Recurrent single-molecule epistates define tumor methylome differences

Huy Q. Dinh

Ben Berman lab @ USC Epigenome Center
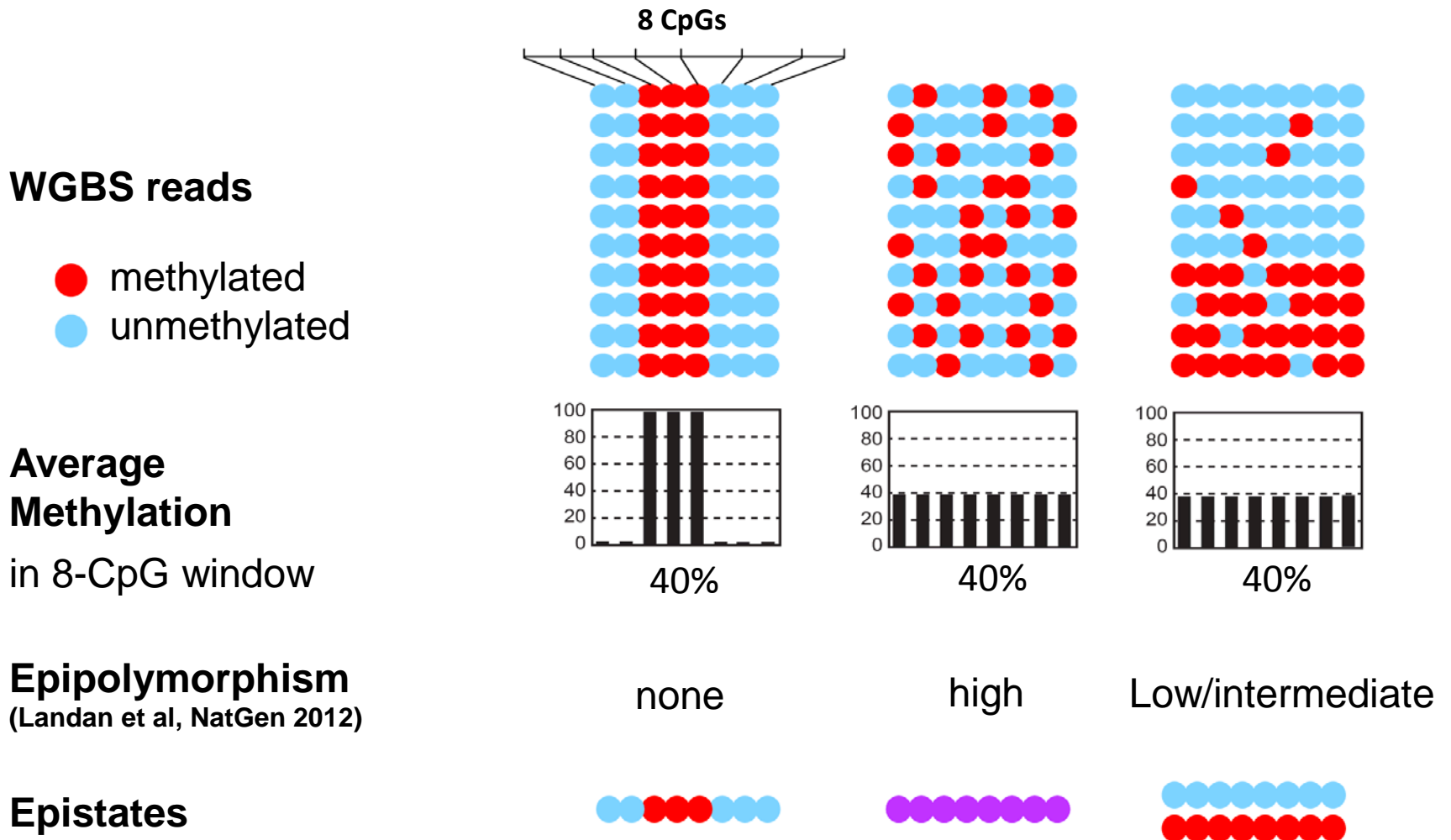
University of Southern California

The Cancer Genome Atlas

Tumor whole genome DNA methylation profiling by bisulfite sequencing

- ~ 10K samples profiled by Illumina Infinium 27K & 450K microarray platforms within TCGA projects.

