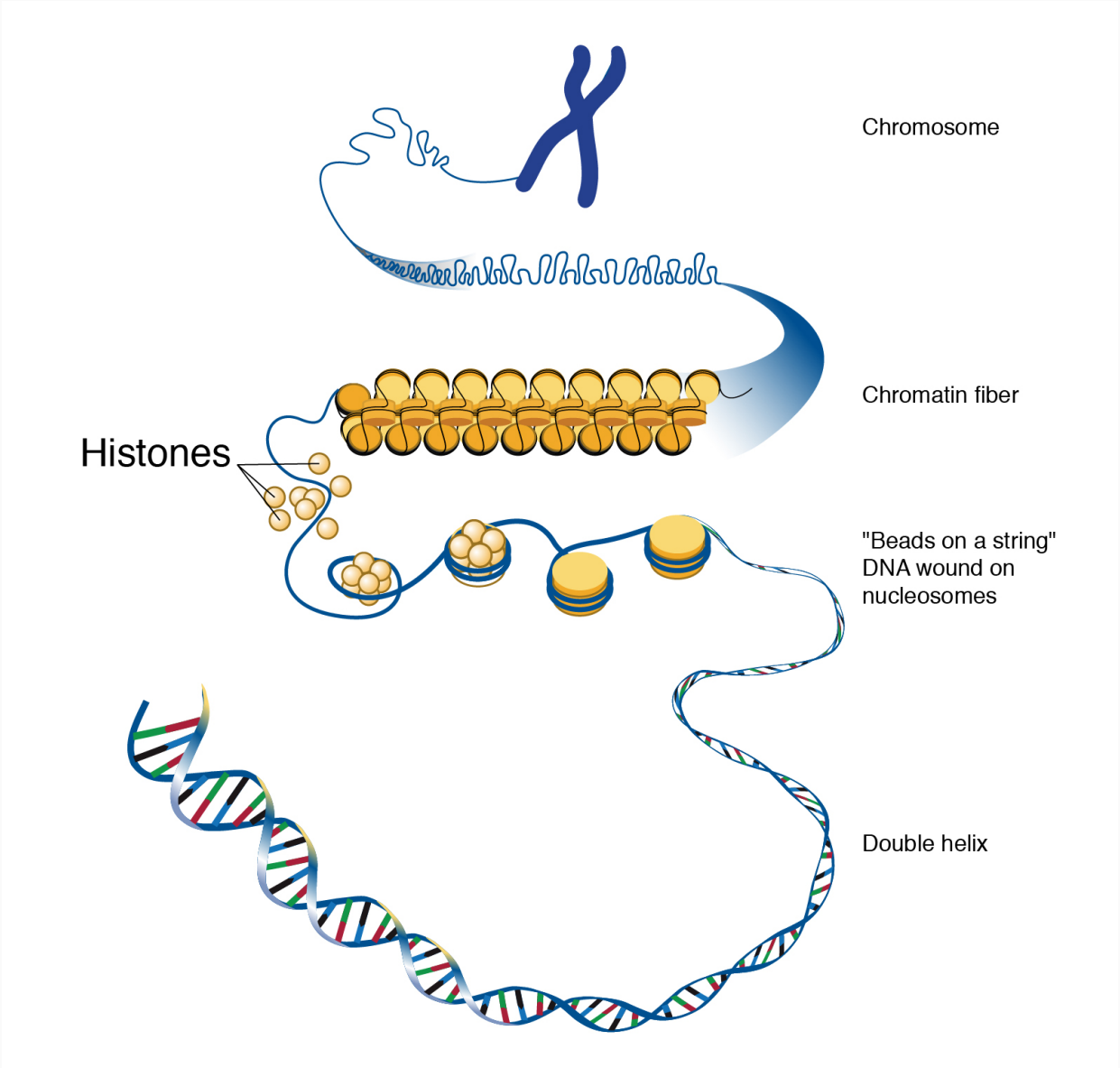
Genome-wide maps of chromatin state in pluripotent and lineage-committed cells

