

## **Supplementary Information**

### **Genome-wide DNA methylation profiling reveals epigenetic changes in the rat nucleus accumbens associated with cross-generational effects of adolescent THC exposure**

Corey T. Watson, Henrietta Szutorisz, Paras Garg, Qammarah Martin, Joseph. A. Landry, Andrew J. Sharp, Yasmin L. Hurd

#### **Supplementary Materials and Methods (pg. 2-3)**

Animals, **pg. 2**

Cross-generational THC paradigm, **pg. 2-3**

Analysis of DMRs in the context of LINE, SINE, and LTR repeat elements, **pg. 3**

Quantitative reverse transcription PCR (qRT-PCR) analysis, **pg. 3-4**

#### **Supplementary Figures (pg. 5-11)**

Figure S1, **pg. 5**

Figure S2, **pg. 6**

Figure S3, **pg. 7**

Figure S4, **pg. 8**

Figure S5, **pg. 9**

Figure S6, **pg. 10**

Figure S7, **pg. 11**

#### **Literature Cited (pg. 12)**

## SUPPLEMENTARY MATERIALS AND METHODS

### Animals

21-day-old male and female Long-Evans rats and lactating Long-Evans females with litter (to be used as surrogate mothers in breeding) were purchased from Charles River Laboratories, Inc. (Wilmington, Massachusetts). Care and handling procedures were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, approved by the Local Animal Care and Use Committee.

### Cross-generational THC paradigm

Male and female rats arrived at the facility at post-natal day (PND) 21 and they were housed in same sex groups. The animals were carried through our established adolescent THC treatment protocol (Ellgren *et al*, 2007; Tomaszewicz *et al*, 2012). Rats were exposed to either a moderate dose of THC (1.5 mg/kg i.p.) or VEH, one injection every third day during PND 28 to 49. VEH-exposed females were mated with VEH-exposed males and THC-exposed females with THC-exposed males in a ratio of 2 females to 1 male during PND64-68. Females were single housed throughout pregnancy. After birth at ~PND2, mixed litters were established combining an approximately equal number (12-14) of pups from THC- and VEH-exposed parents with a balanced proportion of males and females in each litter (Dinieri and Hurd, 2012). The litters were cross-fostered to drug-naïve surrogates, which were used as nursing mothers. F1 offspring were weaned at ~PND24 and groups of 3-4 animals were maintained without any drug treatment on normal 12-h light/dark cycle with *ad libitum* access to food and water. Adult (~PND62) animals were anaesthetized with CO<sub>2</sub>, decapitated, brains were frozen in isopentane, and stored at -80°C until subsequent experiments. Animal care and handling were performed by technicians unfamiliar with parental treatment and all

subsequent tissue processing was carried out in a randomized order with respect to treatment history.

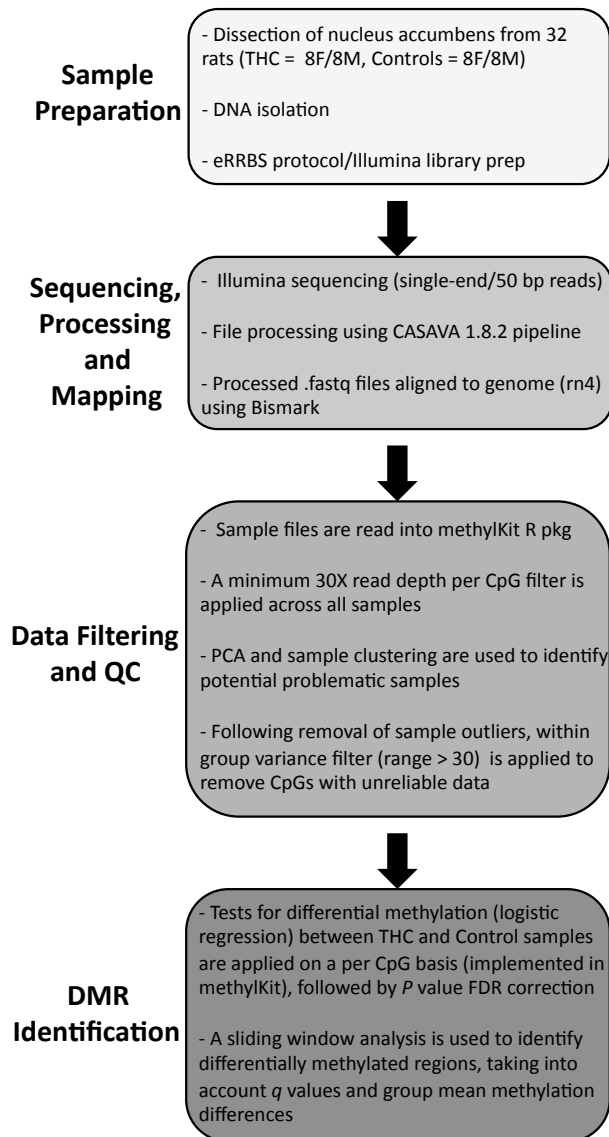
### **Analysis of DMRs in the context of LINE, SINE, and LTR repeat elements**