- Whole Genome Bisulfite Sequencing (WGBS)
  - Methylation profile of 28M CpGs with more sequence variation information.
  - WGBS on 47 TCGA patient samples completed
  - >15x sequence coverage / sample

- Identification of recurrent Differentially Methylated Regions (DMRs) across multiple tumors based on single-molecule analysis of WGBS.

Samples	#	Samples	#
LUAD	6	GBM	6
LUSC	5	BRCA	6
BLCA	7	STAD	5
UCEC	6	COAD	3
		READ	3

USC/JHU Epigenome Characterization Center
Peter Laird & Steve Baylin, PIs

Tumor heterogeneity in WGBS

8 CpGs

WGBS reads

🔴 methylated
🔵 unmethylated

Average Methylation

in 8-CpG window

| 40% | 40% | 40% |

Epipolymorphism
(Landan et al, NatGen 2012)

| none | high | Low/intermediate |

Epistates
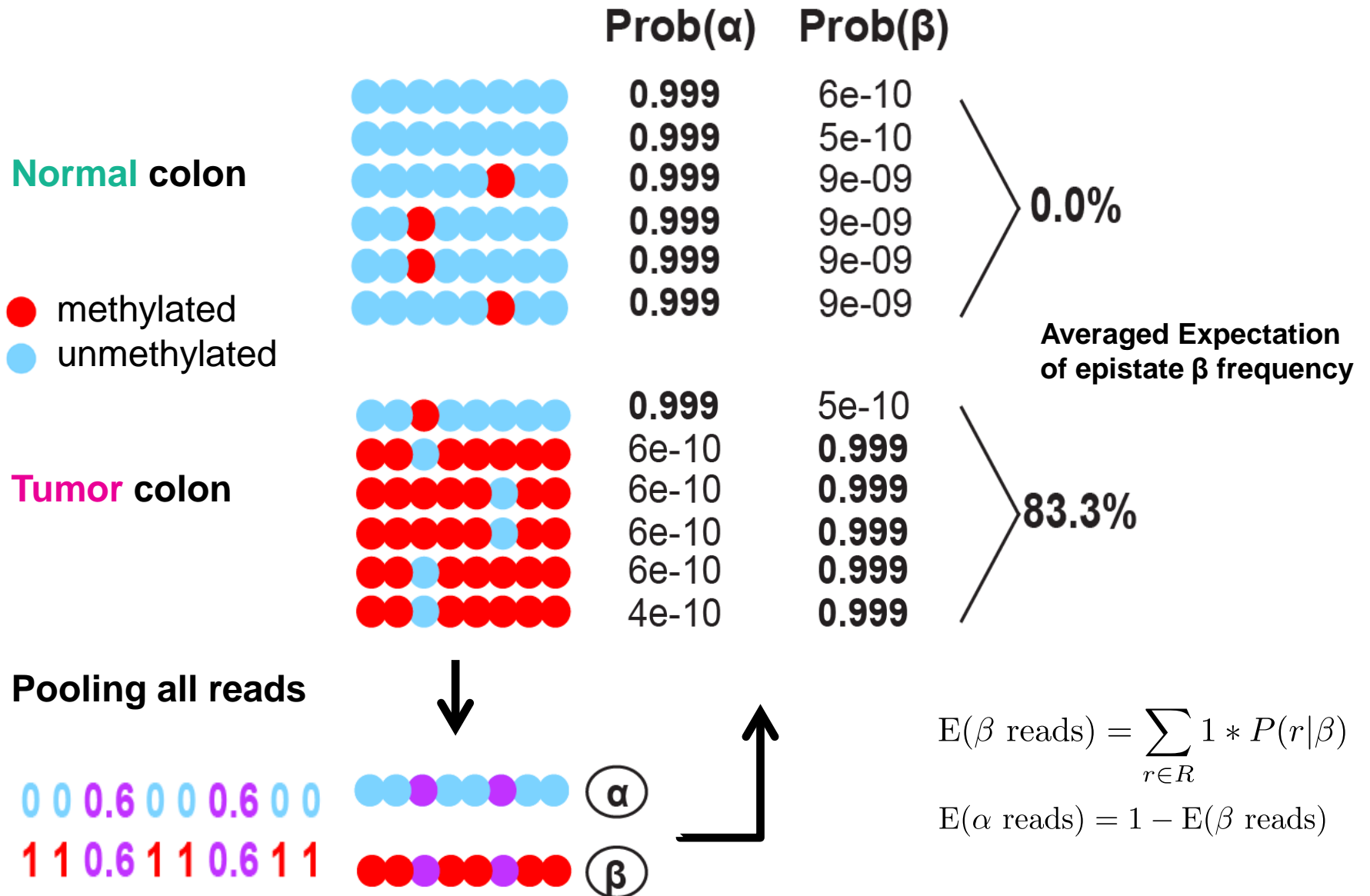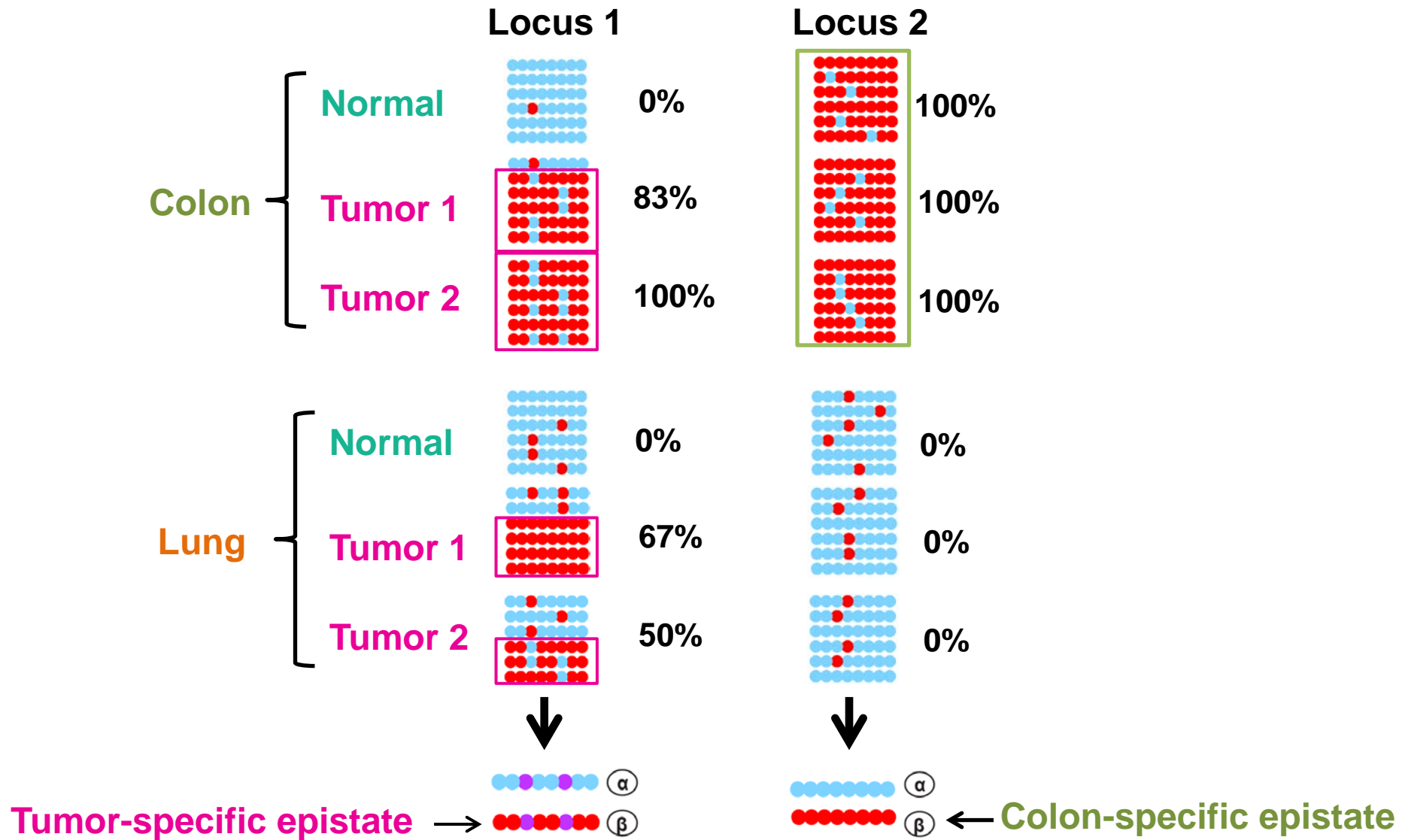
Learning epistate mixture based on Expectation Maximization (EM)
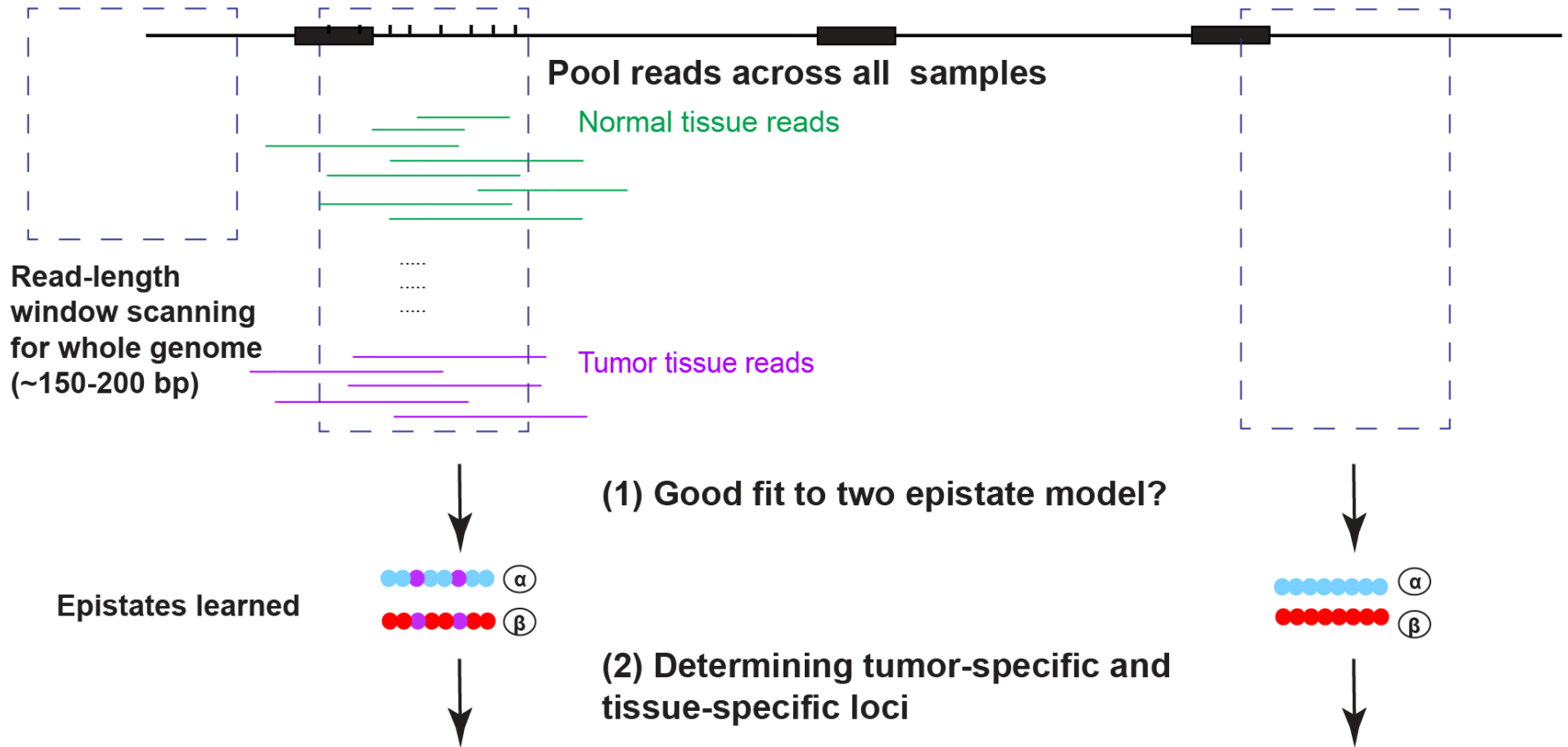Extended from allelic methylation model (Fang et al, PNAS 2012)