Edmund Miller

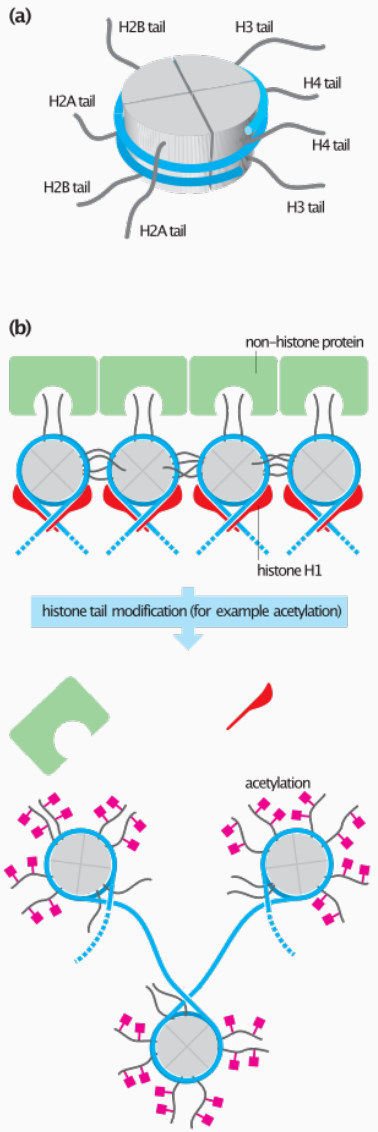
2021-03-09 Wed

Background

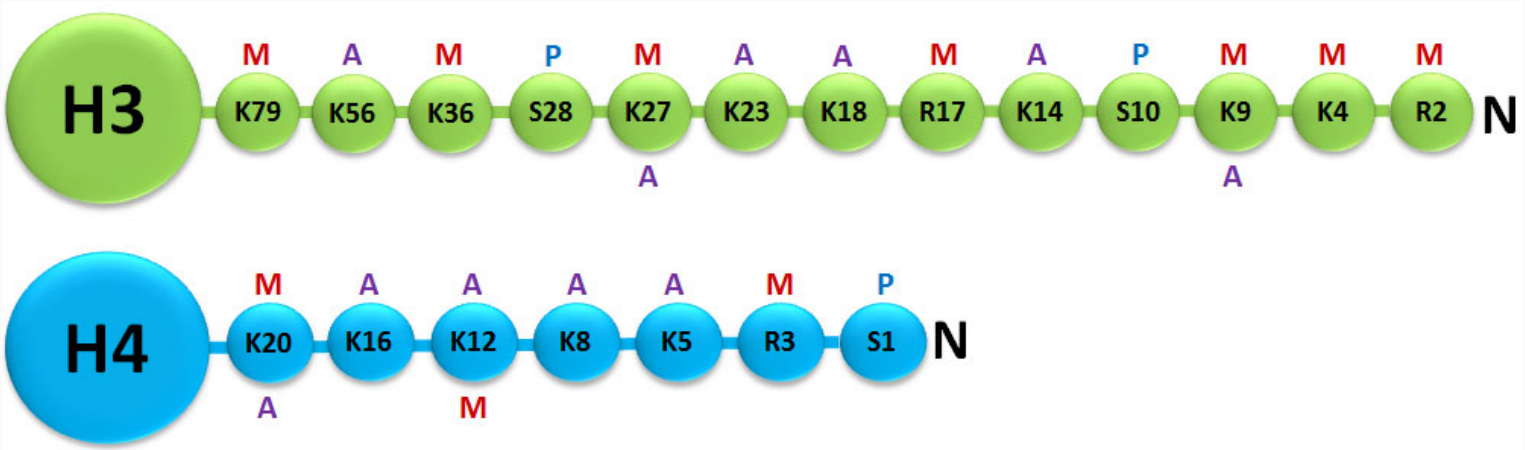
Histones


Chromosome
Chromatin fiber
Histones
"Beads on a string"
DNA wound on
nucleosomes
Double helix

Histone modifications



Histone modifications

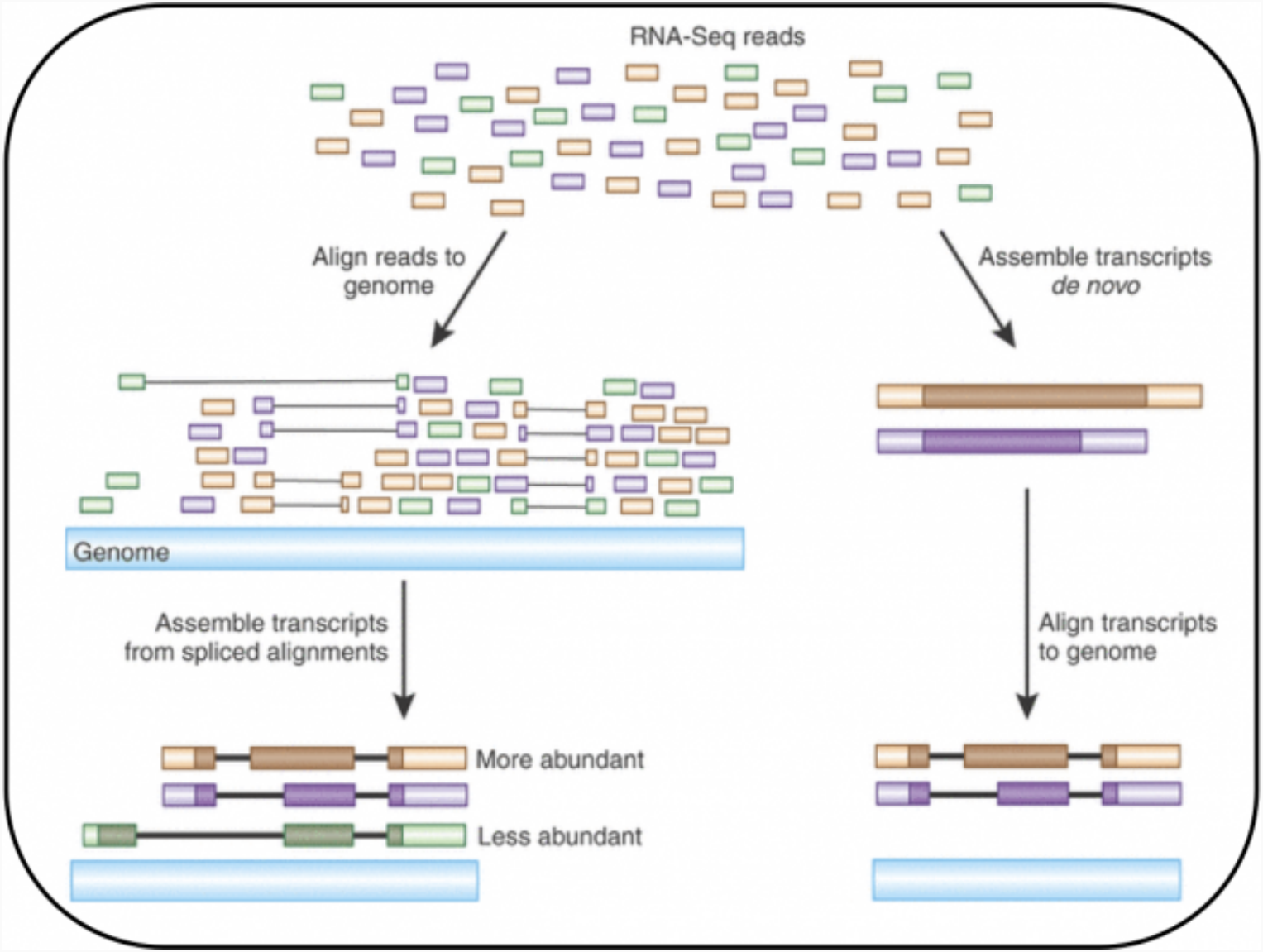


- Bivalent Chromatin

Chromatin immunoprecipitation (ChIP)

