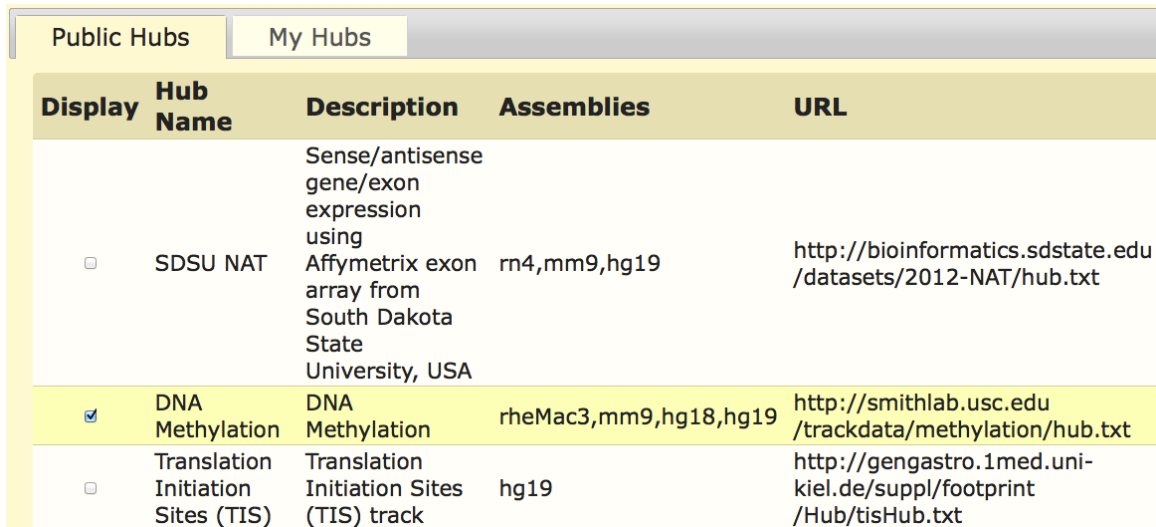# MethBase: a reference methylome database

MethBase is a central reference methylome database created from public BS-seq datasets. It contains hundreds of methylomes from well studied organisms. For each methylome, Methbase provides methylation level at individual sites, hypo- or hyper-methylated regions, partially methylated regions, allele-specifically methylated regions, and detailed meta data and summary statistics. These results are generated with the MethPipe software package, a standalone, comprehensive pipeline for analyzing BS-seq data, both WGBS and RRBS.

1 Access MethBase

MethBase is publicly available to the scientific community as a track hub in the UCSC Genome Browser. If you are using the main site of UCSC Genome Browser, the MethBase track hub is built in by default, you may select the MethBase tracks from the Public Hubs section of the UCSC Genome Browser (Figure 1). If you are using the a mirror site of UCSC Genome Browser, you may make the MethBase tracks available by loading the configuration file `http://smithlab.usc.edu/trackdata/methylation/hub.txt` (Figure 2). Please refer to Using UCSC Genome Browser Track Hubs for how to view unlisted hubs.

Display	Hub Name	Description	Assemblies	URL
☐	SDSU NAT	Sense/antisense gene/exon expression using Affymetrix exon array from South Dakota State University, USA	rn4,mm9,hg19	http://bioinformatics.sdstate.edu /datasets/2012-NAT/hub.txt
☑	DNA Methylation	DNA Methylation	rheMac3,mm9,hg18,hg19	http://smithlab.usc.edu /trackdata/methylation/hub.txt
☐	Translation Initiation Sites (TIS)	Translation Initiation Sites (TIS) track	hg19	http://gengastro.1med.uni-kiel.de/suppl/footprint /Hub/tisHub.txt