Estimating within-sample epistate frequencies
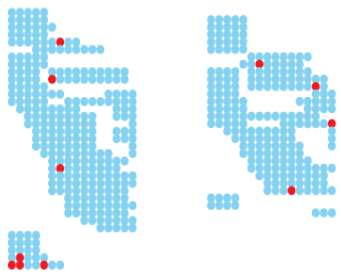
Using epistates to identify DMRs

Genome-wide scanning for epistates

Pool reads across all samples

Normal tissue reads

Read-length window scanning for whole genome (~150-200 bp)

Tumor tissue reads

(1) Good fit to two epistate model?

Epistates learned

α
β

(2) Determining tumor-specific and tissue-specific loci

α
β

Estimating tumor purity based on epistate frequencies
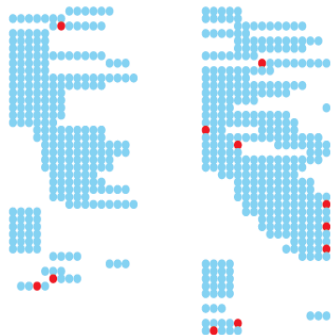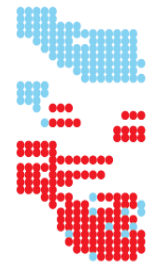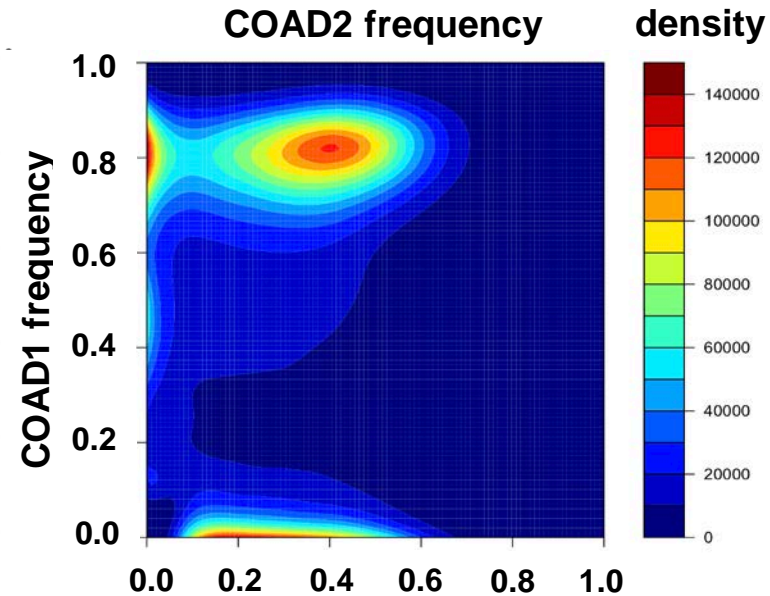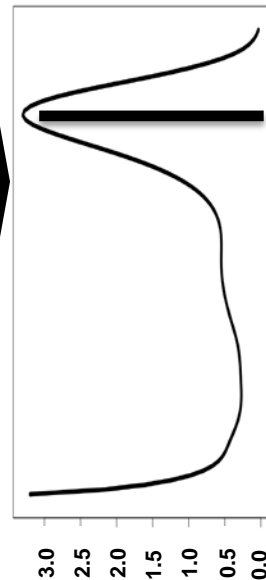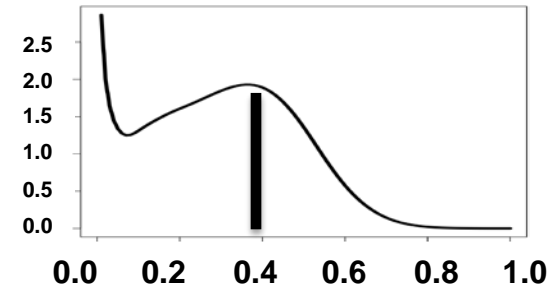
Epistate estimate tumor purity highly correlated with ABSOLUTE estimation
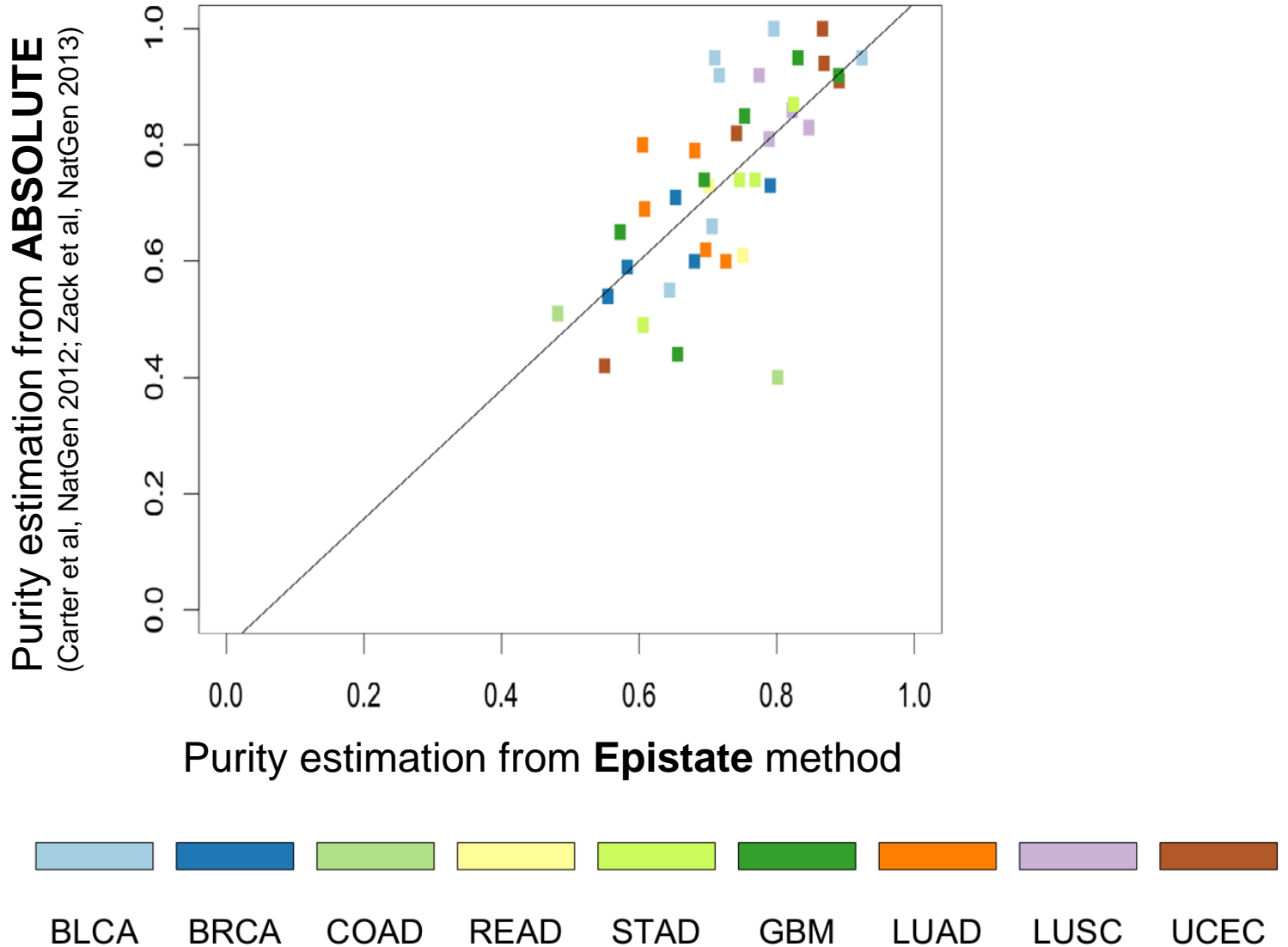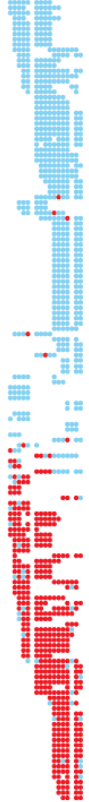
Epistates can be detected at low frequency

