{Y}_{chr} \in \mathbb{R}^{l_{chr} \times l_{chr} \times C}$, methylation matrices $Y^{CG}, Y^{CH} \in \mathbb{R}^{\sum_{chr} l_{chr} \times C}$, scHi-C loci loading rank K_{chr}, $\forall chr \in [Chr]$, data modality common rank R.

Output: scHi-C loci loadings $A_{chr,r}, B_{chr,r} \in \mathbb{R}^{l_{chr} \times K_{chr}}$, methylation loci loadings $v_r^{CG}, v_r^{CG} \in \mathbb{R}^{\sum_{chr} l_{chr}}$, and data modality common cell loading vector $c_r \in \mathbb{R}_+^C$, $\forall chr \in [Chr]$ and $\forall r \in [R]$.

1: Initialize the decomposition objects $\tilde{\mathcal{Y}}_{chr} \leftarrow \mathcal{Y}_{chr}$ $\forall chr \in [Chr]$, $\tilde{Y}^{CG} \leftarrow Y^{CG}$, $\tilde{Y}^{CH} \leftarrow Y^{CH}$.

2: **for** $r = 1$ to R **do**.

3: Initialize $\hat{A}_{chr,r}$, and $\hat{B}_{chr,r}$ by rank … here that the core tensor size is absorb…

Pools all the loci loading information across the data

4: Initialize $\hat{v}_r^k$ by first left singular vector of $\tilde{Y}^k$ (singular value is absorbed), $\forall k \in \{CG, CH\}$.

5: **while** the convergence criterion is not met **do**

6: Update $\hat{c}_r \leftarrow \dfrac{\left(\frac{1}{N_h}\sum_{chr} Y_{chr}^T X_{chr} + \frac{1}{N_m}(\tilde{Y}^{CG})^T v_r^{CG} + \frac{1}{N_m}(\tilde{Y}^{CH})^T v_r^{CH}\right)_+}{\left\|\left(\frac{1}{N_h}\sum_{chr} Y_{chr}^T X_{chr} + \frac{1}{N_m}(\tilde{Y}^{CG})^T v_r^{CG} + \frac{1}{N_m}(\tilde{Y}^{CH})^T v_r^{CH}\right)_+\right\|_2}$,

with $Y_{chr} = \text{unfold}_3(\tilde{\mathcal{Y}}_{chr}) \in \mathbb{R}^{(l_{chr} * l_{chr}) \times C}$ and $X_{chr} = [(\hat{A}_{chr,r} \odot \hat{B}_{chr,r})\mathbf{1}_{K_{chr}}] \in \mathbb{R}^{(l_{chr} * l_{chr})}$.

7: Update $(\hat{A}_{chr,r}, \hat{B}_{chr,r}) \leftarrow \text{Eigen}_{K_{chr}}(\tilde{\mathcal{Y}}_{chr} \times_3 \hat{c}_r)$, $\forall chr \in [Chr]$. Note here that the eigenvalues are absorbed into $\hat{A}_{chr,r}$.

8: Update methylation loci loadings $v_r^k \leftarrow \tilde{Y}^k \hat{c}_r$. $\forall k \in \{CG, CH\}$.

9: **end while**

10: Update $\tilde{\mathcal{Y}}_{chr} \leftarrow \tilde{\mathcal{Y}}_{chr} - (\hat{A}_{chr,r}\hat{B}_{chr,r}^T) \circ \hat{c}_r$ and $\tilde{Y}^k \leftarrow \tilde{Y}^k - \hat{v}_r^k \circ \hat{c}_c$ for $\forall k \in \{CG, CH\}$.

11: **end for**

Data analysis results

Ramani2017

Massively multiplex single-cell Hi-C

Vijay Ramani, Xinxian Deng, Ruolan Qiu, Kevin L Gunderson, Frank J Steemers, Christine M Disteche, William S Noble, Zhijun Duan & Jay Shendure

Nature Methods **14**, 263–266(2017) | Cite this article

nature methods ARTICLES
https://doi.org/10.1038/s41592-019-0547-z

Lee2019

Simultaneous profiling of 3D genome structure and DNA methylation in single human cells

Dong-Sung Lee[1,5], Chongyuan Luo[2,3,5], Jingtian Zhou[2,5], Sahaana Chandran[1], Angeline Rivkin[2], Anna Bartlett[2], Joseph R. Nery[2], Conor Fitzpatrick[4], Carolyn O'Connor[4], Jesse R. Dixon[1*] and Joseph R. Ecker[2,3*]

Large scale efforts

4D Nucleome

PLOS COMPUTATIONAL BIOLOGY

Kim2020

RESEARCH ARTICLE

Capturing cell type-specific chromatin compartment patterns by applying topic modeling to single-cell Hi-C data

Hyeon-Jin Kim[1†], Galip Gürkan Yardımcı[1†], Giancarlo Bonora[1], Vijay Ramani[1,2], Jie Liu[1], Ruolan Qiu[1], Choli Lee[1], Jennifer Hesson[3,5], Carol B. Ware[3,5], Jay Shendure[1], Zhijun Duan[4,5*], William Stafford Noble[1,6*]

nature methods BRIEF COMMUNICATION
https://doi.org/10.1038/s41592-019-0502-z

Tan2021

Cell CellPress

Resource

Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development

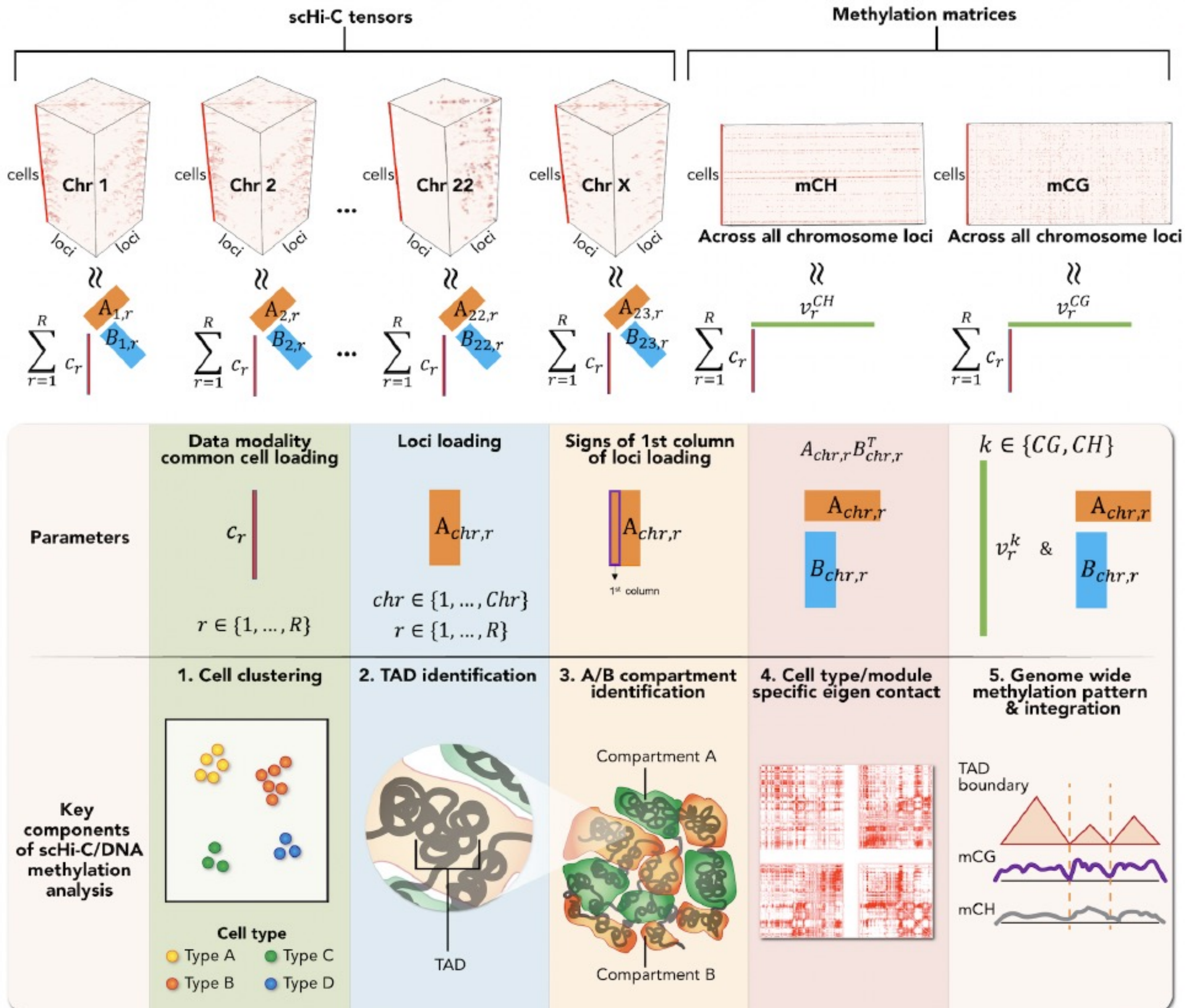
Longzhi Tan,[1,2,*] Wenping Ma,[3,4,5] Honggui Wu,[3,4,6] Yinghui Zheng,[3,4] Dong Xing,[3,4] Ritchie Chen,[1] Xiang Li,[3,4,5] Nicholas Daley,[2,7,10] Karl Deisseroth,[1,8,9] and X. Sunney Xie[3,4,11,*]

Li2019

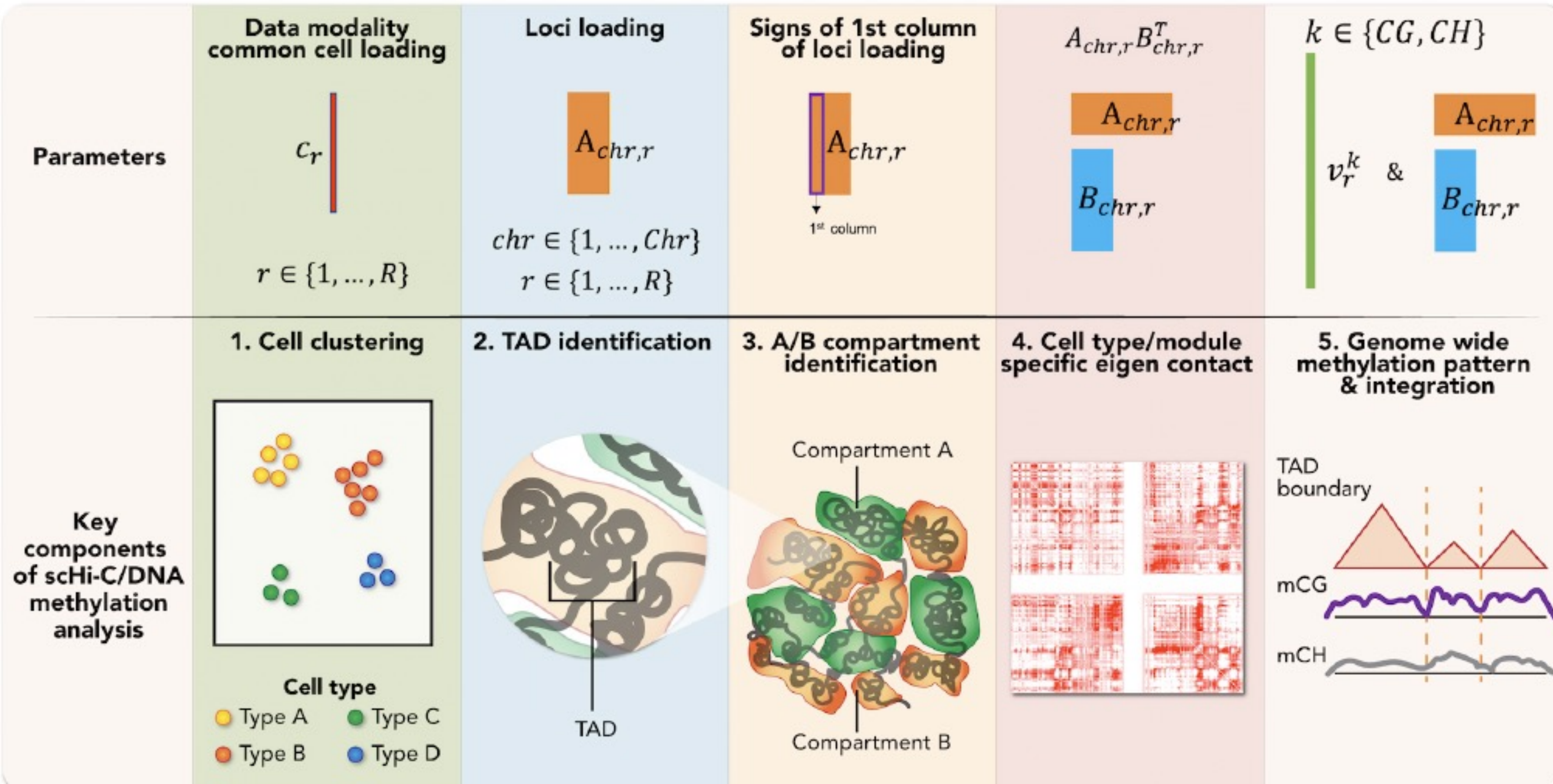
Joint profiling of DNA methylation and chromatin architecture in single cells

Guoqiang Li[1,9], Yaping Liu[2,3,4,9], Yanxiao Zhang[1], Naoki Kubo[1], Miao Yu[1], Rongxin Fang[1,5], Manolis Kellis[6,7] and Bing Ren[1,8*]