RepeatMasker content (Jurka, 2000; Smit *et al*, 1996-2010) was downloaded from the UCSC genome browser (Kent WJ, 2002) for the rn4 assembly. Only LTRs, SINEs, and LINEs were considered for this analysis. To test for an enrichment of repeats in DMRs, we first created a background clustered CpG coordinate set composed of only ERRBS CpGs that were neighbored by at least 2 other CpGs occurring within 500 bp (n = 484,461). The coordinates of individual CpGs in the background list were then merged into genomic coordinate sets, in which a single genomic coordinate comprised  $\geq 3$  CpGs. This clustered list was constructed to mirror the genomic characteristics of the characterized DMRs, which had to meet specific criteria for passing our sliding window analysis. The midpoint of each DMR and background region were extended by 5 kbp on each side to create two sets of 10 Kbp windows (DMR windows = 1,027; background windows = 48,503). These window sets were designed to maintain equal weighting when comparing repeat overlap between DMRs and the background list, in that genomic coordinates from either list were only counted once, and were of equal size. Repeat enrichment was then assessed by: (1) comparing the bp fraction of DMR 10 Kbp windows overlapped by repeats to that for the background windows; and (2) calculating the average counts of repeat subfamilies in windows and comparing between DMR and background coordinate sets.

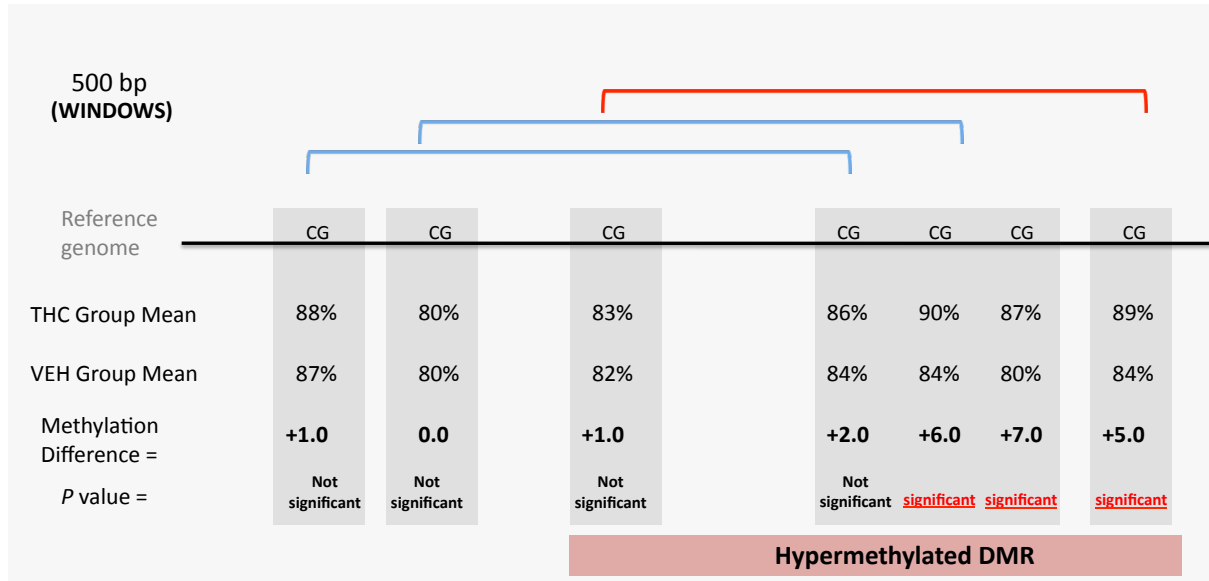
### **Quantitative reverse transcription PCR (qRT-PCR) analysis**

mRNA levels were measured in the NAc of a separate cohort of THC and VEH animals (n=8 males and 8 females/group). RNA was prepared from bilateral tissue punches