chr1:145039592   (239 bps)   145039830

Epistate α
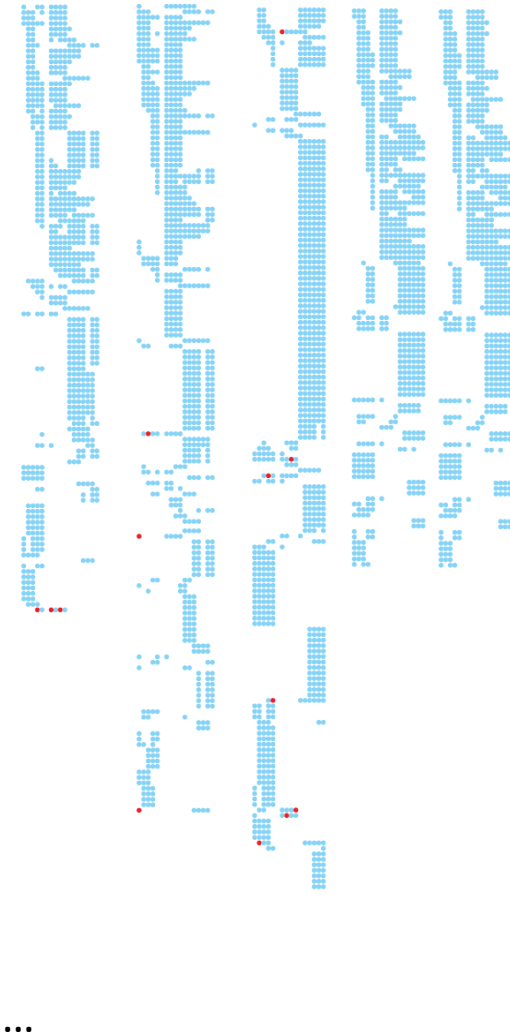Epistate β

COAD1: 40.34%   COAD2: 13.82%   COAD3: 15.79%

Adjacent Normals   **Unmethylated Tumors**
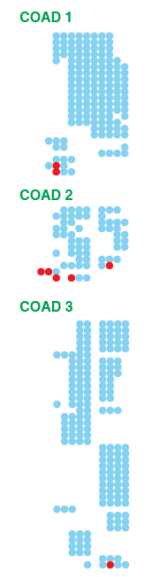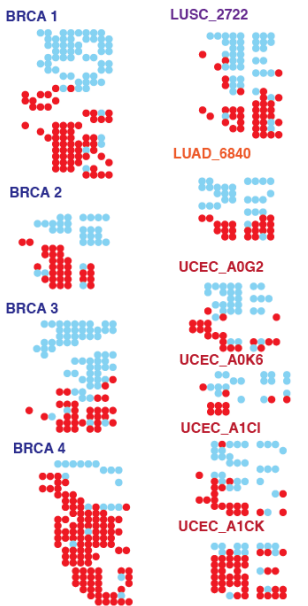
Methylated epistate found in only COAD tumors

