## Supplemental information

## Cytosine modifications modulate the chromatin architecture of transcriptional enhancers

Elise A Mahé, Thierry Madigou, Aurélien A Sérandour, Maud Bizot, Stéphane Avner, Frédéric Chalmel, Gaëlle Palierne, Raphaël Métivier and Gilles Salbert

## Legends to supplemental Figures

Supplemental Figure S1. Relevance of P19 ECCs for the study of the influence of DNA hydroxymethylation in NPC enhancer priming and activation. $(A)$ Graph depicting the percentage of overlap between hydroxymethylated regions (hMRs) identified by hMeDIP-seq in ESC-derived (orange curve) or ECC-derived (gray curve) NPCs and another replicate of ESC-derived NPCs (GSM978376). (B) Integrated genome browser screenshots showing 5hmC signal (hMeDIP) in two regions of ESC- and ECC-derived NPCs. (C) Heatmap of 5hmC signal (hMeDIP) in ECC- and ESC-derived NPCs at 20,492 regions gaining 5hmC (up DhMRs) during P19 NPC commitment. (D) Venn diagram indicating the overlap between ECC-derived NPC hMRs and 48 enhancers shown to be active in neural tissues of E11.5 dpc mouse embryos (VISTA enhancers, http://enhancer.lbl.gov/). (E) IGB screenshots showing activation of selected VISTA enhancers during P19 cells differentiation into NPCs. Images of E11.5 dpc mouse embryos expressing LacZ under the control of the indicated enhancers were downloaded from the VISTA enhancer browser.

Supplemental Figure S2. Dynamics of chromatin features at down DhMRs during P19 cells differentiation. (A) Average kinetic profiles of $5 \mathrm{hmC}, 5 \mathrm{mC}, \mathrm{H} 3 \mathrm{~K} 4 \mathrm{me} 1, \mathrm{H} 3 \mathrm{~K} 27 \mathrm{ac}$ and chromatin opening (FAIRE) at 1,712 regions losing 5hmC (down DhMRs) at 48 hours of P19 ECC differentiation. (B) Average profiles of FOXA1, PBX1 and MEIS1 binding to down DhMRs at 48 hours of differentiation.

Supplemental Figure S3. TALE-HDs are recruited to common sites. (A) Heatmap representations of PBX1 signal at MEIS1-bound regions at 48 hours of P19 ECC differentiation. (B) Part of a STRING (Search Tool for the Retrieval of Interacting Genes; http://string-db.org/) schematic representation of interactions between proteins retrieved from PBX1 ChIP-MS data. Pink lines represent known physical interaction and grey lines functional relationships. The structure of the detected proteins is shown within circles. (C) Validation of MEIS1/2 and PBX1/2/3 binding to differentially modified MEIS1/PBX1 probes by supershift experiments. Ab: antibody; NE: P19 nuclear extract; HA: Human influenza virus hemagglutinin (negative control). Validation was also achieved for PBX1/HOXA9 probes (data not shown). (D) Representative Western blot analysis of a pull-down experiment showing MEIS1 interaction with the PBX1/HOXA9 probe containing various combinations of cytosine modifications at
the CpG dinucleotide (C: unmodified, M: 5mC, H: 5hmC - first letter: upper strand, second letter: lower strand).

Supplemental Figure S4. MEIS1 binds a TGACGG motif. (A) Average profiles of the SCL-exo signal at the full 20,287 bulk MEIS1 binding sites before (grey curve) or after (orange curve) exclusion of the 958 MEIS1 sites with centered 5 hmCpGs . For each region, the " 0 " position corresponds to the peak summit of the MEIS1 ChIP-seq signal. (B) Box plot of the CpG density (number of CpGs per 100 bp ) in the MEIS1 ChIP-seq peaks for 20,287 bulk MEIS1 sites or the 958 sites with centered 5 hmC . The mean value ( m ) is indicated for each category of binding sites. (C) IGB screenshots of MEIS1 ChIP-seq and SCL-exo signals at 5 regions selected from the 958 MEIS1 sites with centered 5 hmC and showing putative MEIS1 binding sites highlighted in red font in their corresponding sequences. Asterisks indicate 5 hmCpGs whereas open circle indicate non-hydroxymethylated CpGs. (D) Results of de novo motif search in the set of 958 MEIS1 binding regions with the SeqPos motif tool from Cistrome. (E) First three ranked de novo motifs found in the 958 MEIS1 sites with SEME (http://biogpu.d1.comp.nus.edu.sg/~chipseq/webseqtools2/). (F) First three ranked de novo motifs found with SEME in all 3,627 MEIS1 motifs devoid of CpGs. (G) Western blot analysis of a pull-down experiment in which the PBX1/MEIS1 biotinylated probe was incubated with nuclear extracts with or without (-) double-stranded DNA competitors including a TGACGG site. The modification status of the CpG dinucleotide within the the competitor oligonucleotides is indicated (C: unmodified $\mathrm{C}, \mathrm{M}: 5 \mathrm{mC}, \mathrm{H}$ : 5 hmC - first letter: upper strand, second letter: lower strand).