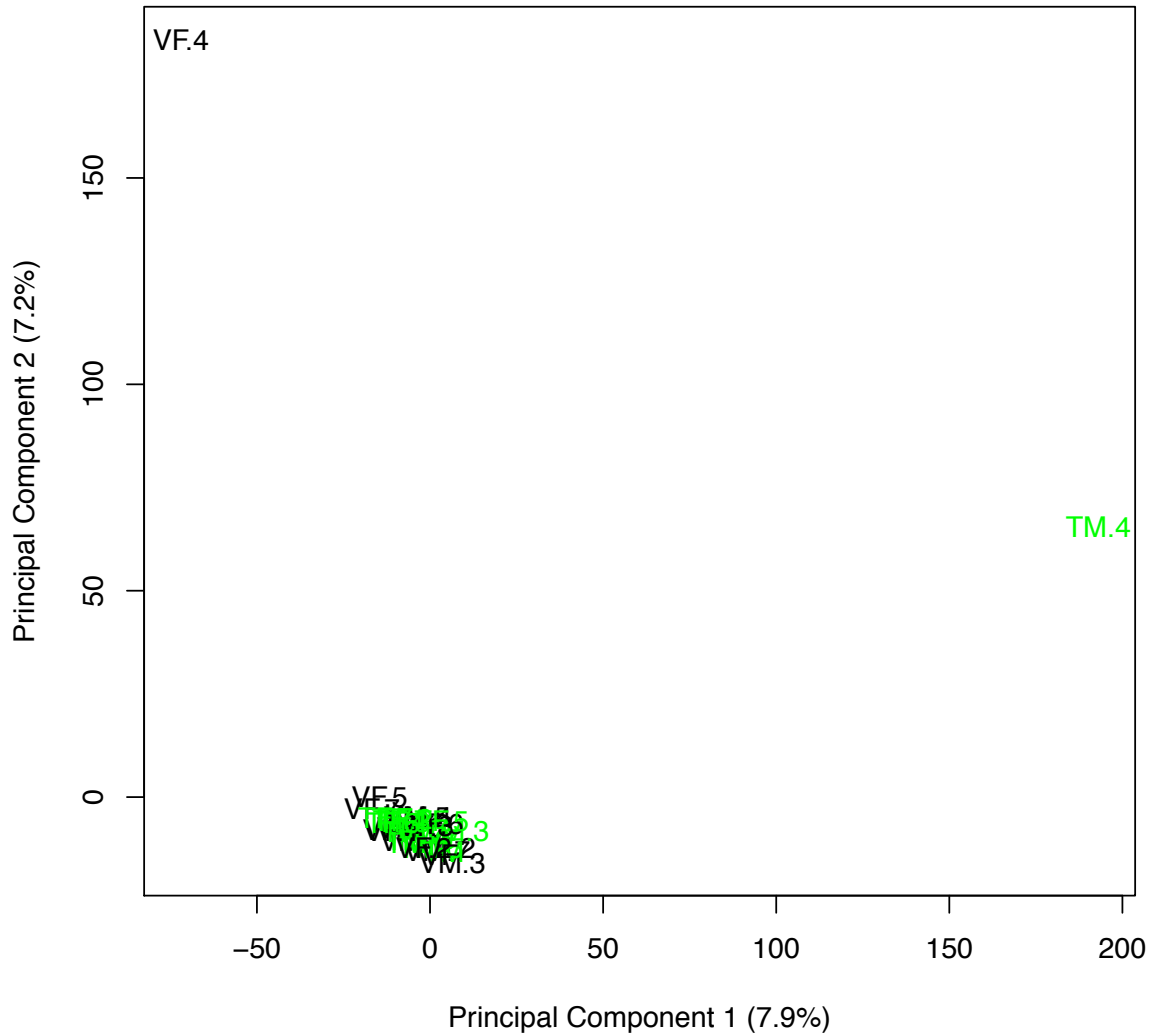
using the RNAqueous-Micro Kit (Ambion) and cDNA was obtained with a first-strand synthesis kit (Quanta Biosciences). Quantitative real-time PCR was performed using the LightCycler480 Probes Master reagent (Roche) and the TaqMan PCR program in a LightCycler 480 instrument (Roche). The following Taqman-based assays (Applied Biosystems) were used in triplicate PCR reactions: *Grin2A*, Rn00561341\_m1; *Cdh15*, Rn01432568\_m1; *Ntrk2*, Rn01441749\_m1; *Shank1*, Rn00582088\_m1; *Shank3*, Rn00572344\_m1; *Grik5*, Rn00568522\_m1; *Scn5a*, Rn00565502\_m1; *Dlgap2*, Rn00588099\_m1; *Dlgap3*, Rn00597609\_m1; *Mta1*, Rn00574899\_m1; *Kcnn1*, Rn00570904\_m1; *Begain*, Rn00577368\_m1; *beta-2-microglobulin*, Rn00560865\_m1). The data were then normalized using the  $\Delta \Delta$  CT method (Livak and Schmittgen, 2001) using *beta-2-microglobulin* as a reference gene. The Welch two-sample t-test was used to assess significance of differential expression.



