

UC Office of the President

Research Grants Program Office (RGPO) Funded Publications

Title

Epigenetic targeting of the Nanog pathway and signaling networks during chemical carcinogenesis

Permalink

<https://escholarship.org/uc/item/2bq7j3rw>

Journal

Carcinogenesis, 35(8)

ISSN

0143-3334

Authors

Tommasi, Stella

Zheng, Albert

Yoon, Jae-In

et al.

Publication Date

2014-08-01

DOI

10.1093/carcin/bgu026

Peer reviewed

Epigenetic targeting of the *Nanog* pathway and signaling networks during chemical carcinogenesis

Stella Tommasi, Albert Zheng, Jae-In Yoon¹ and Ahmad Besaratinia*

Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, M/C 9603, Los Angeles, CA 90033, USA and ¹QIAGEN, 1012 C Bldg, Woolim-Lions Valley, 371-28 Gasandong, Gumchungu, Seoul, Republic of Korea

*To whom correspondence should be addressed. Tel: +1 323 442 7251; Fax: +1 323 865 0103; Email: besarati@med.usc.edu

Chemical carcinogenesis has long been synonymous with genotoxicity, which entails DNA damage, genetic mutations and chromosomal abnormalities. The present study investigates a paradigm-shifting model in which epigenetic changes are key contributors to chemical carcinogenesis. Using genome-wide microarray-based analysis followed by conventional validation assays, we have progressively chronicled changes in the epigenetic landscape, as reflected in the patterns of DNA methylation, in the target organ of tumorigenesis in mice treated *in vivo* with a prototype chemical carcinogen (benzo[a]pyrene). Here, we demonstrate characteristic CpG island gain/loss of methylation and demethylation of repetitive DNA elements in carcinogen-treated mice, dependent on tumor progression. Alterations of the DNA methylome are accompanied by silencing of major DNA methyltransferases. Members of the *Nanog* pathway that establishes and maintains pluripotency in embryonic stem cells and possibly triggers uncontrolled proliferation of neoplastic cells are preferential targets of aberrant DNA methylation and concomitant gene dysregulation during chemical carcinogenesis. Several components of the *MEK/ERK*, *JAK/STAT3*, *PI3K/AKT*, *WNT/β-catenin* and *Shh* signaling cascades, which are known to modulate *Nanog* expression, also show concurrent changes in the patterns of DNA methylation and gene expression. Our data support an epigenetic model of chemical carcinogenesis and suggest that surveillance of the epigenetic landscape, particularly at the loci and in the pathways identified in this study, may have utility for early detection and monitoring of the progression of malignancy.

Introduction

Chemical carcinogenesis has long been ascribed to genotoxic events, which entail DNA damage, genetic mutations and chromosomal abnormalities (1–3). In addition to genotoxicity, chemical carcinogens may also possess a non-genotoxic mode of action, e.g. epigenetic effects, which are emerging as key contributors to carcinogenesis (4–6). Epigenetic effects include aberrant DNA methylation, histone modifications and variants, miRNAs dysregulation, chromatin remodeling and nucleosome positioning (7–9). Of these, aberrant DNA methylation is the most extensively studied epigenetic alteration in carcinogenesis (7,8). Gain of methylation (hypermethylation) at CpG islands, clustered at the promoter, 5'-untranslated region and exon 1 of known genes (promoter CpG islands) or localized within gene bodies (intragenic CpG islands) is a common event in cancer

Abbreviations: B[a]P, benzo[a]pyrene; BW, body weight; COBRA, combined bisulfite restriction analysis; DMSO, dimethyl sulfoxide; DNMT, DNA methyltransferase; ESC, embryonic stem cell; H&E, hematoxylin and eosin; IAP, intracisternal A particle; IPA, Ingenuity Pathway Analysis; LINE, long-interspersed nuclear element; LTR, long terminal repeat; MBD, methyl-CpG-binding domain; MIRA, methylated CpG island recovery assay; qRT-PCR, quantitative real-time PCR; SINE, short interspersed nuclear element.

(9–11). Global loss of methylation (hypomethylation) at repetitive DNA elements, such as long and short interspersed nuclear elements (LINEs and SINEs, respectively), and long terminal repeat (LTR) retrotransposons is also a frequent occurrence in carcinogenesis (12–14).

Although human population studies and animal model experiments have established an association between exposure to chemical carcinogens and epigenetic effects, a direct cause and effect relationship has yet to be established (4,5). Suggestive evidence exists in support of a causal link between epigenetic effects and chemical carcinogenesis. For example, the reactive metabolite of a prototypical chemical carcinogen, benzo[a]pyrene (B[a]P) (15), binds preferentially to methylated CpG sites (16), thereby, possibly impeding the establishment and/or maintenance of DNA methylation patterns by DNA methyltransferases (DNMTs) and methyl-CpG-binding domain (MBDs) proteins (17–19). It is also plausible that carcinogen-induced DNA damage may cause mutations in genes maintaining the epigenetic state, including those regulating the DNA methylation machinery and chromatin-modifying enzymes (20,21).

The objectives of the present study were 2-fold: (i) to investigate whether epigenetic effects occur in response to exposure to chemical carcinogens and, if so, (ii) to determine whether epigenetic effects are predictors of chemically induced carcinogenesis. To attain these objectives, we have performed genome-wide DNA methylation analysis in the target organ of tumorigenesis, i.e. seminal vesicles, in mice treated *in vivo* with B[a]P both before and after tumor development. For verification purposes, we have confirmed the data obtained by our genome-wide microarray-based analysis (22) using the conventional single-gene methylation detection assays (23,24). Furthermore, we have used a bisulfite sequencing-based assay (25) to determine the methylation status of major repetitive DNA elements, including LINE L1, intracisternal A particle (IAP) LTR and SINE B1 (26–28), in B[a]P-treated mice both before and after tumor development. Here, we demonstrate a relationship between carcinogen exposure, alterations of the epigenome and gene expression changes, particularly in the *Nanog* and its interconnected signaling pathways. Our data indicate that carcinogen exposure results in epigenetic modifications, specifically at loci that control key signaling pathways required for normal cell proliferation during development.

Materials and methods

B[a]P treatment of mice

Thirty adult male mice (6–8 weeks old) on a C57BL/6 genetic background (Stratagene, La Jolla, CA) were randomly divided into two groups: (i) experimental (B[a]P treatment; $n = 15$) and (ii) control (solvent treatment; $n = 15$), each subdivided into three categories, including (i) 6 week treatment (T0), (ii) 6 week treatment + 6 week latency (T1: early lesion formation) and (iii) 6 week treatment + ≥ 10 week latency (T2: tumor development). A flowchart of the study design is shown in Figure 1A. The mice assigned to each experimental or control group were kept in polypropylene cages in groups of two to three animals per cage and housed in an air-conditioned animal room with controlled ambient temperature, relative humidity and 12 h light/dark cycle. The mice had access to food (PicoLab Rodent Diet 20; PMI Nutrition International, LLC; Brentwood, MO) and water *ad libitum* at all times. All experiments were approved by the Institutional Animal Care and Use Committee in accordance with the recommendations of the National Institutes of Health provided in the Guide for the Care and Use of Laboratory Animals.

The experimental mice received intraperitoneal injections of B[a]P once per week for a duration of 6 weeks using the following protocol: first week: 25 mg/kg body weight (BW); second week: 50 mg/kg BW; third week: 75 mg/kg BW and fourth week to sixth week: 100 mg/kg BW of B[a]P. The specified doses of B[a]P were prepared fresh on the day of administration by dissolving the chemical in dimethyl sulfoxide (DMSO) (B[a]P and DMSO; Sigma–Aldrich, St Louis, MO). The incremental doses of B[a]P were delivered to the mice by intraperitoneal injection (100 μ l) on the lower right or left quadrant of the