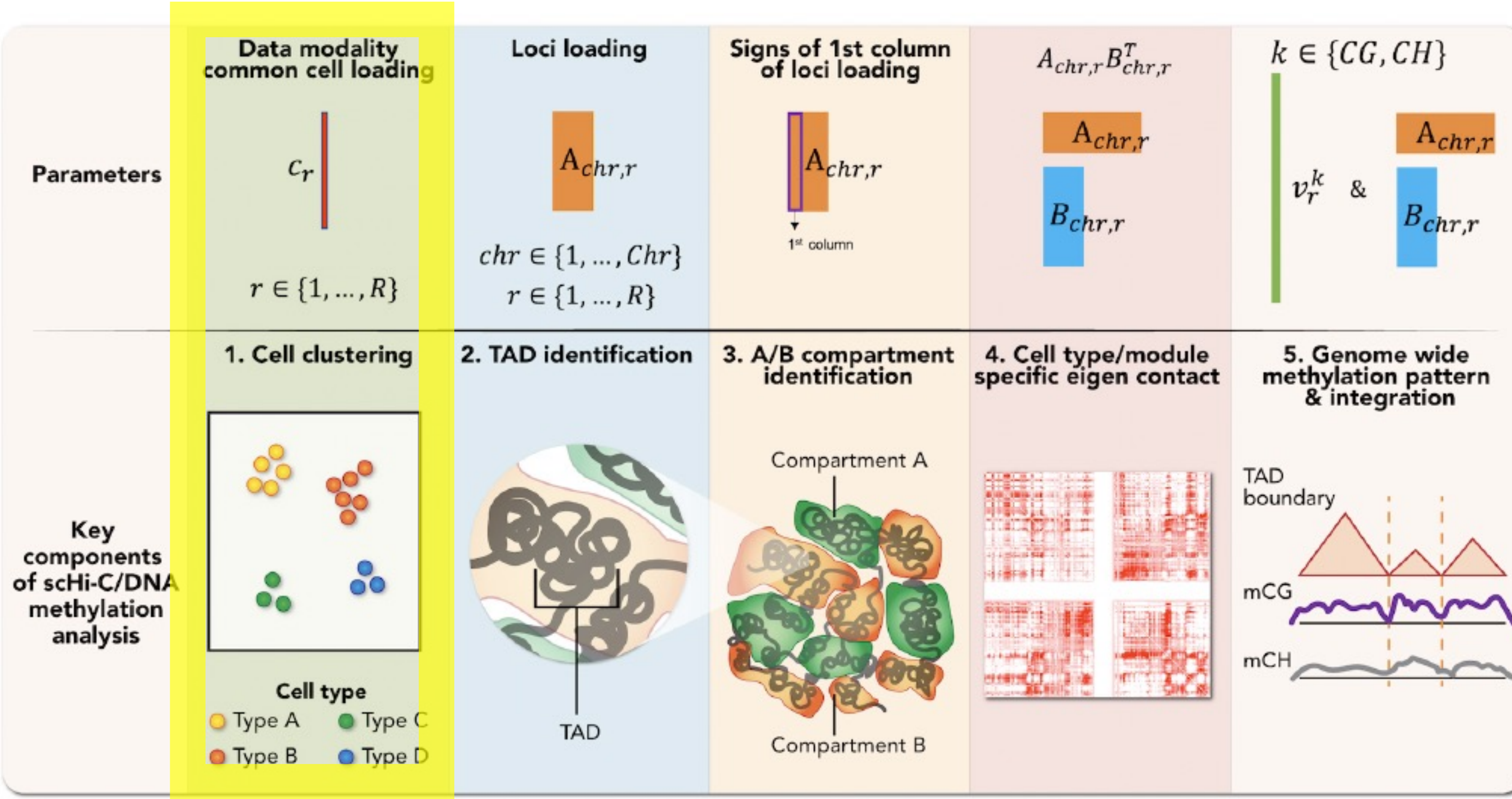
Muscle provides …



Muscle provides …

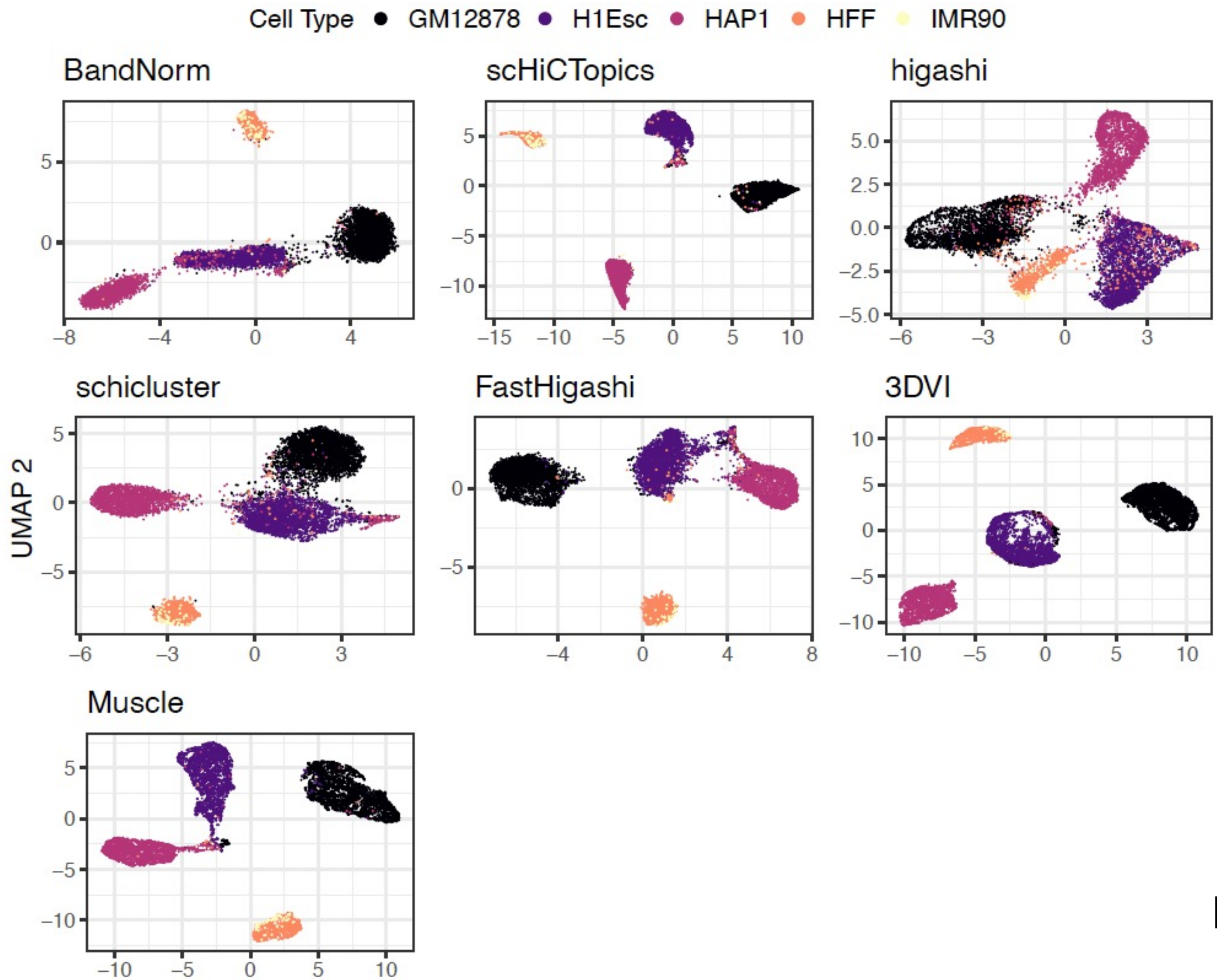


Muscle provides …



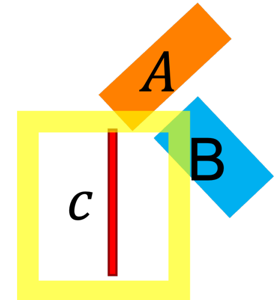
Data analysis results

- **Cell type clustering (only scHi-C tensors)**

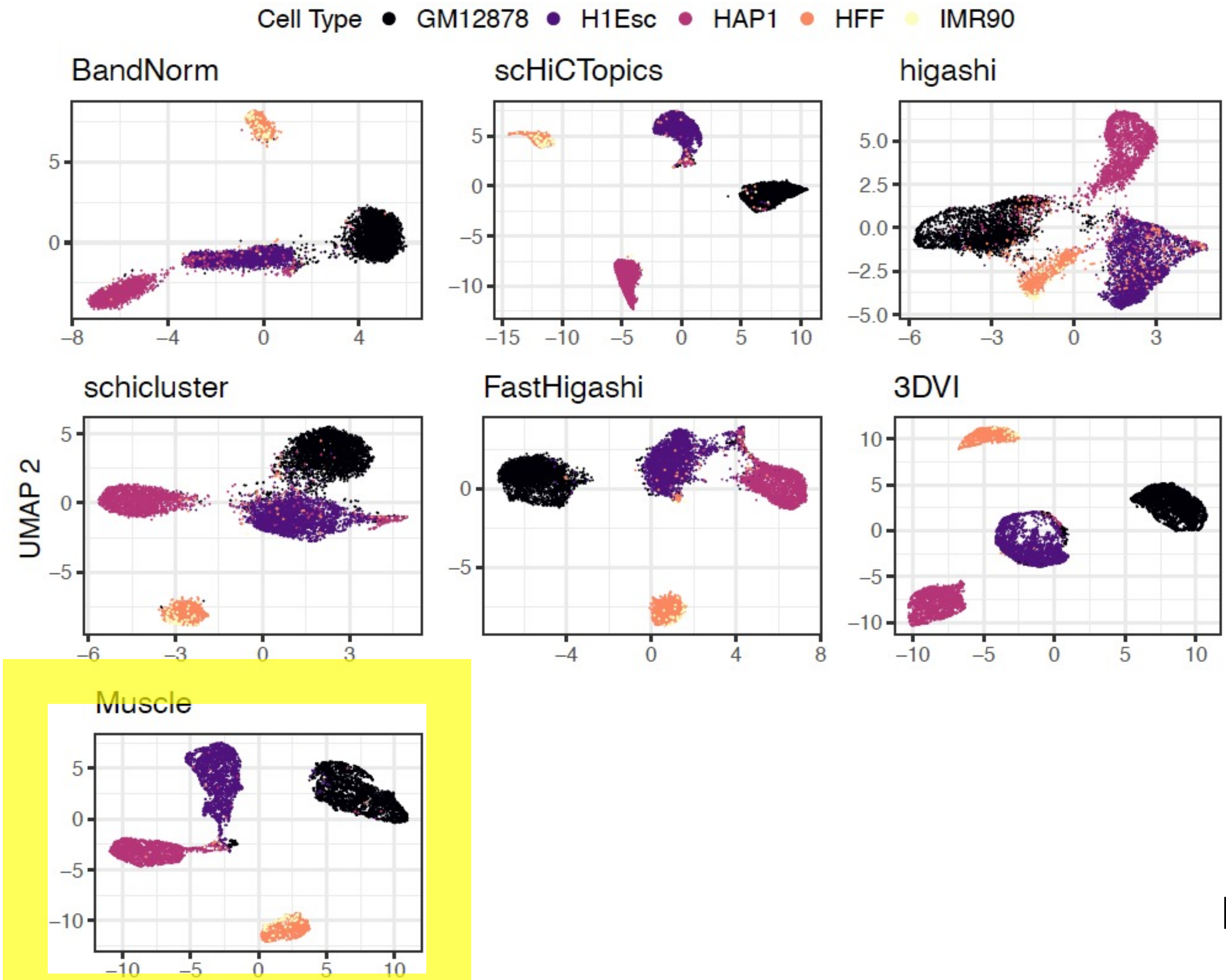


Kim 2020

Data analysis results



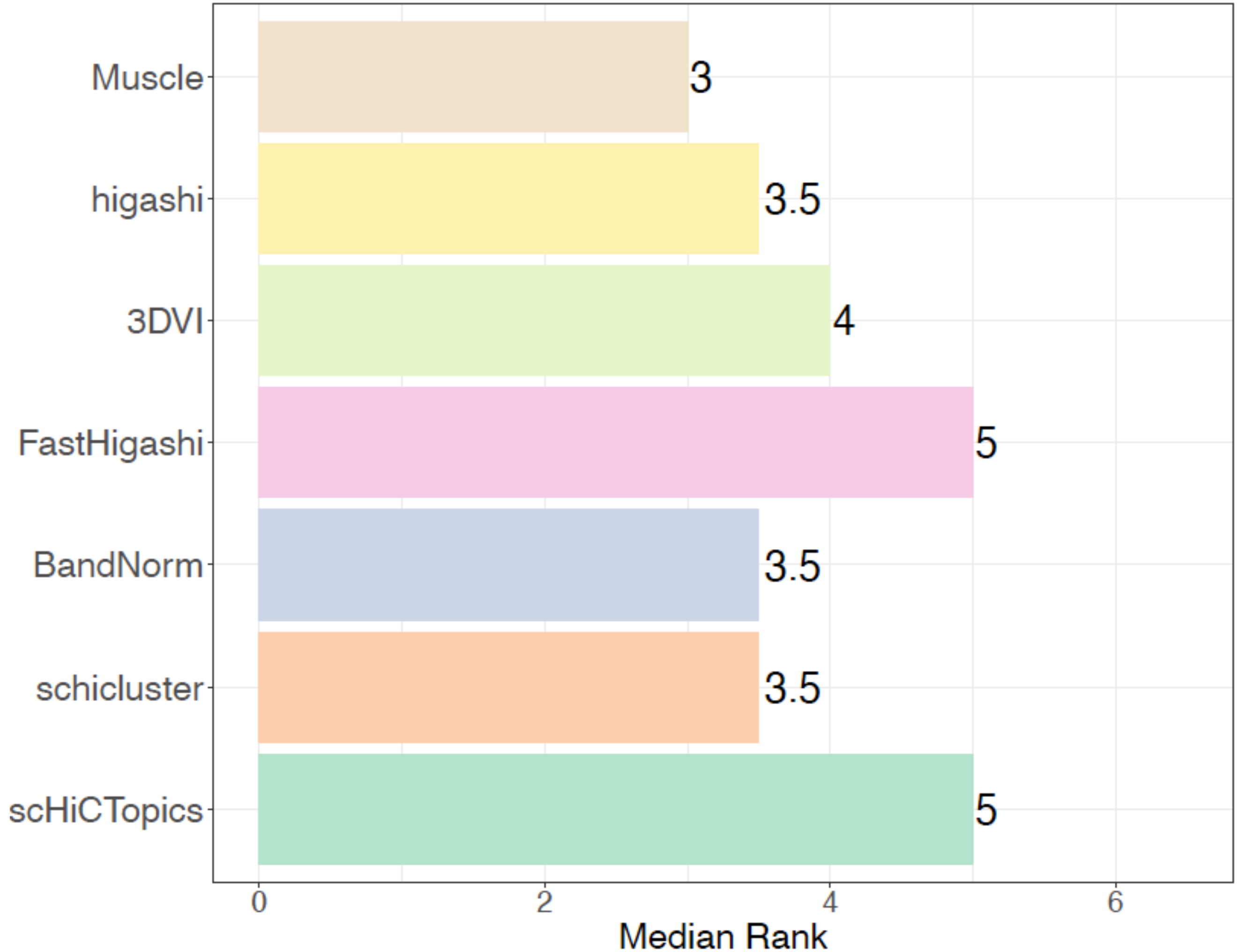
- **Cell type clustering (only scHi-C tensors)**



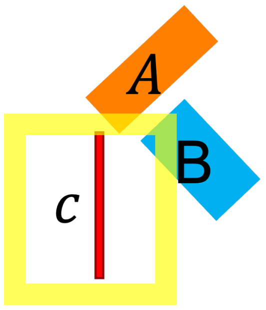
Kim 2020

Data analysis results

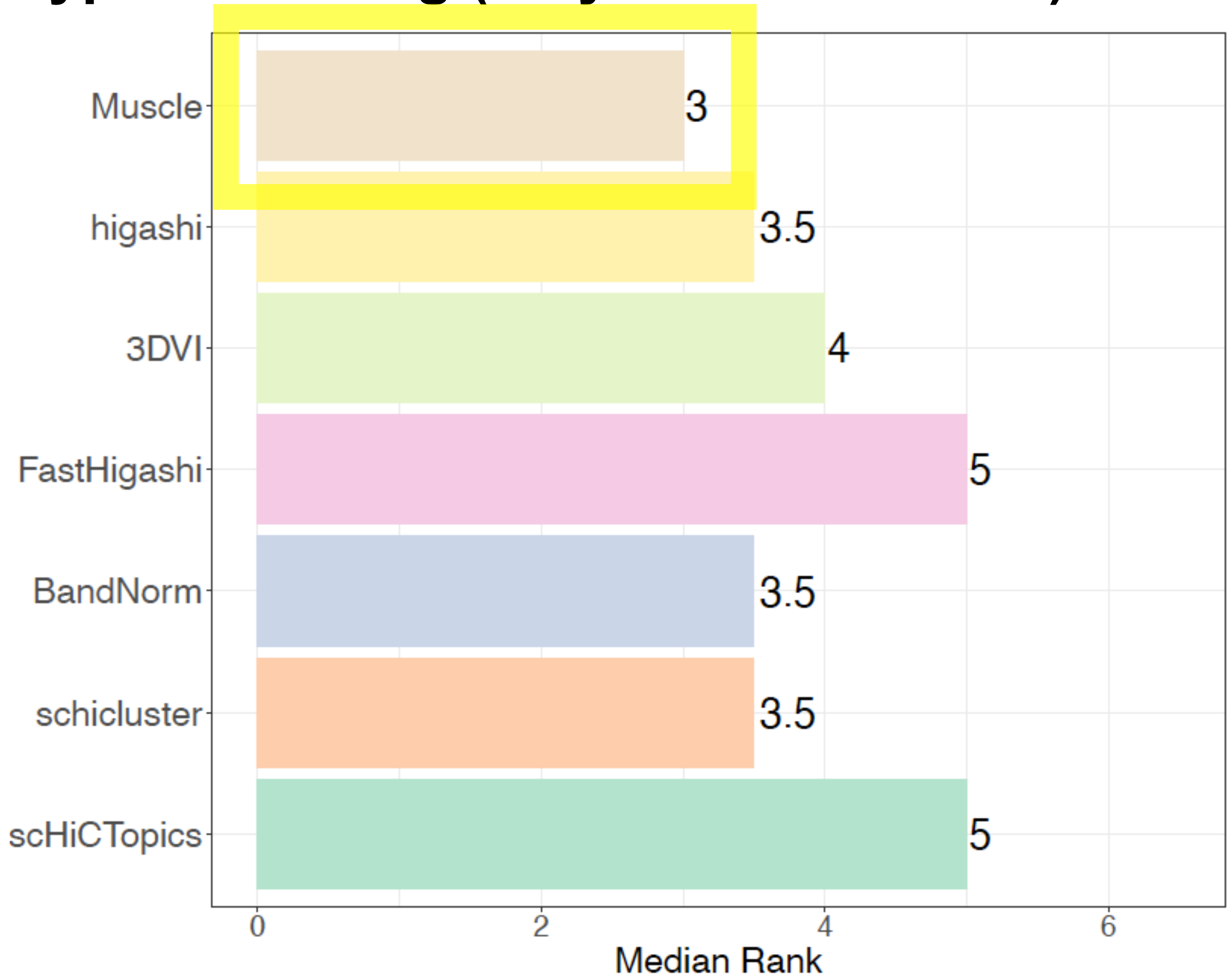
- **Cell type clustering (only scHi-C tensors)**



Data analysis results

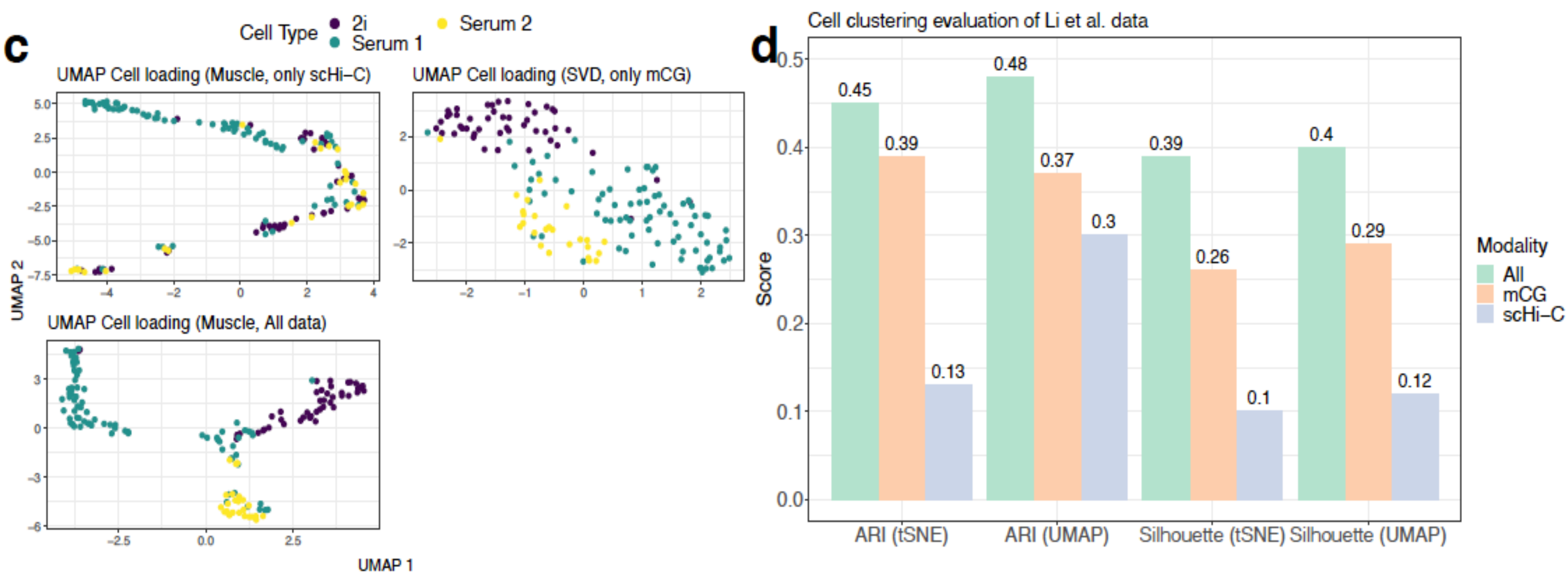


- **Cell type clustering (only scHi-C tensors)**



Data analysis results

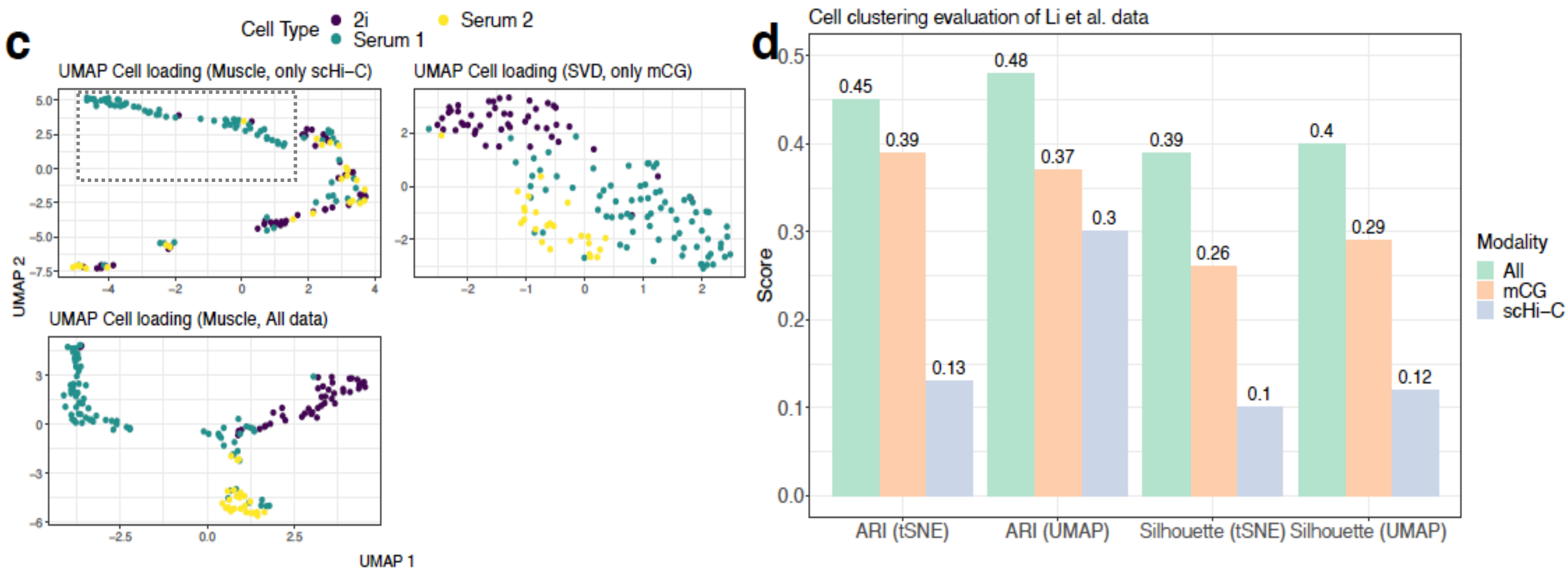
- **Cell type clustering (Multi-modality)**



Li 2019

Data analysis results

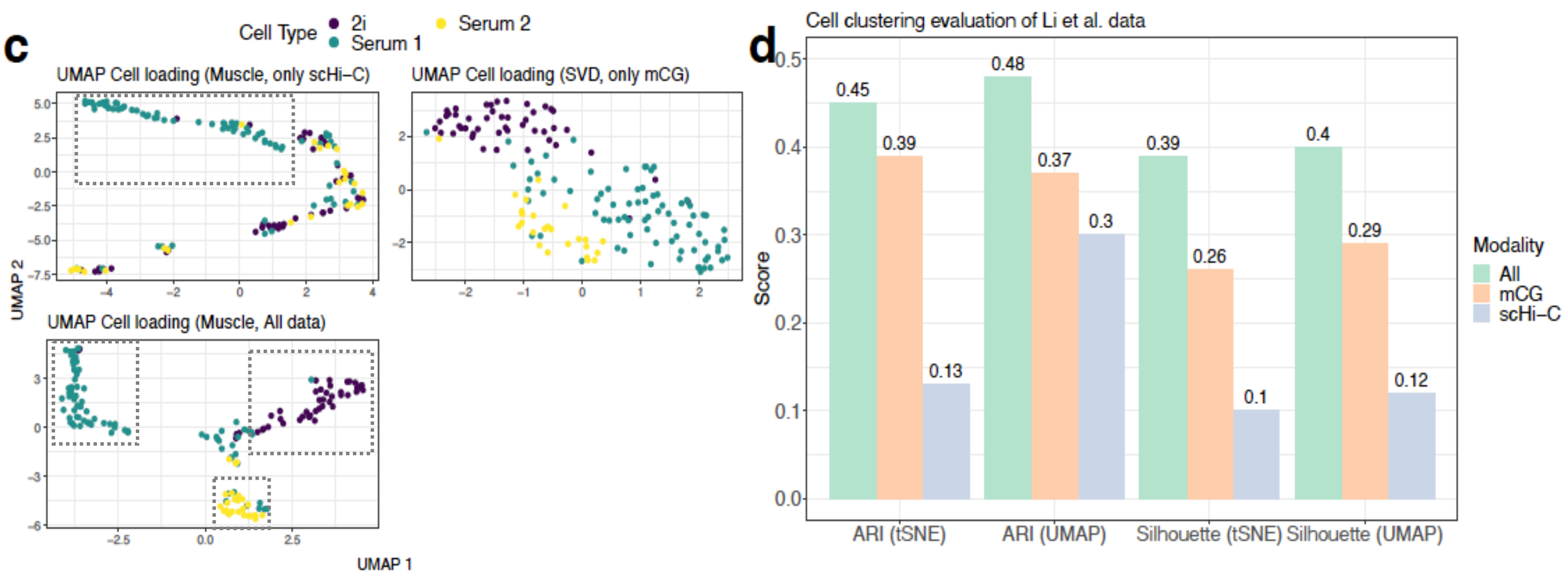
- **Cell type clustering (Multi-modality)**