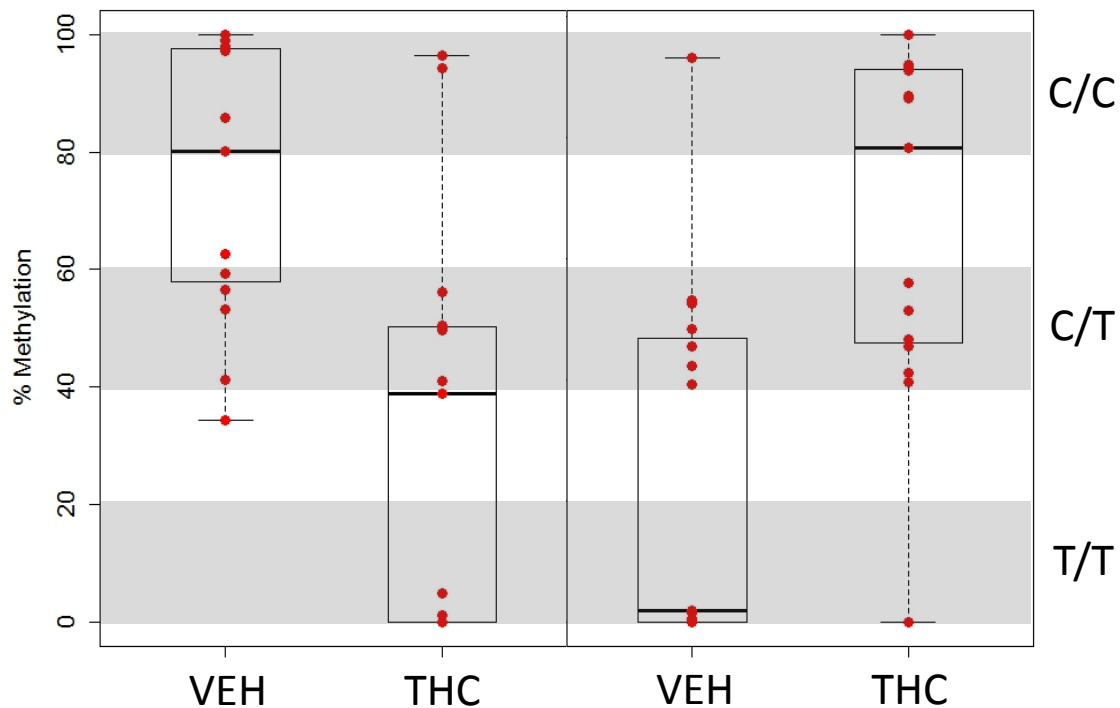
**Figure S1.** Overview of analysis pipeline from methods used for sample preparation to the identification of differentially methylated regions associated with cross-generational effects of THC exposure. Methods and approaches from each stage of analysis are outlined in brief.