Supplemental Figure S5. Efficiency of 5-azadC and DMOG treatments. ( $A-B$ ) Global 5 hmC and 5 mC levels of three DNA quantities (300, 150 or 75 ng ) from P19 cells either undifferentiated (-RA) or differentiated during 48 hours with RA $(+R A)$ and treated with or without $5-\operatorname{azadC}(A-n=6)$ or DMOG ( $B-n=9$ ) were evaluated by dot-blot. Quantified signals for either 5 mC of 5 hmC were normalized to total DNA signals obtained by methylene blue staining. Statistical significance was tested with a Mann and Whitney test (5-azadC: \# for $\mathrm{p}<0.05$, \#\# for $\mathrm{p}<0.01$; DMOG: ${ }^{* * *}$ for $\mathrm{p}<0.005$ ). ( $C$ ) Example of a dot blot assay for 5 hmC (upper panel) and 5 mC (lower panel) in the experimental conditions depicted in (B).

Supplemental Figure S6. Perturbing cytosine modification levels delays differentiation of neural progenitors. (A) Kinetic analysis of the expression of Oct4, Nanog, Tshz1, Pax6, Irx3, FoxA1, Meis1 and Pbx1 genes by RT-qPCR ( $n=6$ ) during the first 48 hours of RA-induced differentiation of P19 ECCs (control: black lines, DMOG: red lines, 5-azadC: dashed grey lines). Statistical significance was tested with a Mann and Whitney test (DMOG: * for $\mathrm{p}<0.05,{ }^{* *}$ for $\mathrm{p}<0.01$; 5 -azadC: \# for $\mathrm{p}<0.05$, \#\# for
p<0.01). (B) Western blot quantification of MEIS1, PBX1 and FOXA1 proteins for 5-azadC (left panel) or DMOG (right panel) treated samples compared to control $(\mathrm{n}=3)$.

Supplemental Figure S7. Specificity of 5-azadC and DMOG treatments. (A-B) Impact of 5-azadC (A) or DMOG (B) treatments on a control site having no CpG and bound by MEIS1 and FOXA1 proteins. Histograms correspond to DIP-, ChIP- and FAIRE-qPCR results obtained from P19 cells undifferentiated (-RA) or differentiated during 48 hours with RA (+RA) with or without inhibitor treatment ( $5-m C$ : $n=9$; $5 \mathrm{hmC}, \mathrm{H} 3 \mathrm{~K} 4 \mathrm{me} 1, \mathrm{H} 3 \mathrm{~K} 27 \mathrm{ac}$ and FAIRE: $\mathrm{n}=6$; MEIS1 and FOXA1: $\mathrm{n}=9$ ). Analysis of the opening of chromatin was performed by Formaldehyde Assisted Isolation of Regulatory Elements (FAIRE). (C) ChIP-qPCR analysis of H3K27me3 enrichment at the 15 enhancers of interest in DMOG treated cells versus control cells $(n=6)$. Note that the signals for 5 mC and 5 hmC at this control region without CpGs are below the thresholds used to consider enrichment as significant (red lines).

Supplemental Figure S8. Impact of DNMT inhibition on characteristic features of enhancers. Histograms correspond to DIP-, ChIP- and FAIRE-qPCR results obtained from undifferentiated (-RA) or differentiated P19 cells (+RA, 48 hours) with or without 5 -azadC treatment ( $5-\mathrm{mC}$ : $\mathrm{n}=9$; 5 hmC , H3K4me1, H3K27ac and FAIRE: $n=6$; MEIS1 and FOXA1: $n=9$ ). Enrichment values were reported to the -RA condition and are expressed as fold change. Statistical significance was tested with Mann and Whitney (\# for $\mathrm{p}<0.05$, \#\# for $\mathrm{p}<0.01$, \#\#\# for $\mathrm{p}<0.005$ ). ( $A$ ) Enhancers bound by MEIS1 but not FOXA1 (M1 to M5). (B) Enhancers bound by both MEIS1 and FOXA1 (MF1 to MF5). (C) Enhancers bound by FOXA1 but not MEIS1 (F1 to F5).