Li 2019

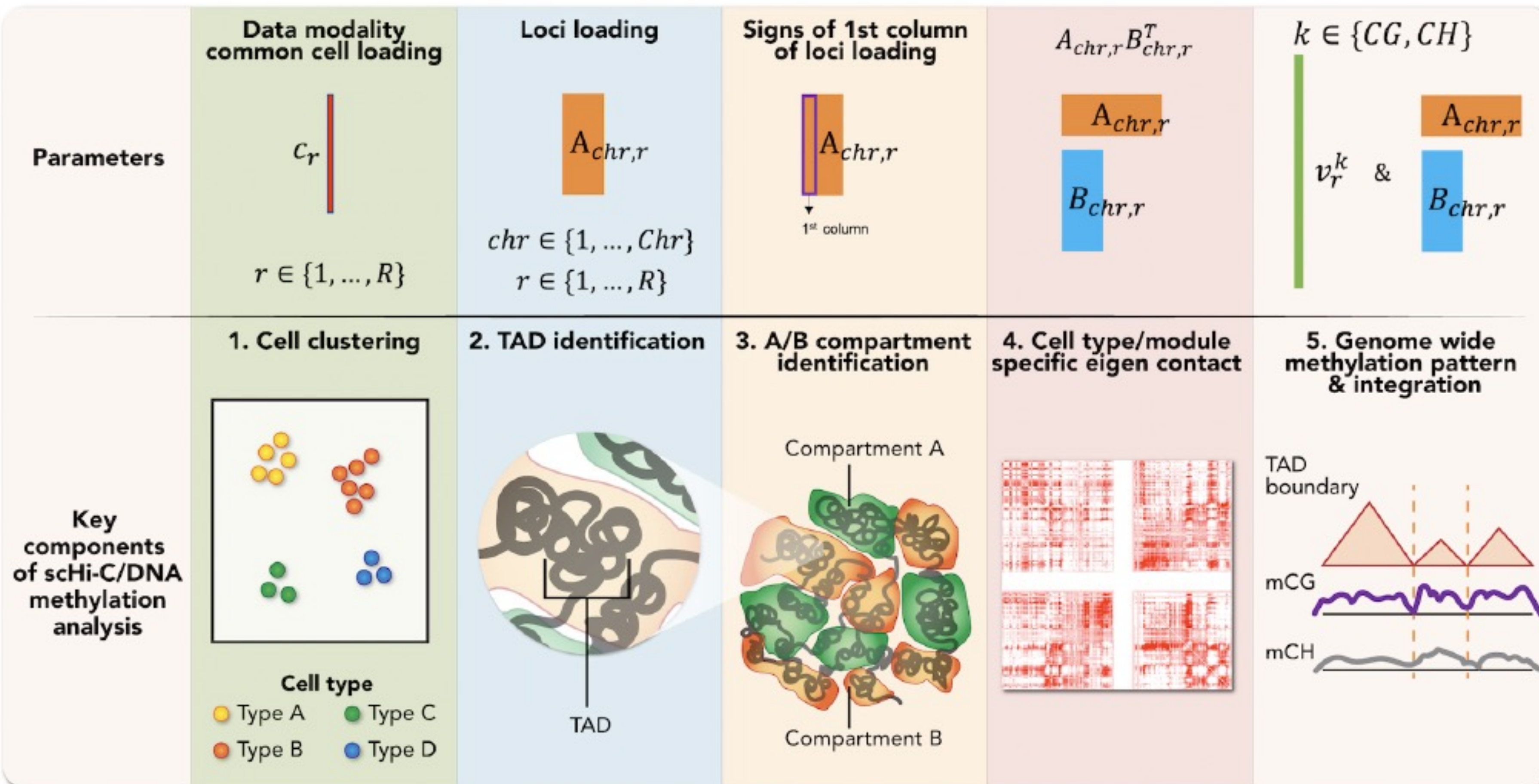
Data analysis results

- **Cell type clustering (Multi-modality)**

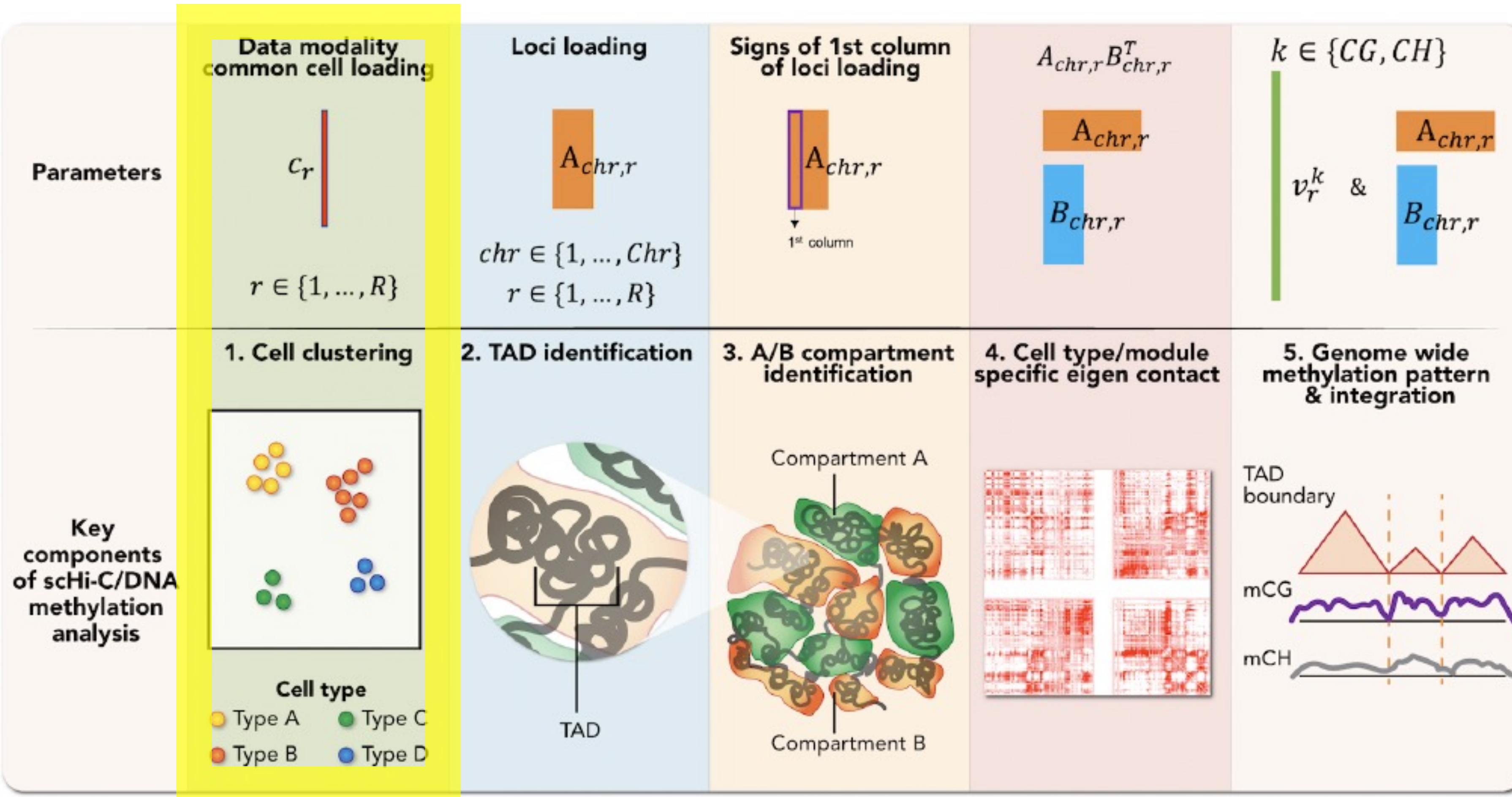


Li 2019

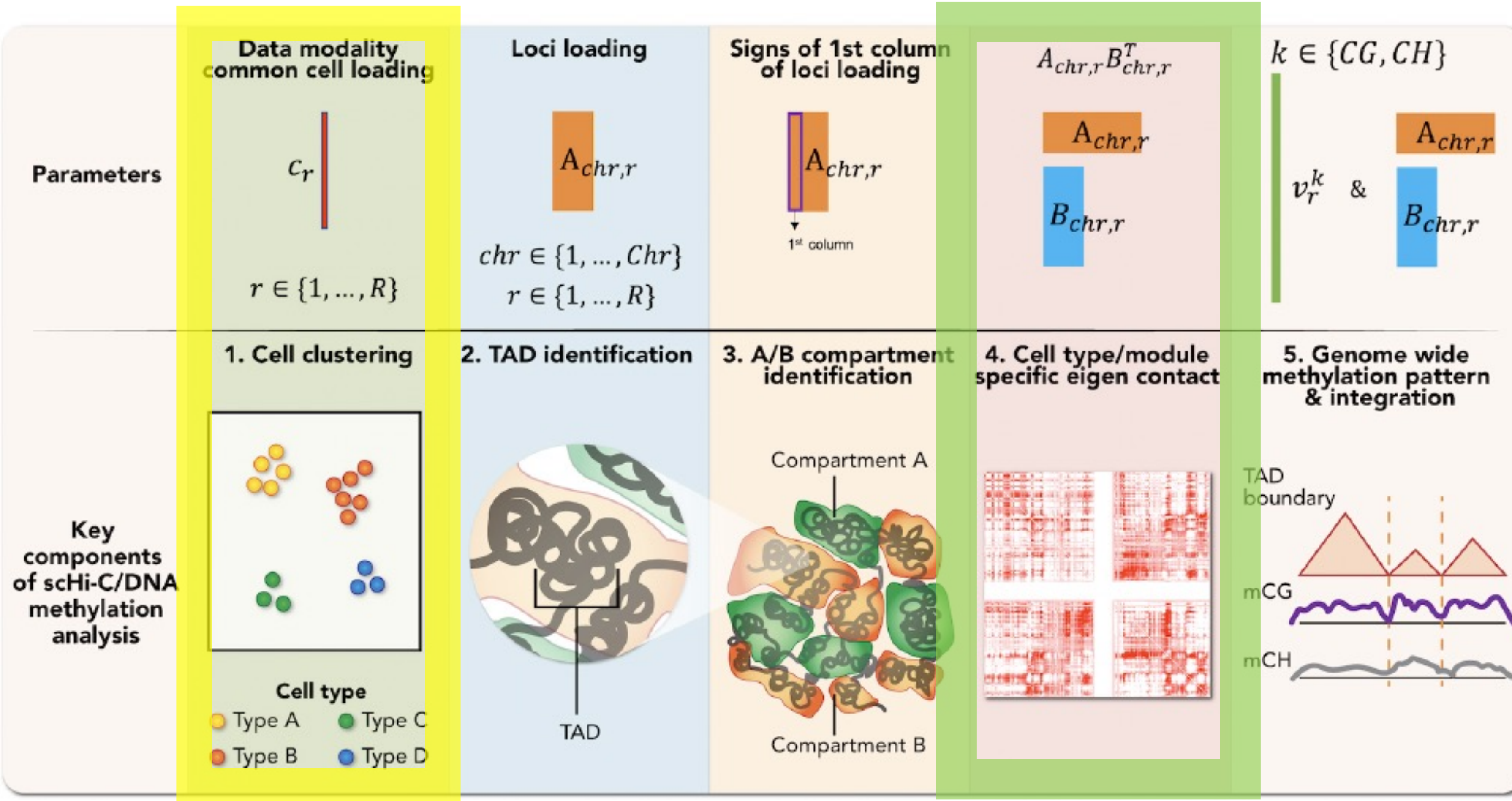
Muscle provides …



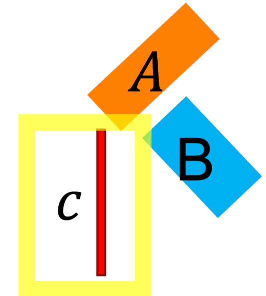
Muscle provides …



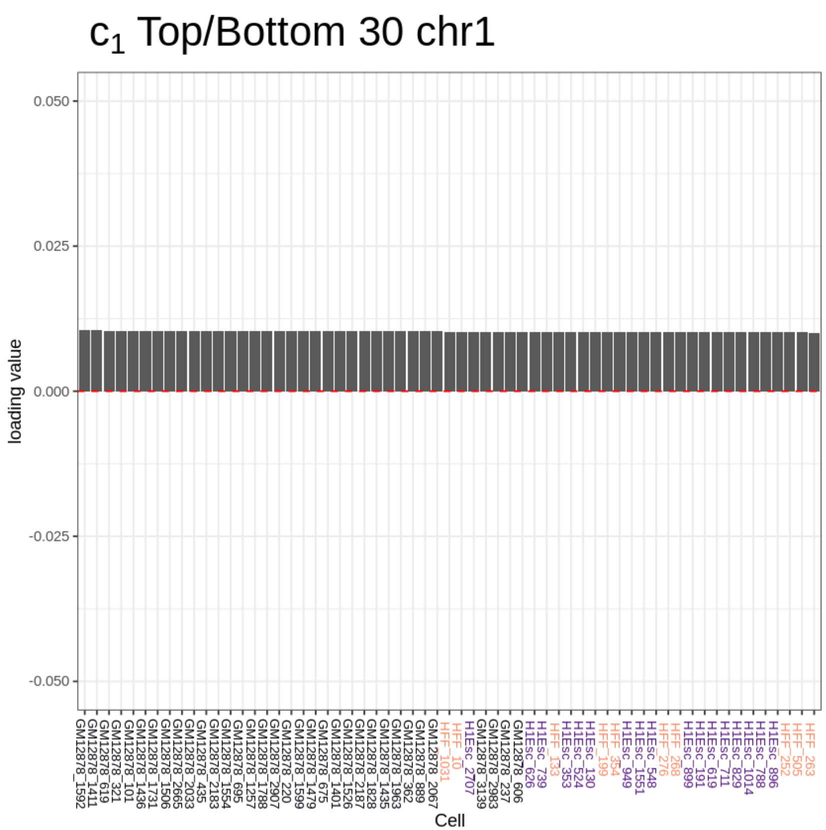
Muscle provides …



Data analysis results



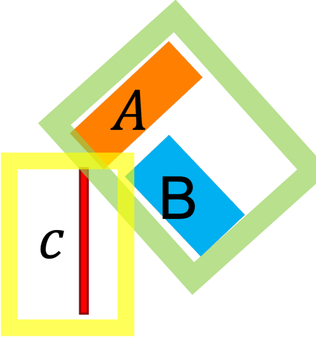
- **Module(cell type) specific "Eigen" Matrix**



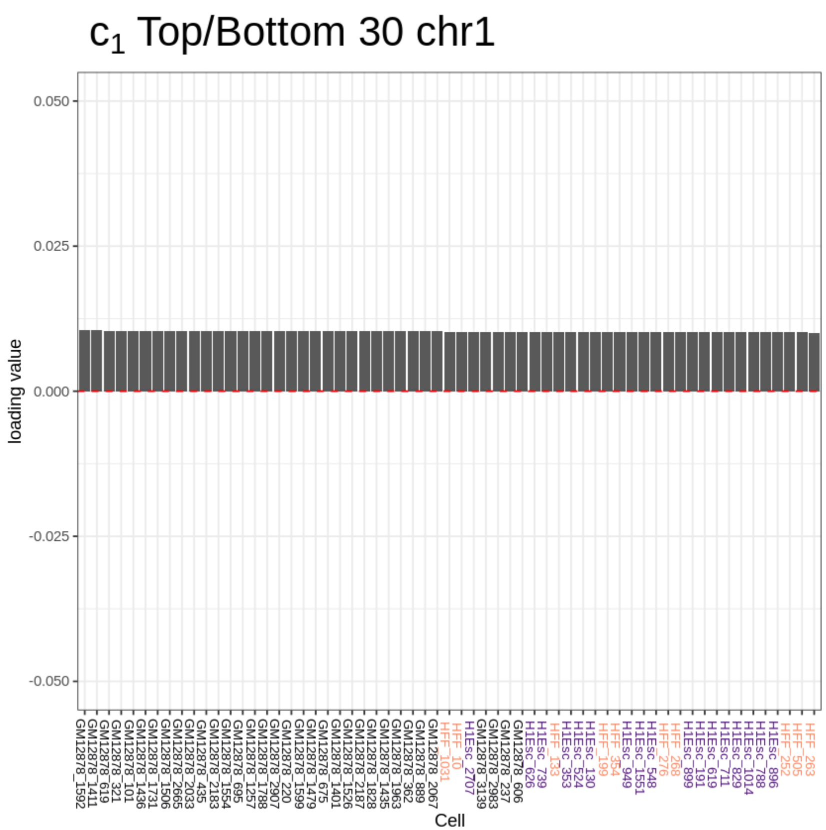
Grand Mean

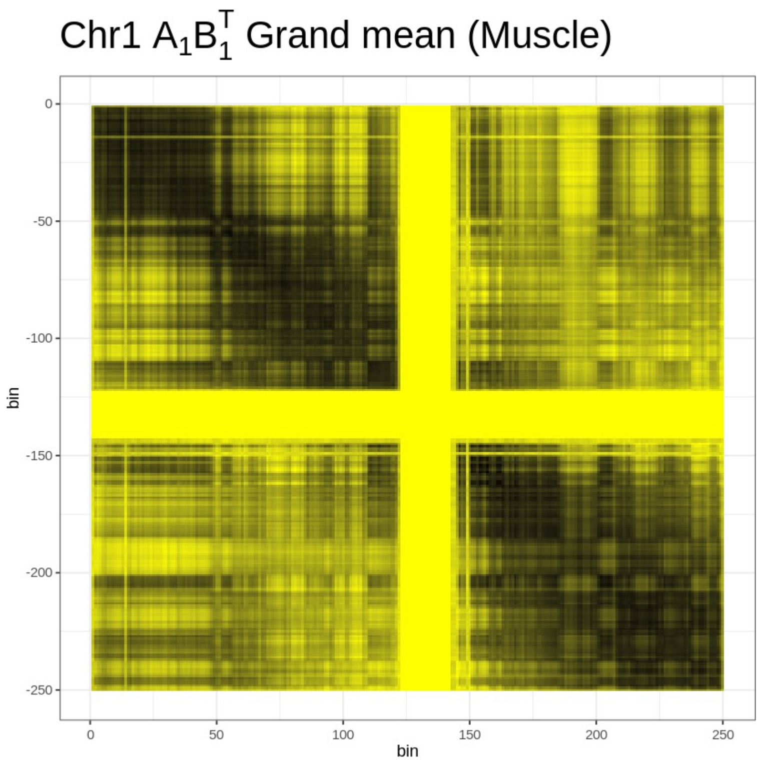
Kim 2020

Data analysis results



- **Module(cell type) specific "Eigen" Matrix**

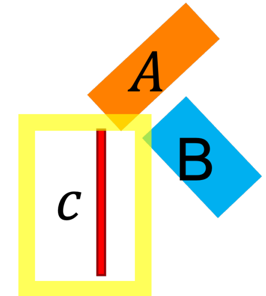




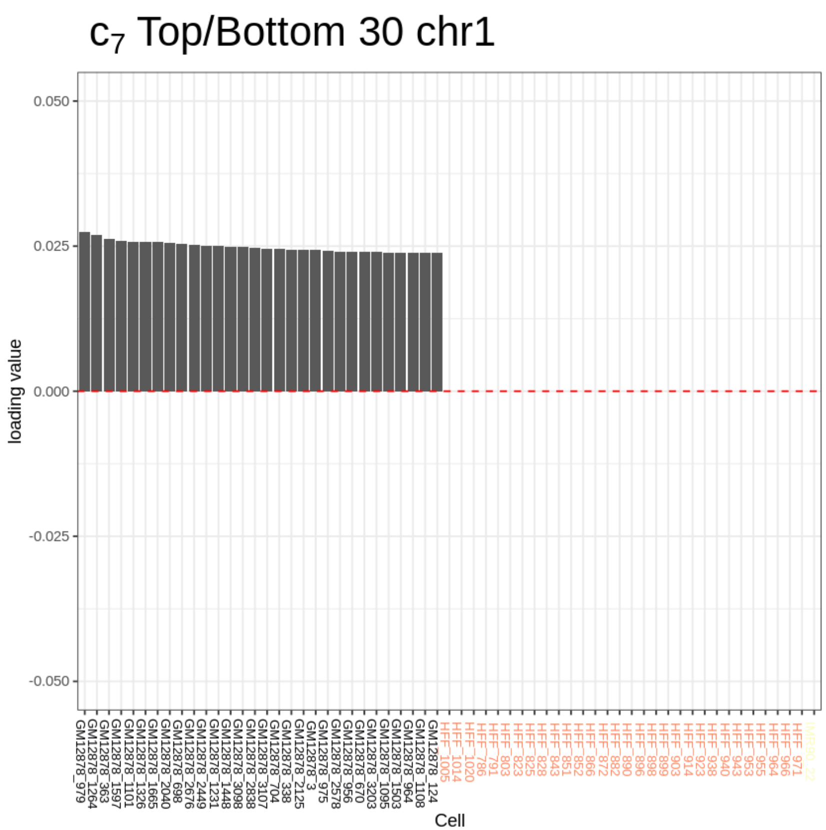
Grand Mean

Kim 2020

Data analysis results



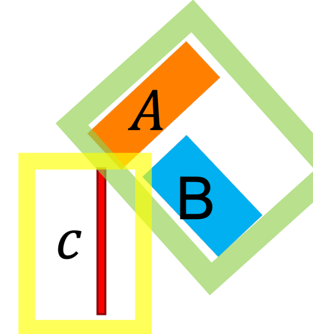
- **Module(cell type) specific "Eigen" Matrix**



GM12878

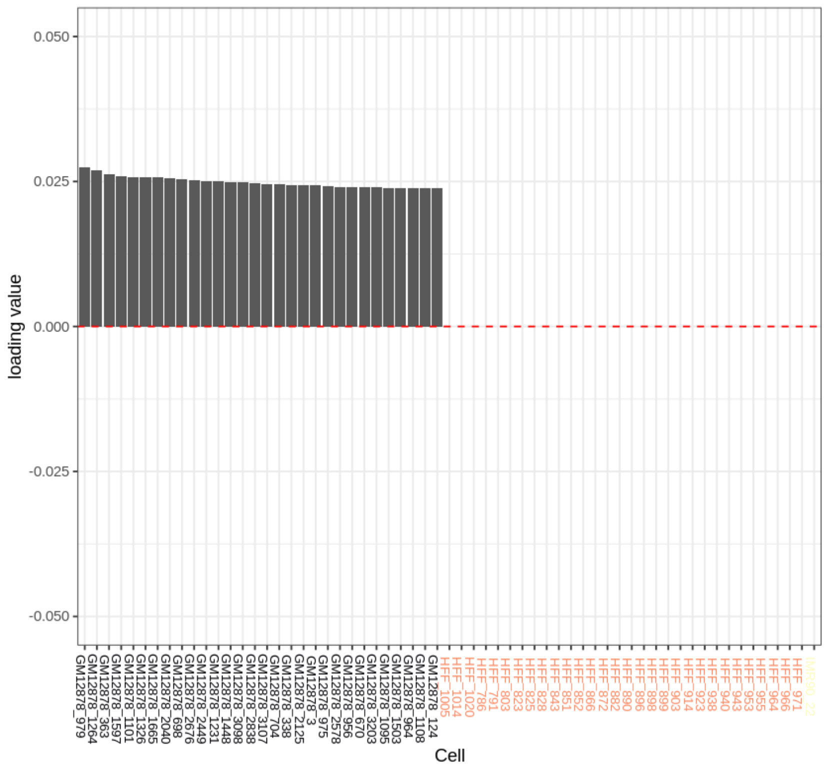
Kim 2020

Data analysis results

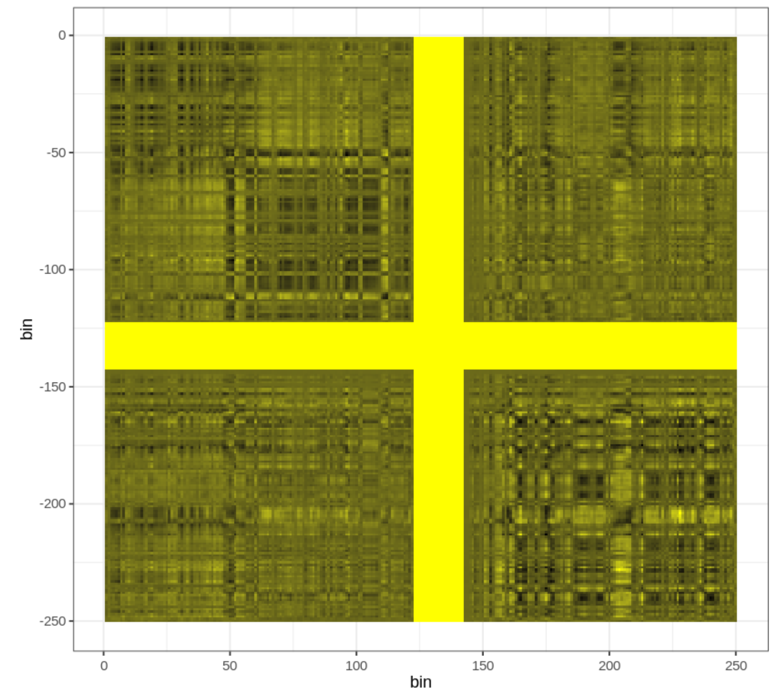


- **Module(cell type) specific "Eigen" Matrix**

c_7 Top/Bottom 30 chr1



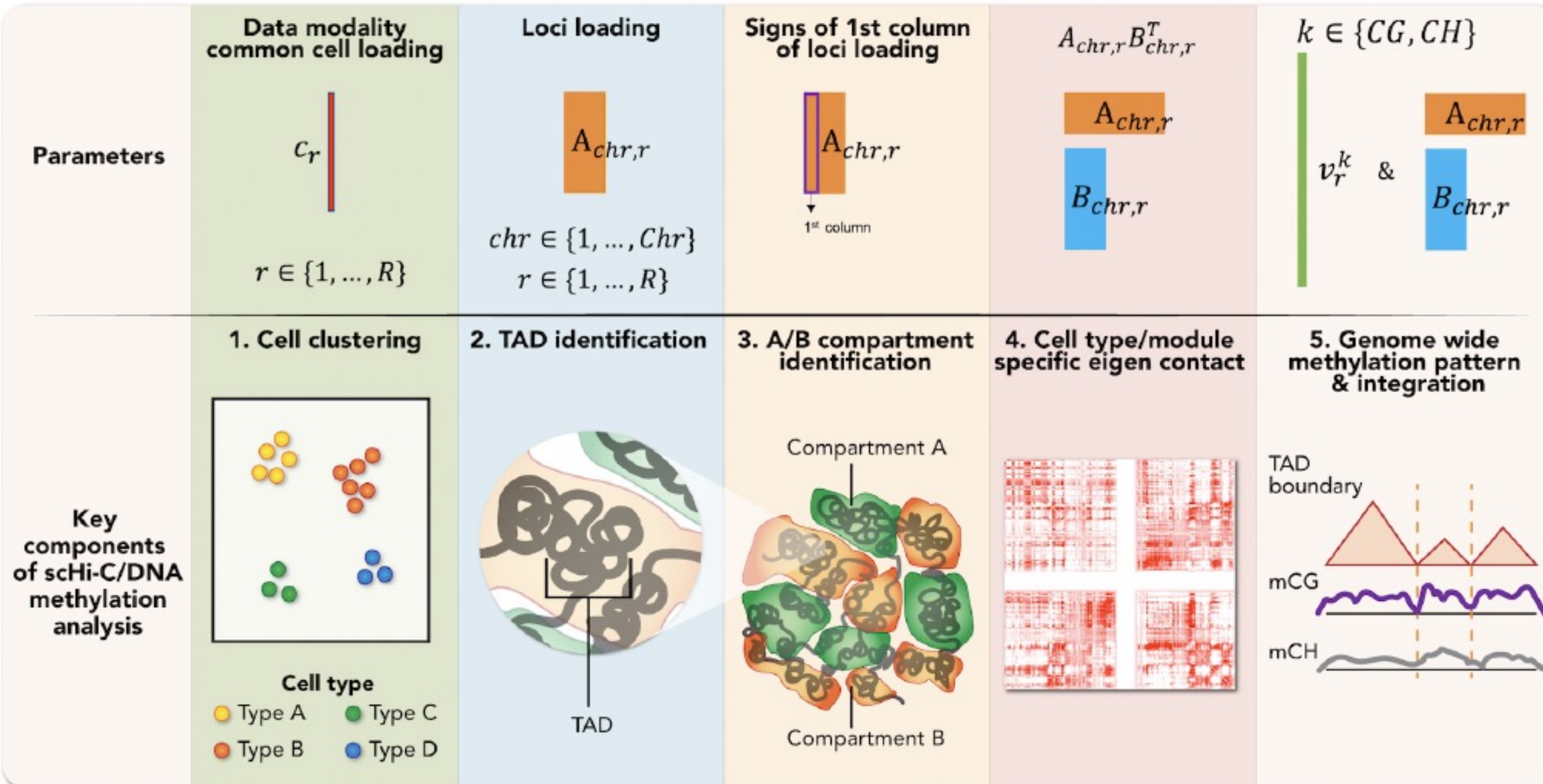
Chr1 $A_7B_7^T$ GM12878 (Muscle)



