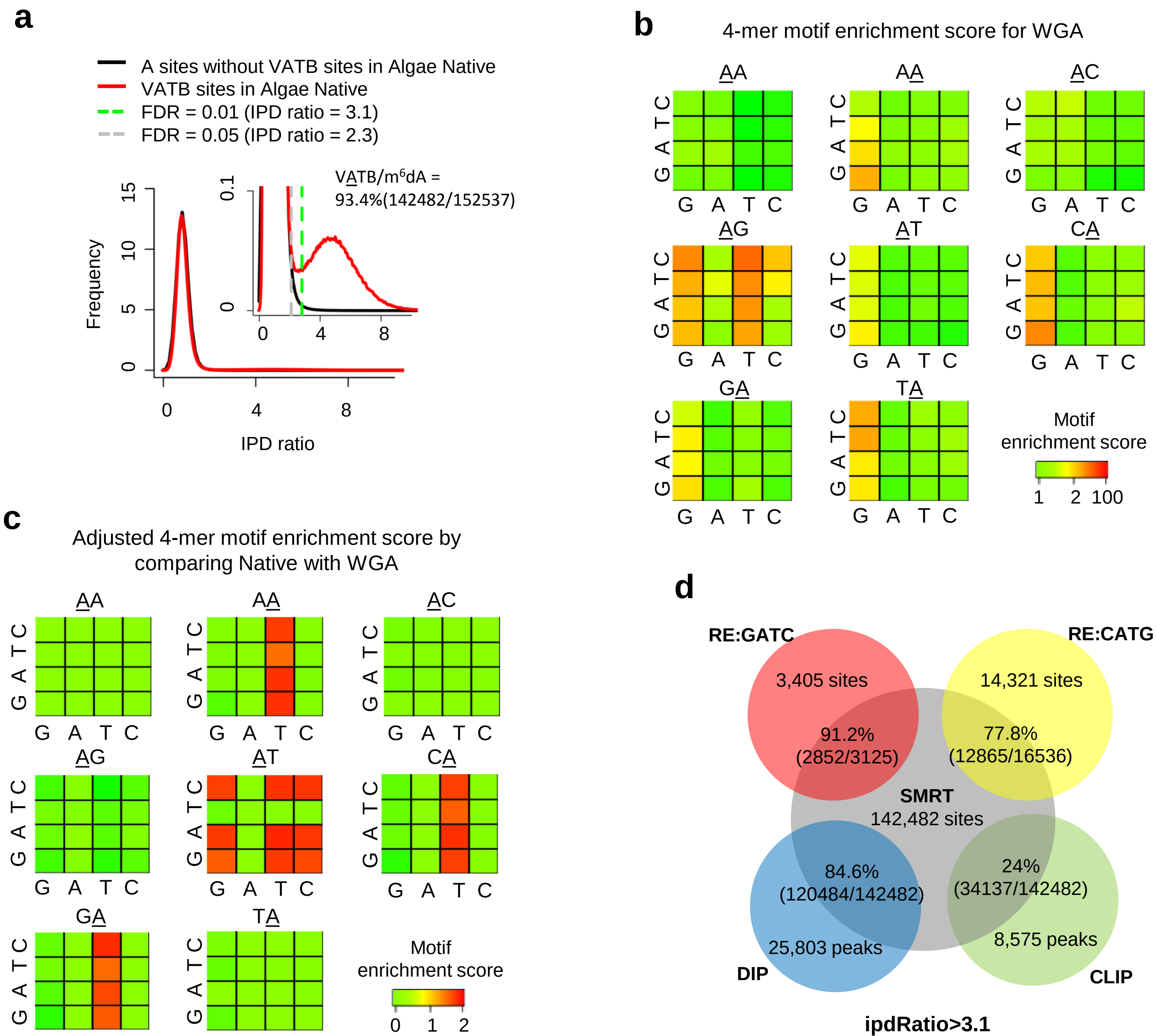
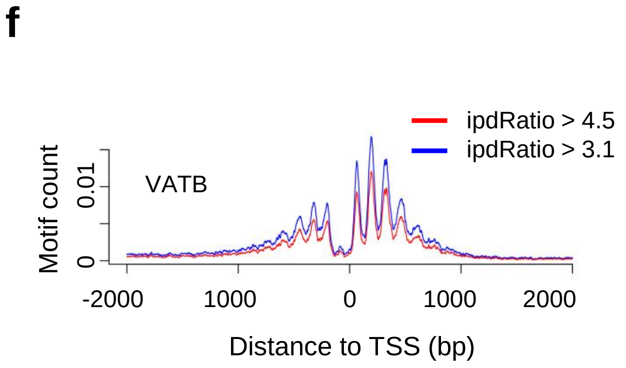
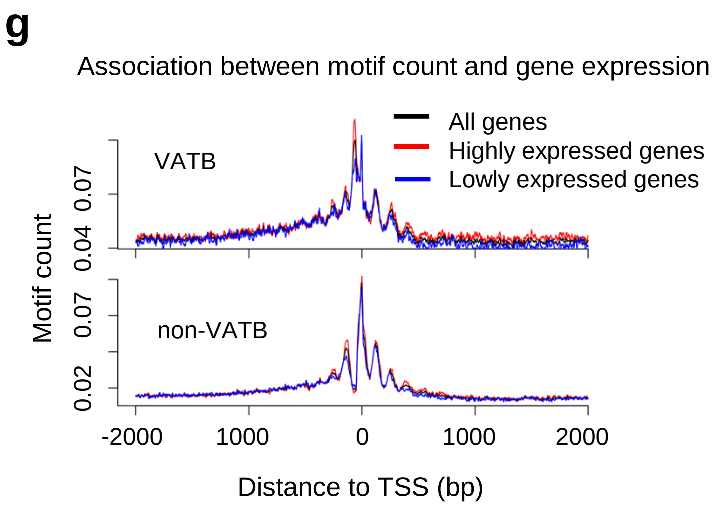
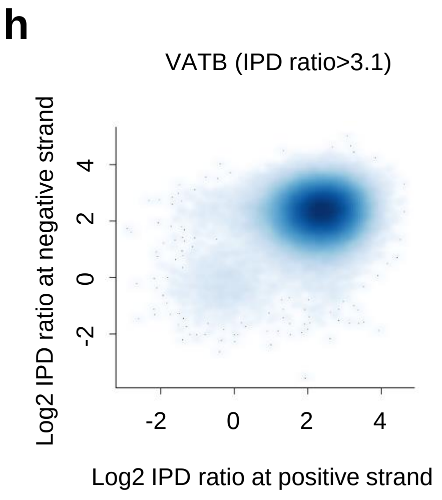