1. DNA and proteins are crosslinked
2. The DNA-protein complexes are sheared (sonication or nuclease digestion)
3. Proteins are pulled out using antibodies, along with the cross-linked DNA fragments
4. DNA fragments are purified

Sequencing technology

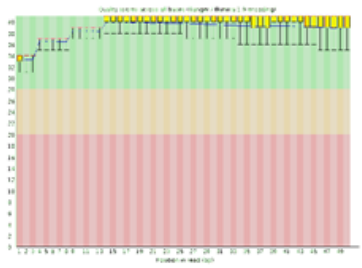


(Hass and Zody, Advancing RNA-Seq analysis, Nature

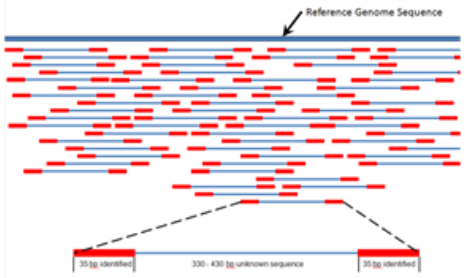
ChIP-seq

FASTQ data

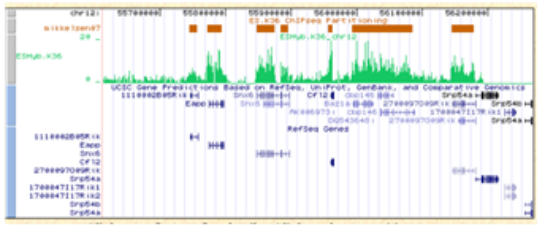
Quality assessing



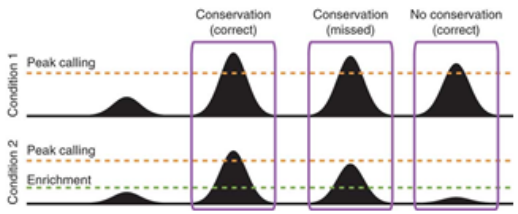
Mapping



Generate wiggle and coverage tracks



Call peaks



Custom analysis

- Enriched Motifs
- Annotation
- Overlapping or differentially bound peaks
- Etc...

CpG Sites

Introduction

History

- Mikkelson et al. from the Broad submitted to Nature and was published in August 2007.
- Johnson et al. from Stanford submitted to Science and was published in June 2007.
- Barski et al. from NHLBI, NIH submitted to Cell and was published in May 2007.
- Barski was the third most cited biology paper in 2009
- All of these were after Frank Pugh's work on H2A.Z in yeast

Snarky Scientist

Goals

- Report the development of a method for mapping ChIP enrichment by sequencing (ChIP-Seq)
- Define Broad categories of promoters based on their chromatin state and create maps
- Demonstrate using ChIP for genome-wide annotation of novel promoters

Histone Marks

Actively transcribed genes

- H3K4me3 - Promoters
- H3K36me3 - Body of active genes

Repressed Genes

- H3K27me3 - Gene repression
- H4K20me3 - Tightly associated with heterochromatin
- H3K9me2/3 - Heterochromatin and Gene repression