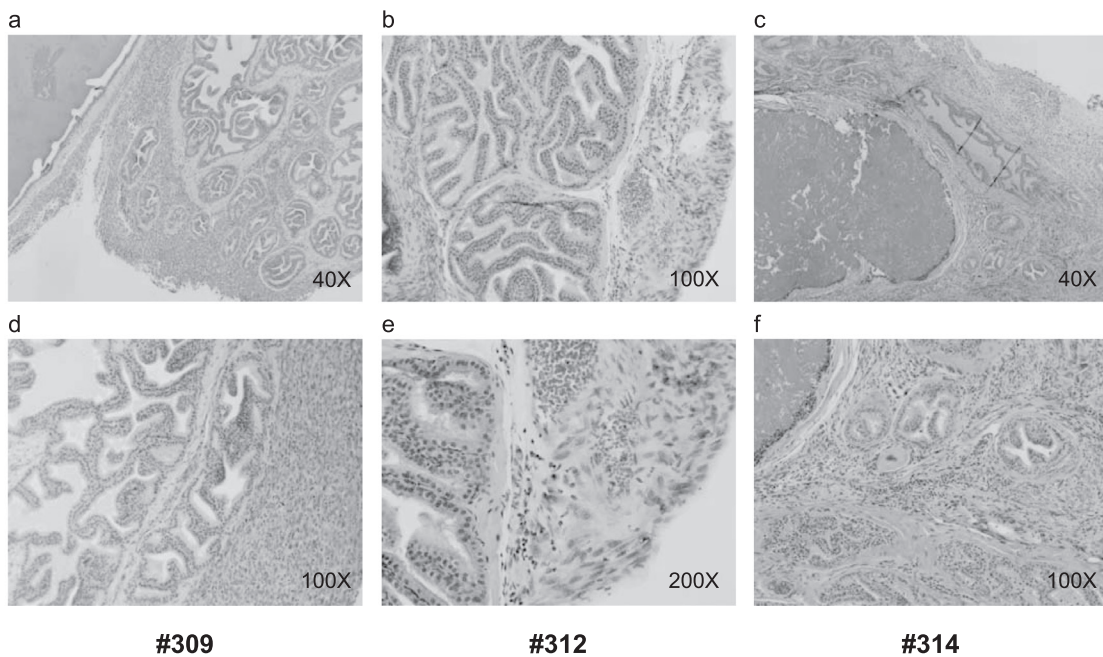
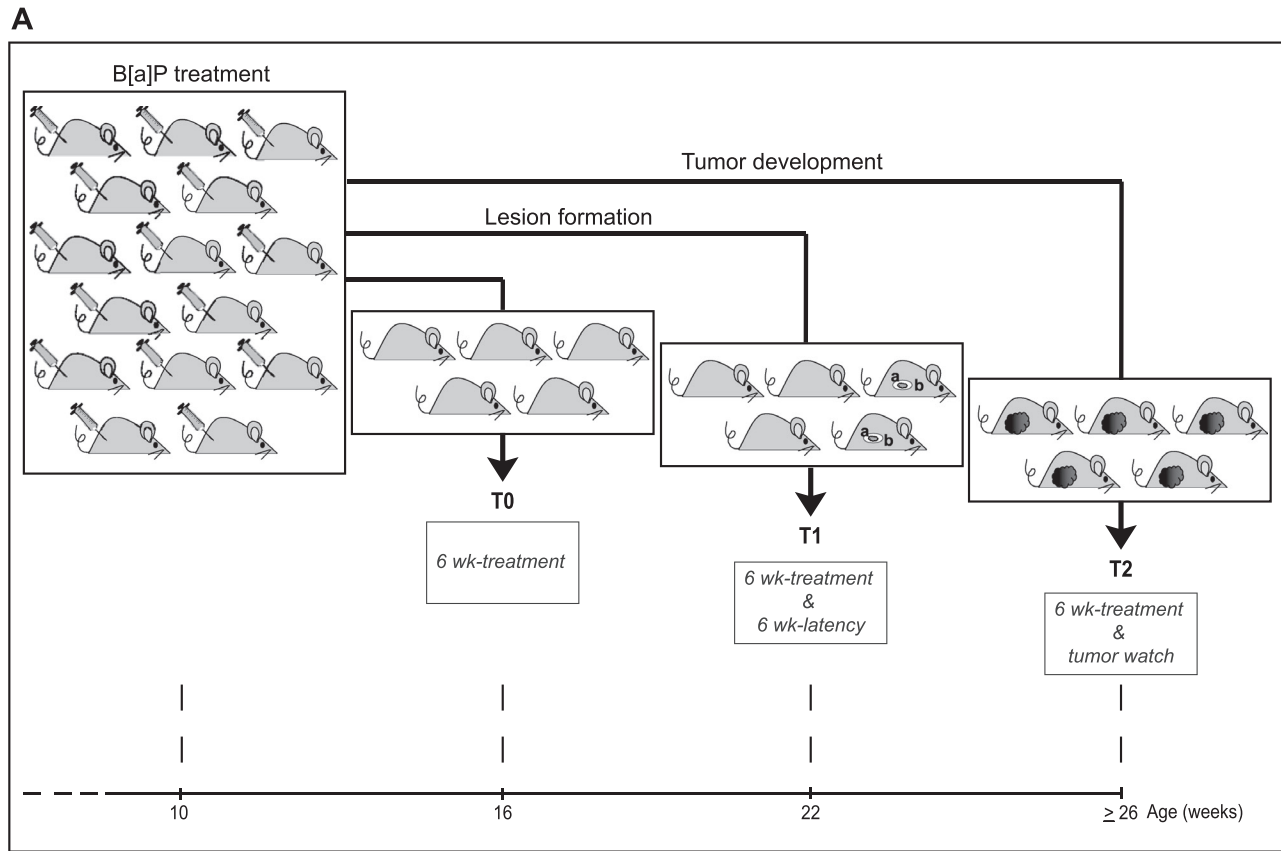


Fig. 1. Outline of the experimental design and histological analysis of tumors in B[a]P-treated mice. **(A)** Adult male mice were treated intraperitoneally with weekly doses of B[a]P over a period of 6 weeks. Subgroups of animals were euthanized immediately after the last B[a]P injection and at ensuing intervals posttreatment. Subsequently, the target organ of tumorigenesis (seminal vesicles) were collected for histological and biochemical analyses. Mice treated with solvent DMSO were used as control (a) precursor lesions; (b) adjacent normal tissues. **(B)** Tumors developed in the seminal vesicles of B[a]P-treated mice were collected in formalin and processed for H&E staining. Sections were photographed at lower (panels a–c) and higher (panels d and e) power magnifications.

abdomen in alternate weeks. Control mice received similar injections of solvent DMSO using the same dosing regimen, as described for B[a]P. All mice were monitored closely for development of any unusual symptoms. At the end

of all experiments, the B[a]P-treated mice and controls were euthanized by CO₂ asphyxiation and subjected to necropsy and macroscopic examination. For biochemical assays, the accessory sex organs, including seminal vesicles

and prostate glands or tumors, were harvested, snap frozen and preserved at -80°C until further analysis. Alternatively, the harvested tissues or tumors were fixed in formalin, embedded in paraffin and used for hematoxylin and eosin (H&E) slide preparation according to standard protocols.

Genome-wide DNA methylation profiling

We used the methylated CpG island recovery assay (MIRA) in combination with microarray analysis (22) to catalogue the DNA methylation profile, on a genome-wide scale, in the target organ of tumorigenesis, i.e. seminal vesicles, in mice treated *in vivo* with B[a]P both before and after tumor development. As a pull-down assay for enrichment of the methylated CpG content of DNA, the MIRA is based on the ability of the MBD2b protein to bind methylated CpG dinucleotides, while this reaction is enhanced in the presence of MBD3L1 protein (22). Briefly, genomic DNA was isolated from either seminal vesicles or tumors developed at the same organ site and subjected to MIRA enrichment, as described previously (22). Subsequently, the MIRA-enriched DNA and input DNA fractions were amplified by PCR and labeled and hybridized to the Roche NimbleGen Mouse DNA Methylation 3x720K CpG Island Plus RefSeq Promoter Arrays (Roche NimbleGen, Indianapolis, IN). This set of microarrays covers 20 404 RefSeq gene promoters and 15 988 annotated CpG islands of the mouse genome. The raw microarray data were processed and analyzed using a standard bioinformatics approach, as described in ref. (22). The microarray data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus database (Accession No: GSE41422).

Single-gene DNA methylation analysis by combined bisulfite restriction analysis and bisulfite sequencing

We used the combined bisulfite restriction analysis (COBRA) (24) and sodium bisulfite sequencing (23) to verify the methylation status of individual target loci/genes identified by MIRA-microarray analysis in the target organ of tumorigenesis in B[a]P-treated mice both before and after tumor development. Briefly, 1 μg of genomic DNA was treated with sodium bisulfite using the Qiagen EpiTect kit according to the manufacturer's instructions (Qiagen, Valencia, CA). Subsequently, the bisulfite-treated DNA was assayed by standard COBRA (24) using sets of primers specifically designed for each target CpG island. The primer sequences used for PCR amplification of all the analyzed targets are available upon request. Mouse genomic DNA was methylated *in vitro* with M. SssI CpG methyltransferase (New England Biolabs, Ipswich, MA) and served as positive control. For genomic sequencing, the PCR products obtained after bisulfite conversion of DNA were cloned into the TOPO-TA cloning vector according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). Up to 10 randomly selected clones from each experimental and control group were sequenced using an ABI-3730 DNA Sequencer (ABI Prism; PE Applied Biosystems, Foster City, CA).