**Figure S2.** Schematic of the sliding-window algorithm used to identify DMRs associated with effects of cross-generational THC exposure. A sliding window was moved along the genome, extending 500 bp out from each CpG in a forward direction, as depicted by blue and red brackets at the top of the figure. In each 500 bp window, several statistics were calculated for each CpG and used as criteria for defining DMRs. Specifically, we required DMRs to contain at least 3 significant CpGs ( $q < 0.01$ ) with concordant between group differences ( $> 2\%$ ) in mean methylation, and these CpGs had to represent at least 50% of CpGs within the window. In this example, the first 2 windows shown (blue brackets), each containing 4 CpGs, do not meet these criteria. The third window shown (red bracket) includes 5 CpGs, 3 of which are statistically significant, and are on average more highly methylated in the THC group, qualifying the window as a hypermethylated DMR. The genomic coordinates of the DMR are then delineated by the bp positions of the first and last CpGs in the window, including those that are not individually significant based on logistic regression.

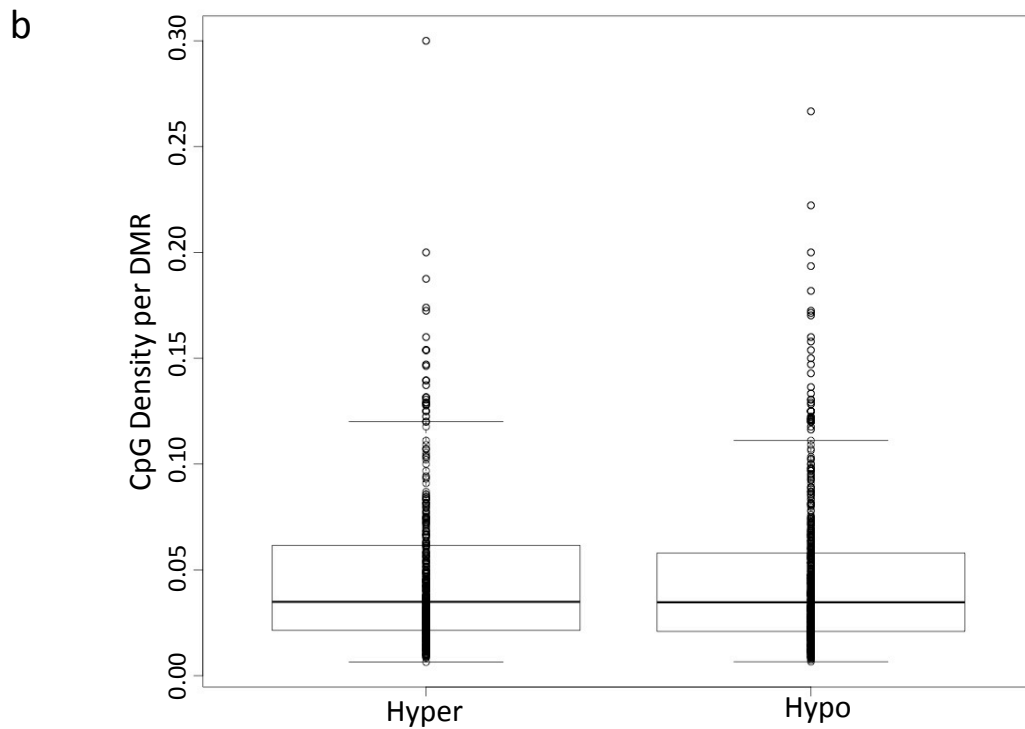
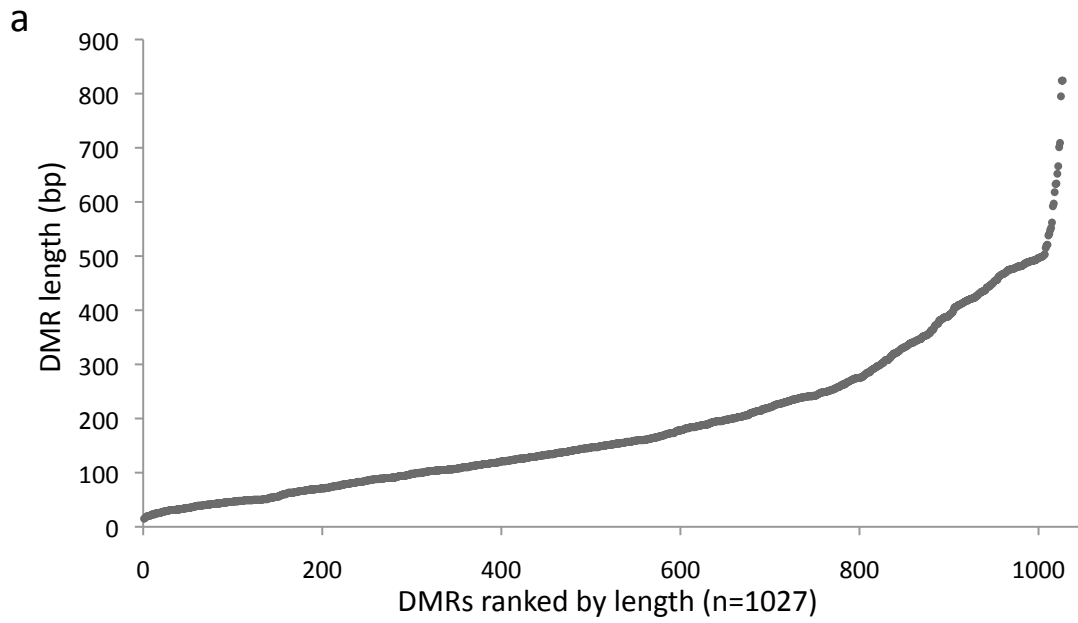