Bivalent Promoters

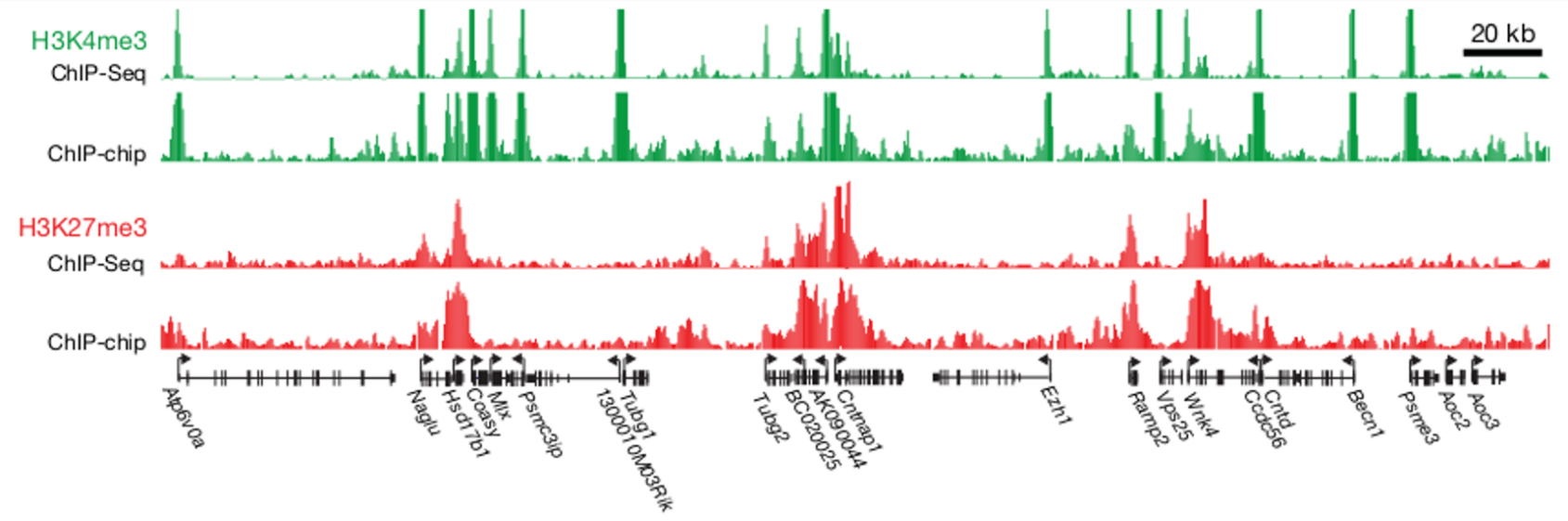
- Both H3K4Me3 and H3K27Me3

Genome-wide chromatin-state maps

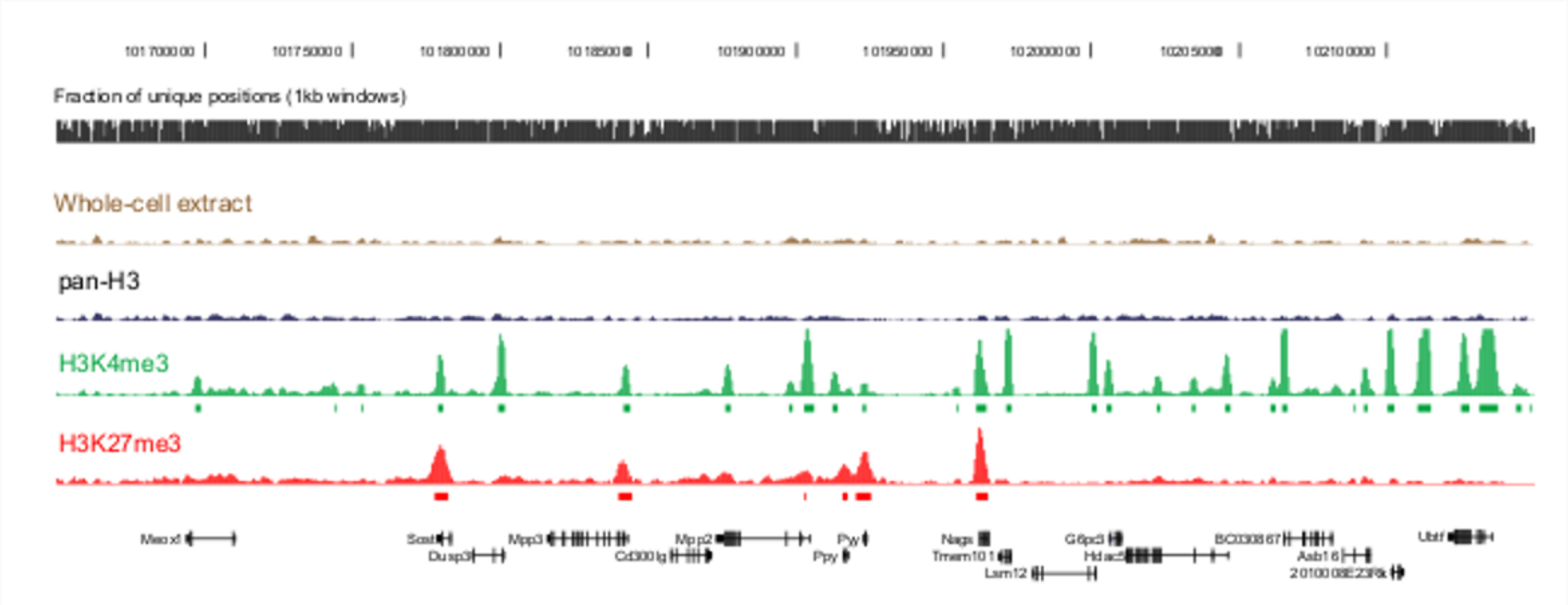
Supplementary Table 1

Cell type1	Epitope	Replicates	Uniquely aligned reads
ES cells	pan-H3	1	4,490,474
	H3K4me3	2	8,398,790
	H3K9me3	2	4,411,447
	H3K27me3	2	7,211,279
	H3K36me3	2	7,217,118
	H4K20me3	2	5,139,339
	RNAP II	1	2,736,500
ES cells- hybrid	H3K4me3	2	11,471,501
	H3K9me3	1	3,741,367
	H3K36me3	2	7,826,112
Neural progenitor cells (NPCs)	H3K4me3	2	6,995,068
	H3K9me3	2	4,614,191
	H3K27me3	2	8,166,774
	H3K36me3	2	7,899,115
Embryonics fibroblasts (MEFs)	H3K4me3	2	11,371,374
Lineage-commited	H3K9me3	2	4,468,908
	H3K27me3	2	12,208,145
	H3K36me3	2	10,315,848