Supplemental Figure S9. Impact of TET inhibition on characteristic features of enhancers. Histograms correspond to DIP-, ChIP- and FAIRE-qPCR results obtained from undifferentiated (- RA) or differentiated P19 cells (+RA, 48 hours) with or without DMOG treatment ( 5 hmC : $\mathrm{n}=9$, excepted 48 hours $n=15$; $5-m C$ : $n=6$; H3K4me1: $n=6$, excepted 48 hours $n=9$; H3K27ac: $n=12$; FAIRE, MEIS1 and FOXA1: $n=9$ ). Enrichment values were reported to the -RA condition and are expressed as fold change. Statistical significance was tested with Mann and Whitney (* for $p<0.05$, ** for $p<0.01$, *** for $p<0.005$ ). (A) Enhancers bound by MEIS1 but not FOXA1 (M1 to M5). (B) Enhancers bound by both MEIS1 and FOXA1 (MF1 to MF5). (C) Enhancers bound by FOXA1 but not MEIS1 (F1 to F5).

Supplemental Figure S10. Principal Component Analysis (PCA) and Hierarchical Clustering on Principal Component (HCPC) of the impact of DNMT or TET inhibition on undifferentiated P19 ECCs. PCA and HCPC results obtained from the $\log _{2}$-transformed fold-change between 5 -azadC $(A-D)$ - or DMOG ( $E-H$ )treated versus control cells for $5 \mathrm{hmC}, 5 \mathrm{mC}, \mathrm{H} 3 \mathrm{~K} 4 \mathrm{me} 1, \mathrm{H} 3 \mathrm{~K} 27 \mathrm{ac}$ and FAIRE at the 15 enhancers studied.
$(A)$ and $(E)$ Scree plots of information rate of the principal components. Most of the variance in the data set is contained in the first two $(A)$ or four $(E)$ principal components. $(B)$ and $(F)$ Individuals factor maps showing the coordinates of the individuals on the two first principal components. $(C)$ and $(G)$ Correlation loading plots of the variables with the components. The eigenvalues of the correlation matrix are symbolized as vectors representing traits that influence the principal components 1 and 2 . $(D)$ and $(H)$ Hierarchical Clustering on Principal Components map.

Supplemental Figure S11. Principal Component Analysis (PCA) and Hierarchical Clustering on Principal Component (HCPC) of the impact of DNMT or TET inhibition on differentiated P19 ECCs. PCA and HCPC results obtained from the $\log _{2}$-transformed fold-change between 5 -azadC $(A-D)$ - or DMOG $(E-H)$ treated versus control cells for $5 \mathrm{hmC}, 5 \mathrm{mC}, \mathrm{H} 3 \mathrm{~K} 4 \mathrm{me}, \mathrm{H} 3 \mathrm{~K} 27 \mathrm{ac}$ and FAIRE at the 15 enhancers studied. $(A)$ and $(E)$ Scree plots of information rate of the principal components. Most of the variance in the data set is contained in the first three principal components. $(B)$ and $(F)$ Individual factor map showing the coordinates of the individuals on the two first principal components. $(C)$ and $(G)$ Correlation loading plots of the variables with the components. The eigenvalues of the correlation matrix are symbolized as vectors representing traits that influence the principal components 1 and 2. (D) and (H) Hierarchical Clustering on Principal Components map.

Supplemental Figure S12. 5-azadC and DMOG treatments affect the methylation or hydroxymethylation status of selected CpGs in eight enhancers of interest. Epimark ${ }^{\circledR}$ 5-hmC and 5-mC analysis of the impact of 5-azadC and DMOG treatments of undifferentiated or differentiated P19 cells, in the context of CCGG sequences. This assay is based on the lack of activity of the Mspl restriction enzyme towards CCGG restriction sites containing glucosylated 5 hmCs . Hence, sites containing 5 hmCs can be discriminated from sites containing 5 mCs after glucosylation of genomic DNA with betaglucosyltransferase. After digestion with MspI, DNA was amplified by PCR using primers pairs specific for each tested regions (see Supplemental Materials and Methods). The levels of unmodified cytosines were also analysed by using Hpall, another restriction enzyme cutting the sequence CCGG only when non-methylated. Results are expressed as percentages for each cytosine form.