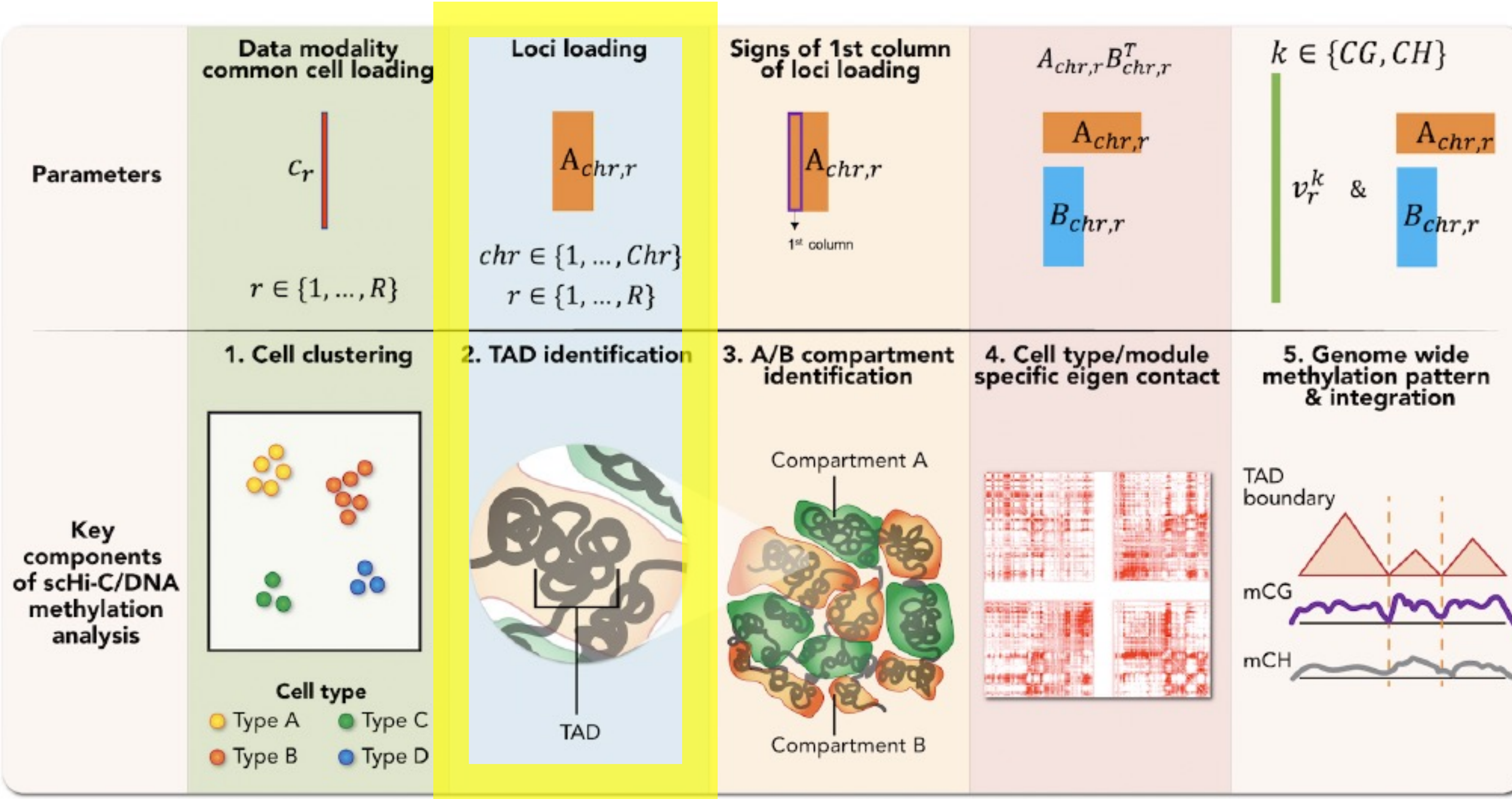
GM12878

Kim 2020

Muscle provides …

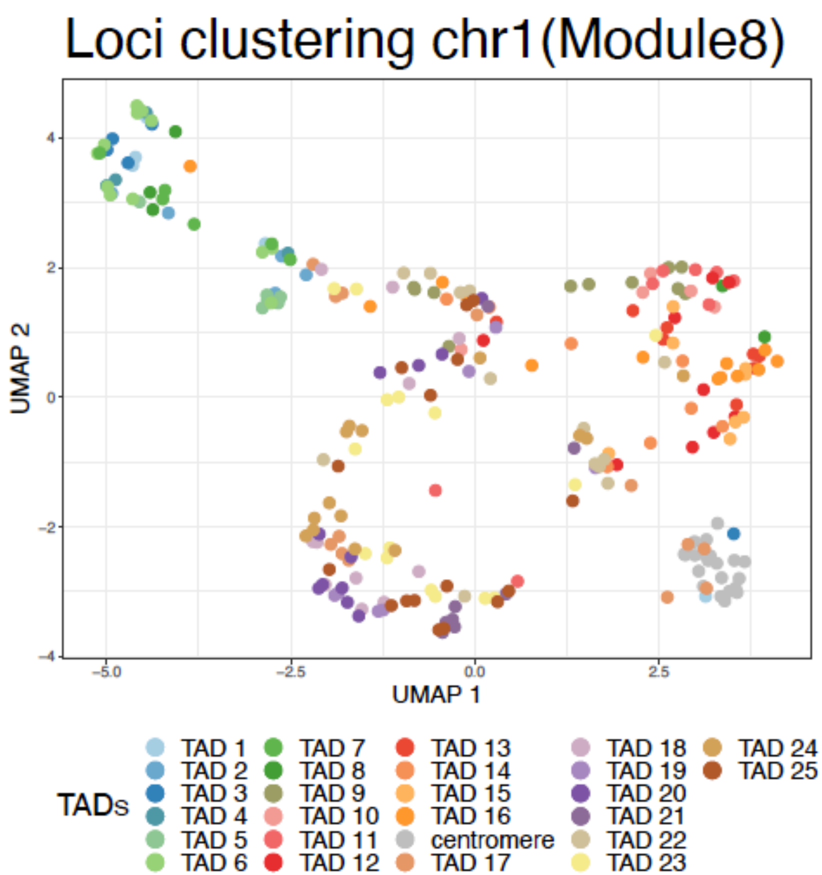


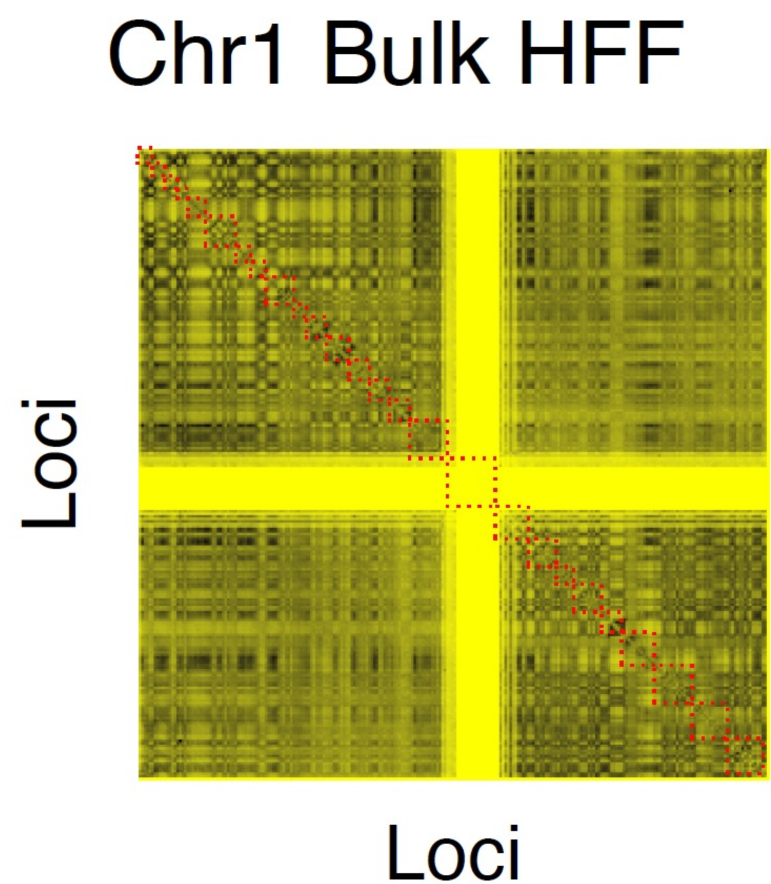
Muscle provides …



Data analysis results

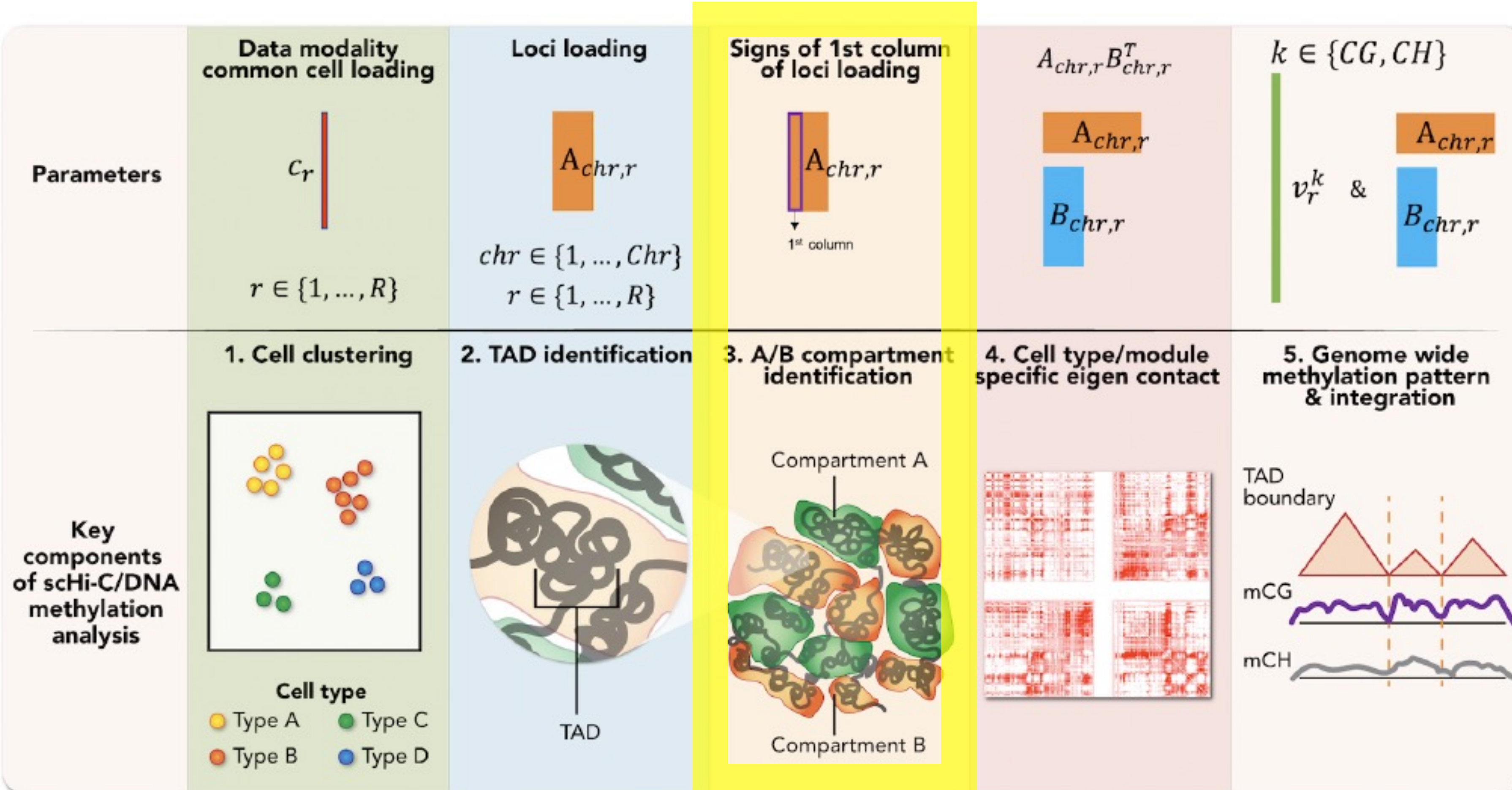
- **Loci clustering (TAD)**





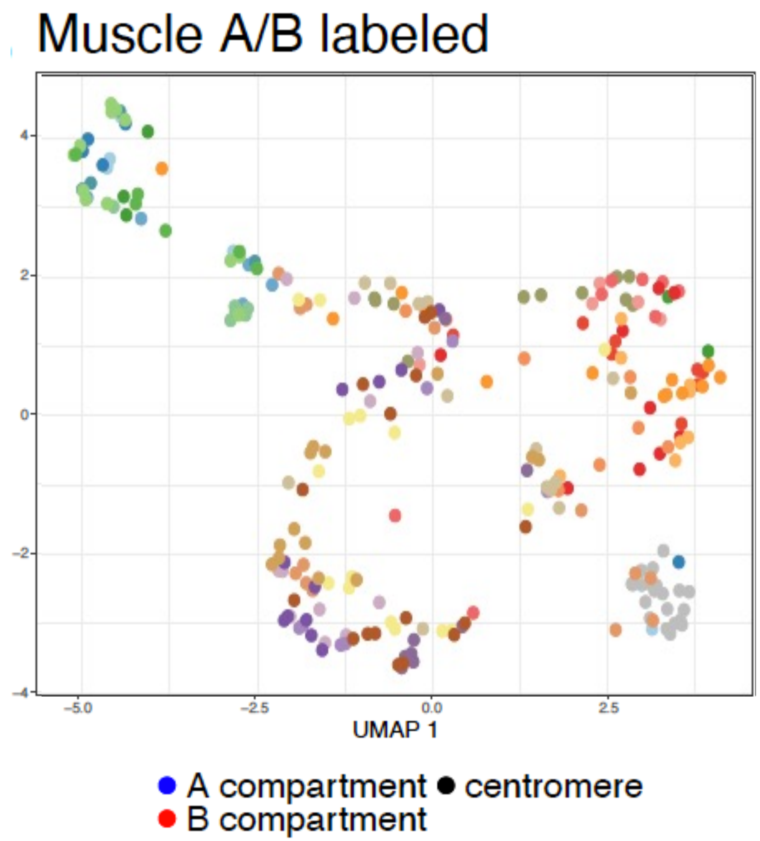
Kim 2020

Muscle provides …



Data analysis results

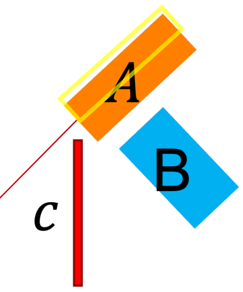
- **A/B compartments**



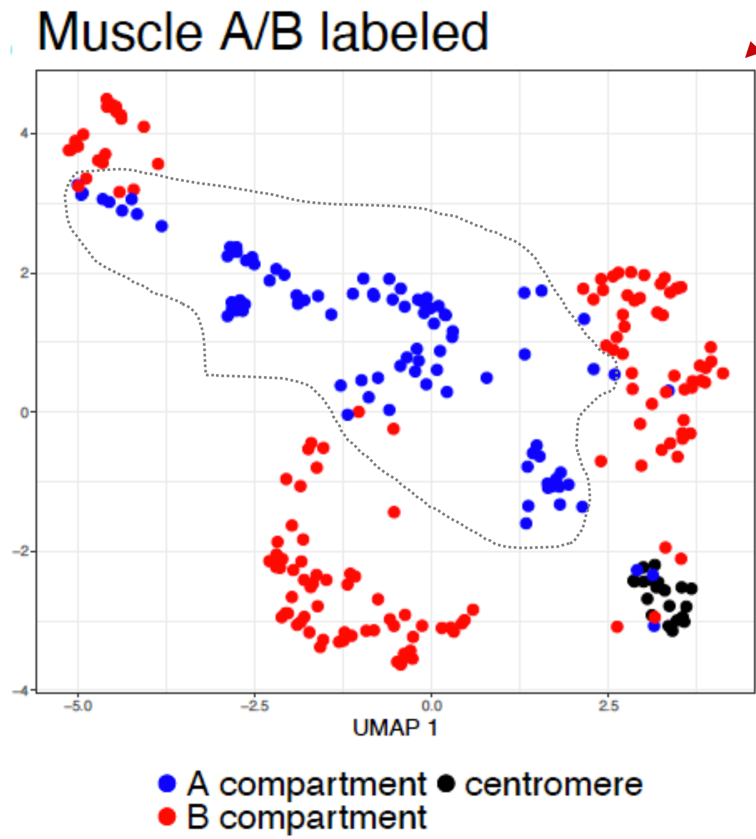
Tensor Module 8 (HFF)

Kim 2020

Data analysis results



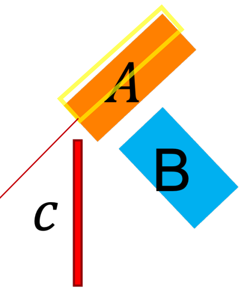
- **A/B compartments**



Tensor Module 8 (HFF)

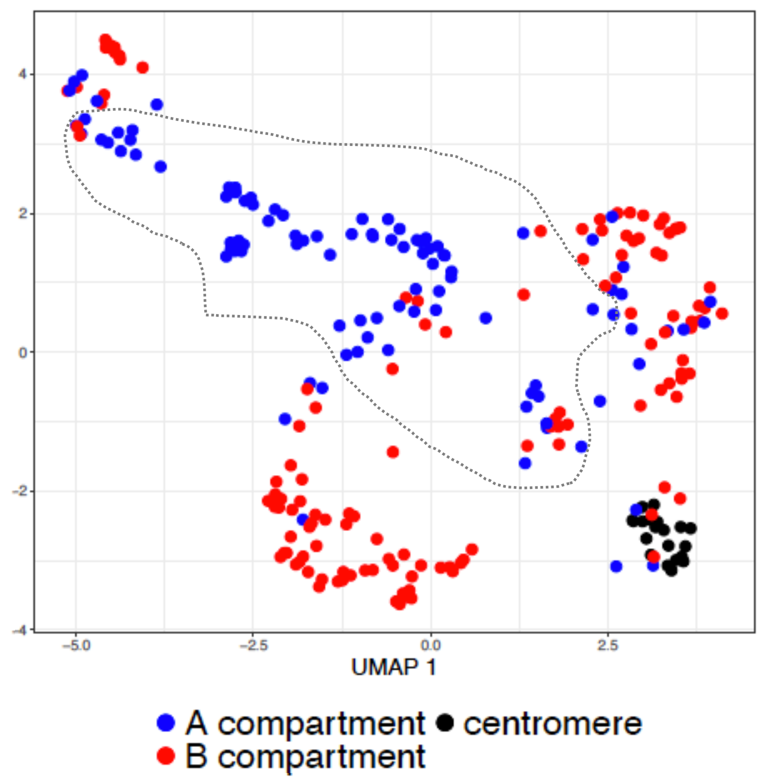
Kim 2020

Data analysis results



- **A/B compartments**

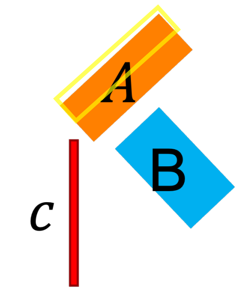
Bulk A/B labeled



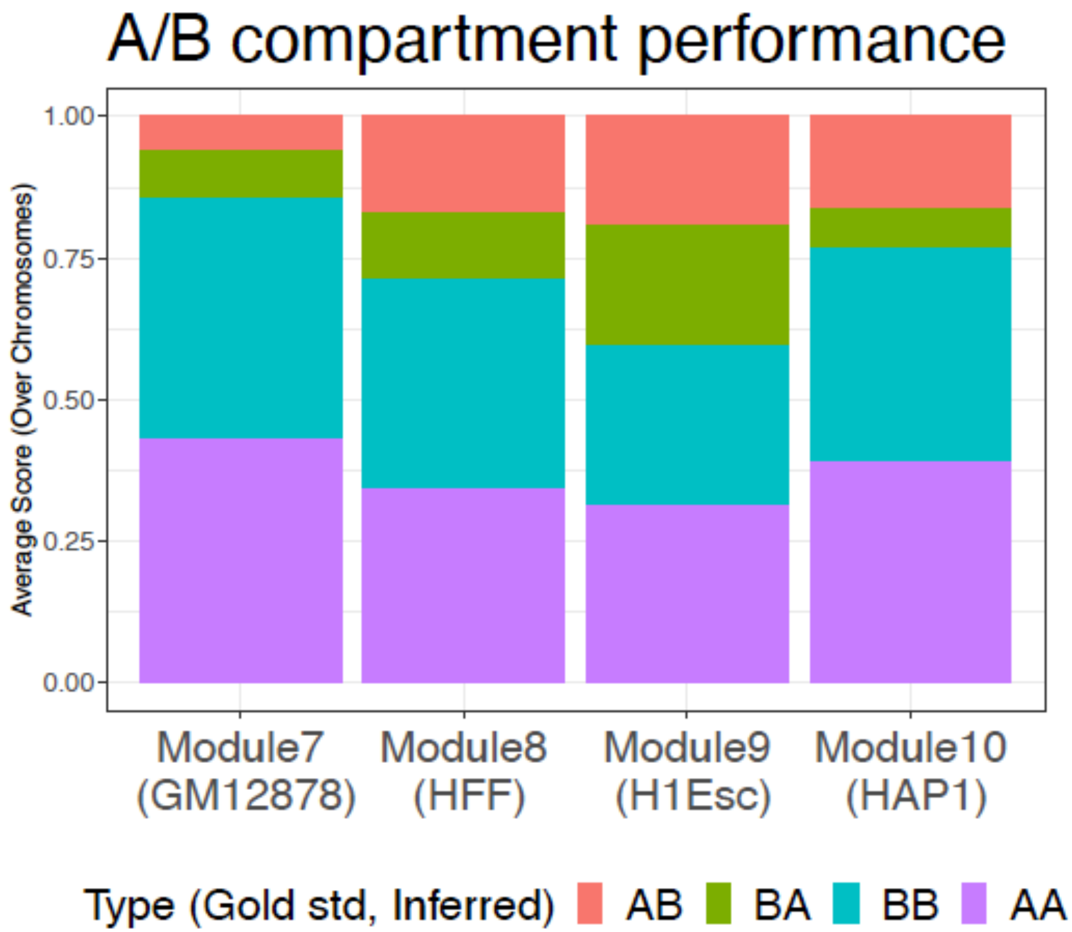
Tensor Module 8 (HFF)