a
A sites without VATB sites in Algae Native
VATB sites in Algae Native
FDR = 0.01 (IPD ratio = 3.1)
FDR = 0.05 (IPD ratio = 2.3)
VATB/m^{6}dA =
93.4%(142482/152537)
Frequency
0
5
10
15
0
4
8
IPD ratio
0.1
0
0
4
8
b
4-mer motif enrichment score for WGA
AA
AA
AC
AG
AT
CA
GA
TA
G A T C
G A T C
Motif
enrichment score
1
2
100
c
Adjusted 4-mer motif enrichment score by
comparing Native with WGA
AA
AA
AC
AG
AT
CA
GA
TA
G A T C
G A T C
Motif
enrichment score
0
1
2
d
RE:GATC
RE:CATG
3,405 sites
14,321 sites
91.2%
(2852/3125)
77.8%
(12865/16536)
SMRT
142,482 sites
84.6%
(120484/142482)
24%
(34137/142482)
25,803 peaks
8,575 peaks
DIP
CLIP
ipdRatio>3.1


e
Chr1
5,633,800 bp
5,634,000 bp
5,634,200 bp
5,634,400 bp
5,634,600 bp
5,634,800 bp
5,635,000 bp
m6dA-DIP-seq
m6dA-RE-seq
SMRT-seq
CATG
CATG
Chr1
5,735,000 bp
5,736,000 bp
m6dA-DIP-seq
m6dA-RE-seq
SMRT-seq
CATT
CATG
GATT
CATC
AATT
AATT


f
ipdRatio > 4.5
ipdRatio > 3.1
VATB
Motif count
0
0.01
-2000
1000
0
1000
2000
Distance to TSS (bp)


g
Association between motif count and gene expression
All genes
Highly expressed genes
Lowly expressed genes
VATB
non-VATB
Motif count
0.04
0.07
0.02
0.07
-2000
1000
0
1000
2000
Distance to TSS (bp)


h
VATB (IPD ratio>3.1)
Log2 IPD ratio at negative strand
4
2
0
-2
-2
0
2
4
Log2 IPD ratio at positive strand

Supplemental Fig S4. Additional data in the characterization of a complete m^{6}dA methylome of *C. reinhardtii*. **(a)** FDR estimation by comparing the IPD ratio distribution of *C. reinhardtii* of VATB site (red) with all other A sites (black) within the native sample. The inset provides an enlarged view. **(b)** A rigorous motif enrichment analysis for WGA as a negative control. Each 4x4 heatmap corresponds to all sixteen 4-mer motifs, for which 2nd and 3rd bases are fixed at the center/title. The rows and columns in the heatmaps represent the first and last bases of 4-mer motifs. Each cell in the following 4x4 heatmaps shows the motif enrichment scores for the WGA DNA sample. **(c)** The ratio of the motif scores between the native DNA and the WGA control. **(d)** Putative m^{6}dA sites called by SMRT-seq (IPD ratio > 3.1) are highly consistent with those detected by independent techniques: m^{6}dA-DIP-seq (DIP), m^{6}dA-CLIP-exo-seq (CLIP) and m^{6}dA-RE-seq (RE). **(e)** Illustrative examples showing m^{6}dA events called by m^{6}dA-DIP-seq, m^{6}dA-RE-seq, and SMRT-seq. The motif content is annotated for each m^{6}dA event. An m^{6}dA event will be missed by m^{6}dA-RE-seq if the event resides in a motif context not recognized by the RE. m^{6}dA-DIP-seq can miss an m^{6}dA event due to certain bias and lack of sensitivity commonly associated with antibody based approaches. **(f)** The distribution of putative methylated V$\underline{\text{A}}$TB motif count around TSS (red curve with IPD ratio > 4.5 and blue curve with IPD ratio > 3.1). The area covered by the peak is approximated as the multiplication of peak height and the number of nucleotide in the linker region (~70nt): area under curve is ~1 for IPD ratio > 4.5 and ~1.2 for IPD ratio > 3.1, respectively. (**g**) No clear association between gene expression and VATB motif count (top panel) as well as non-VATB (i.e. TATN/NATA). Motif counts are compared between three groups of genes: all genes (black), highly expressed (red, FPKM>1) and lowly expressed genes (blue, FPKM<1). (**h**) Single molecule, strand-specific analysis of SMRT-seq data to examine full-, non- or hemi-methylation status at putative m^{6}dA sites (VATB with IPD ratio > 3.1). The X- and Y- axes denote the single molecule, strand-specific IPD ratio of each pair of reverse-complementary VATB sites at the two strands of each single molecule.