**Figure S3.** Principal component analysis of 16 THC (green) and 16 VEH (black) animals, using methylation values at 776,220 autosomal CpGs. Shown are the first and second principal components, accounting for 7.9% and 7.2% of the variance in the plotted dataset. Samples “VF.4” (VEH\_Female4) and “TM.4” (THC\_Male4) were identified as outliers and were not considered in subsequent analyses.

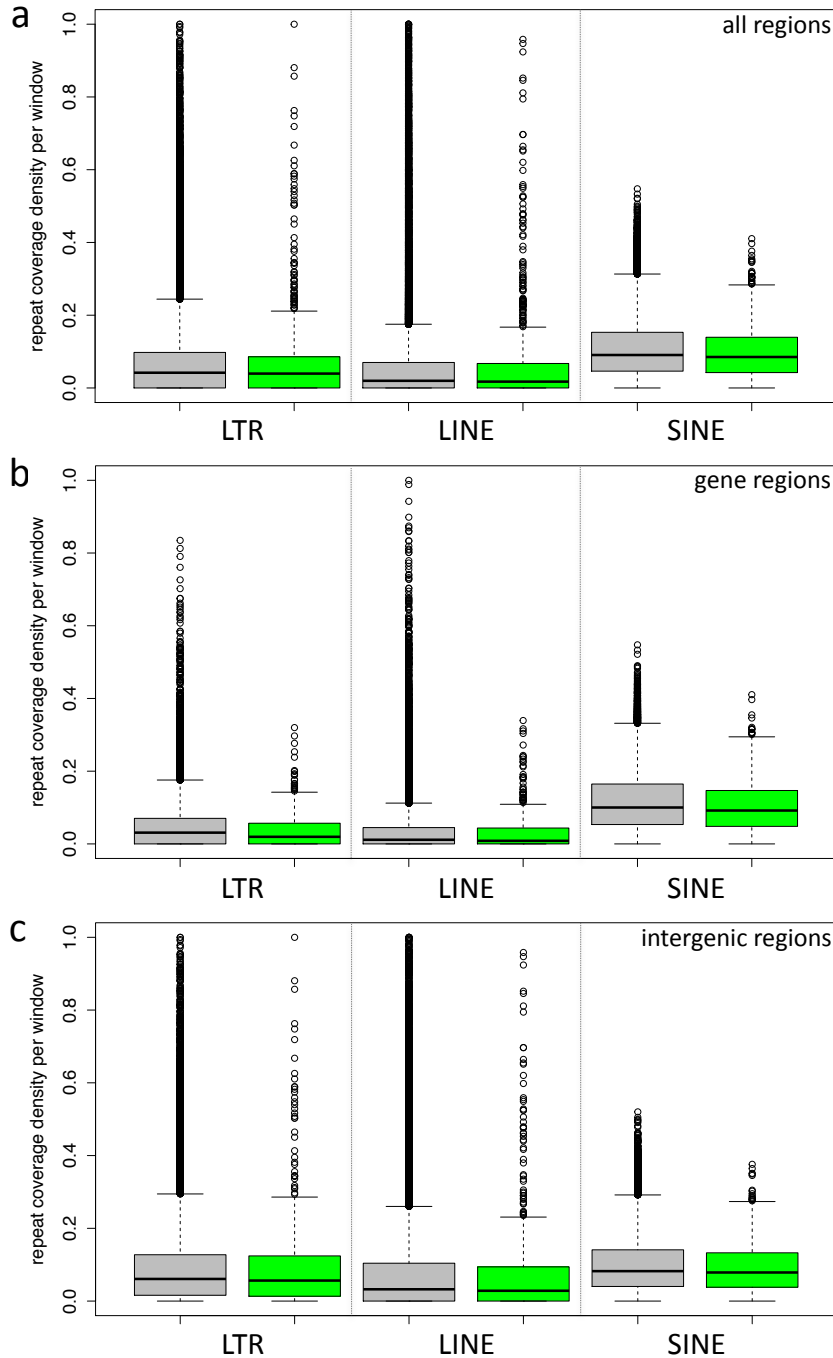


**Figure S4.** Examples of CpGs exhibiting extreme within-group variation, a signature consistent with underlying genetic polymorphism. Shown are boxplots of percent methylation values in THC and VEH groups for two CpGs (split by gray vertical line), with each individual (n=30) represented by a red dot. On the right are hypothetical DNA genotypes for a C/T single nucleotide polymorphism. For each CpG, both large within group ranges and between group differences in methylation were observed, as was a clear clustering of individuals around 100%, 50%, and 0% methylation highlighted by gray horizontal shaded bars. Both of the CpGs shown were significantly different between THC and VEH groups by logistic regression ( $p < 2.2 \times 10^{-16}$ ). However, it is likely that the observed signatures at these CpGs are at least in part associated with underlying genetic variation. Thus, we integrated a filter into our pipeline to remove sites that had high within group variance in order to limit the inclusion of potential false-positives driven by underlying inter-individual genetic differences.