Supplemental Figure S13. Basal levels of enhancer features at the 15 selected enhancers. Fold enrichment compared to a negative control of $5 \mathrm{hmC}, 5 \mathrm{mC}, \mathrm{H} 3 \mathrm{~K} 4 \mathrm{me} 1, \mathrm{H} 3 \mathrm{~K} 27 \mathrm{ac}$ and chromatin opening (FAIRE) in undifferentiated P19 cells for the five MEIS1 sites (M), the five MEIS1-FOXA1 sites (MF), the five FOXA1 sites (F) and an enhancer having no CpG ("No CpG "). Red lines correspond to the thresholds determined for each mark and above which enrichment was considered as significant. Thresholds were set at: 4-fold enrichment compare to negative control for 5 hmC , 2 -fold enrichment for 5 mC and 3 -fold
enrichment for H3K4me1 and H3K27ac. No threshold was set for chromatin opening as this feature corresponds to levels of opening rather than presence or absence of a specific mark.

Supplemental Figure S14: Knock-down of TETs by siRNAs in P19 cells reduces global 5hmC levels. (A) RT-qPCR analysis of TET1,2,3 expression upon transfection of control siRNA (siCTRL) or individual TET siRNAs in P19 cells treated or not with RA $(\mathrm{n}=6)$. (B) Quantification of a dot blot analysis of global 5 hmC levels in untransfected (-Si) or siRNA-transfected P19 cells treated or not with RA ( $\mathrm{n}=6$ ). In $(A)$ and $(B)$, statistical significance was tested with a Mann and Whitney test (* for $p<0.05$, ** for $p<0.01$, *** for $\mathrm{p}<0.005$ ).

Supplemental Figure S15: hMeDIP- and FAIRE-qPCR analysis of 8 selected enhancers upon TET siRNA transfection in P19 cells treated or not with retinoic acid (RA). Enhancers showing marked variations in 5 hmC content upon DMOG treatment were selected for this study. Statistical significance was tested with a Mann and Whitney test (* for $p<0.05$, ** for $p<0.01$, ${ }^{* * *}$ for $p<0.005, n=4$ to 6 )

Supplemental Figure S16: (A) Heatmaps of MEIS1 ChIP-seq and hMeDIP-seq (5hmC) signals in clusterized MEIS1 binding regions outside CpG islands. (B) Average profiles of 5hmC (hMeDIP), FOXA1, TET1 and TET2 recruitment at the C1, C2, C3, C4 and C5 clusters shown in (A).

Supplemental Figure S17. Model for cytosine modification-promoted events during enhancer priming and activation. RA-mediated neural differentiation of P19 ECCs is associated with loss of expression of the pluripotency factors OCT4 and NANOG and the sequential induction of FOXA1 and MEIS1 pioneer factors. At an early stage of differentiation, FOXA1 binding to inactive enhancers associates with 5 mC oxidation which indirectly promotes monomethylation of H3K4 by histone methyltranferases (HMTs), while 5 hmC formation promotes chromatin opening. Both processes allow the establishment of a "primed enhancer" chromatin state. Later during differentiation (i.e. when MEIS1 levels raise) MEIS1 recruitment to primed enhancers is favored by the presence of modified cytosines either fully methylated or methylated on one DNA strand and hydroxymethylated on the other strand and correlates with H3K27 acetylation by histones acetyl-transferases (HATs), resulting in enhancer activation.

Supplemental Table S1. Proteomic data for selected proteins, including HOX and TALE-HD proteins, found associated with PBX1 on chromatin. Exclusive Unique Peptide Count and percentage of coverage for ChIP-MS data are shown. Results are presented for three replicates of PBX1 immunoprecipitation compared to three replicates of $\operatorname{IgG}$ negative control. Proteins were retained only if at least one reliable
peptide (95\% confidence) was found in two out of the three PBX1 replicates and in none of the $\operatorname{lgG}$ control replicates.