DNA methylation analysis in repetitive DNA elements

We used a bisulfite sequencing-based assay (25) to analyze the methylation status of major repetitive DNA elements, including LINE L1, IAP-LTR and SINE B1, in the target organ of tumorigenesis in B[a]P-treated mice both before and after tumor development. The LINE L1, IAP-LTR and SINE B1 comprise 18.78, 3.13 and 2.66%, respectively, of the mouse genome (26–28). The assay is based on sodium bisulfite treatment of the genomic DNA, followed by primer amplification of the consensus sequences of the respective elements and direct sequencing, thereafter (25). Detailed information on the genomic sequence and primer design of the repetitive DNA elements analyzed in the present study is available in ref. (22).

Motif discovery and canonical pathway analysis

Differentially methylated CpG islands were analyzed for shared *de novo* motifs and known transcription factor recognition sites, using Partek® Genomics Suite™ software (Partek® 6.6). Novel motifs of length 6–16bp were identified based on their highest score (score = log ratio of the probability that the sequence was generated by the motif versus that by background distribution). Binding sites for known transcription factors were selected from the JASPAR database in Partek based on their high probability of occurrence (*P* value). Functional identification of gene networks and canonical pathways analysis were performed using the Ingenuity Pathway Analysis® program (IPA®: v 9.0) and the GO Enrichment Analysis in Partek® Genomics Suite™.

Quantitative real-time PCR

Standard quantitative real-time reverse transcription-PCR (qRT-PCR) was used to determine the level of transcription of individual target genes identified by MIRA-microarray analysis, as described previously (22). Briefly, total RNA was extracted from the seminal vesicles of mice treated *in vivo* with B[a]P both before and after tumor development, using the RNeasy kit (Qiagen). DNase-treated RNA (0.5 μg) was reverse transcribed into cDNA using SuperScript® VILO™ cDNA Synthesis kit (Invitrogen). The mRNA expression level of

target genes was determined by qRT-PCR using the EXPRESS SYBR® GreenER™ qPCR SuperMix (Invitrogen) and the CFX96 Touch™ Real-Time PCR detection system (Bio-Rad). All reactions were performed in triplicate and fold changes were determined using the $2^{-\Delta\Delta\text{C}_t}$ method (22). The primer sets used for qRT-PCR are listed in Supplementary Table S1, available at Carcinogenesis Online.

Results

Mice survival and tumorigenicity

We have treated adult male mice with progressively increasing doses of B[a]P weekly for a duration of 6 weeks, as described in Materials and methods. For control purposes, we have treated counterpart mice with solvent DMSO using the same protocol as described for B[a]P. Subgroups of animals were euthanized immediately after treatment (T0), 6 week posttreatment (T1) or after tumor development (T2), as outlined in Figure 1A. The mice generally well tolerated B[a]P treatment and had a survival rate of >75%, as determined in two independent sets of experiments. All surviving B[a]P-treated mice developed large aggressive tumors in the seminal vesicles as early as 10 week posttreatment. Small lesions and other abnormalities were also detectable in the secondary sex organs of B[a]P-treated mice at T1. Control mice treated with DMSO showed 100% survival at the end of all experiments and were lesion free at all organ sites.

Histological analysis of B[a]P-induced tumors

The accessory sex organs, including the target organ of tumorigenesis, i.e. seminal vesicles, were collected from all B[a]P-treated and control mice at the time of necropsy, fixed in formalin and subjected to H&E staining. Consistent with landmark studies (29), all surviving B[a]P-treated mice developed large solid masses (Δ : $2.5 \times 2 \times 2$ cm) in the seminal vesicles and prostate glands. Histopathological examination identified the tumors as high-grade sarcomas, mostly leiomyosarcoma. Prominent features of the tumors included characteristic spindle-like cells, abundance of apoptotic cells and mitotic figures, as well as areas of pleomorphism and infiltration (Figure 1B, panels a–f). One mouse from group T0 and all mice from group T1 had macroscopic abnormalities in the genitourinary tract. These aberrations often consisted of (uni)-testicular atrophy and/or hyperplasia of seminal vesicles/prostate (mostly on the left side), which are likely to have preceded the tumors developed in older mice. Hyperplasia and other abnormalities were also observed sporadically in the digestive tract of several mice at the end of treatment (T0), and thereafter (T1 and T2). One mouse from group T2 (#313) also developed an aggressive sarcoma on the intestinal wall. Representative H&E sections of tissues/tumors collected from B[a]P-treated mice are shown in Figure 1B, panels a–f.

DNA methylation profiling during B[a]P-induced tumorigenesis

We have used a genome-wide microarray-based approach (22) to detect aberrant DNA methylation in the target organ of tumorigenesis, i.e. seminal vesicles, in mice treated *in vivo* with B[a]P immediately after treatment (T0) and following tumor development (T2). For control purposes, we performed similar analysis on counterpart organs from solvent-treated mice. Due to small size of lesions in mice at T1, the seminal vesicles of these animals were excluded from the analysis. Briefly, genomic DNA was isolated from either seminal vesicles or tumors developed at the same organ site and subjected to MIRA enrichment, as described previously (22). Following pull-down of the methylated CpG islands by MIRA, the enriched- and input DNA fractions were labeled, mixed and hybridized to the mouse CpG island plus promoter tiling arrays (Roche NimbleGen). The microarray data were normalized and analyzed using rigorous algorithms for peak calling, as described in ref. (22). Two independent lists of differentially methylated targets were generated based on the level of stringency used for data analysis, as follows: (i) stringent: methylation differences detected in four out of four biological replicates within a treatment group and absent in all respective controls and (ii) relaxed:

methylation differences detected in three out of four biological replicates within a treatment group and absent in all respective controls. When the stringent criteria were used for analysis, we identified 372 CpG targets that displayed aberrant methylation in tumors from B[a]P-treated animals relative to controls (251 hypermethylated and 121 hypomethylated CpG islands) (Figure 2A). A lower yet significant number of differentially methylated CpG islands was detected in the seminal vesicles of apparently asymptomatic mice treated with B[a]P at T0 (immediately after treatment), i.e. 123 aberrantly methylated targets, of which 60 were hypermethylated and 63 were hypomethylated (Figure 2A). Seventy-two percent of these 123 CpG islands coincide with those identified in tumors (T2), suggesting that aberrant DNA methylation in these targets occurs early in the process of tumorigenesis. Principal component analysis of the differentially methylated CpG islands confirmed that all samples within a treatment group clustered together, with a significant overlap between the T0 and T2 groups (Figure 2B and C). In tumors (T2), hypermethylated CpG islands map mostly within annotated genes (41%), whereas loss of methylation occurs predominantly at upstream regulatory regions (77%) (Figure 2D). Likewise, a characteristic pattern for genomic distribution of the aberrantly methylated targets was found in seminal vesicles of mice at T0, with 46% of hypermethylated CpG islands being intragenic and 76% of hypomethylated CpG islands mapping

to the 5'-end of annotated genes (Figure 2D). The latter reaffirms that aberrant DNA methylation patterns are established immediately after carcinogen exposure, at a time preceding lesion formation (Figure 2D). We also observed that aberrant DNA methylation preferentially targets genes that are normally associated with either active histone mark (H3K4me3) or bivalent marks (H3K4me3 and H3K27me3) in murine embryonic stem cells (ESCs), as established by comparison of our data to published databases (30) (Figure 2E).