Kim 2020

Data analysis results

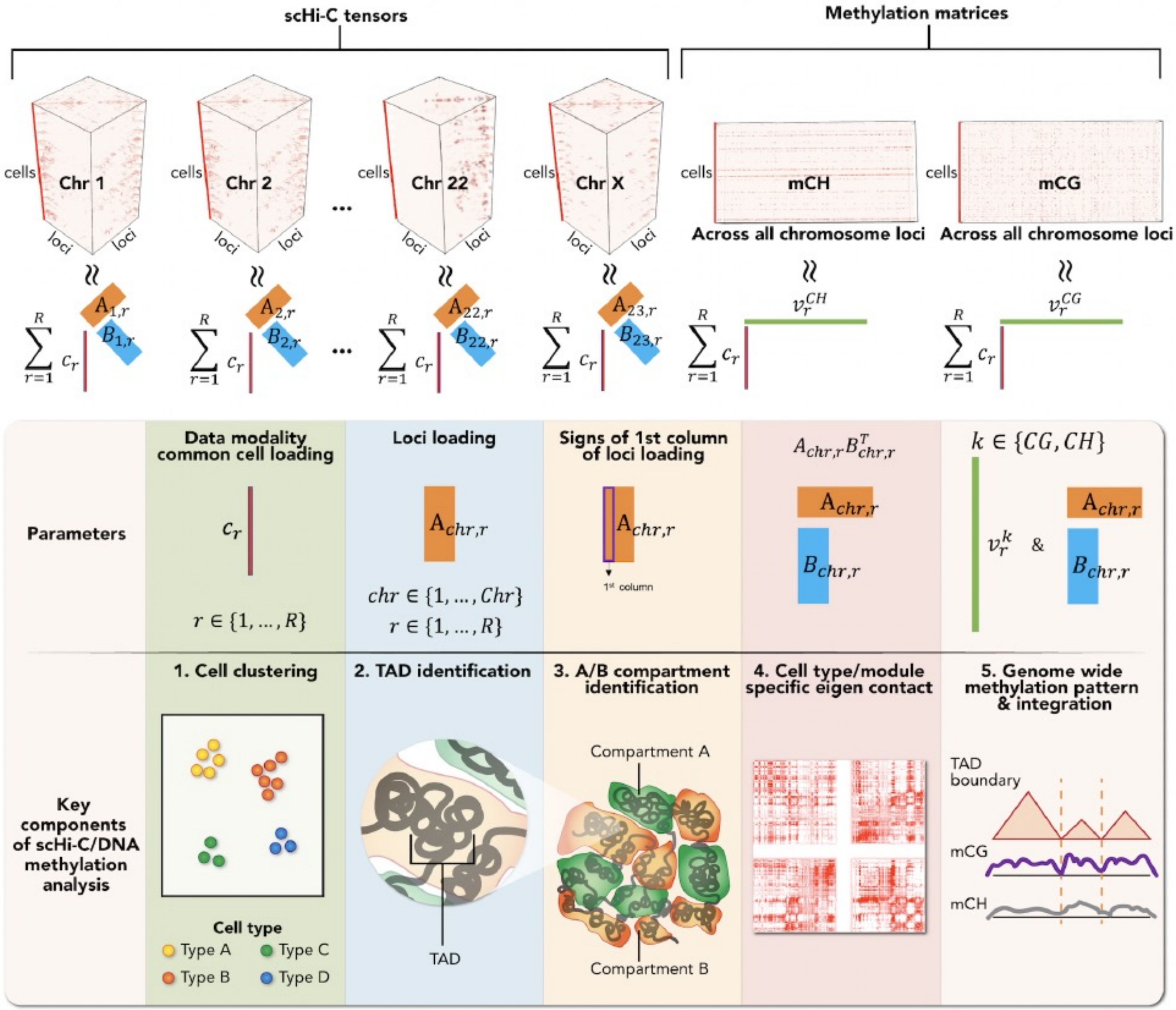


- **A/B compartments and Loci clustering**

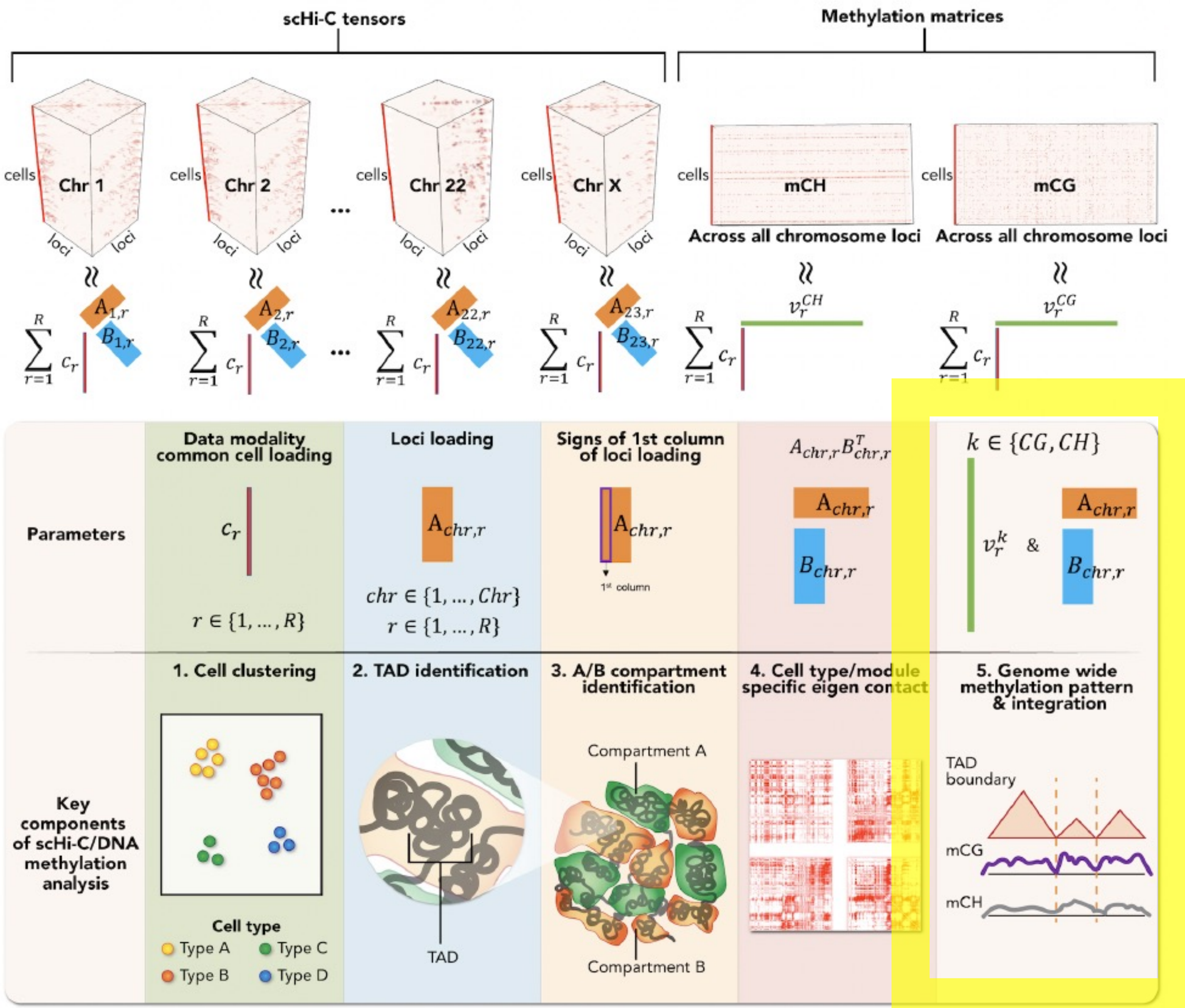


Kim 2020

Muscle provides …

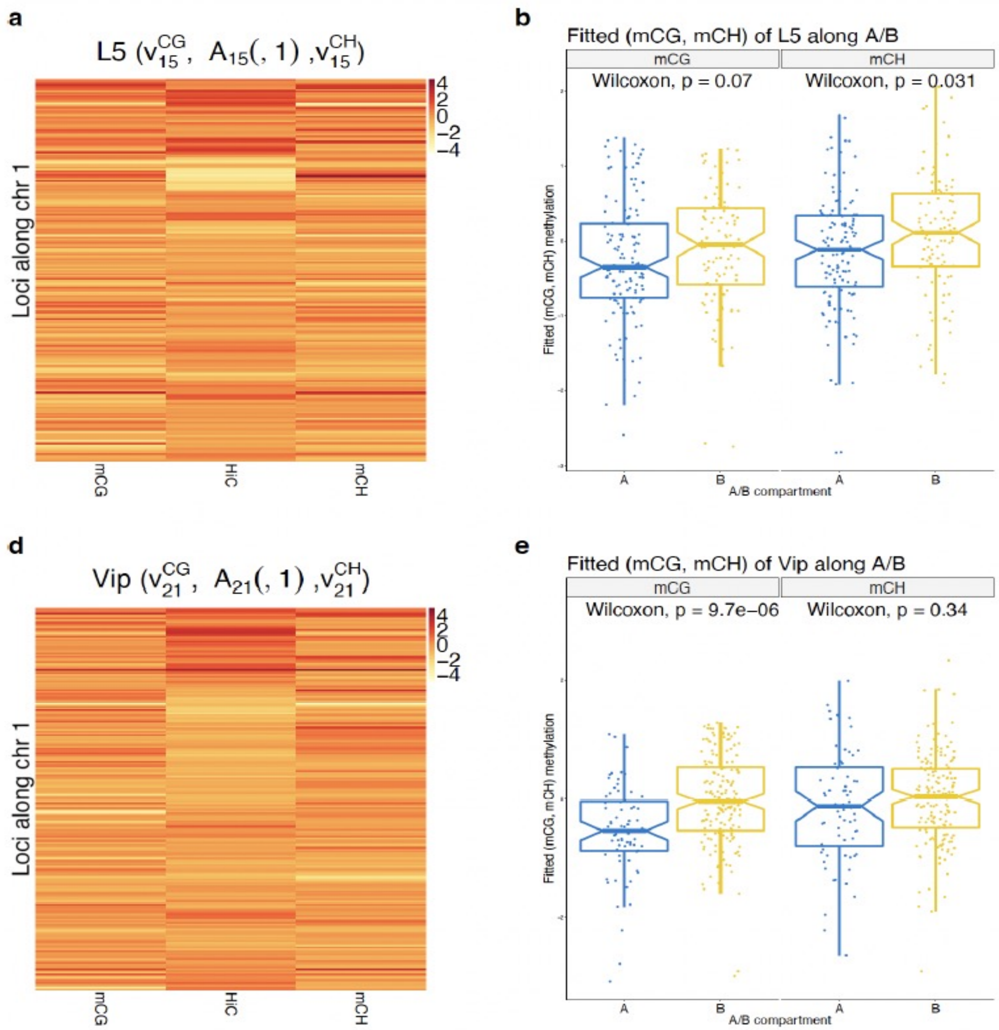


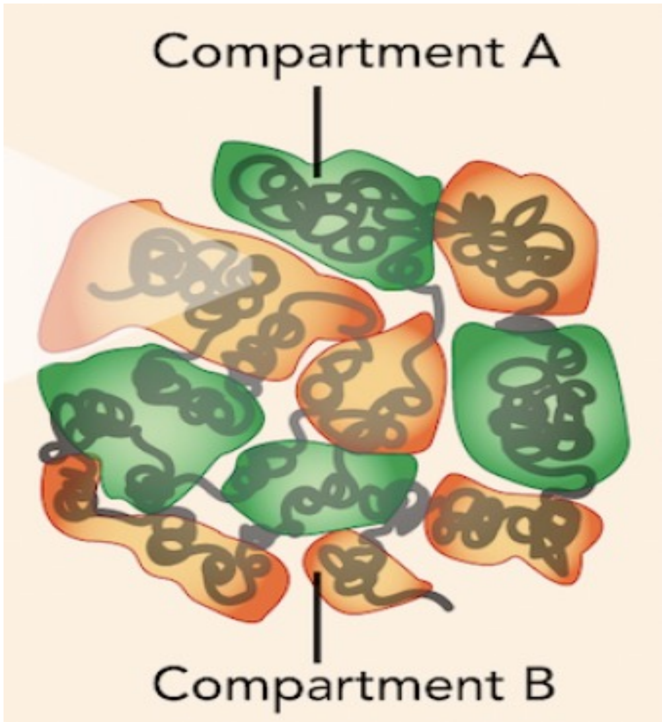
Muscle provides …



Data analysis results

- **Integrative inference**



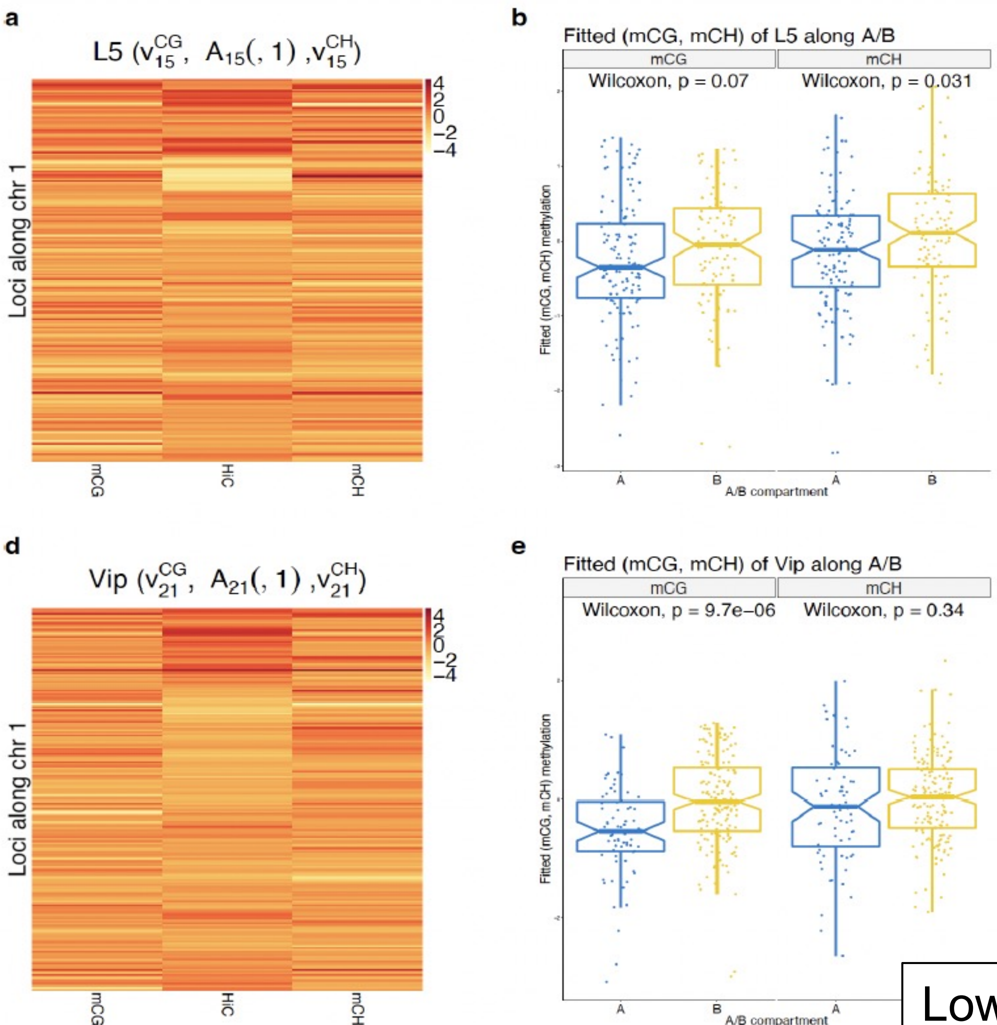


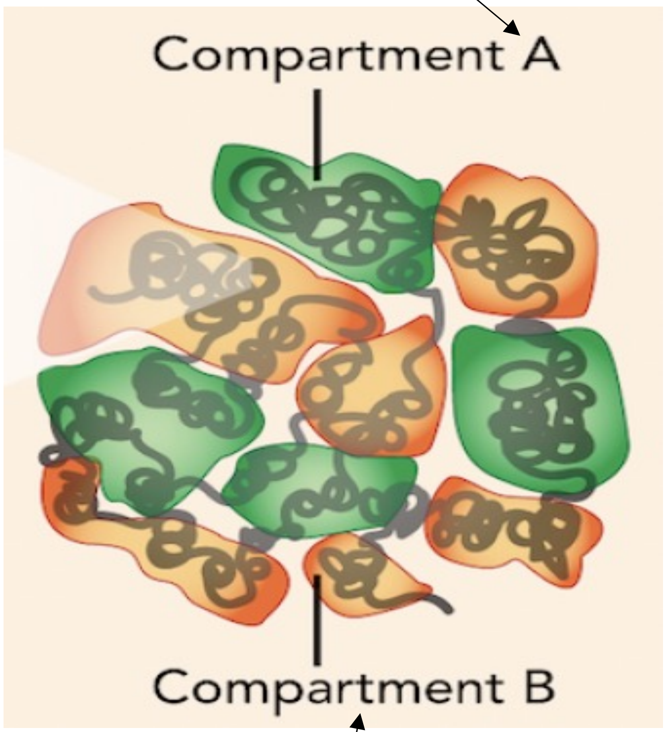
Lee 2019

Data analysis results

- **Integrative inference**

High gene expression / Low methylation



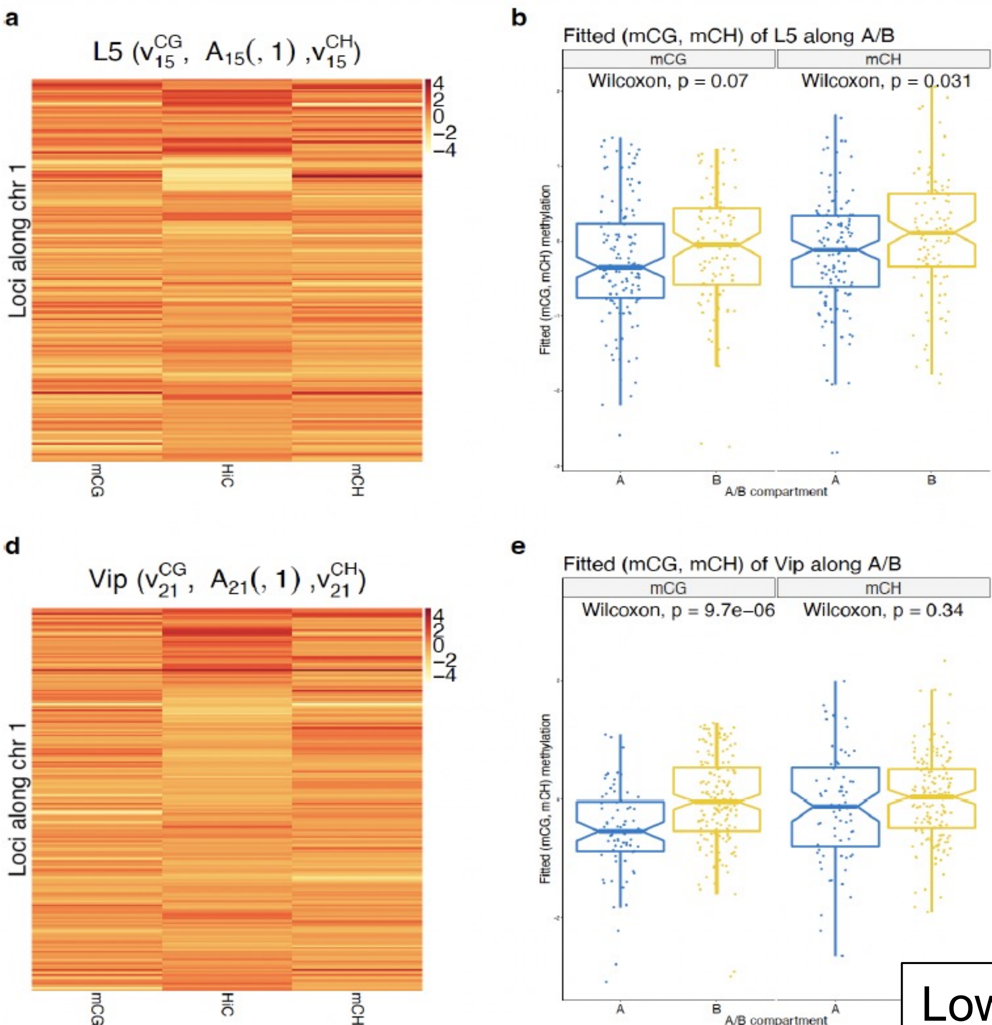


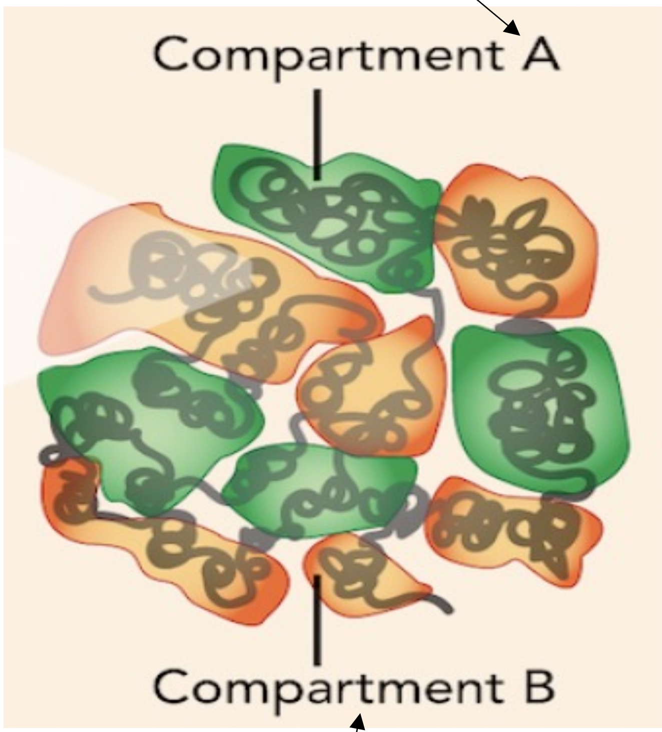
Low gene expression / High methylation

Data analysis results

- **Integrative inference**

High gene expression / Low methylation





Low gene expression / High methylation

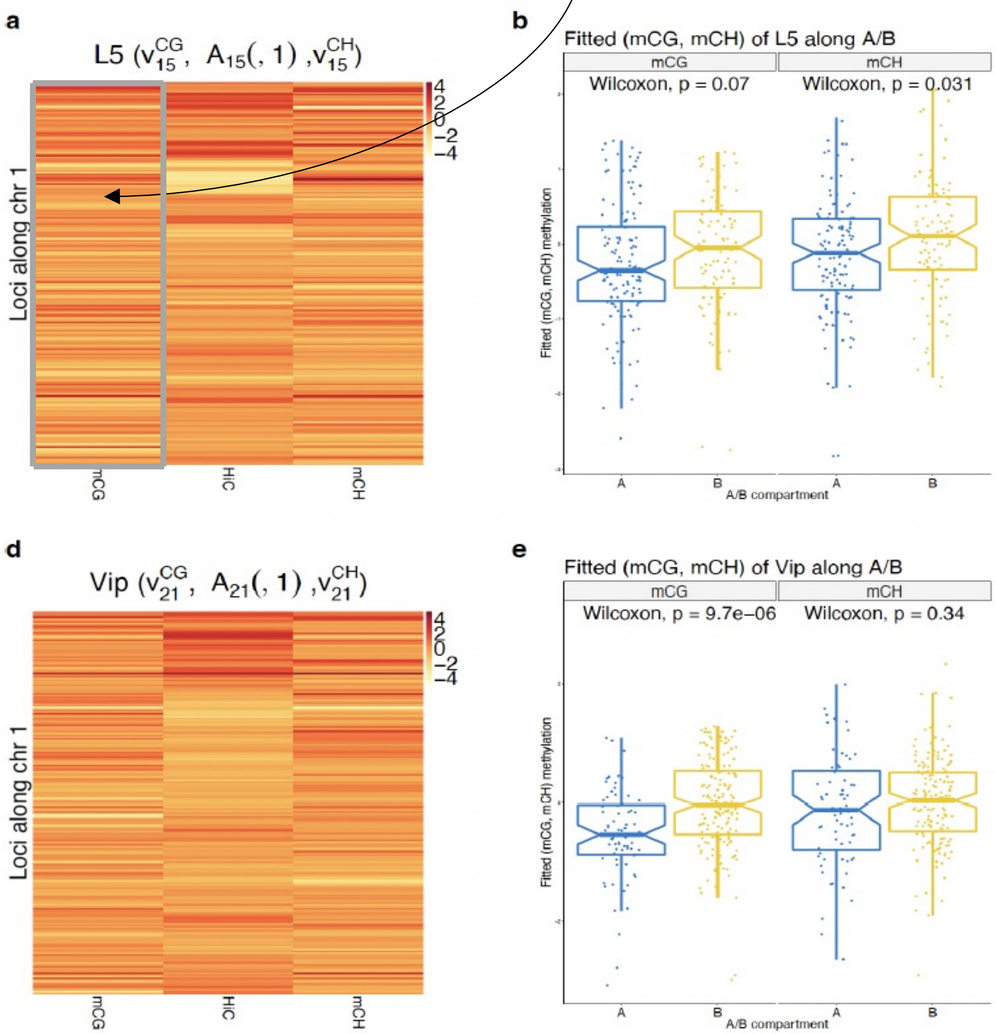
Data analysis results

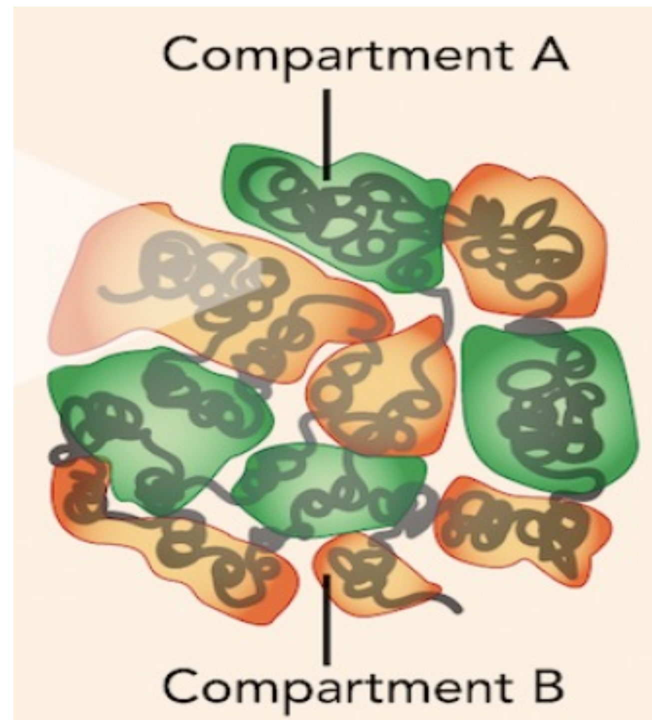
CG

v^k & A

$k \in \{CG, CH\}$

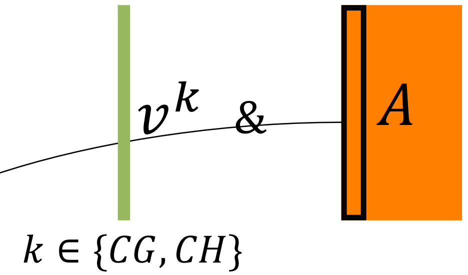
- **Integrative inference**



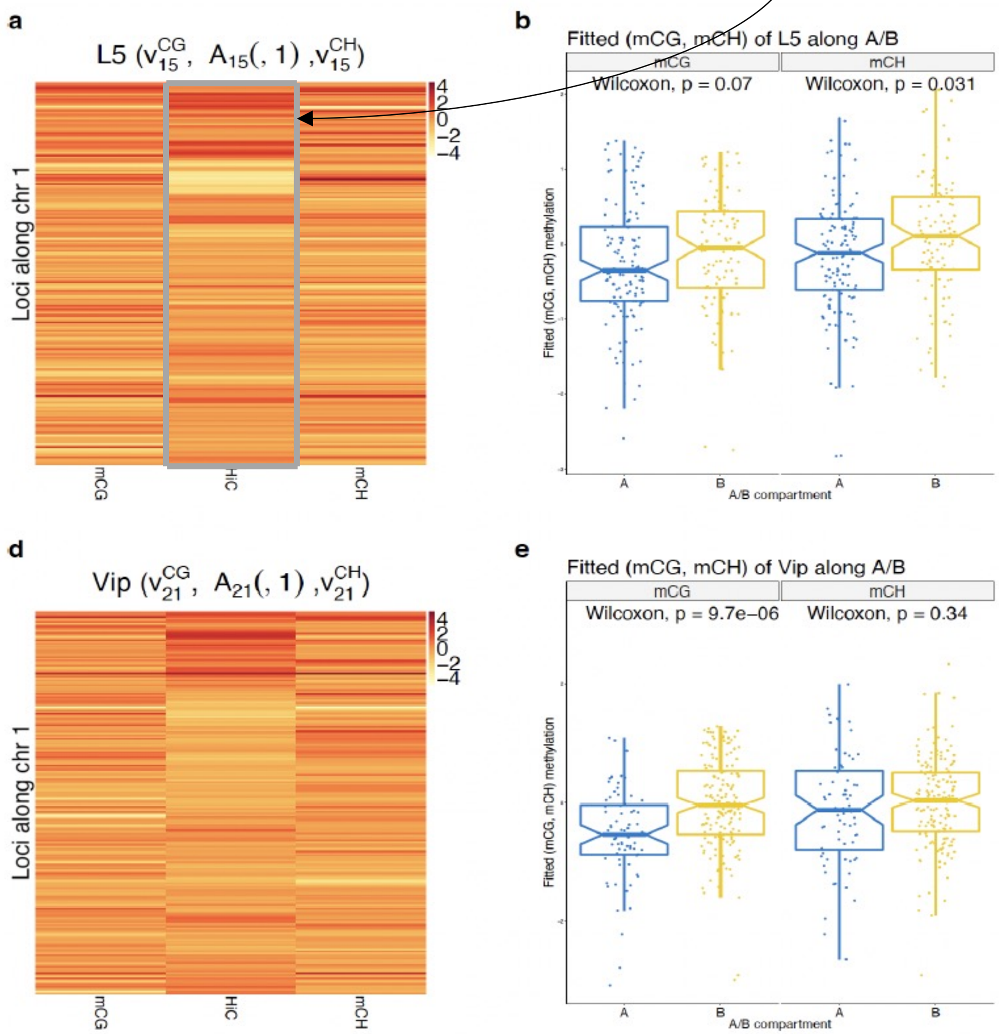


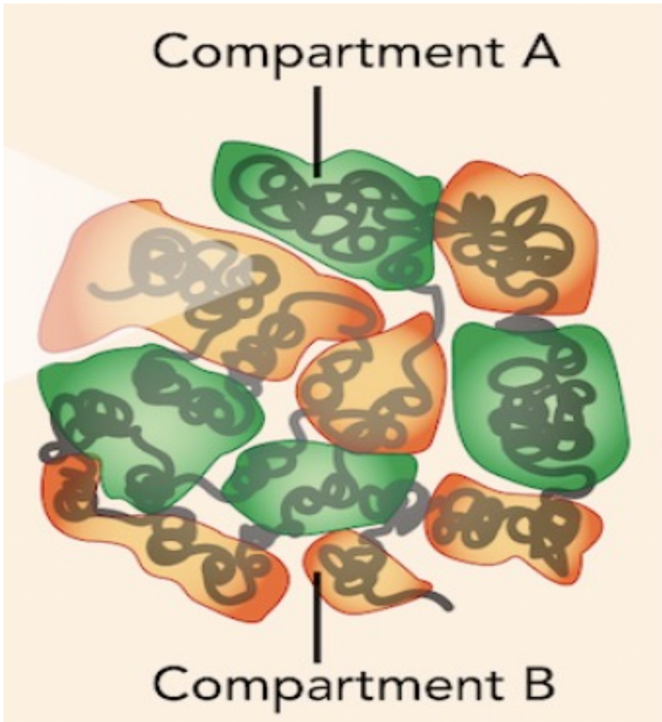
Lee 2019

Data analysis results



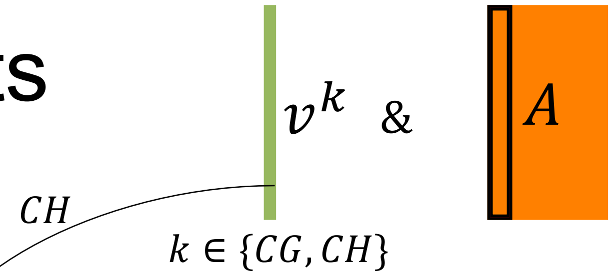
- **Integrative inference**



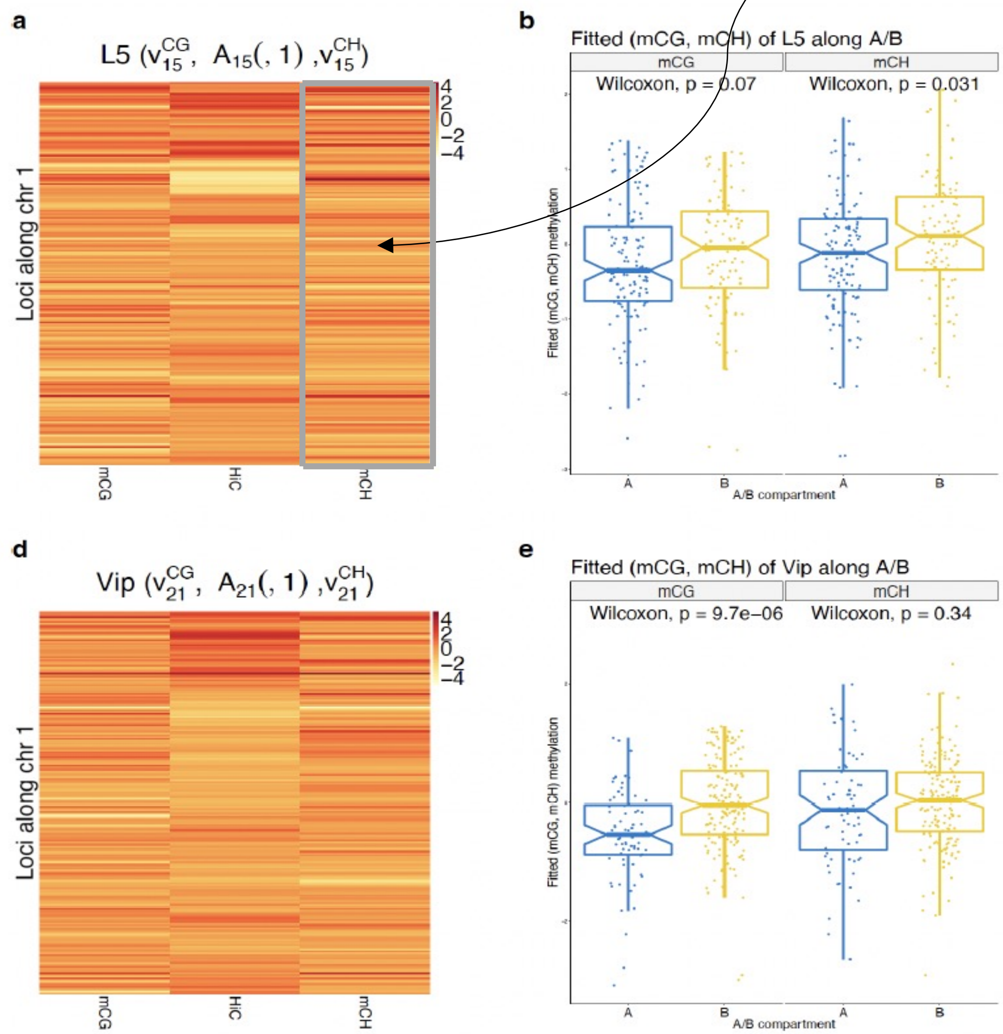


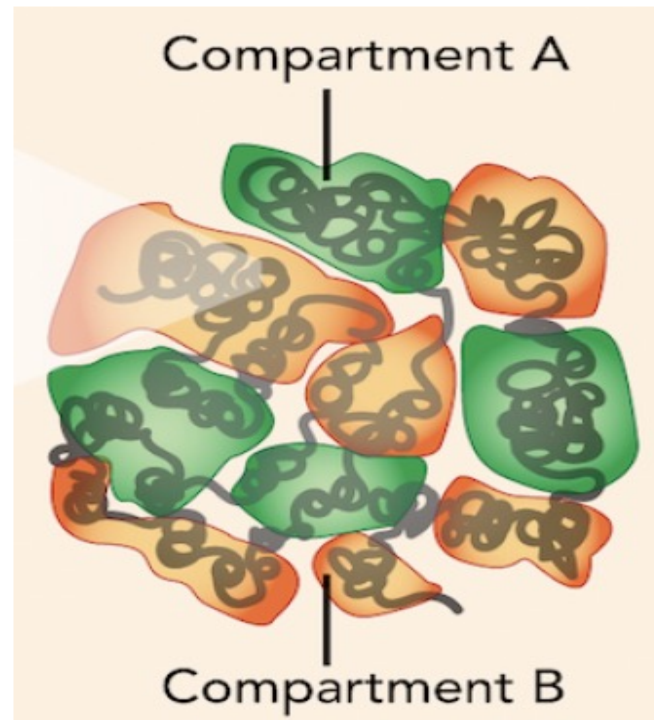
Lee 2019

Data analysis results



- **Integrative inference**



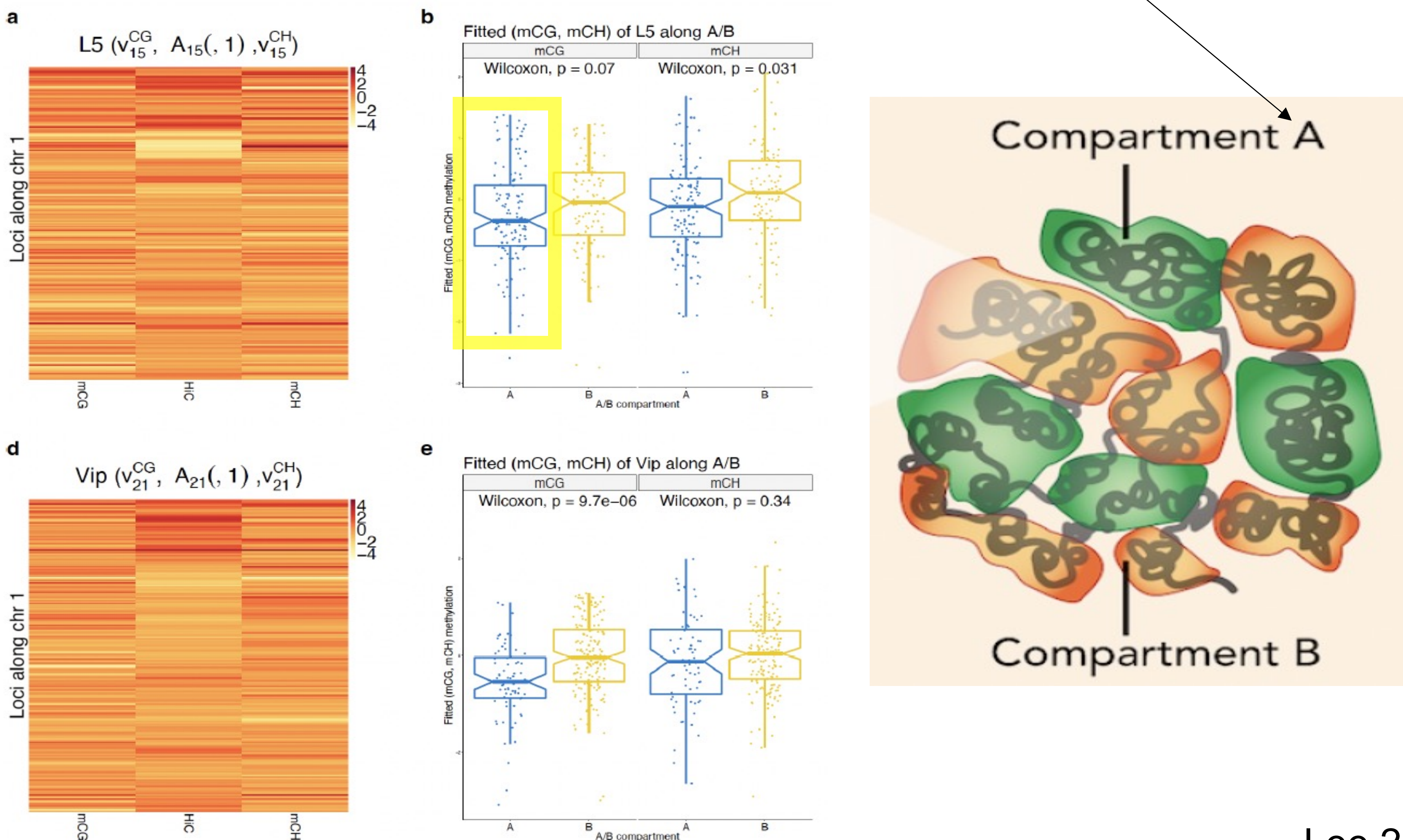


Lee 2019

Data analysis results

- **Integrative inference**

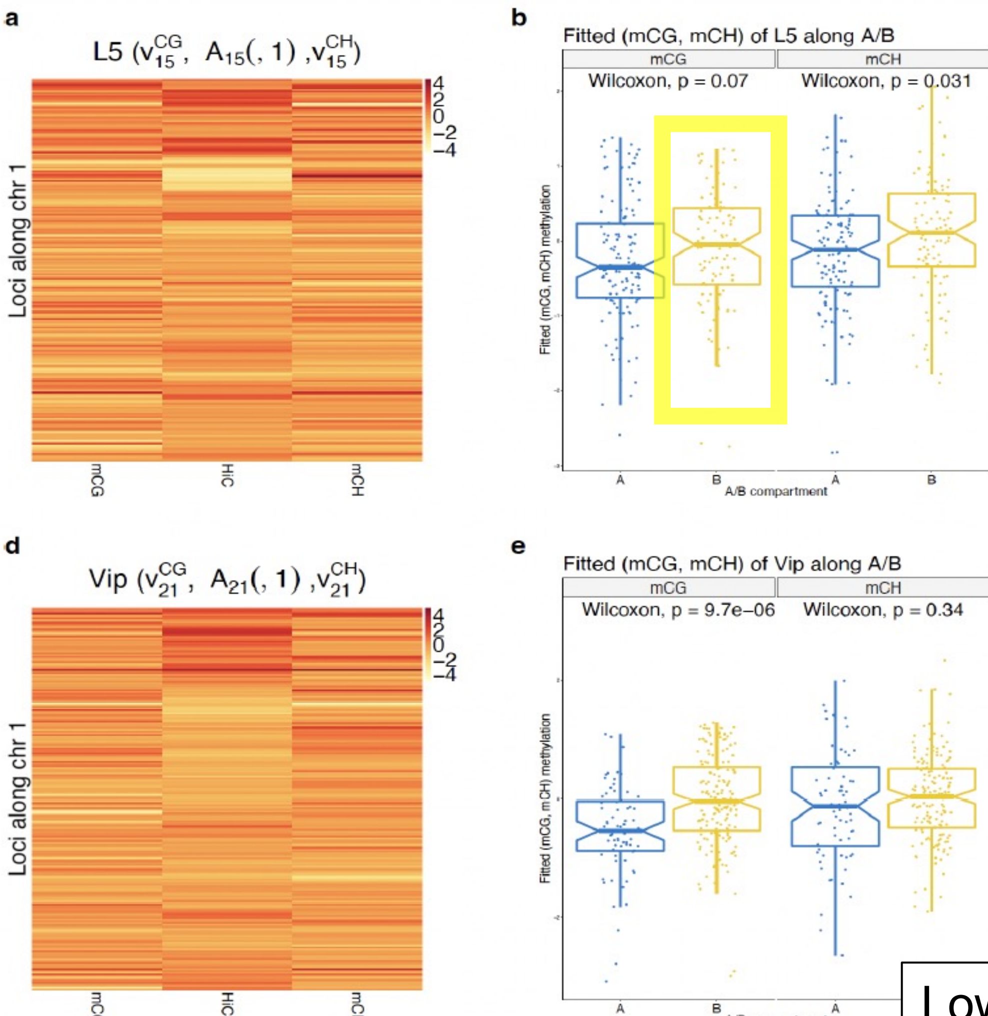
High gene expression / Low methylation

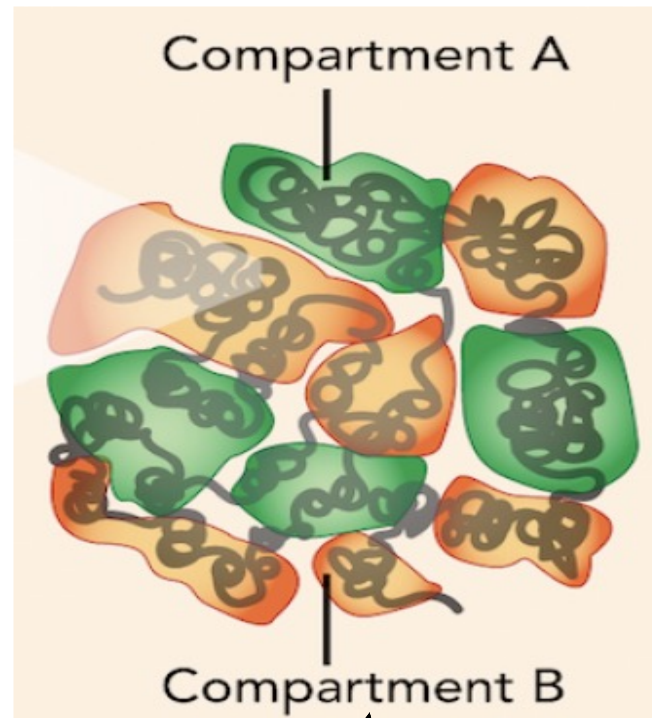


Lee 2019

Data analysis results