Epistate DMR example: Epigenetically silenced distal element (SMAD3)
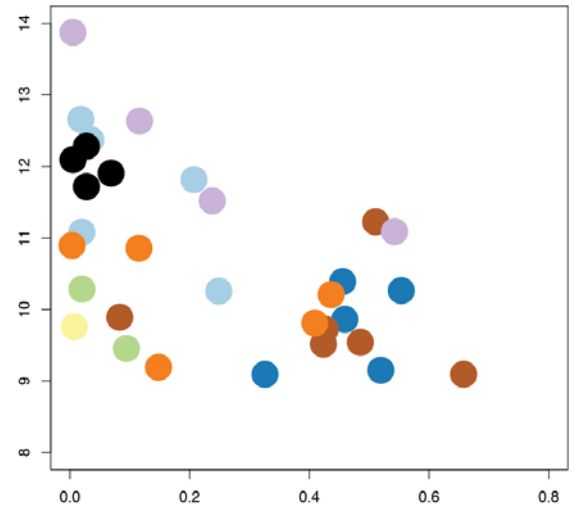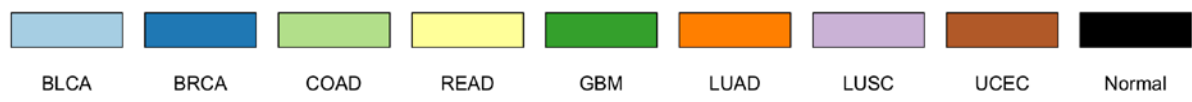
Summary
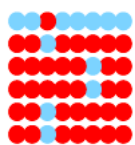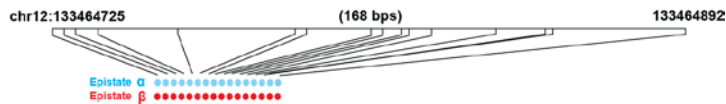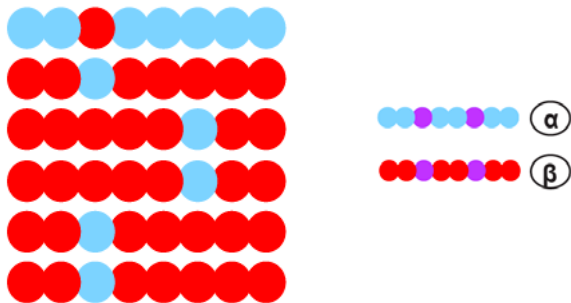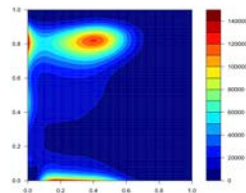
- We developed **single-molecule epistate method** for analyzing multiple WGBS samples/tumors.

- **Epistate applications:**
  - Methylation-based tumor purity estimation.
  - Differentially Methylated Regions finding reveals epigenetically silenced (regulatory) regions at (low) epistate frequencies.

Acknowledgements

- **Berman Lab**
  - Yaping Liu
  - Lijing Yao
  - **Ben Berman**

- **JHU**
  - Steve Baylin

- **USC Epigenome Center**
  - **Peter Laird**
  - **Hui Shen**
  - Toshi Hinoue
  - Dan Weisenberger
  - Charlie Nicolet
  - Zack Ramjan
  - Kim Siegmund
  - Natalia Shatokhina
  - Selene Tyndale
  - Helen Truong