Validation of differentially methylated targets by COBRA and bisulfite sequencing

To validate the DNA methylation microarray results, we have randomly selected and analyzed several targets identified by MIRA-microarray analysis using the conventional COBRA (24) and bisulfite DNA sequencing (23). Representative COBRA results for hyper- and hypomethylated targets are shown in Figure 3A and B (upper panels) and Supplementary Figure S1, available at *Carcinogenesis* Online. In all cases, we confirmed DNA methylation differences between control and experimental samples collected at different time points posttreatment, which mirrored the differences observed in the array analysis. These confirmatory results indicate that epigenetic marks, i.e. CpG methylation, are established or lost during B[a]P-induced

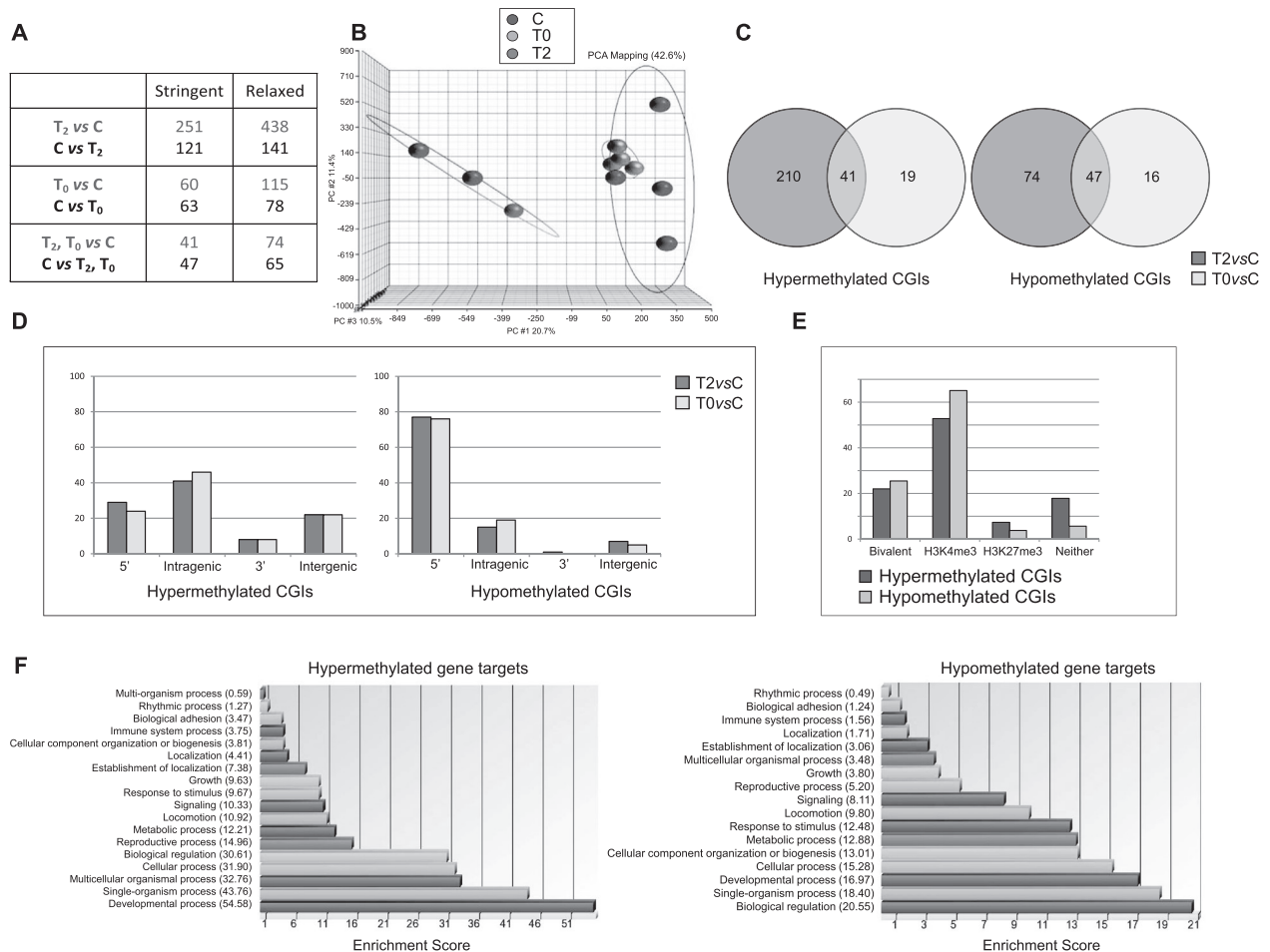


Fig. 2. Identification of differentially methylated targets in B[a]P-treated mice. (A) Aberrantly methylated CpG Islands were identified by MIRA-microarray analysis. Two independent lists of differentially methylated targets were generated based on the level of stringency used for data analysis, as described in the text. The minimum difference between groups was ≥ 2 -fold. The numbers of methylation peaks within each dataset are shown (hypermethylated CpG Islands are in red). (B) Principal component analysis (PCA) of methylation data obtained in B[a]P-treated mice at T0 and T2 and control (C) using the Partek® software. (C) Venn diagrams of aberrantly methylated targets identified by MIRA-microarray analysis in B[a]P-treated mice at T0 (yellow) and T2 (blue) versus control. (D) Genomic localization of aberrantly methylated targets identified by MIRA-microarray analysis. (E) Colocalization of aberrantly methylated targets identified by MIRA-microarray analysis with histone marks established in ref. (30). (F) Gene ontology analysis by Partek® of relevant biological groups associated with hyper- and hypomethylated gene targets, respectively. Enrichment score: negative log of the enrichment *P* value. A high score indicates that the genes of a functional group are overrepresented in the gene list.