Supplemental Materials and Methods

Antibodies

|  | Target | Reference |
| :--- | :--- | :--- |
|  | 5hmC | Diagenode, CS-HMC-100 |
|  | 5mC | Diagenode, Mab-081-500 |
|  | PBX1 | Abnova, H00005087-M01 |
|  | H3K4me1 | Abcam, ab8895 |
| Immunoprecipitations | H3K27ac | Abcam, ab4729 |
|  | MEIS1 | Abcam, ab19867 |
|  | FOXA1 | Abcam, ab5089 |
|  | TET1 | Millipore, 09-872 |
|  | TET2 | ProteinTech, 21207-1-AP |
|  | H3K27me3 | Upstate, 07-449 |
|  | 5hmC | Active Motif, \#39769 |
| Blots | 5mC | Diagenode, Mab-081-500 |
|  | MEIS1 | Abcam, ab124686 |
|  | PBX1/2/3/4 | Santa Cruz, sc-25411X |
|  | FOXA1 | Abcam, ab5089 |
|  | Sp1 | Santa Cruz, sc-59 |
|  | Actin | Sigma, A1978 |
|  | MEIS1/2 | Santa Cruz, sc-10596X |
| Supershift | PBX1/2/3 | Santa Cruz, sc-888 |
|  | HA | Santa Cruz, sc-805 |
|  |  |  |

Oligonucleotides
Oligonucleotides were designed under Primer 3 interface (http://frodo.wi.mit.edu/primer3/)

| Primer type | Name | Forward primer | Reverse primer |
| :---: | :---: | :---: | :---: |
| RT-qPCR | RPS28 | GGTGACGTGCTCACCCTATT | TTCCGTGGGCTAAGTAGTGG |
|  | OCT4 | TTGCAGCTCAGCCTTAAGAA | TGGTCTCCAGACTCCACCTC |
|  | NANOG | AAGCAGAAGATGCGGACTGT | GCACTTCATCCTTTGGTTTTG |
|  | FOXA1 | GAAGATGGAAGGGCATGAAA | GCCTGAGTTCATGTTGCTGA |
|  | MEIS1 | CAGAAAAAGCAGTTGGCACA | TGCTGACCGTCCATTACAAA |
|  | PBX1 | GCCAATATTTATGCTGCCAAA | AATCCCCATTGAGTGACTGC |
|  | TSHZ1 | CCGAGGAGGAAGCAGCAG | TGGGCTTCTTTGATCTCTGTC |
|  | IRX3 | ACAGCACCTGCTGGGACTT | TTCTCCACTTCCAAGGCACT |
|  | PAX6 | ATGGCAAACAACCTGCCTAT | CATGGGCTGACTGTTCATGT |
|  | p21 | CACTTCTTACCTGGGGGTGA | AGCTGGCCTTAGAGGTGACA |
|  | GATA4 | TCCAGCCTGAACATCTACCC | GTGTGAAGGGGTGAAAAGGA |
|  | TET1 | GAGCCTGTTCCTCGATGTGG | CAAACCCACCTGAGGCTGTT |
|  | TET2 | GTTGTTGTCAGGGTGAGAATC | TCTTGCTTCTGGCAAACTTACA |
|  | TET3 | CCGGATTGAGAAGGTCATCTAC | AAGATAACAATCACGGCGTTCT |
|  | p53 | TGGAAGACTCCAGTGGGAAC | TCTTCTGTACGGCGGTCTCT |
|  | NESTIN | CCAAAGAGGTGTCCGATCAT | TCATCAGCAAACCCATCAGA |
| ChIP-qPCR | Negative control | TCCTTTGACCTCCACACACA | GCAACAGAAGATGCGATGAA |
|  | M1 | AGAGCGAACCGAGAAATCAA | AGAGCCTCTATTTGCCACGA |
|  | M2 | GACCTGAGCGCACAGTAACA | ATGAAGCTGCCTTTTGAGGA |
|  | M3 | CCGCATTATGAACACACTCG | GCTCTCCCAGGGGTAATAGG |
|  | M4 | ACATTCCCCAGACAAGCAAG | CACTGACTGTCATCCCCTCA |
|  | M5 | CCGAGTGTTCCTCCACCTAA | ATTTGCATACAGGCGGAGTC |
|  | MF1 | TGAGGTGCTCAGTGGAGATG | CTCCAGAGTGAGCGTTGACA |
|  | MF2 | CAGTCCCATCGATCACTTACA | TTCAAAGGCTTGCTTGACCT |
|  | MF3 | GCAGCCACTGAAGAGACAAA | TGGTTGTGGGCTTGTGTTTA |
|  | MF4 | GACAAGCAGAACAGAACGGG | TGTCACTGCCCCTTTCATCT |
|  | MF5 | TGCTTTACTCTCTGCAGCCT | ATATGGCCCTGTCAGTTCCC |
|  | F1 | GAAAAGCAAAGCAAAGGGGC | GAAGGCGACACTGCTCTGTA |
|  | F2 | CACATGCTAAGCCAGGAAGC | CCTGAGTTTGACCCCTGGAT |
|  | F3 | GCCGATGAGTAAACAGCTGA | CCGTTTCAAGCTGGTGTACA |
|  | F4 | CAGGTGGCTGAGACTGAGTA | TGGGAAATTTCAGGTTGGCC |
|  | F5 | TGAAGTCCTCCAGTGGCTTT | CCCCAAGGCAAGAGTAAATACA |
|  | No CpG | ACTATCTATCAAGGGAGCAAACA | TGAGCAGTAAAACCAGGTCTG |
| Epimark ${ }^{\text {® }}$ | M2 | CACATGTCCCTTTTGCTCCT | AGAGCGAGCACCTCACAAAA |
| 5-hmC | M3 | CCGCATTATGAACACACTCG | GCTCTCCCAGGGGTAATAGG |
| and 5-mC | M5 | TCTTTCAGCGGGACTGAGAT | GGTTGATTGCCAGCTTTGTT |
| Analysis | F4 | ATGACCAGTATGCACCACCA | TCTTCCCTGACCCTACATGG |