Figure 1 | Comparison of ChIP-Seq and ChIP-chip data



Supplementary Figure 2



Promoter state in ES and lineage-committed cells

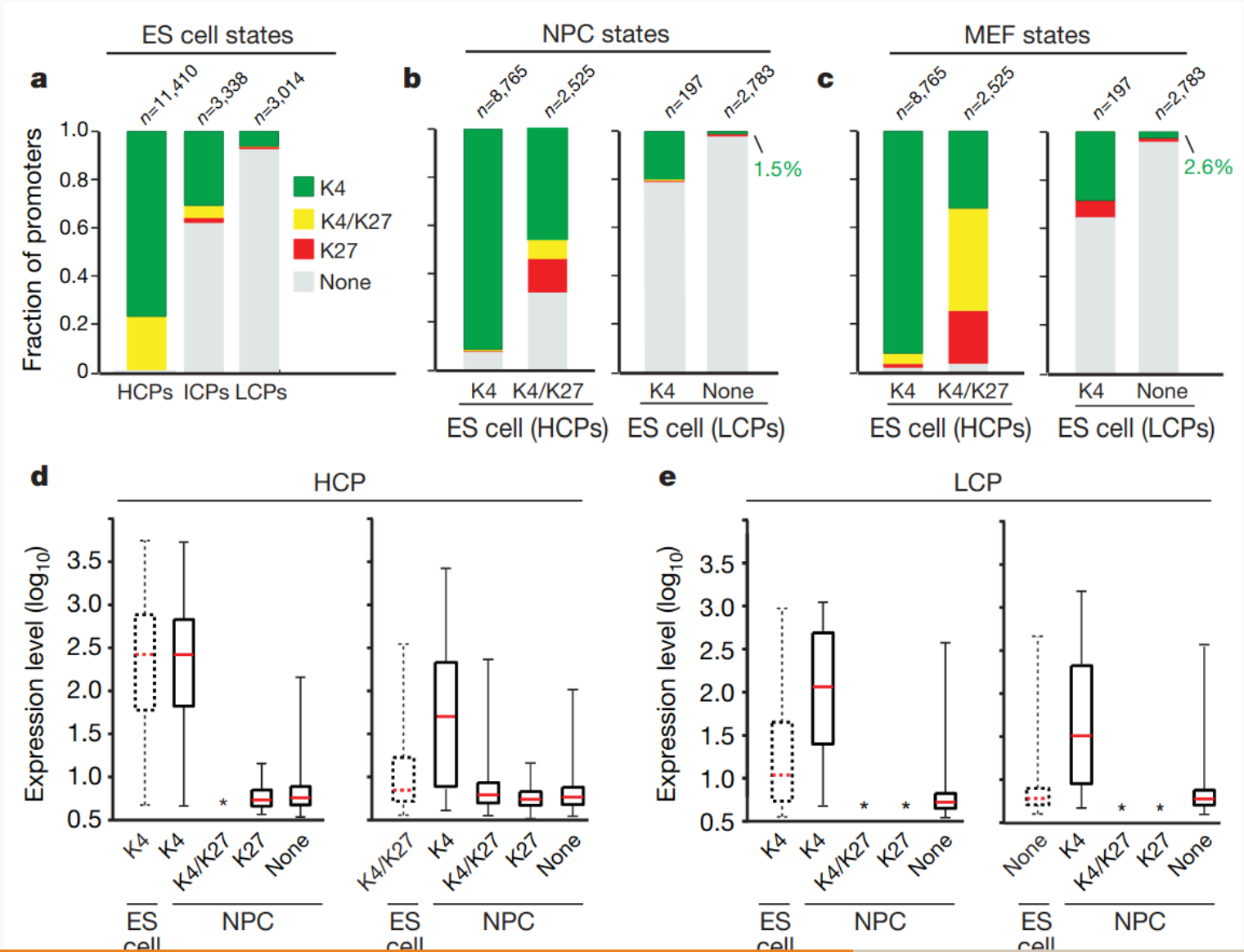
High CpG promoters in ES cells

- Divided promoters into 'high' CpG promoters(HCP) and 'low' groups, and an intermediate. 11K high, 3K low and 3.3K intermediate
- HCP Associated with significant H3K4me3 within a 1-2kb window
- There is a strong correlation between H3K4me3 intensity and the expression level of associates genes
- Not all promoters associated with H3K4me3 are active.

High CpG promoters in ES cells

- ~22% of HCPs are bivalent and show both H3K4me3 and H3K27me3 (Fig 2b)
  - The majority are 'narrow' with more punctate H3K27me3
- Bivalent Promoters show low activity despite H3K4me3
  - This suggests that the represseive PcG activity dominates the ubiquitous trxG activity
- Monovalent promoters (H3K4me3) regulate housekeeping genes, they're just always on
- Bivalent promoters are associated with genes with more complex expression patterns
  - Developmental transcription factors
  - Morphogens
  - Cell surface molecules
  - Lineage-specific microRNAs

Figure 2 | Histone trimethylation state predicts expression of HCPs andLCPs



High CpG promoters in NPCs and MEFs

- Most HCPs marked with H3K4me3 alone in ES cells retain this mark both in NPCs and MEFs
  - This sub-class of promoters regulates ubiquitous house-keeping genes
- Small proportion (~4%) of these promoters have H3K27me3 in MEFs
  - Correlates with lower expression levels