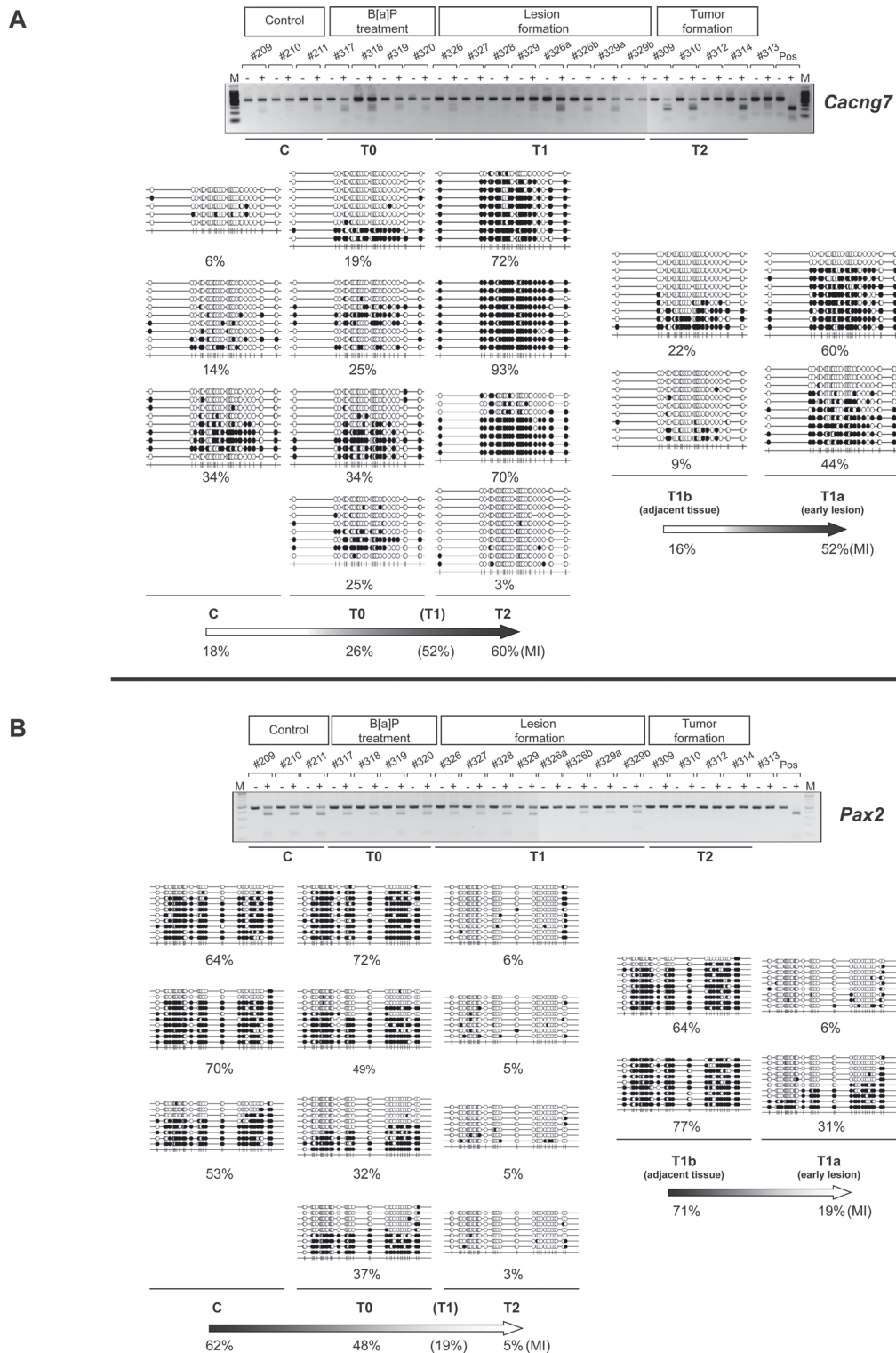


Fig. 3. Representative gene targets identified by MIRA-microarray analysis. **(A)** Top: genomic DNA from B[a]P-treated and control mice was treated with sodium bisulfite, and the hypermethylated CpG island downstream of the *Cacng7* gene was amplified with gene-specific primers and subjected to COBRA. Digested fragments on the gel are indicative of methylated restriction sites within the CpG island. *In vitro* methylated mouse genomic DNA served as positive control (Pos). The symbols (+) and (–) show the presence and absence, respectively, of the restriction enzyme in reaction mix. M = 100bp ladder DNA marker. Bottom: the extent of CpG methylation in the *Cacng7*-associated CpG island was determined by sodium bisulfite sequencing in B[a]P-treated mice at T0 and T2 and control (C). The methylation status of precursor lesions (a) and adjacent normal tissues (b) in B[a]P-treated mice at T1 was also compared. The sequencing results of up to 10 independent clones and the respective percentage of methylation per sample are shown. Open and closed circles represent unmethylated and methylated CpG dinucleotides, respectively. MI = methylation index. **(B)** Top: COBRA analysis of the hypomethylated CpG island located downstream of the *Pax2* gene in B[a]P-treated mice. Digested fragments on the gel are indicative of methylated restriction sites within the CpG island. Bottom: the extent of CpG methylation within the *Pax2* CpG island was determined by sodium bisulfite sequencing in B[a]P-treated mice at T0 and T2 and control (C). The methylation status of precursor lesions (a) and adjacent normal tissues (b) in B[a]P-treated mice at T1 was also compared. The sequencing results of several independent clones and the respective percentage of methylation per sample are shown.

tumorigenesis. One of such targets is the *Cacng7* gene, whose product plays a role in the synaptic expression of cerebellar alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid-type glutamate receptors (31). A normally unmethylated CpG island located downstream of the murine *Cacng7* gene becomes increasingly methylated immediately after treatment with B[a]P (T0), and thereafter (T1 and T2) (Figure 3A, top). A detailed methylation map for each CpG within the *Cacng7* CpG island was also constructed using the bisulfite sequencing technique (Figure 3A, bottom). A steady increase in the median percentage of CpG methylation at this locus was detectable in B[a]P-treated mice, as reflected in the methylation indices of 18, 26, 52 and 60% for samples of control, T0, T1 and T2 mice, respectively. Despite the substantial methylation increase in the *Cacng7* gene in precursor lesions of B[a]P-treated mice at T1, the methylation indices of this gene in the adjacent normal appearing tissues were comparable with that in solvent-treated controls. We have also verified B[a]P treatment-associated loss of methylation at several loci using both the COBRA and bisulfite sequencing. Figure 3B combines the COBRA and bisulfite sequencing results for a CpG island located 3'-end to the murine *Pax2*, a key gene involved in normal prostate development (32). DNA methylation at this downstream CpG island decreases immediately after the cessation of B[a]P treatment (from 62% in control animals to 48% in T0 mice to 19% in T1 mice) and is virtually lost in tumor-bearing animals at T2 (5%) (Figure 3B, bottom). Notwithstanding the loss of methylation in *Pax2* in early lesions of B[a]P-treated mice at T1, the surrounding apparently normal tissues had comparable methylation indices to that of DMSO-treated controls (Figure 3B, bottom). We have also detected agglomerate loss of methylation within gene families, e.g. the *Hoxa* and *Hoxb* homeobox gene clusters, a phenomenon commonly found in various types of human cancer (Supplementary Figure S2, available at *Carcinogenesis* Online) (33). Supplementary Table S2, available at *Carcinogenesis* Online, lists the hypermethylated and hypomethylated CpG islands identified in samples of mice at T0 and T2.

Methylation profiling of major repetitive DNA elements during B[a]P-induced tumorigenesis

To determine whether *in vivo* treatment of mice with B[a]P can induce global hypomethylation events, we have analyzed the methylation status of major repetitive DNA elements in the target organ of tumorigenesis in B[a]P-treated mice both before and after tumor development. Using a bisulfite sequencing-based approach (25), we determined the status of CpG methylation in LINE L1, IAP-LTR and SINE B1, which are routinely used as surrogate markers to estimate the overall DNA methylation level in the mouse genome (26–28). As shown in Supplementary Figure S3, available at *Carcinogenesis* Online, there was a significant reduction in CpG methylation in the IAP-LTR between experimental mice and controls ($P = 0.02$). However, no appreciable differences in CpG methylation were observed in the SINE B1 and LINE L1 between experimental and control mice.

Motif discovery

We next used the Partek® Genomics Suite™ Software (v 6.6) to search for common non-redundant sequence instances across the differentially methylated CpG islands in samples of B[a]P-treated mice at T0 and T2. The top-scored *de novo* motifs, identified by Partek analysis, are illustrated in Figure 4A and show an overall prevalence of purine residues (mostly guanines, followed by adenines), often in the context of CpG dinucleotides. Of note, B[a]P epoxy diols (B[a]PDE) are known to react with DNA to form covalent adducts preferentially at the N2 position of guanines, and to a lesser extent, at N6 position of adenines. It is well established that B[a]P-N2-dG adducts form more efficiently at methylated CpGs than non-methylated CpGs (18,19). We have also explored the occurrence of known transcription factor recognition sites within genomic loci targeted by aberrant DNA methylation during chemical carcinogenesis. A list of potential transcription factors with their respective consensus binding sites is shown in Figure 4B. Of interest, the non-histone HMG-IY transcription factor

binds to A-T rich DNA sequences and participates in enhanceosome formation, chromatin remodeling and regulation of transcription, with a crucial role in many cellular processes, including cell growth and differentiation (34). Other potentially relevant transcription factors include the zinc finger protein *Mzfl*, which plays an important role in cell proliferation and tumorigenesis (35), and the forkhead box protein *Foxd3*, a novel epigenetically regulated tumor suppressor gene that controls ESC self-renewal and pluripotency as well as cell growth (Figure 4B) (36). Enrichment of consensus binding sites within the differentially methylated CpG islands indicates that aberrant DNA methylation can also interfere with *in vivo* binding of key transcription factors and/or recruitment of methyl-binding proteins (37).

Functional pathway analysis of differentially methylated genes

Using a combination of the Ingenuity Pathway Analysis® (IPA®: v 9.0) and the GO Enrichment Analysis in Partek® Genomics Suite™ (v 6.6), we obtained gene ontology information for the annotated genes identified as aberrantly methylated by MIRA-microarray analysis. Functional annotation analysis revealed that gene targets involved in connective tissue development and function, embryonic development, cell-to-cell signaling and interaction and nervous system development and function were particularly enriched (Figure 2F and IPA results). These target genes are members of the frequently disrupted signaling networks in cancer, including the *MEK/ERK*, *JAK/STAT3*, *PI3K/AKT*, *WNT/β-catenin* and *Shh* cascades (38–40) (Figure 5B and Supplementary Figure S4, available at *Carcinogenesis* Online). These signaling networks cross-talk to and modulate the *Nanog* pathway, which is overall the most represented canonical pathway, according to IPA (Figure 5A). The significance of the *Nanog* pathway in the establishment and maintenance of the pluripotent state and carcinogenesis is increasingly appreciated (see Discussion).

Correlation between DNA methylation and gene expression during B[a]P-induced carcinogenesis

To shed light into the underlying mechanisms of aberrant DNA methylation, we next measured the mRNA levels of major murine methyltransferases including the 'maintenance' *Dnmt1* and the 'de novo' *Dnmt3a* and *Dnmt3b* methyltransferases. As shown in Figure 6A, we detected cumulative loss of expression of *Dnmt3a* and *Dnmt3b* in samples of B[a]P-treated mice both before (T0) and after tumor development (T2), whereas no significant changes were detectable in the expression level of *Dnmt1*. Misregulation of *Dnmt3a* and *Dnmt3b* can interfere with the establishment and maintenance of normal patterns of DNA methylation and, in turn, lead to tumorigenesis (41).

To investigate the impact of DNA methylation on gene expression during chemical carcinogenesis, we have also quantified the transcription level of several functionally important genes that were aberrantly methylated, as identified by our MIRA-microarray analysis. The examined genes are known components of the signaling cascades linked to the *Nanog* pathway and are potentially relevant for the establishment of a malignant phenotype (Figure 5B). Figure 6 shows the mean normalized expression levels of the *Wnt4* and *Fzd3* (Figure 6B), *Mapk3* (*Erk1*) and *Mapk11* (Figure 6C), *Foxd3*, *Nanog* and *Gata6* genes (Figure 6D) in samples of B[a]P-treated mice at T0 and T2 relative to control. With the exception of the *Gata6* gene, which is upregulated, most of the above genes show overall reduction in transcription levels during B[a]P-induced carcinogenesis.