Figure 1: Select MethBase hub track from UCSC Genome Browser main site

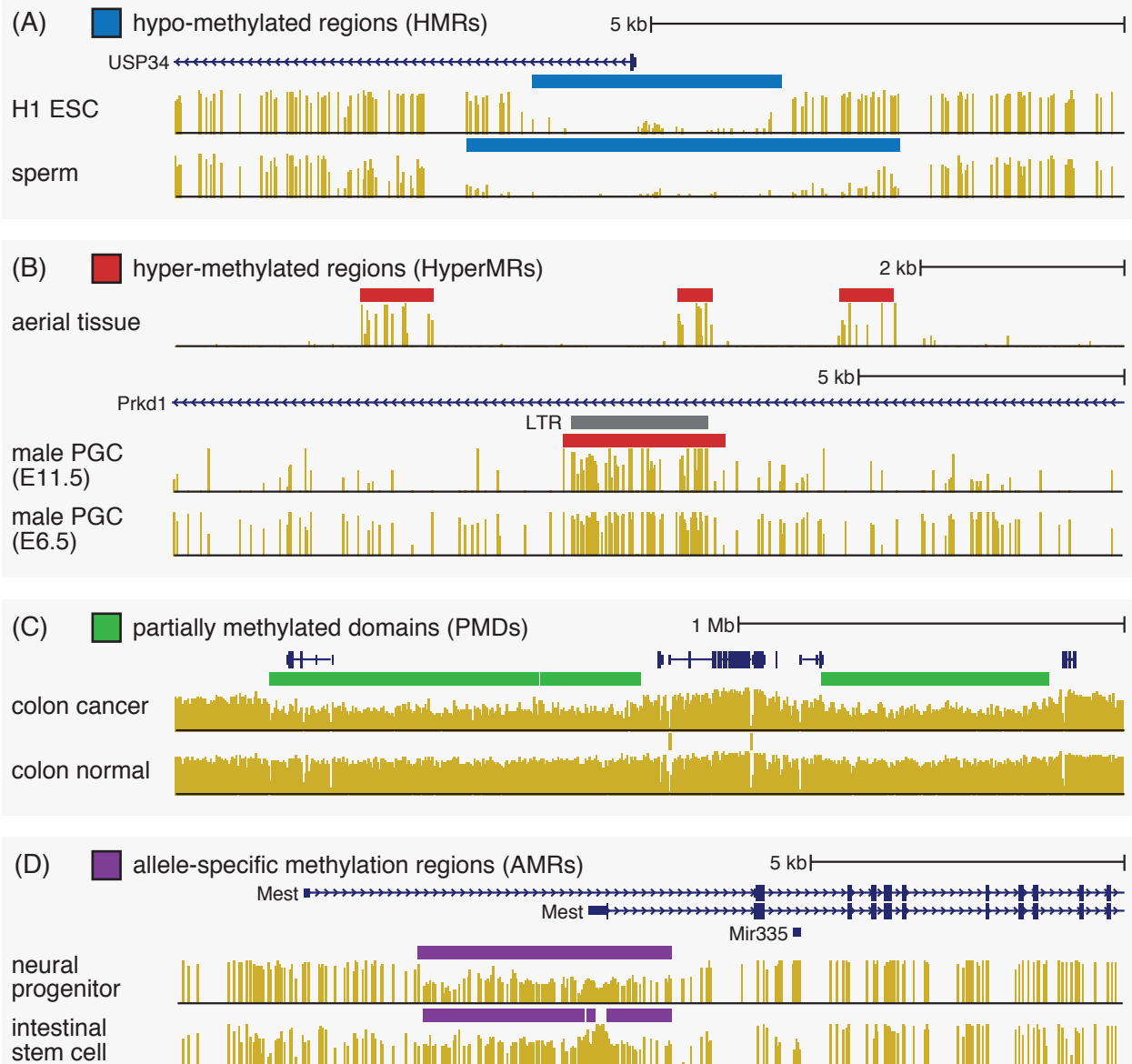Figure 2: Add MethBase track hub in mirror sites of UCSC Genome Browser

Figure 3: Examples of high-level methylation features available in MethBase through the UCSC Genome Browser track hub: (A) hypo-methylated regions (HMRs); (B) hyper-methylated regions (HyperMRs); (C) partially methylated domains (PMDs), and (D) allele-specific methylated regions (AMRs).

2 Select tracks for viewing

After you load the MethBase track hub setting file and go to the Genome Browser page, it shows the methylation level tracks and the HMR tracks of typical methylomes (Figure 3A).

Besides the default methylation tracks and HMR tracks, MethBase also provides a variety of other tracks, including coverage, hyper-methylated regions (Figure 3B), partially methylated domains (Figure 3C) and allele-specific methylated AMRs (Figure 3D). To display those tracks, you may go to the track setting page

in one of the three ways: 1) click the project name form the list of projects in Figure 5; 2) right click on of the viewable tracks, and select "Configure track settings" from the manu; or 3) go to the track description page and click the link "Go to track controls". In the track setting page (Figure 4), you can select additional samples from that project and modify the display settings of more types of data.