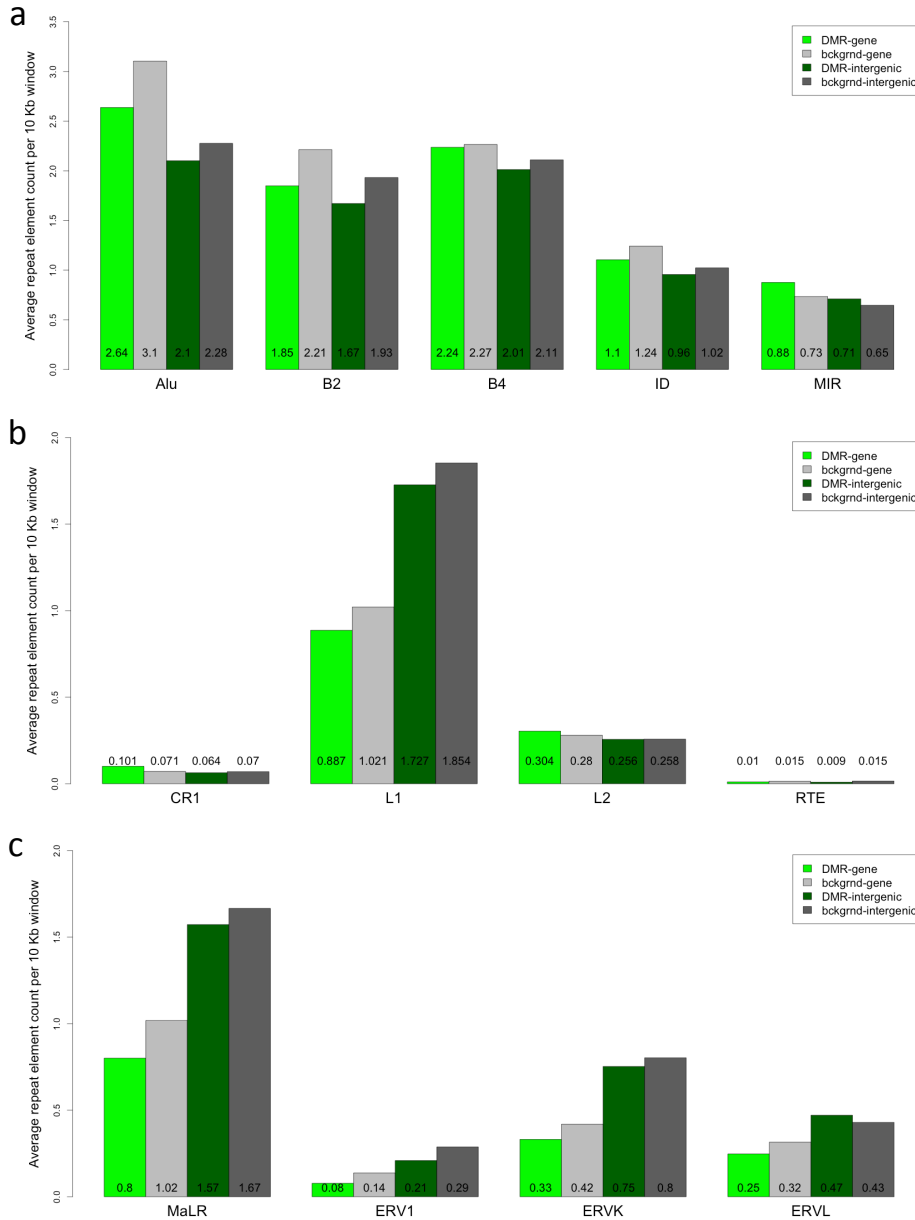




**Figure S5.** Basic features of characterized DMRs associated with effects of cross-generational THC exposure. (a) The lengths in bp of 1,027 DMRs. (b) Boxplot of CpG density (number of CpGs/DMR length) for hyper- and hypomethylated DMRs.



**Figure S6.** Assessment of the potential for preferential overlap of DMRs with various repeat element classes (LTRs, SINEs, and LINEs). Using the sets of 10 Kbp windows described above in the Supplementary Materials and Methods, we compared the bp fraction of DMR 10 Kbp windows ( $n=1,027$ ) overlapped by repeats to that for the background windows ( $n=484,461$ ). Boxplots show comparisons between DMR (green) and background (gray) windows for each repeat class considering (a) all windows genome-wide, (b) only windows overlapping genes, and (c) only windows falling within intergenic space. Outliers are shown as open circles in each boxplot.



**Figure S7.** Assessment of the increased occurrence of SINE, LINE, and LTR repeat subfamilies in regions surrounding DMRs in the rat genome (rn4). Using the sets of 10 Kbp windows described above in the Supplementary Materials and Methods, we calculated the average counts of repeat subfamilies in windows and compared these counts between DMR and background coordinate sets, again considering genic and intergenic windows separately. Bar charts are shown for average counts per window for (a) SINE subfamilies (Alu, B2, B4, ID, and MIR), (b) LINE subfamilies (CR1, L1, L2, and RTE), and (c) LTR subfamilies (MaLR, ERV1, ERVK, and ERVL). Average values are shown for each bar. Only five of the twenty-six comparisons revealed elevated average counts in DMR windows; SINE-MIR, LINE-CR1, and LINE-L2 elements, and SINE-MIR and LTR-ERVL element had higher counts in genic and intergenic DMR windows, respectively. However, all of these differences were modest.

## References

Dinieri JA, Hurd YL (2012). Rat models of prenatal and adolescent cannabis exposure. *Methods in molecular biology* **829**: 231-242.

Ellgren M, Spano SM, Hurd YL (2007). Adolescent cannabis exposure alters opiate intake and opioid limbic neuronal populations in adult rats. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **32**(3): 607-615.

Jurka J (2000). Repbase Update: a database and an electronic journal of repetitive elements. *Trends Genet* **16**(9): 418-420.

Kent WJ SC, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002). The human genome browser at UCSC. *Genome Res* **12**(6): 996-1006.

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**(4): 402-408.

Smit A, Hubley R, Green P (1996-2010). RepeatMasker Open-3.0.

Tomasiewicz HC, Jacobs MM, Wilkinson MB, Wilson SP, Nestler EJ, Hurd YL (2012). Proenkephalin mediates the enduring effects of adolescent cannabis exposure associated with adult opiate vulnerability. *Biological psychiatry* **72**(10): 803-810.