Discussion

Chemical carcinogenesis has historically been accounted for by genotoxicity, which entails DNA damage, genetic mutations and chromosomal abnormalities (1–3). However, an emerging model of carcinogenicity recognizes that epigenetic alterations can also play an important role in the initiation and progression of cancer, although the underlying molecular mechanisms remain to be elucidated (4–6). To investigate whether epigenetic alterations are a determinant of

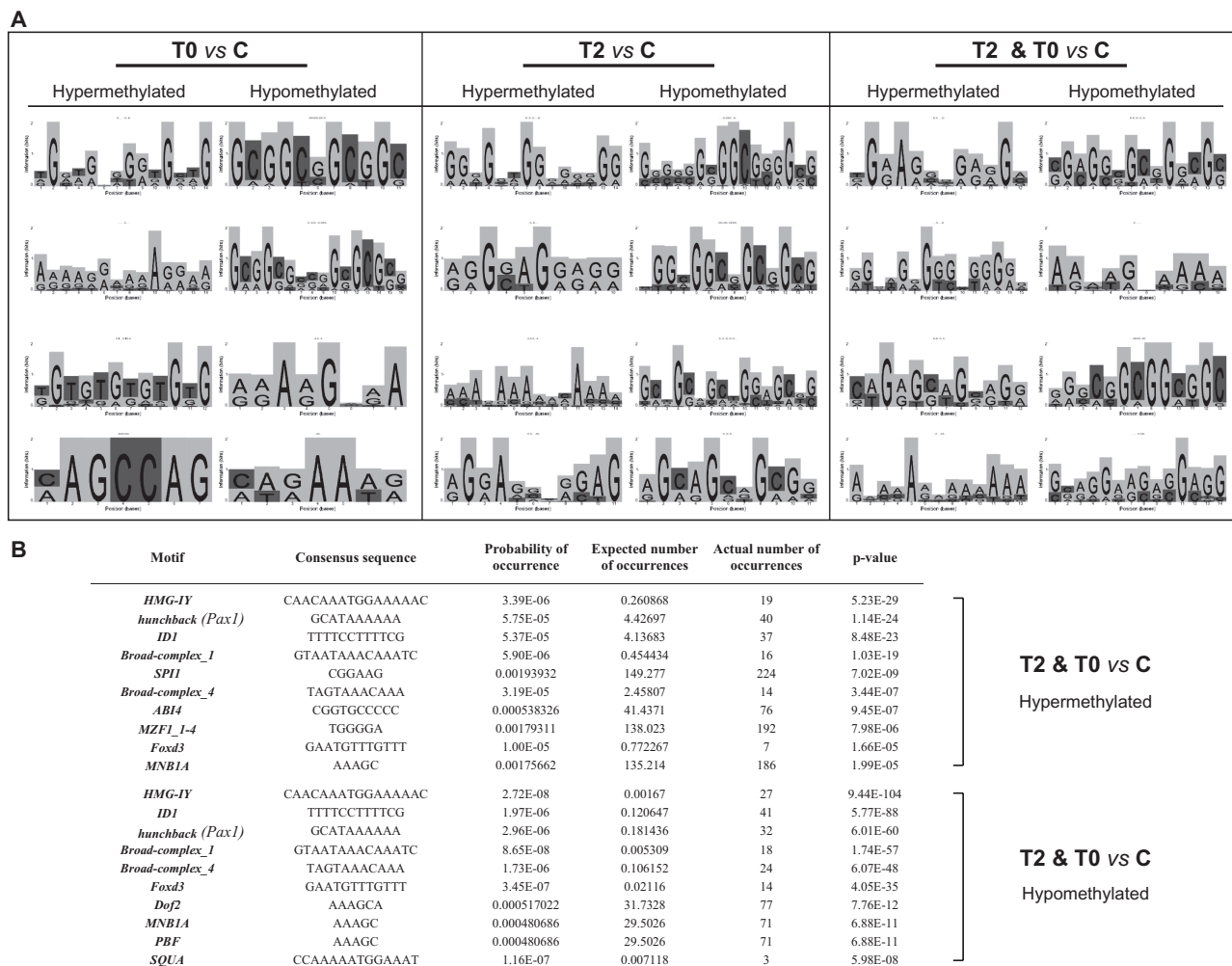


Fig. 4. Summary of motif occurrences across differentially methylated CpG islands in B[a]P-treated mice at T0 and T2. (A) The top *de novo* motifs were selected using the Partek® Genomics Suite™ software and show high enrichment of guanines, often in the context of CpGs. The ‘Sequence Logo’ windows graphically display the best motifs found in the hyper- and hypomethylated CpG islands in samples of B[a]P-treated mice at T0 and T2. The height of each position is the relative entropy (in bits) and indicates the importance of a base at a particular location in the binding site. (B) Known motifs were identified by using the JASPAR database in Partek®, based on the number of occurrences above the threshold.

chemical carcinogenesis, we have chronicled changes in the epigenetic landscape during chemically induced carcinogenesis. Accordingly, we have assessed the alterations of the epigenome, as reflected in the patterns of DNA methylation, in mice treated *in vivo* with a representative chemical carcinogen, B[a]P (15), before and after tumor development. It is well established that rodents exposed to B[a]P develop large aggressive tumors (including sarcoma) at the accessory sex organs (29). Thus, we constructed the whole DNA methylome in the seminal vesicles or tumors formed at this organ site in B[a]P-treated mice at different intervals prior to (T0 and T1) and after tumor development (T2) (Figure 1A). Genome scale analysis, supported by standard validation assays, showed locus-specific hyper- and hypomethylation in B[a]P-treated mice, even at a time preceding detectable lesion formation. Of the 123 aberrantly methylated loci identified in the normal appearing seminal vesicles in B[a]P-treated mice at T0, a significant portion (72%) overlapped with those identified as tumor-specific differentially methylated targets in T2 mice (Figure 2A–C).

De novo motif discovery analysis of the aberrantly methylated CpG islands, in samples of B[a]P-treated mice at T0 and T2, indicates a high enrichment of guanine residues, often in the context of CpG dinucleotides (Figure 4A). This sequence specificity of the aberrantly methylated targets is consistent with preferential binding of the reactive metabolite of B[a]P, B[a]PDE, to guanines in the DNA (19). Given the

fact that these epigenetic changes manifest early and in the absence of evident morphological abnormalities, it is tempting to speculate that these alterations are the initiating oncogenic events, directly related to the effect of carcinogen exposure, although a causal relationship remains to be established. At this time, we cannot exclude that other carcinogen-induced epigenetic effects, such as histone modifications, chromatin remodeling and/or microRNA gene modulation, may have triggered the aberration of DNA methylation patterns observed in the present study. Alternatively, it is plausible that the detected epigenetic changes are the result of B[a]P genotoxicity, e.g. mutations in crucial genes that can directly or indirectly influence key pathways involved in the establishment and maintenance of DNA methylation patterns. Of relevance, we observed downregulation of the ‘*de novo*’ *Dnmt3a* and *Dnmt3b* methyltransferases (Figure 6A), which can lead to global loss of DNA methylation. In confirmation, we have found demethylation of the IAP-LTR repetitive elements in tumors of B[a]P-treated mice (Supplementary Figure S3, available at *Carcinogenesis* Online) (41). The IAP retrotransposons are often associated with ‘metastable epialleles’ in mouse, with varying methylation at specific CpG sites that can be influenced by *in utero* and/or early life exposure to environmental toxicants (6). Furthermore, we have detected agglomerates of differentially methylated CpG islands along extended chromosomal regions containing groups of consecutive genes, such as the *Hoxa* and *Hoxb* gene clusters, in B[a]P-treated mice both before