Hansen-Human-2011 Track Settings

Hansen-Human-2011

Maximum display mode: [full ▾] [Submit] [Cancel] Reset to defaults

Select dataType (help):
partially methylated domain [hide ▾]   **hypo methylated regions** [dense ▾]   methylation level [full]

Select subtracks by celltype:

Celltype	+ −
Human AdenoPolyp	☑
Human ColoCancer	☐
Human ColonicMucosa	☐

List subtracks: ○ only selected/visible   ⦿ all   (2 of 17 selected)

		Celltype↓1	DataType↓2	Track Name↓3	
☑	hide	Human AdenoPolyp	partially methylated domain	Human_AdenoPolyp_PMD	
☑	dense	Human AdenoPolyp	hypo methylated regions	Human_AdenoPolyp_HMR	
☑	full	🔧 Human AdenoPolyp	methylation level	Human_AdenoPolyp_Meth	
☑	hide	🔧 Human AdenoPolyp	coverage	Human_AdenoPolyp_Read	
☑	hide	🔧 Human AdenoPolyp		Human_AdenoPolyp_Allelic	
☑	hide	Human AdenoPolyp	allele specific methylated regions	Human_AdenoPolyp_AMR	
☐	dense	Human ColoCancer	hypo methylated regions	Human_ColoCancer_HMR	
☐	dense	Human ColoCancer	hypo methylated regions	Human_ColonCancer_HMR	
☐	full	🔧 Human ColoCancer	methylation level	Human_ColonCancer_Meth	
☐	hide	🔧 Human ColoCancer	coverage	Human_ColonCancer_Read	
☐	hide	🔧 Human ColoCancer		Human_ColoCancer_Allelic	
☐	hide	Human ColoCancer	allele specific methylated regions	Human_ColoCancer_AMR	
☐	dense	Human ColonicMucosa	hypo methylated regions	Human_ColonicMucosa_HMR	
☐	full	🔧 Human ColonicMucosa	methylation level	Human_ColonicMucosa_Meth	
☐	hide	🔧 Human ColonicMucosa	coverage	Human_ColonicMucosa_Read	
☐	hide	🔧 Human ColonicMucosa		Human_ColonicMucosa_Allelic	
☐	hide	Human ColonicMucosa	allele specific methylated regions	Human_ColonicMucosa_AMR	
2 of 17 selected					

[Submit]

Figure 4: Track setting interface for selectting additional samples and tracks

Below the Browser, you will a list of public methylomes in MethBase under the **DNA Methylation** section (Figure 5). From the list, you may modify the display settings of several samples from a project.

Figure 5: List of projects available in the MethBase, from the display settings of all tracks from that sample can be modified

3 View meta data and summary statistics

You can click on any track to go to the Description page for that track, which gives detailed meta and summary statistics about that methylome (Figure 6).

Human_ColonicMucosa_Meth (hub_12967_HansenHuman2011_HumanColonicMucosaMeth)

Position: chr20:31,349,033-31,349,565
Total Bases in view: 533
Statistics on: 9 items covering 9 bases (1.69% coverage)
Average item spans 1.00 **bases.**
Average value 0.295745 **min** 0 **max** 0.6 **standard deviation** 0.253729

Go to Colorectal Cancer and Adenomatous Polyp track controls

Data last updated: 2013-06-05 10:33:23

Description

Sample		BS rate	Methylation	Coverage	%CpGs	#HMR	#AMR	#PMD
Human_ColonCancer	colorectal cancer	0.987	0.602	6.163	0.840	39735	963	682
Human_AdenoPolyp	adenomatous polyp colon cancer	0.982	0.628	3.706	0.794	33521	183	1556
Human_ColonicMucosa	normal colonic mucosa	0.987	0.706	7.498	0.854	39409	666	3222

Methods

All analysis was done using a bisulfite sequencing data analysis pipeline MethPipe developed in the Smith lab at USC.

Figure 6: Detailed meta data and summary statistics of methylomes

4 Download data for further analysis

There are two ways to download data files from MethBase. First, you may go to the Table Browser to download the data files of any tracks for additional analysis (Figure 7).

Figure 7: Download data from the Table Browser

Alternatively, you may click the **Download** link from the Description page of each methylome (Figure 6), which will direct you to the track data directory. The *.meth.bw and *.read.bw files contain methylation level and coverage information respectively at each CpG site. The *.hmr.bb, *.pmd.bb, *.amr.bb files contain HMRs, PMDs and AMRs, and the *.allelic.bw file gives the allelic score at each CpG site. To extract text files from these bigWig and bigBed files, you may use the bigWigToBedGraph and bigBedToBed utilities available from UCSC server.

5   Contact

Please send emails to the methpipe mailing list at **methpipe@googlegroups.com** if you have any questions, suggestions or comments.

6   References

Song Q, Decato B, Hong E, Zhou M, Fang F, Qu J, Garvin T, Kessler M, Zhou J, Smith AD (2013) A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. PLOS ONE (in press) [PDF]