- **Integrative inference**

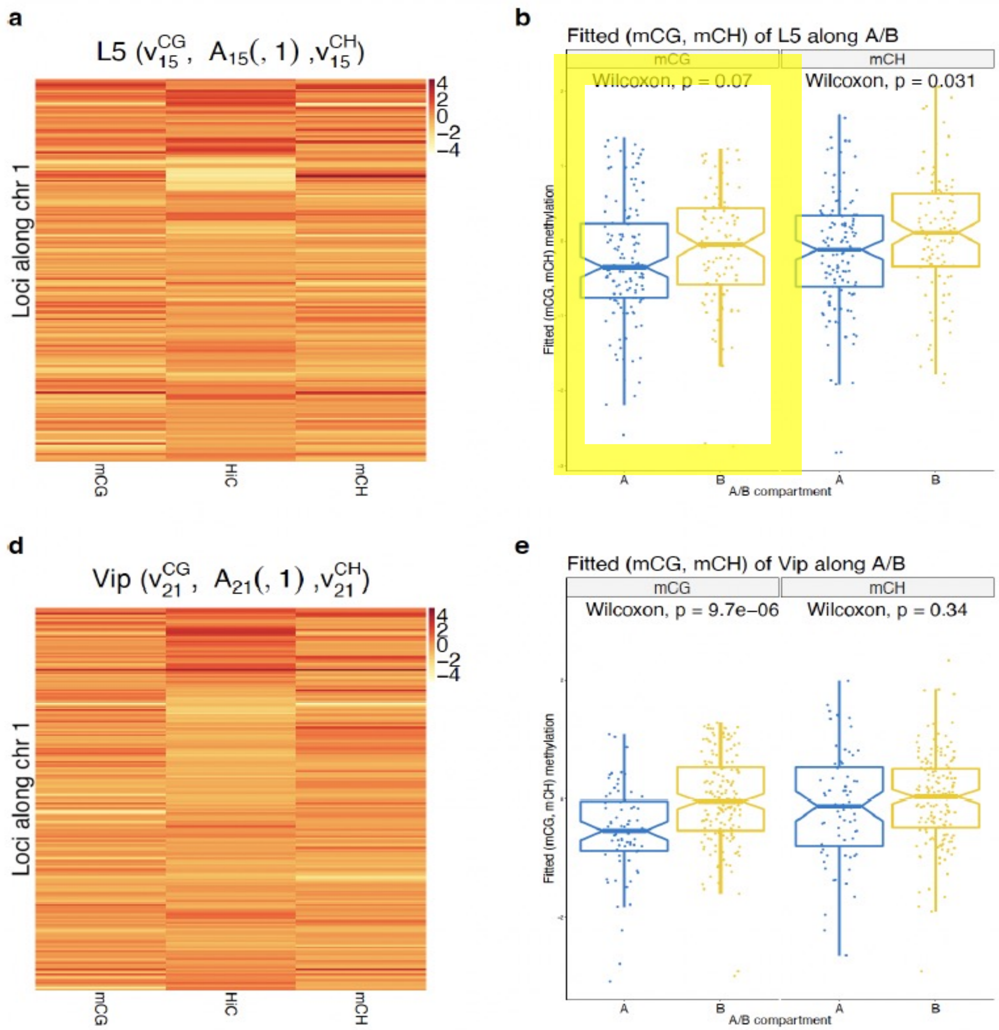


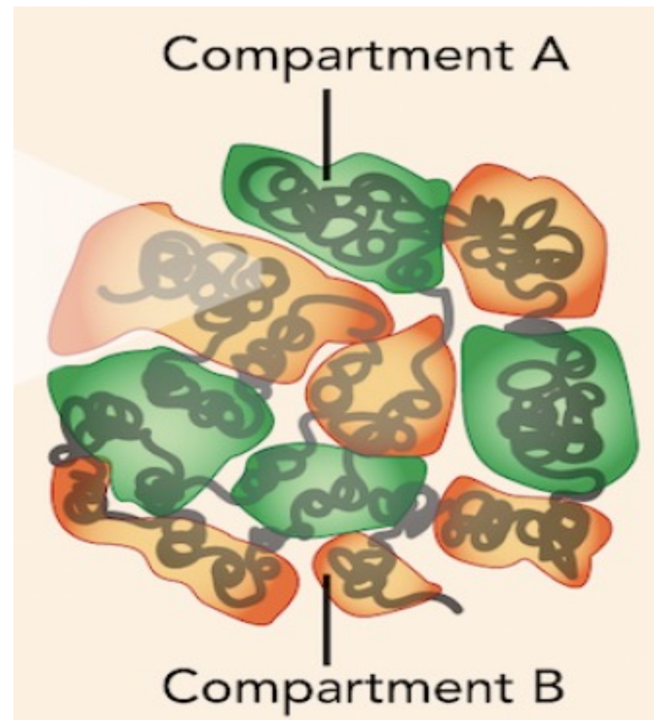


Low gene expression / High methylation

Data analysis results

- **Integrative inference**

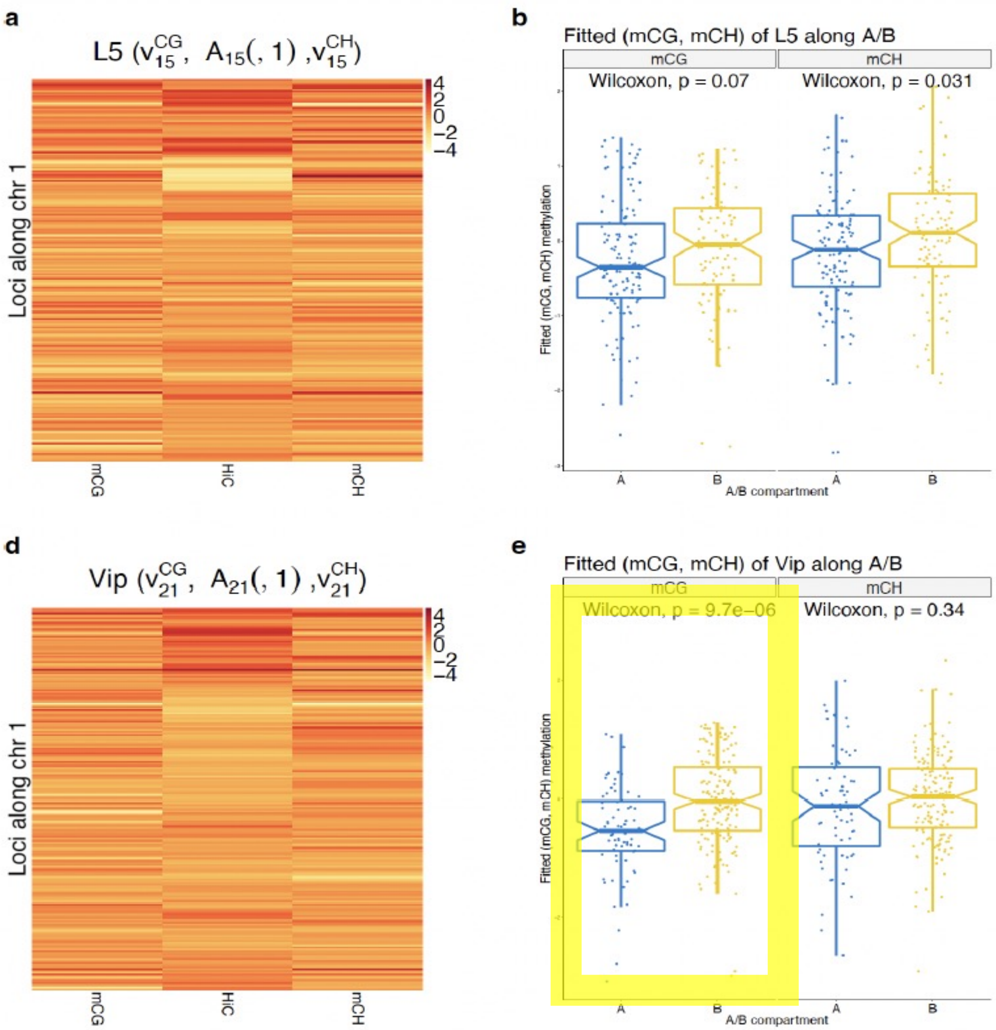


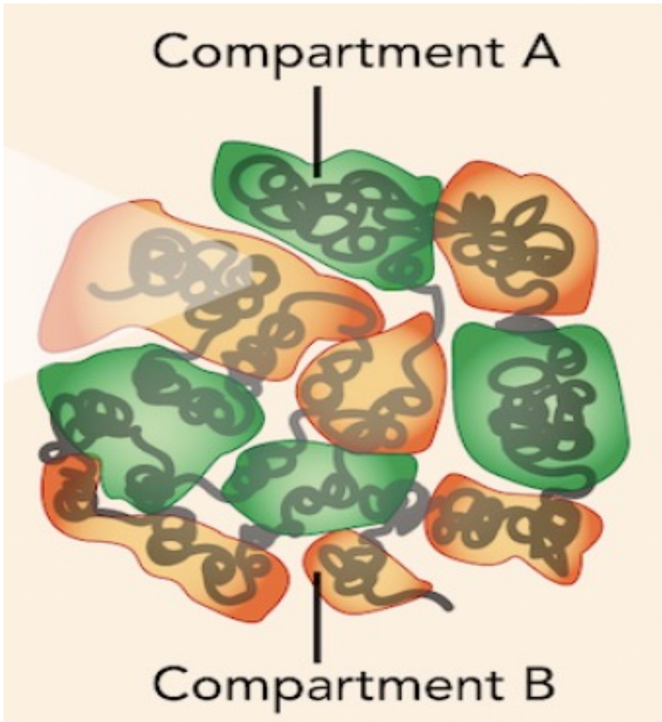


Lee 2019

Data analysis results

- **Integrative inference**

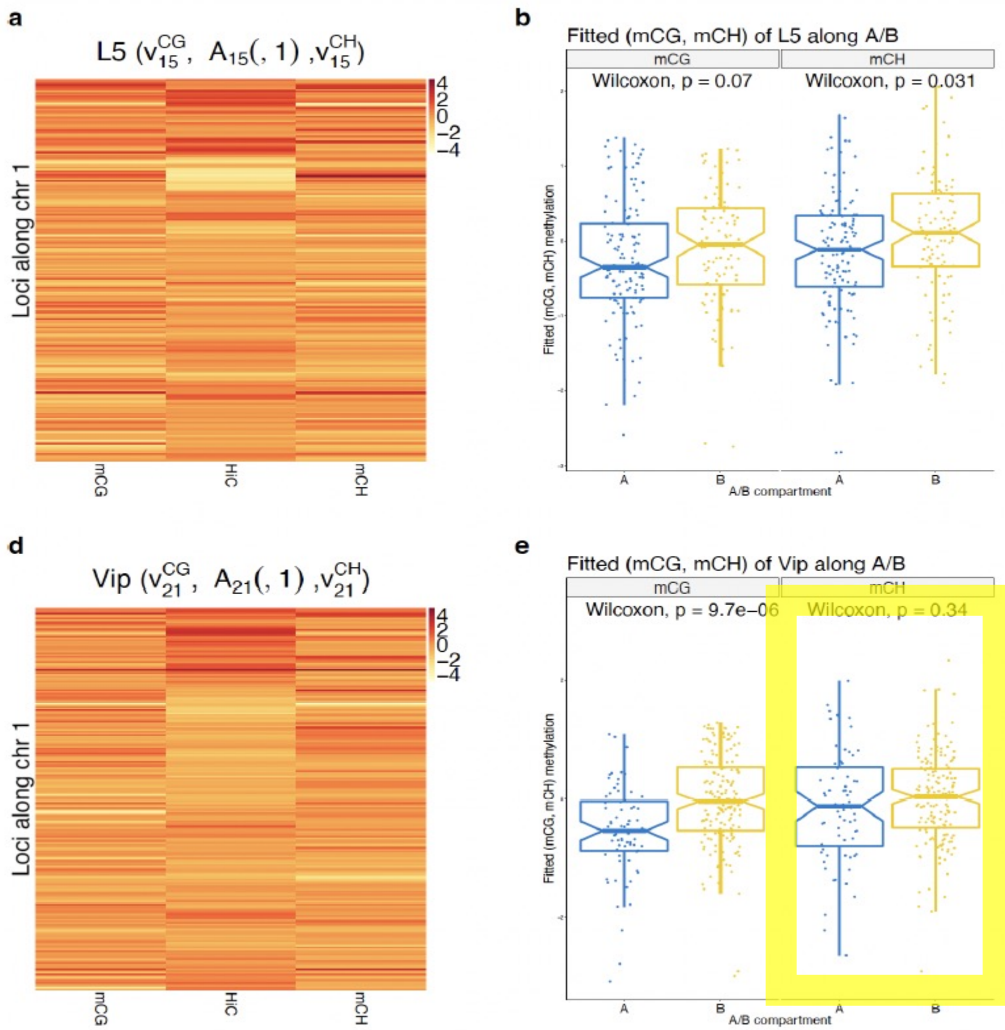


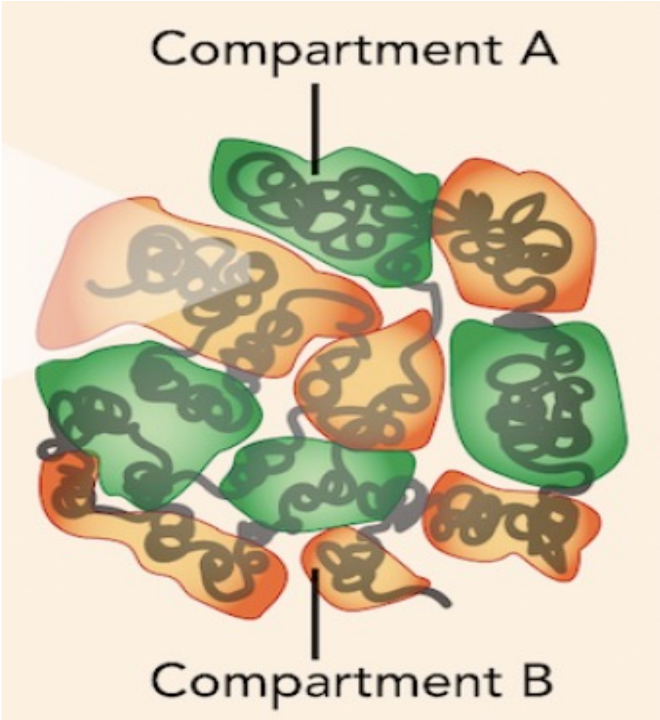


Lee 2019

Data analysis results

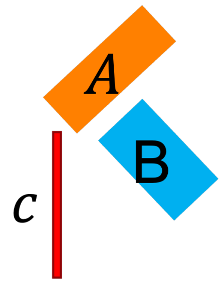
- **Integrative inference**





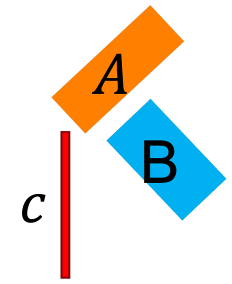
Lee 2019

Take aways and discussion



- Muscle provides cell type clustering ( )

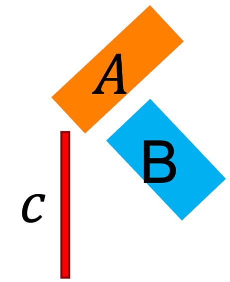
Take aways and discussion



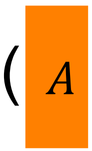
- Muscle provides cell type clustering ( )

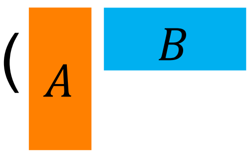
  provides TAD, A/B compartments (A)

Take aways and discussion



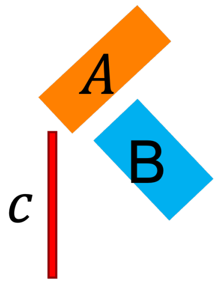
- Muscle provides cell type clustering ()

  provides TAD, A/B compartments ()

  provides cell type specific contact pattern ()

  as **model parameters.**

Take aways and discussion



- Muscle provides cell type clustering ( )

  provides TAD, A/B compartments (A)

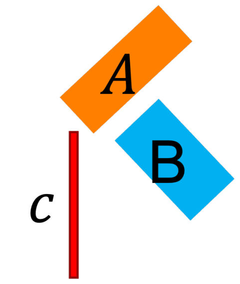
  provides cell type specific contact pattern (A B)

  as **model parameters.**

➡ Understanding of **physical proximity** of genome

cell type specific manner

Take aways and discussion



- Muscle provides cell type clustering ( )

  provides TAD, A/B compartments (A)

  provides cell type specific contact pattern (A B)