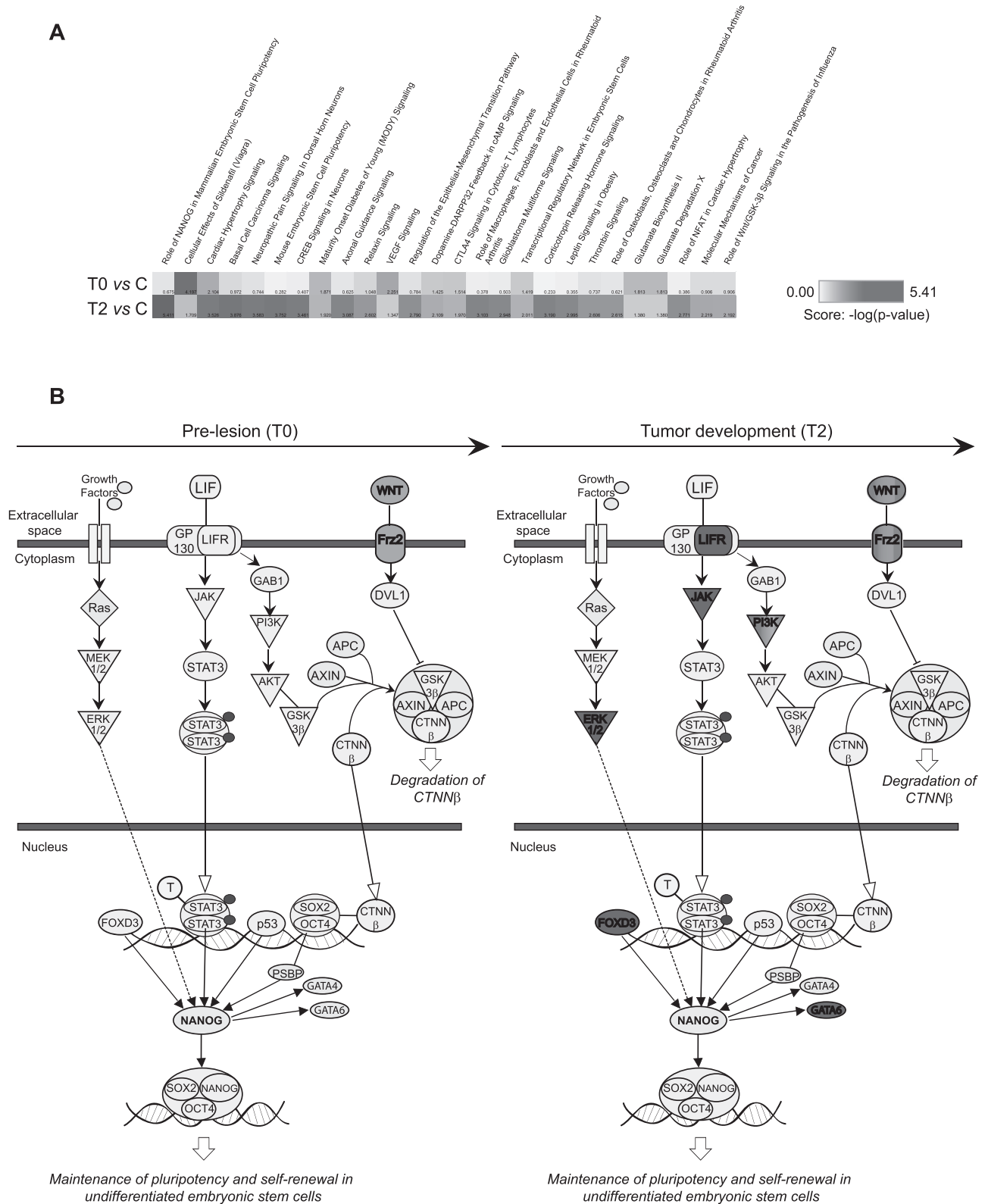


Fig. 5. Functional pathway analysis of aberrantly methylated genes in B[a]P-treated mice prior to (T0) and after (T2) tumor development. **(A)** The Canonical Pathway Heat Map was generated using the Comparison Analysis in IPA®. The heat map visualizes the enriched canonical pathways simultaneously in T0 and T2, allowing a direct comparison between the two datasets. The pathway scores are displayed using a blue color gradient, where darker blue corresponds to higher scores. The score represents the negative log of the P value derived from the Fisher's exact test. **(B)** The top canonical pathway (*Nanog* pathway) is crucial in mammalian ESC pluripotency and appears to play a role in chemical carcinogenesis. Aberrant DNA methylation targets several components of the signaling cascades that modulate *Nanog* expression and increases progressively from T0 to T2. Red and green nodes represent hyper- and hypomethylated gene targets identified by MIRA-microarray analysis, respectively. Adapted from IPA with some modifications.

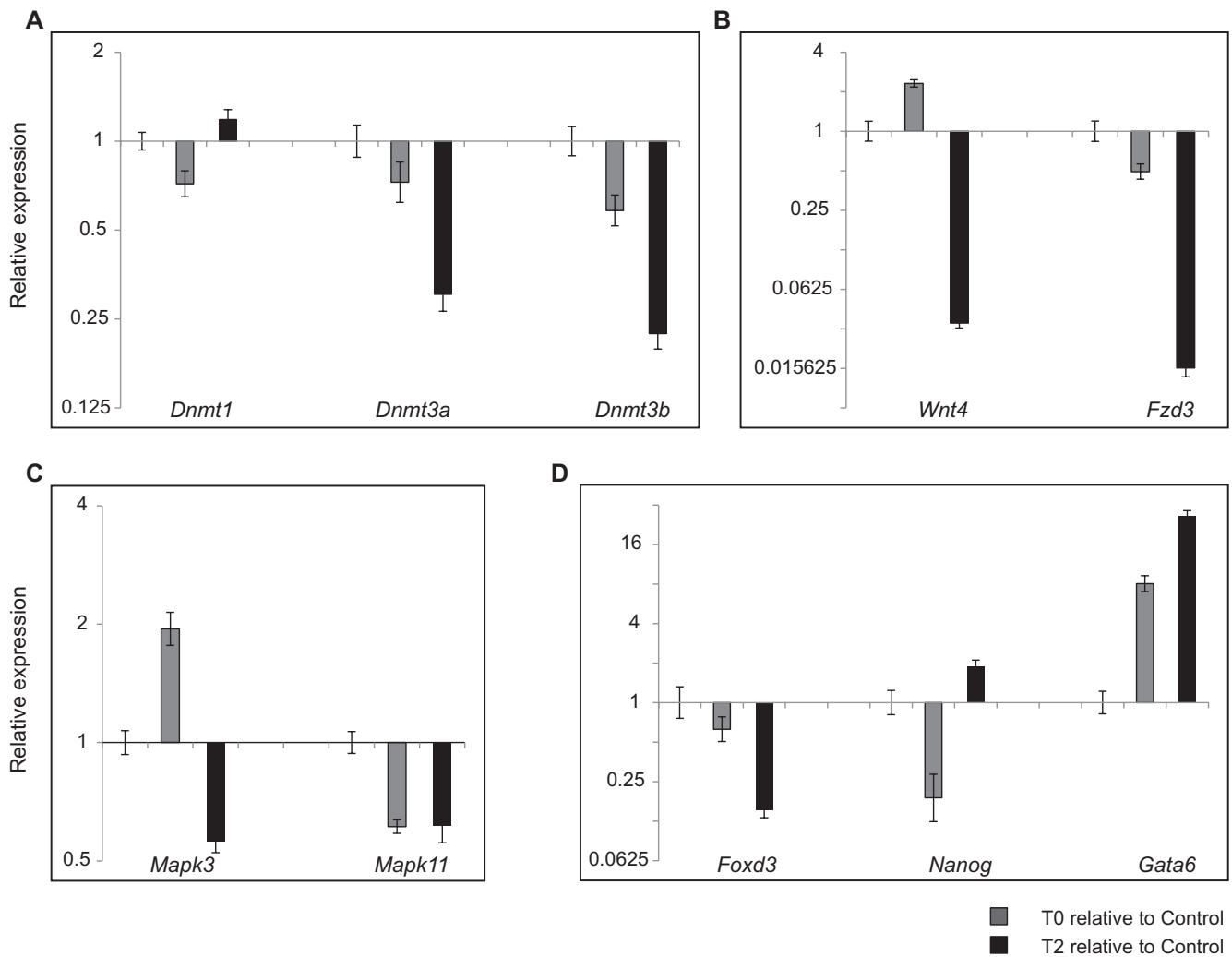


Fig. 6. Relative quantification of gene expression by standard qRT-PCR. The expression status of major methyltransferases and that of individual target genes identified by MIRA-microarray analysis was analyzed by standard qRT-PCR using the $2^{-\Delta\Delta C_t}$ method. Bars represent the mean normalized expression (\pm SD) of three replicates in samples of B[a]P-treated mice at T0 (gray) and T2 (black) relative to control. Data were normalized using an endogenous housekeeping gene as the reference (*Gapdh*) and untreated control as the calibrator (with expression equal to 1). Relative transcription levels of *Dnmt1*, *Dnmt3a* and *Dnmt3b* (A), *Wnt4* and *Fzd3* (B), *Mapk3* (*Erk1*) and *Mapk11* (C) and *Foxd3*, *Nanog* and *Gata6* genes (D) are shown.

and after tumor development (Supplementary Figure S2, available at *Carcinogenesis* Online). Agglomerative aberrant DNA methylation is an epigenetic signature commonly found in several types of human malignancy (33), cancer cell lines and other toxicant-induced models of cell transformation (42). Our overall data confirm that known hallmarks of human cancer, including locus-specific gain/loss of CpG methylation and global loss of methylation in repetitive DNA elements, are key components of chemically induced carcinogenesis.