Figure 3 | Cell-type-specific chromatin marks at promoter

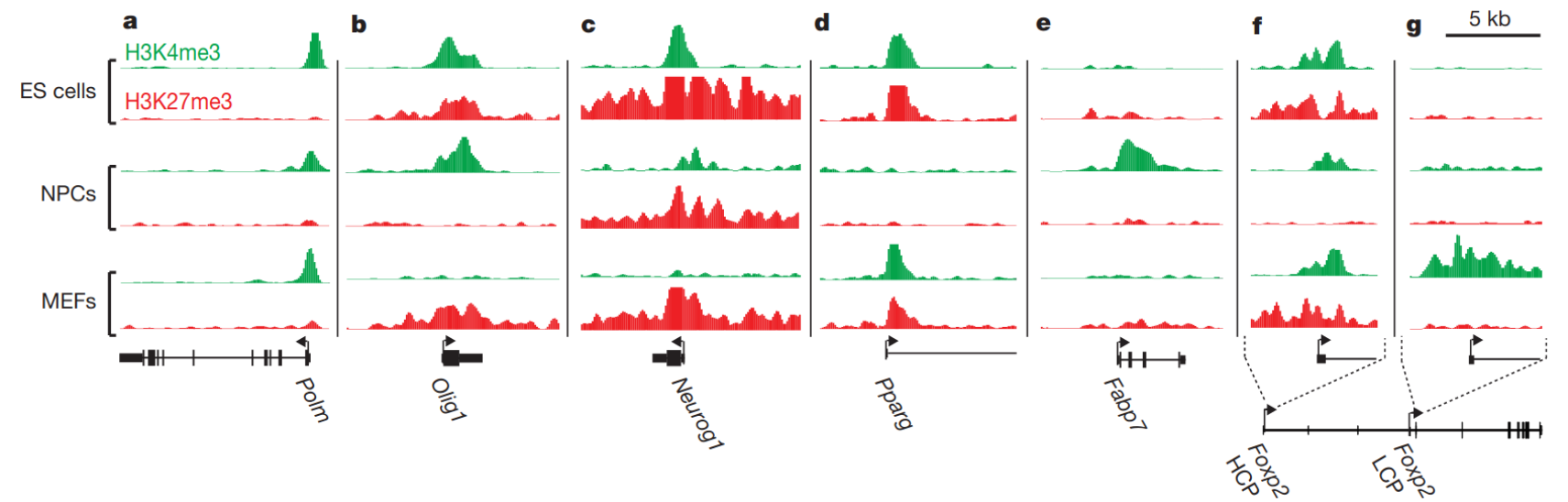
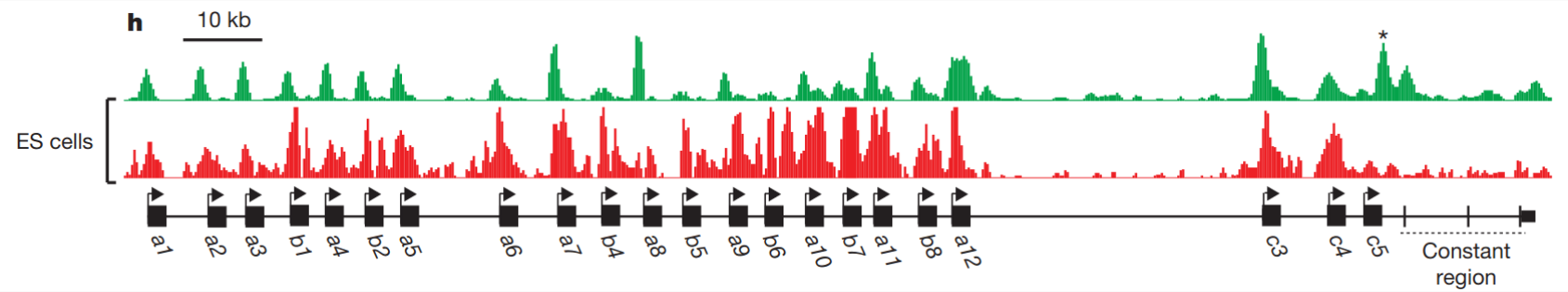
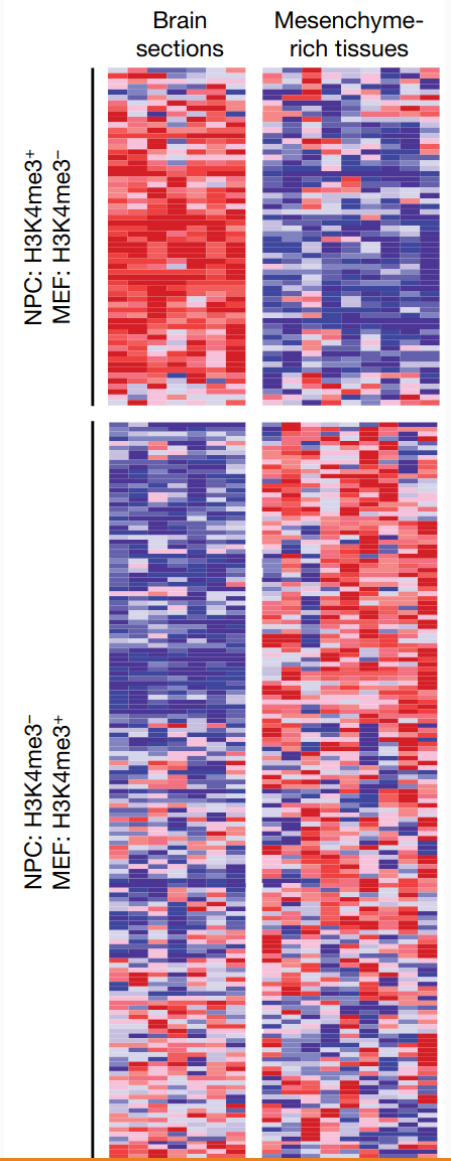


Figure 3 | Cell-type-specific chromatin marks at promoter



Promoter state reflects lineage commitment and potential

Figure 4 | Correlation between chromatin-state changes and lineage expression



Genome-wide annotation of promoters and primary transcripts

Figure 5 | H3K4me3 and H3K36me3 annotate genes and non-coding RNAtranscripts.