  as **model parameters.**

  ⇒ Understanding of **physical proximity** of genome

  cell type specific manner

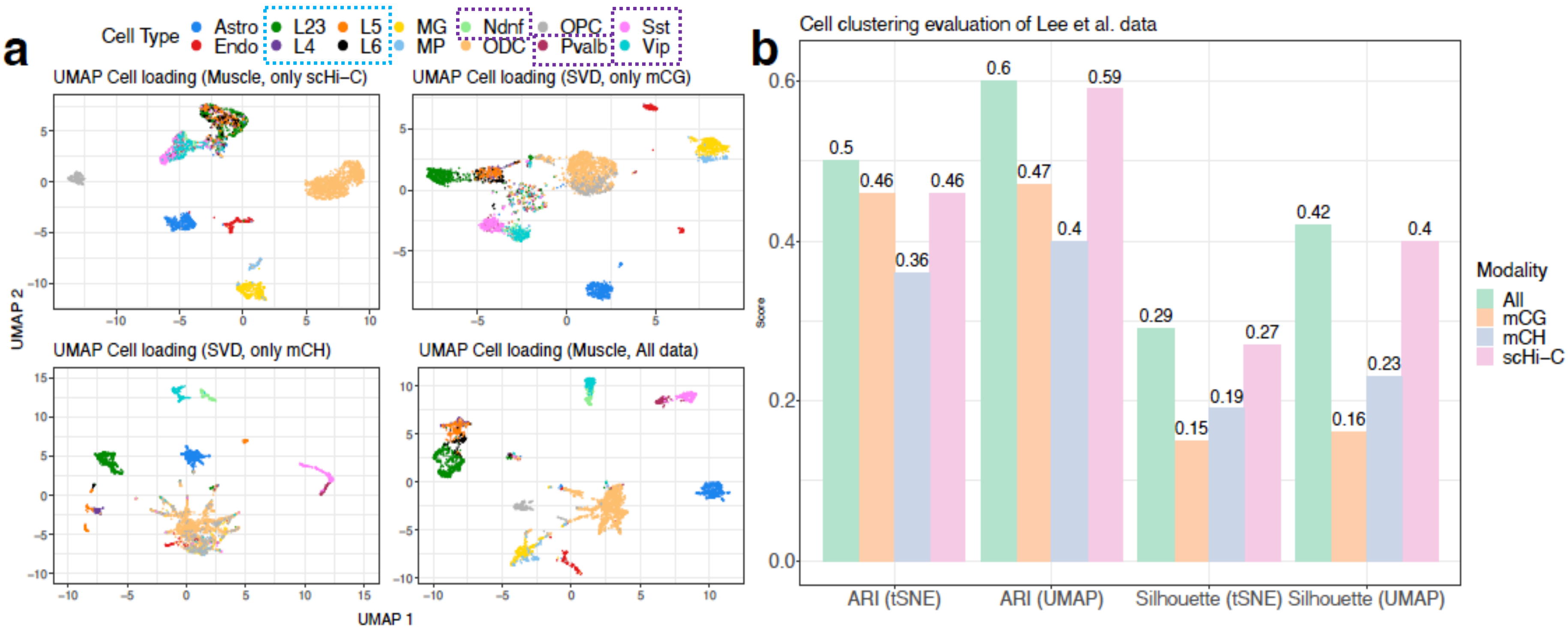
- Integration with other datasets (Methylation)

References

- *Lee, D. S. et al. Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. Nat. Methods 16, 999–1006 (2019).T. Nagano, Y. Lubling, T.J. Stevens, S.*
- *Tan, Longzhi, et al. "Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development." Cell 184.3 (2021): 741-758.*
- *Kim, Hyeon-Jin, et al. "Capturing cell type-specific chromatin compartment patterns by applying topic modeling to single-cell Hi-C data." PLoS computational biology 16.9 (2020): e1008173.*
- *Zheng et al. Normalization and De-noising of Single-cell Hi-C Data ith BandNorm and 3DVI. BioRxiv. 2021*
- *Rao, Suhas SP, et al. "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping." Cell 159.7 (2014): 1665-1680.*
- *Peng, Rui. "Predicting High-order Chromatin Interactions from Human Genomic Sequence using Deep Neural Networks." ml.cmu.edu*
- *https://www.labclinics.com/2018/11/08/role-dna-methylation-disease/?lang=en*
- *https://kkorthauer.org/fungeno2019/methylation/slides/1-intro-slides.html#5*

Data analysis results

- **Cell type clustering (Multi-modality)**

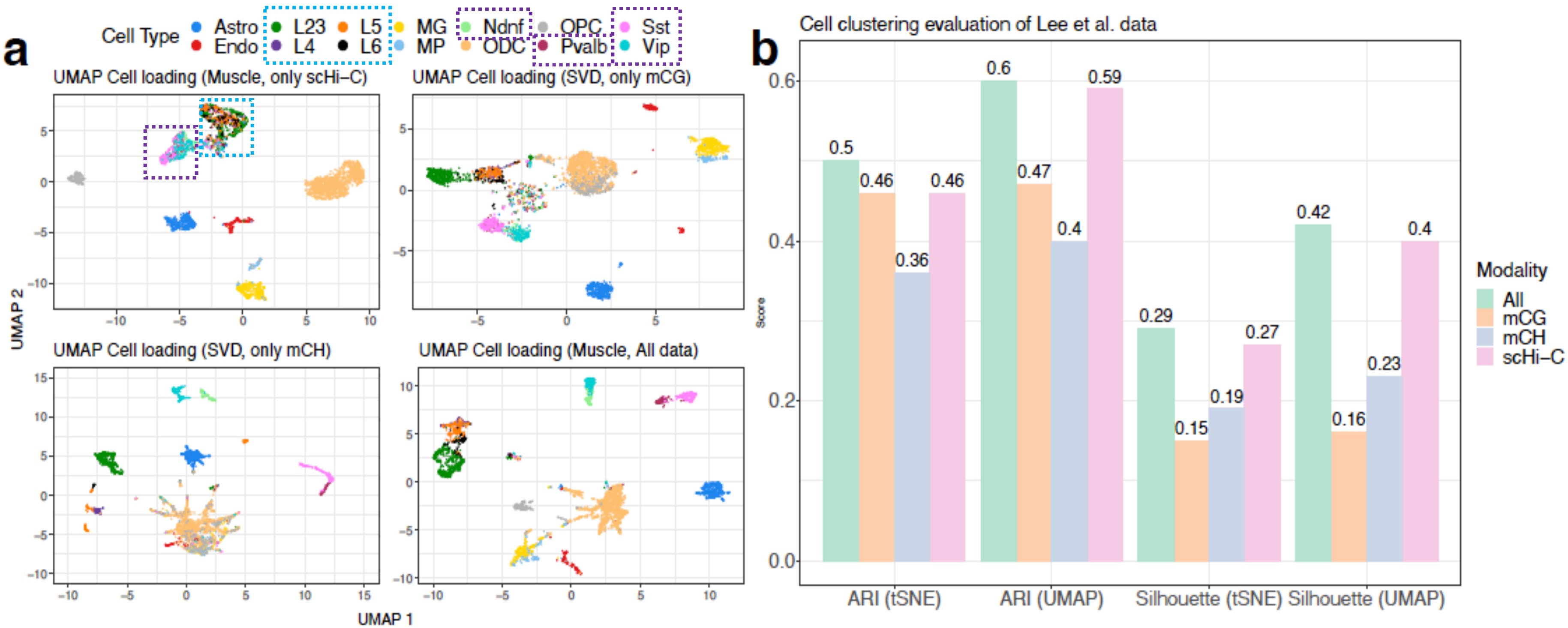


Lee 2019

Data analysis results

- **Cell type clustering (Multi-modality)**

excitatory neurons inhibitory neurons

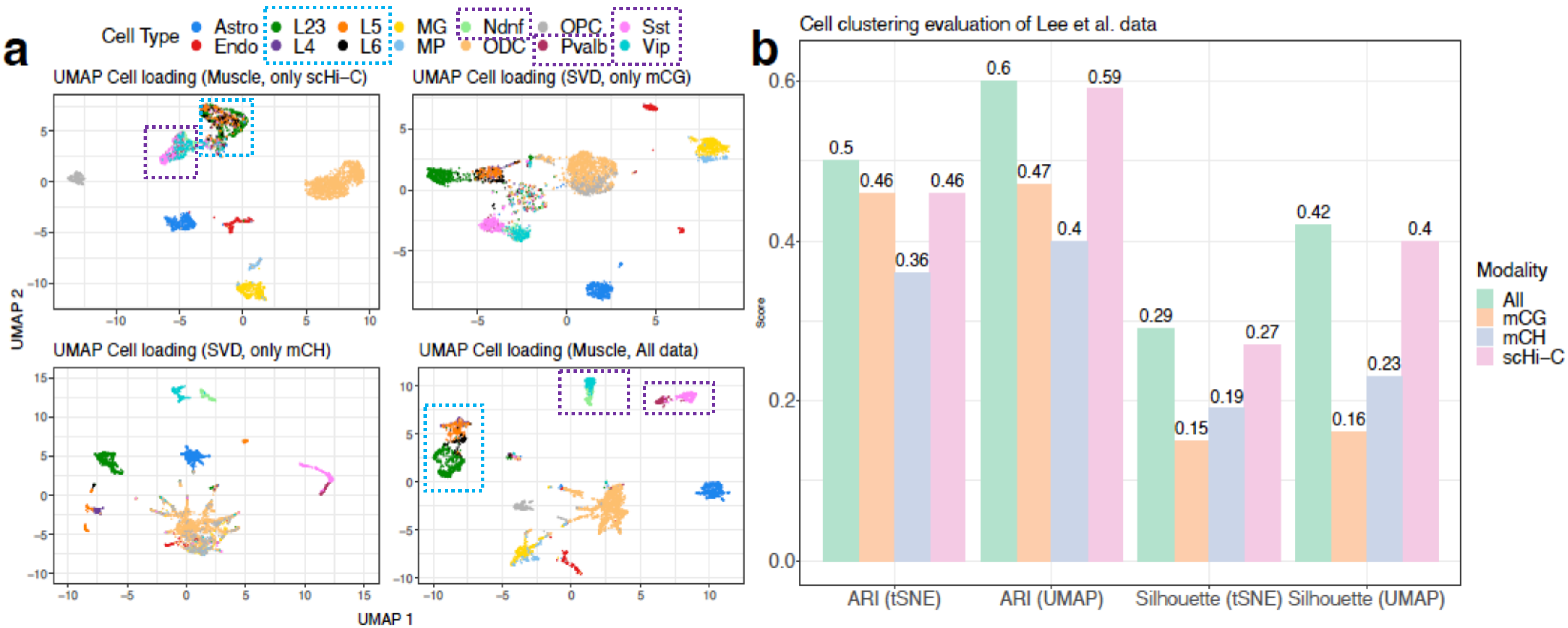


Lee 2019

Data analysis results

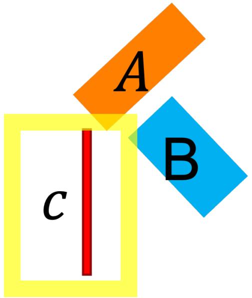
- **Cell type clustering (Multi-modality)**