E



Mahé et al._Supplemental Fig. S3
A MEIS1 PBX1


Signal
Low High


D

A

B

C
Example \#1: ATGATGTCTGGAGAAAGTGAČ́GGTGTCACG̊ATTGA TGA $T$

|  | logo | hits | 2 score | pral |
| :---: | :---: | :---: | :---: | :---: |
| de novo 1 | ExCACG | 563 | 23.9 | 1.0e-30 |
| de novo 2 | \& CGTs | 463 | -23.4 | 1.0e-30 |
| de novo 3 | CGIs C- | 414 | -18;8 | -30 |
| de novo 4 | $x \pm A c G$ | 464 | -18.0 | 1.0e-30 |
| de novo 5 | $\mathrm{C}_{2} \mathrm{C}_{5}$ | 208 | -16.4 | 1.0e-30 |
| de novo 6 | cemgrga | 202 | -12.0 | 1.0e-30 |
| de novo 7 | AGoges. Sm | 602 | -10.8 | 1.2e-27 |
| de novo 8 | G CAsma | 403 | -10.5 | 6.5e-26 |
| de novo 9 | ${ }_{80}$ TsAcom ${ }^{\text {a }}$ | 797 | -9.7 | 1.9e-22 |
| de novo 10 | TGC CAS | 498 | -6.8 | 5.1e-12 |
| de novo 11 | $g=c, C C$ | 855 | -4.8 | 8.3e-7 |
| de novo 12 | -TGACAE | 808 | ${ }^{-3.6}$ | 1.3e-4 |
| de novo 13 | Tearag mat | 151 | ${ }^{-3.1}$ | 8.2e-4 |

E

Empirical score: 17.1

Empirical score: 11.6
wouts "poctiTC
Empirical score: 7.9
F
Motif 1
(Pbx1)

Empirical score: 35.4

Empirical score: 31.7

G











H3K4me1



FAIRE


C H3K27me3 $\quad \begin{aligned} & \square-R A \\ & \square-R A+\text { DMOG } \quad \square+R A \\ & +R A+D M O G\end{aligned}$












思: RA
RA + DMOG

$$
\begin{array}{ll}
\square+\mathrm{RA} \\
\square+\mathrm{RA}+\mathrm{DMOG} & -+\mathrm{RA} \\
\square+\mathrm{RA}+\mathrm{DMOG}
\end{array}
$$












$$
\begin{array}{ll}
\square+\mathrm{RA} & --+\mathrm{RA} \\
\square+\mathrm{RA}+\mathrm{DMOG} & -+-\mathrm{RA}+\mathrm{DMOG}
\end{array}
$$