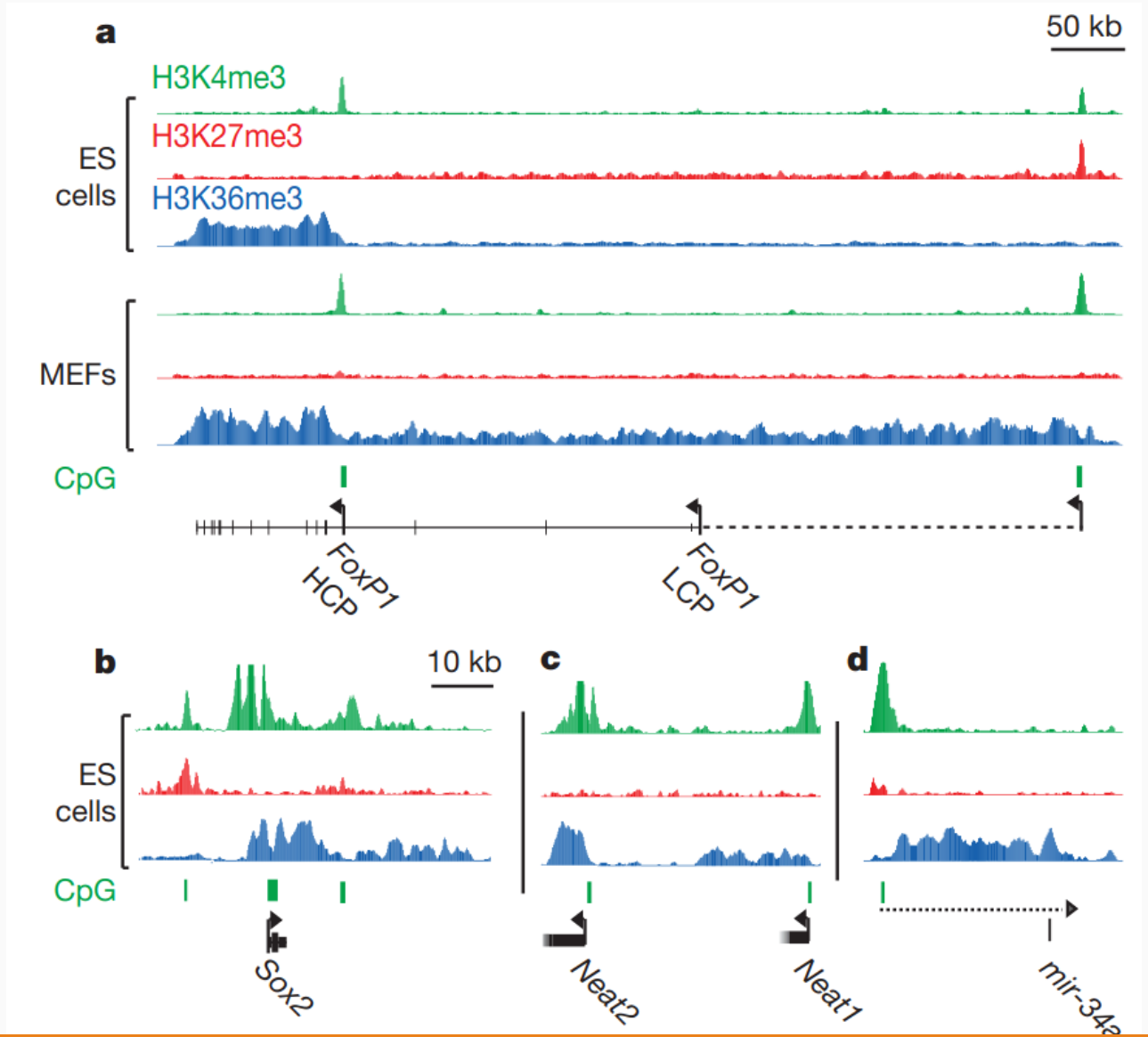
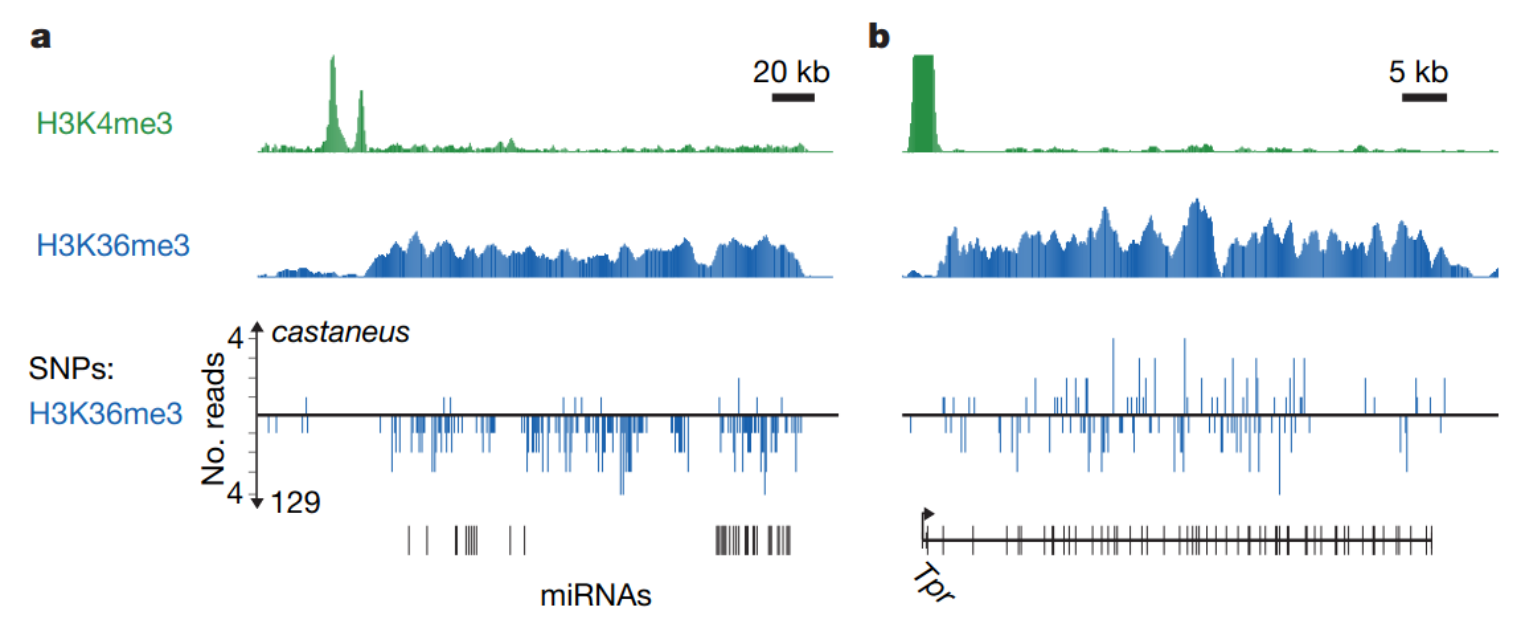
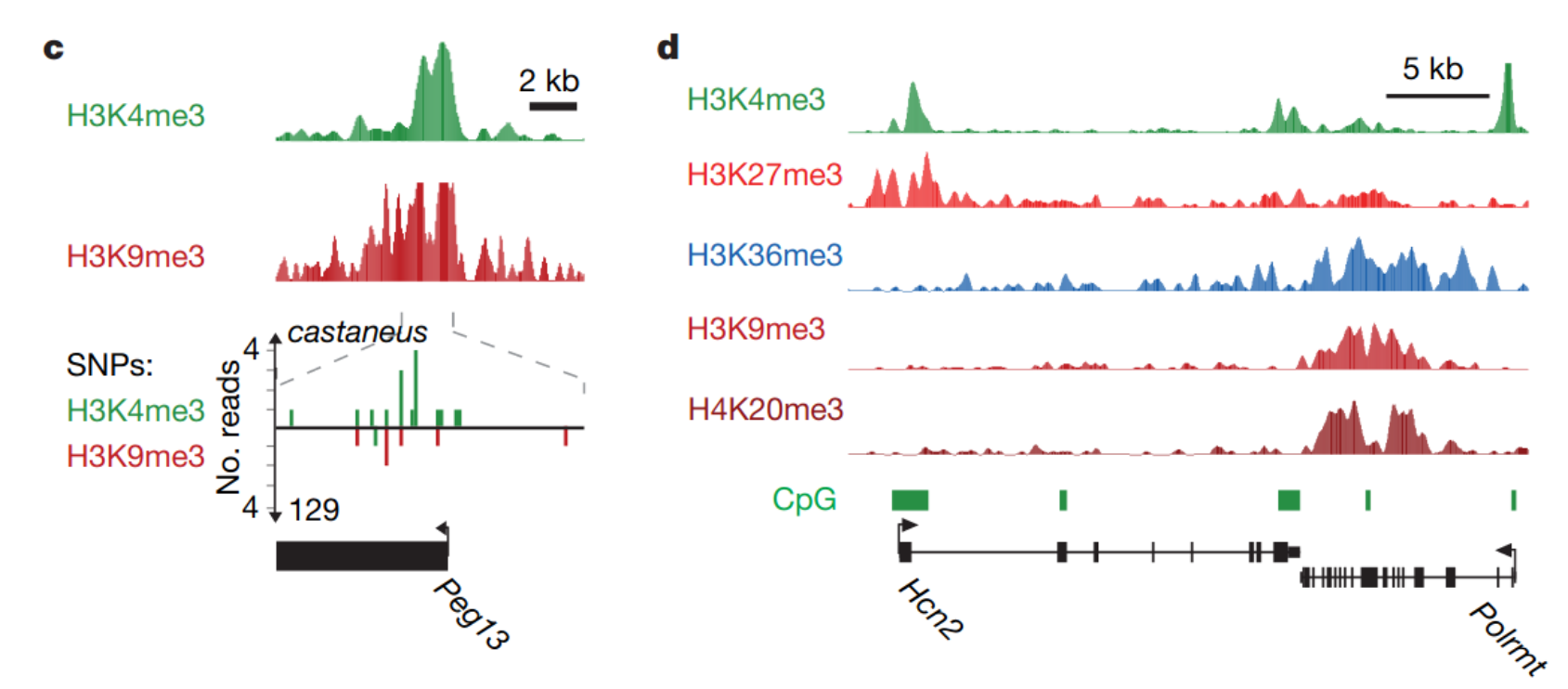


Figure 6 | Allele-specific histone methylation and genic H3K9me3/H4K20me3



H3K9 and H4K20 trimethylation mark specific repetitive elements

Figure 6 | Allele-specific histone methylation and genic H3K9me3/H4K20me3



Discussion

Discussion

- H3K4me3 and H3K36me3 allows recognition of promoters together with their complete transcription units
  - Should help to define alternative promoters and their usage in specific cell types
  - Identify the primary structure of genes encoding non-coding RNAs
  - Detect gene expression
  - Detect allele-specific transcription
- H3K9me3 and H4K20me3 should facilitate the study of heterochromatin, spreading and imprinting mechanisms.

Discussion

- H3K4me3 and H3K27me3 provides a rich description of cellular state.
    - Promoters may be classified as active, repressed or poised for alternative developmental fates
- Given the technical features of ChIP-Seq (high throughput, low cost and input requirement), it is now appropriate to contemplate projects to generate catalogues of chromatin-state maps representing a wide range of human and mouse cell types.