excitatory neurons inhibitory neurons

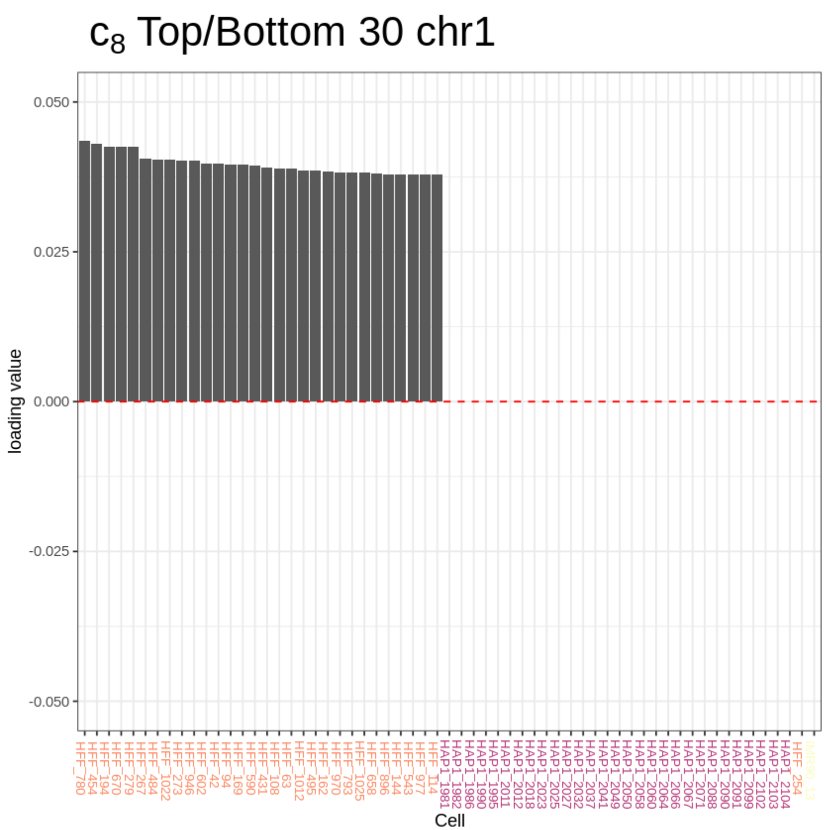


Lee 2019

Data analysis results



- **Module(cell type) specific “Eigen” Matrix**

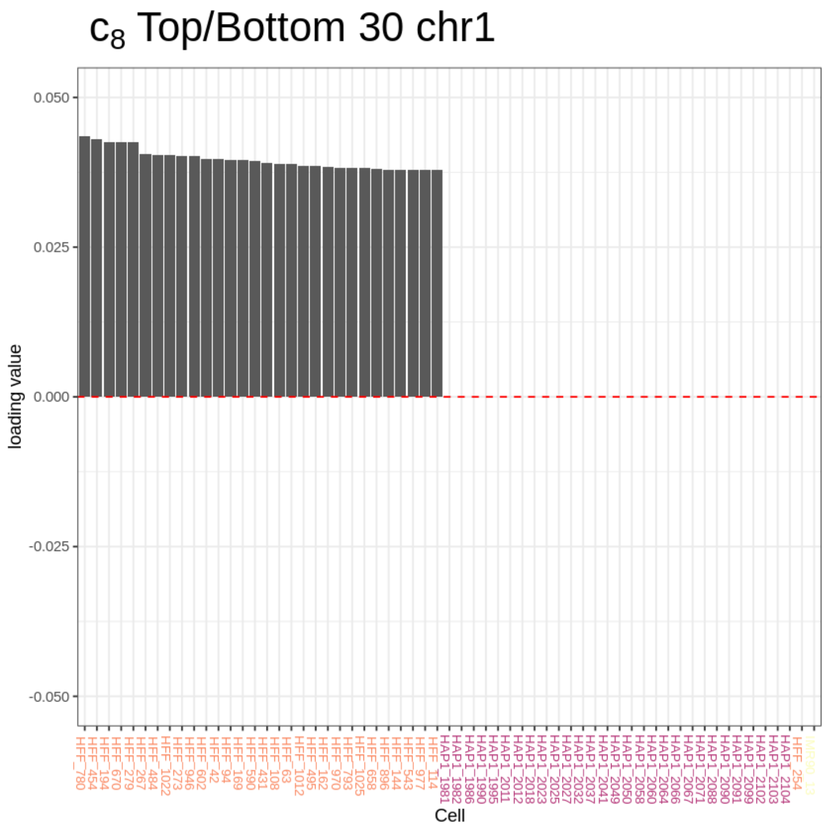


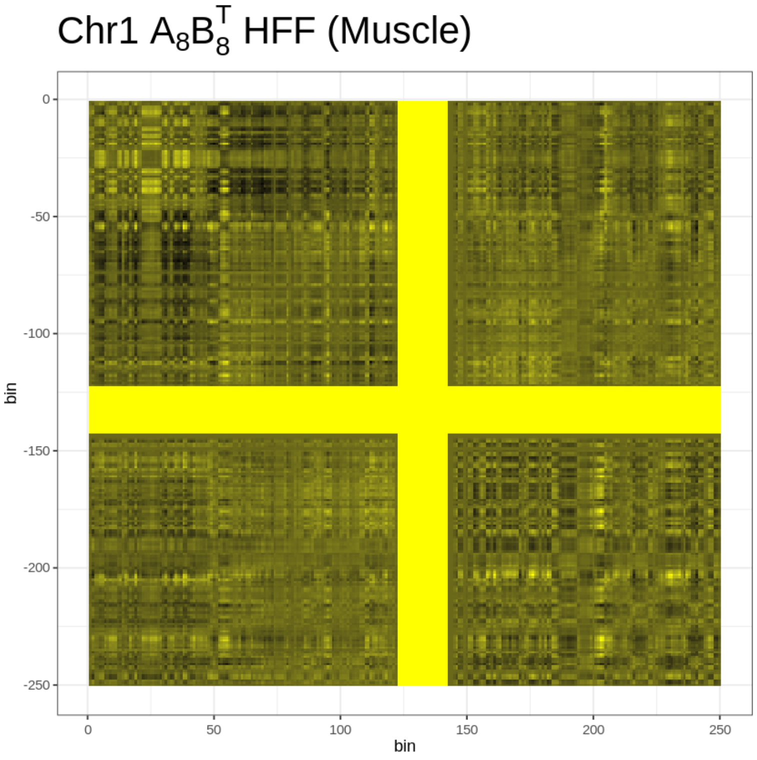
HFF

Kim 2020

Data analysis results

- **Module(cell type) specific “Eigen” Matrix**





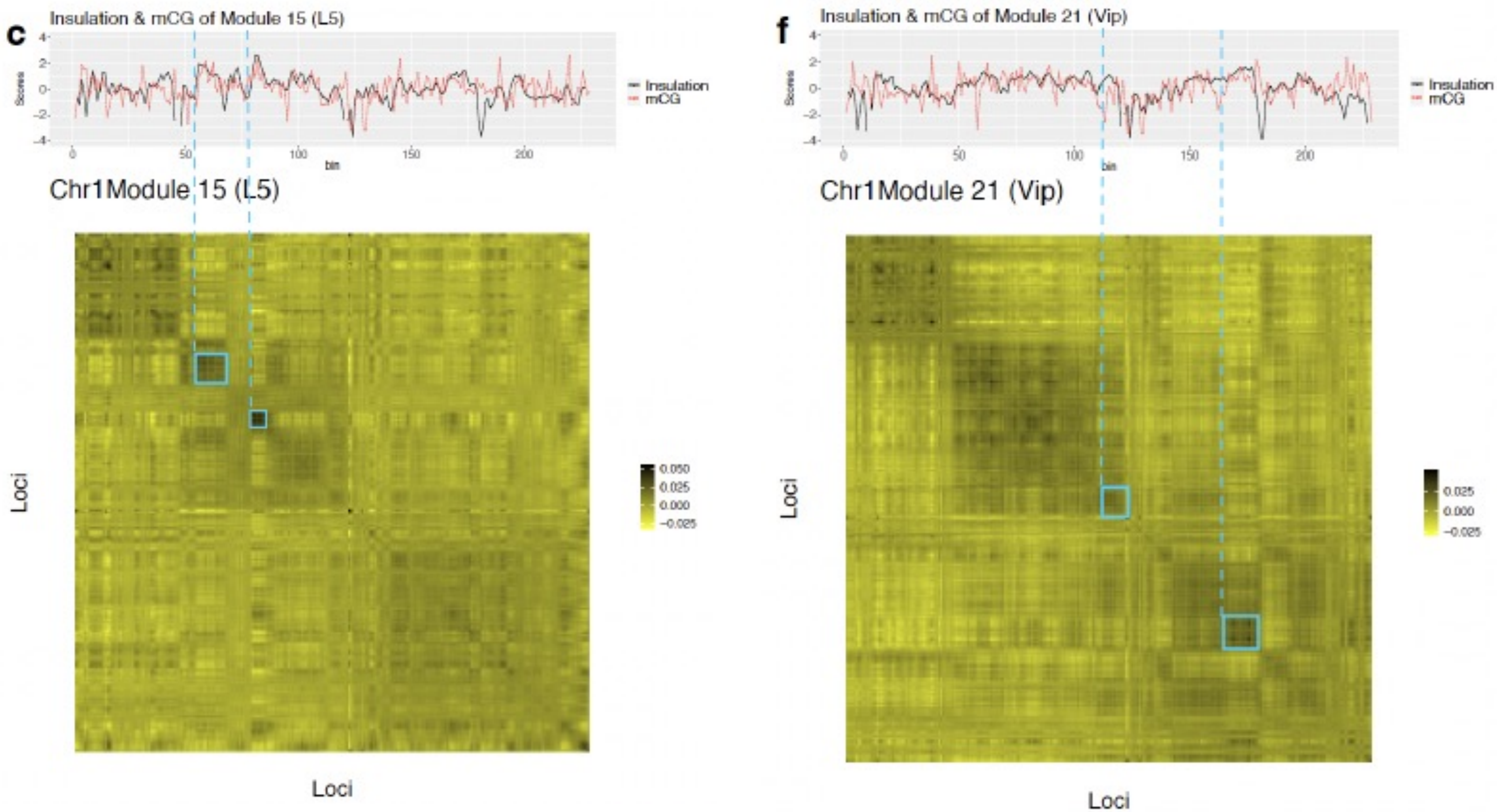
HFF

Kim 2020

Data analysis results

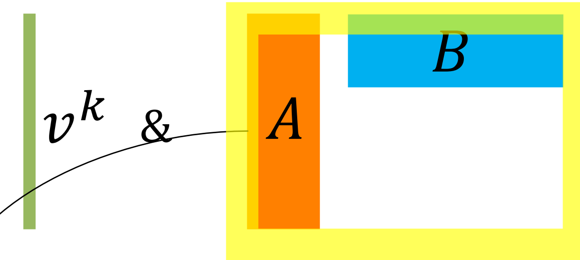
v^k & A B

- **Integrative inference**

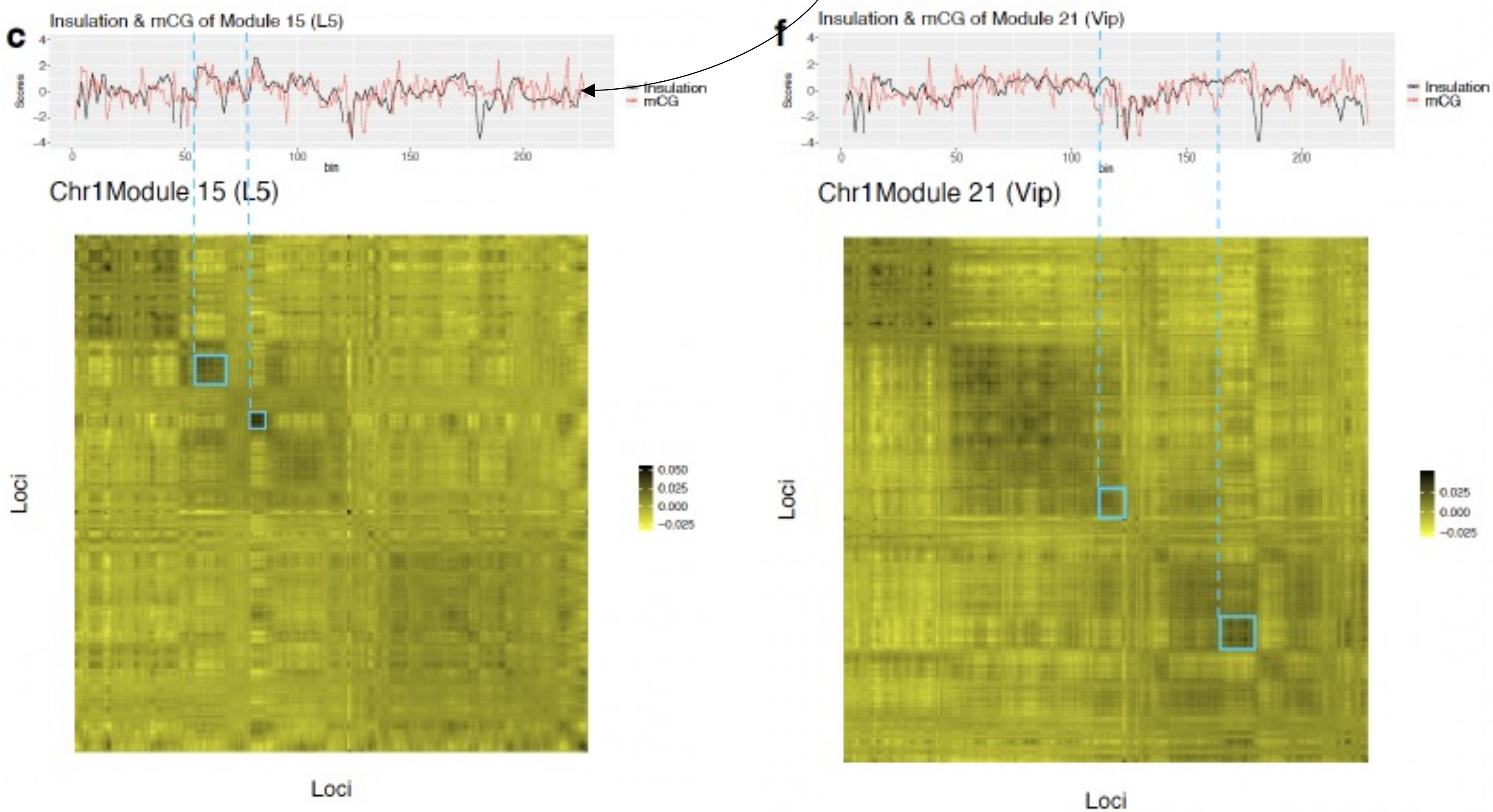


Lee 2019

Data analysis results

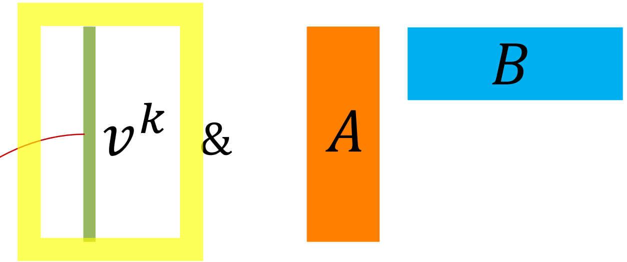


- **Integrative inference**

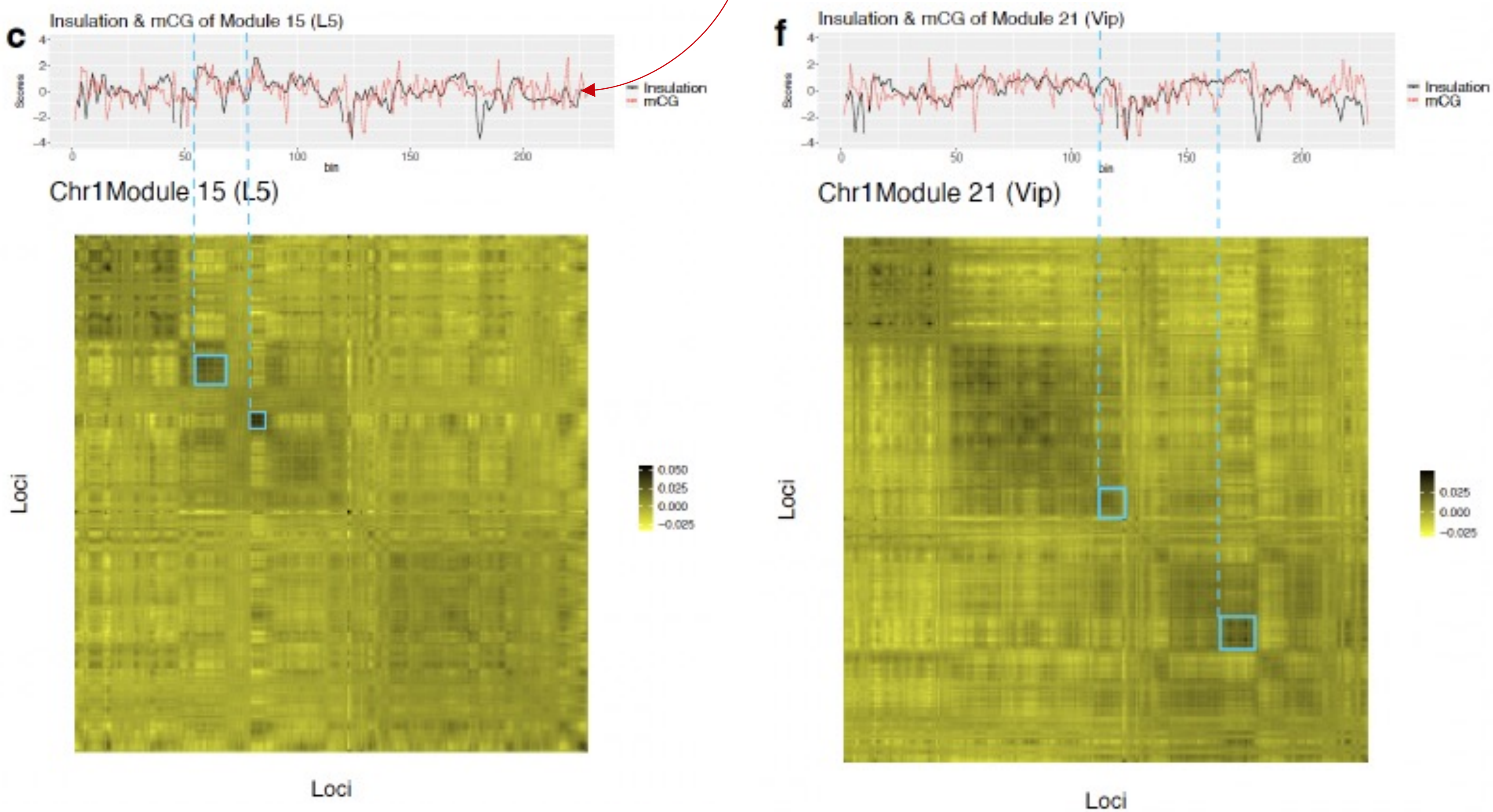


Lee 2019

Data analysis results

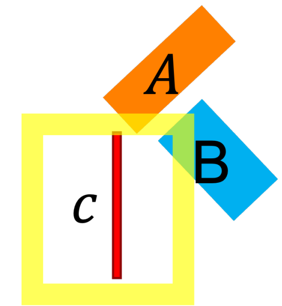


- **Integrative inference**

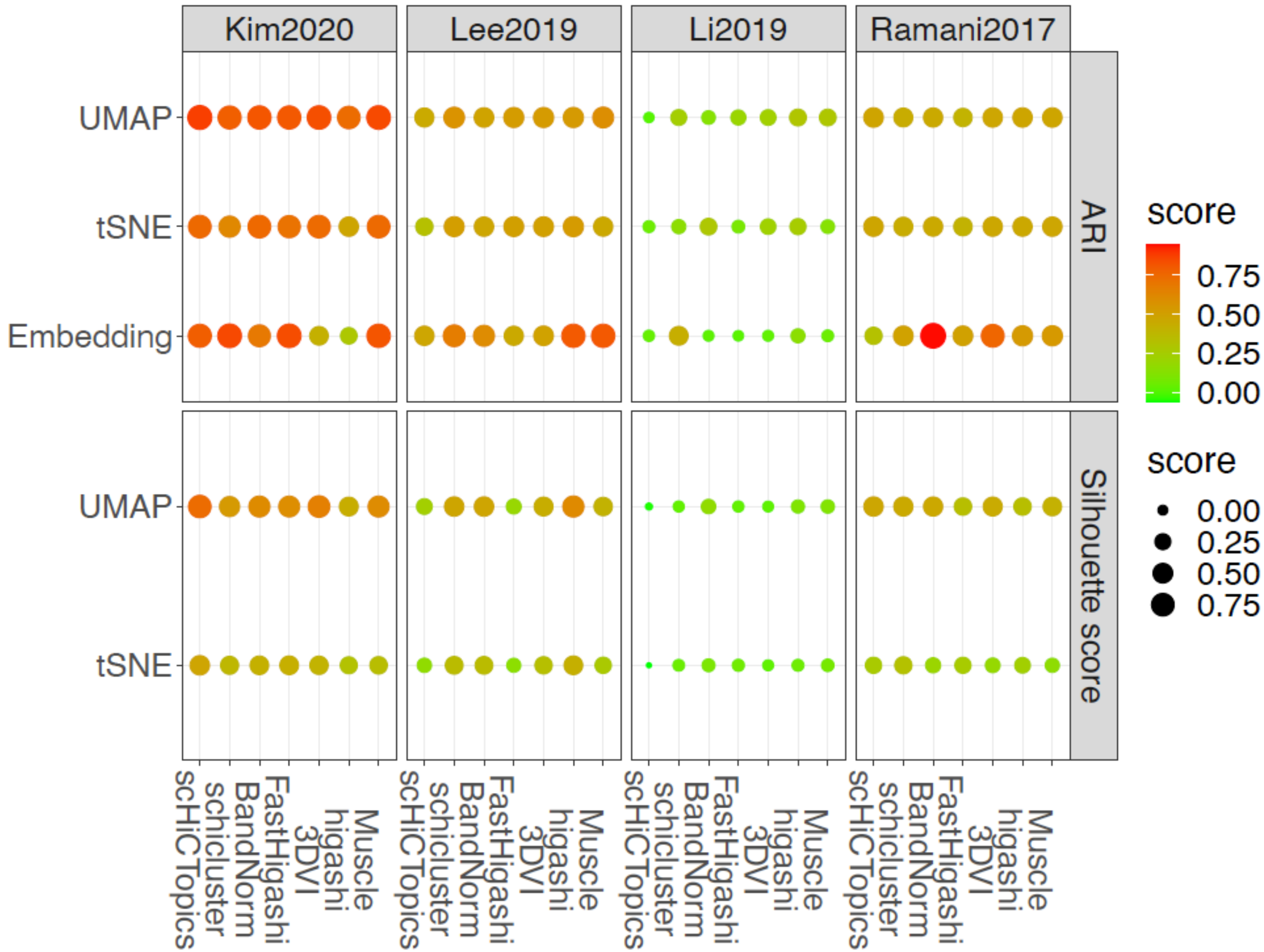


Lee 2019

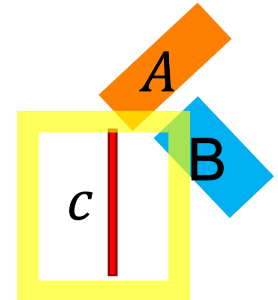
Data analysis results



- **Cell type clustering (only scHi-C tensors)**



Data analysis results



- **Cell type clustering (only scHi-C tensors)**