-     + RA
-     +         + RA + DMOG


























| MF3 | $\begin{aligned} & \nabla-\mathrm{RA} \\ & \Delta-\mathrm{RA}+\mathrm{DMOG} \end{aligned}$ | $+\mathrm{RA}$ <br> + RA + DMOG | $\begin{aligned} & -+ \text { RA } \\ & - \text { + RA + DMOG } \end{aligned}$ |
| :---: | :---: | :---: | :---: |
|  | 5 hmC | 5 hmC | 5 mC |

$$
\begin{array}{lll}
\text { MF4 } & \square \text { - RA } & \square+\mathrm{RA} \\
& \square-R A ~+~ D M O G ~ & \square+\mathrm{RA}+\mathrm{DMOG} \\
\square-++R A \\
& \square+D M O G
\end{array}
$$


















C Variables factor map (PCA)



G Variables factor map (PCA)


B Individuals factor map (PCA)


D Hierarchical clustering on the factor map

H Hierarchical clustering on the factor map



## C Variables factor map (PCA)




G Variables factor map (PCA)


B Individuals factor map (PCA)


D Hierarchical clustering on the factor map



H Hierarchical clustering on the factor map




MF3


DMOG

F2


5-azadC
DMOG


F4


F5


## 5hmC



H3K4me 1


5mC


H3K27ac


FAIRE






ChIP-MS_PBX1_mouse

| Identified Proteins | Accession Number | Molecular Weight | Sample | Exclusive <br> Unique <br> Peptide <br> Count | \% Coverage |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Pre-B-cell leukemia transcription factor 1 OS=Mus musculus $\mathrm{GN}=\mathrm{Pbx} 1 \mathrm{PE}=1 \mathrm{SV}=2$ | PBX1_MOUSE | 47 kDa | PBX1_rep1 | 19 | 48\% |
|  |  |  | PBX1_rep2 | 15 | 40\% |
|  |  |  | PBX1_rep3 | 20 | 51\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Pre-B-cell leukemia transcription factor 2 OS=Mus musculus $\mathrm{GN}=\mathrm{Pbx} 2 \mathrm{PE}=1 \mathrm{SV}=1$ | PBX2_MOUSE | 46 kDa | PBX1_rep1 | 14 | 62\% |
|  |  |  | PBX1_rep2 | 12 | 52\% |
|  |  |  | PBX1_rep3 | 15 | 63\% |
|  |  |  | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Meis1 OS=Mus musculus GN=Meis1 $P E=1 S V=1$ | MEIS1_MOUSE | 43 kDa | PBX1_rep1 | 15 | 54\% |
|  |  |  | PBX1_rep2 | 13 | 44\% |
|  |  |  | PBX1_rep3 | 14 | 49\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Pre-B-cell leukemia transcription factor 3 OS=Mus musculus $\mathrm{GN}=\mathrm{Pbx} 3 \mathrm{PE}=2 \mathrm{SV}=1$ | PBX3_MOUSE | 47 kDa | PBX1_rep1 | 8 | 50\% |
|  |  |  | PBX1_rep2 | 7 | 31\% |
|  |  |  | PBX1_rep3 | 9 | 50\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Hox-A5 OS=Mus musculus GN=Hoxa5$P E=1 S V=1$ | HXA5_MOUSE | 29 kDa | PBX1_rep1 | 10 | 49\% |
|  |  |  | PBX1_rep2 | 5 | 38\% |
|  |  |  | PBX1_rep3 | 10 | 49\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Meis2 OS=Mus musculus GN=Meis2 $P E=1 S V=2$ | MEIS2_MOUSE | 52 kDa | PBX1_rep1 | 4 | 25\% |
|  |  |  | PBX1_rep2 | 4 | 25\% |
|  |  |  | PBX1_rep3 | 5 | 30\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Chromodomain-helicase-DNAbinding protein 4 OS=Mus musculus GN=Chd4 $\mathrm{PE}=1 \mathrm{SV}=1$ | CHD4_MOUSE | 218 kDa | PBX1_rep1 | 7 | 5.7\% |
|  |  |  | PBX1_rep2 | 2 | 1.4\% |
|  |  |  | PBX1_rep3 | 6 | 3.8\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein PKNOX2 OS=Mus musculus GN=Pknox2 $P E=2 S V=1$ | PKNX2_MOUSE | 52 kDa | PBX1_rep1 | 2 | 3.6\% |
|  |  |  | PBX1_rep2 | 2 | 3.6\% |
|  |  |  | PBX1_rep3 | 3 | 11\% |
|  |  |  | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |


| Metastasis-associated protein MTA2 OS=Mus musculus GN=Mta2 PE=1 SV=1 | MTA2_MOUSE | 75 kDa | PBX1_rep1 | 3 | 4.9\% |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | PBX1_rep2 | 2 | 3\% |
|  |  |  | PBX1_rep3 | 4 | 9\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| FACT complex subunit SPT16 OS=Mus musculus GN=Supt16h $P E=1 S V=2$ | SP16H_MOUSE | 120 kDa | PBX1_rep1 | 3 | 3.6\% |
|  |  |  | PBX1_rep2 | 1 | 1.5\% |
|  |  |  | PBX1_rep3 | 3 | 3.6\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | lgG_rep2 | 0 | - |
|  |  |  | lgG_rep3 | 0 | - |
| Homeobox protein Hox-B5 OS=Mus musculus GN=Hoxb5$P E=2 S V=3$ | HXB5_MOUSE | 29 kDa | PBX1_rep1 | 3 | 25\% |
|  |  |  | PBX1_rep2 | 2 | 15\% |
|  |  |  | PBX1_rep3 | 3 | 25\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Metastasis-associated protein MTA1 OS=Mus musculus GN=Mta1 PE=1 SV=1 | MTA1_MOUSE | 81 kDa | PBX1_rep1 | 3 | 10\% |
|  |  |  | PBX1_rep2 | 3 | 7.3\% |
|  |  |  | PBX1_rep3 | 3 | 10\% |
|  |  |  | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Hox-B6 OS=Mus musculus GN=Hoxb6$P E=2 S V=2$ | HXB6_MOUSE | 25 kDa | PBX1_rep1 | 1 | 12\% |
|  |  |  | PBX1_rep2 | 1 | 8\% |
|  |  |  | PBX1_rep3 | 2 | 19\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | lgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Hox-B8 OS=Mus musculus GN=Hoxb8 $P E=2 S V=2$ | HXB8_MOUSE (+1) | 28 kDa | PBX1_rep1 | 0 | 3.3\% |
|  |  |  | PBX1_rep2 | 3 | 13\% |
|  |  |  | PBX1_rep3 | 1 | 8.2\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Histone deacetylase 1 OS=Mus musculus GN=Hdac1 PE=1 SV=1 | HDAC1_MOUSE | 55 kDa | PBX1_rep1 | 2 | 6.4\% |
|  |  |  | PBX1_rep2 | 0 | - |
|  |  |  | PBX1_rep3 | 2 | 6.4\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Hox-C5 OS=Mus musculus GN=Hoxc5$P E=2 S V=3$ | HXC5_MOUSE | 25 kDa | PBX1_rep1 | 1 | 21\% |
|  |  |  | PBX1_rep2 | 1 | 15\% |
|  |  |  | PBX1_rep3 | 2 | 27\% |
|  |  |  | IgG_rep1 | 0 | - |
|  |  |  | lgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Homeobox protein Hox-C4 OS=Mus musculus GN=Hoxc4$\mathrm{PE}=2 \mathrm{SV}=1$ | HXC4_MOUSE | 30 kDa | PBX1_rep1 | 2 | 16\% |
|  |  |  | PBX1_rep2 | 1 | 13\% |
|  |  |  | PBX1_rep3 | 1 | 2.7\% |
|  |  |  | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |


| FACT complex subunit SSRP1 OS=Mus musculus GN=Ssrp1 $P E=1 S V=2$ | SSRP1_MOUSE |  | PBX1_rep1 | 1 | 1.6\% |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | PBX1_rep2 | 1 | 1.6\% |
|  |  | 81 kDa | PBX1_rep3 | 1 | 1.6\% |
|  |  | 81 kDa | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | IgG_rep3 | 0 | - |
| Methyl-CpG-binding domain protein 3 OS=Mus musculus GN=Mbd3 PE=1 SV=1 | MBD3_MOUSE | 32 kDa | PBX1_rep1 | 0 | - |
|  |  |  | PBX1_rep2 | 1 | 5.3\% |
|  |  |  | PBX1_rep3 | 1 | 3.5\% |
|  |  |  | lgG_rep1 | 0 | - |
|  |  |  | IgG_rep2 | 0 | - |
|  |  |  | lgG_rep3 | 0 | - |