In the present study, the identified tumor-specific aberrantly methylated genes colocalize mostly with the murine stem cell active histone marker (H3K4me3), and to a lesser extent, with the bivalent histone marker (H3K4me3 and H3K27me3), as deduced by comparing our methylation data with the published mouse databases (30) (Figure 2E). Nearly half of the annotated targets associated with the H3K4me3 active histone mark are phosphoproteins, mostly involved in intracellular signaling cascades. This is consistent with the functional pathway analysis that shows the convergence of top aberrantly methylated targets at the *MEK/ERK*, *JAK/STAT3*, *PI3K/AKT*, *WNT/ β -catenin* and *Shh* signaling cascades. The annotated targets associated with bivalent marks are mainly involved in the control of transcription and pattern formation and include several developmental regulators. In agreement with findings by others (43,44), we have also found that 10% of the tumor-associated hypermethylated CpG islands overlap

with loci that are bound by components of the polycomb repressive complex 1 (PRC1) (Rnf2 and Phc1) and PRC2 (Eed and Suz12) that are associated with transcriptionally silent chromatin in ESCs (45). Together, these data indicate an interplay between aberrant DNA methylation and histone modifications that shapes the epigenetic landscape during chemical carcinogenesis.

Gene ontology analysis of tumor-specific differentially methylated targets shows enrichment of genes involved in connective tissue development and function, embryonic development and organ development (Figure 2F and IPA results). The vast majority of these genes are normally expressed in the brain, a trend described in several human malignancies. The top canonical pathway whose members are preferential targets of aberrant DNA methylation is the *Nanog* pathway (Figure 5A). *Nanog* is a well-studied transcription factor that plays a crucial role in the maintenance and self-renewal of undifferentiated ESCs (46). *Nanog* expression is normally restricted to pluripotent cells and is downregulated upon differentiation; however, *Nanog* has also been found to be dysregulated in cancer cell lines and tumors (47). Accumulating evidence supports that *Nanog* and other pluripotency genes can function as neoplastic engines to drive tumorigenesis, probably promoting infinite self-renewal of a distinct subset of stem-like cells within tumors (46,48). Interestingly, we have observed methylation defects in genes upstream of *Nanog*, particularly in

components of the oncogenic signaling cascades, including the *MEK/ERK*, *JAK/STAT3*, and *WNT/β-catenin*, etc., which modulate *Nanog* expression.

Consistent with DNA methylation data, standard qRT-PCR analysis confirmed abnormal and deregulated expression of several components of the signaling cascades interconnected with the *Nanog* pathway, which may altogether cause perturbations in cell fate decision and differentiation and, ultimately, lead to cancer (Figure 6B–D). Apart from altering the expression of key genes (i.e. transcription factors), aberrant DNA methylation can also disrupt their respective genomic binding sites, thus, interfering with regulation of downstream effectors. For instance, epigenetic downregulation of the *Foxd3* gene (Supplementary Table S2, available at *Carcinogenesis* Online and Figure 6D) and functional inactivation of *Foxd3* recognition sites by aberrant DNA methylation (Figure 4B) can both contribute to loss of *Foxd3* tumor suppressive function. Of note, ~10% of the cancer-related hypermethylated genes identified in this study are potential binding targets of *Nanog* in murine ESCs (49). Our findings accord with a recent report by Varley *et al.* (50) who demonstrated that cancer-specific hypermethylation is enriched at sites bound by *Nanog* in human ESCs.

In summary, our study indicates that *in vivo* exposure of mice to a prototype chemical carcinogen can alter the epigenetic landscape in a similar fashion to that found in human cancer. The alterations of the epigenome, as reflected in the patterns of DNA methylation, progressively increase during chemical carcinogenesis. Members of the *Nanog* pathway, which establishes and maintains pluripotency in the mammalian ESCs and possibly triggers uncontrolled proliferation of neoplastic cells (46), are preferential targets of aberrant DNA methylation during chemical carcinogenesis. Moreover, top gene networks targeted by aberrant DNA methylation are components of the signaling cascades that cross-talk to *Nanog* and are frequently disrupted in human cancer. Altogether, our data show the predictive value of aberrant DNA methylation in carcinogenesis and the potential utility of this epigenetic mark for early detection and monitoring of the progression of malignancy. Given the reversibility of this epigenetic change, e.g. through pharmacologic interventions, mitigating aberrant DNA methylation, particularly at the loci and in the pathways identified in the present study, may serve as a therapeutic approach for cancer.

Supplementary material

Supplementary Tables S1 and S2 and Figures S1–S4 can be found at <http://carcin.oxfordjournals.org/>

Funding

American Cancer Society (RSG-11-083-01-CNE); University of California Tobacco Related Disease Research Program (20XT-0116 to A.B.).

Acknowledgements

We thank Meng Li, from the NML Bioinformatics Service Program of the University of Southern California, for help with the Partek® Genomics Suite™ software. We also thank the City of Hope Bioinformatics Core Facility for initial microarray data analysis and Dr Sharon Wilczynski, from the Department of Anatomic Pathology, for help with H&E analysis.

Conflict of Interest Statement: None declared.

References

- Luch,A. (2005) Nature and nurture - lessons from chemical carcinogenesis. *Nat. Rev. Cancer*, **5**, 113–125.
- Loeb,L.A. *et al.* (2008) Advances in chemical carcinogenesis: a historical review and prospective. *Cancer Res.*, **68**, 6863–6872.
- Wogan,G.N. *et al.* (2004) Environmental and chemical carcinogenesis. *Semin. Cancer Biol.*, **14**, 473–486.
- Cortessis,V.K. *et al.* (2012) Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. *Hum. Genet.*, **131**, 1565–1589.
- Pogribny,I.P. *et al.* (2013) Environmental toxicants, epigenetics, and cancer. *Adv. Exp. Med. Biol.*, **754**, 215–232.
- Bollati,V. *et al.* (2010) Environmental epigenetics. *Heredity (Edinb.)*, **105**, 105–112.
- Portela,A. *et al.* (2010) Epigenetic modifications and human disease. *Nat. Biotechnol.*, **28**, 1057–1068.
- Baylin,S.B. *et al.* (2011) A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer*, **11**, 726–734.
- Jones,P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.*, **13**, 484–492.
- Deaton,A.M. *et al.* (2011) CpG islands and the regulation of transcription. *Genes Dev.*, **25**, 1010–1022.
- Maunakea,A.K. *et al.* (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, **466**, 253–257.
- Ehrlich,M. (2009) DNA hypomethylation in cancer cells. *Epigenomics*, **1**, 239–259.
- Nagarajan,R.P. *et al.* (2009) Epigenetic mechanisms in glioblastoma multiforme. *Semin. Cancer Biol.*, **19**, 188–197.
- Wilson,A.S. *et al.* (2007) DNA hypomethylation and human diseases. *Biochim. Biophys. Acta*, **1775**, 138–162.
- Besaratinia,A. *et al.* (2005) DNA damage and mutagenesis induced by polycyclic aromatic hydrocarbons. In Luch,A. (ed.) *The Carcinogenic Effects of Polycyclic Aromatic Hydrocarbons*. Imperial College Press, London, UK, pp. 171–210.
- Weisenberger,D.J. *et al.* (1999) Cytosine methylation in a CpG sequence leads to enhanced reactivity with Benzo[a]pyrene diol epoxide that correlates with a conformational change. *J. Biol. Chem.*, **274**, 23948–23955.
- Wojciechowski,M.F. *et al.* (1984) Inhibition of DNA methyltransferases *in vitro* by benzo[a]pyrene diol epoxide-modified substrates. *J. Biol. Chem.*, **259**, 9711–9716.
- Wilson,V.L. *et al.* (1983) Inhibition of DNA methylation by chemical carcinogens *in vitro*. *Cell*, **32**, 239–246.
- Stone,M.P. *et al.* (2011) Chemistry and structural biology of DNA damage and biological consequences. *Chem. Biodivers.*, **8**, 1571–1615.
- Huidobro,C. *et al.* (2013) The role of genetics in the establishment and maintenance of the epigenome. *Cell. Mol. Life Sci.*, **70**, 1543–1573.
- Dolinoy,D.C. *et al.* (2008) Environmental epigenomics in human health and disease. *Environ. Mol. Mutagen.*, **49**, 4–8.
- Tommasi,S. *et al.* (2013) Mammalian cells acquire epigenetic hallmarks of human cancer during immortalization. *Nucleic Acids Res.*, **41**, 182–195.
- Frommer,M. *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl Acad. Sci. USA*, **89**, 1827–1831.
- Xiong,Z. *et al.* (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
- Yang,A.S. *et al.* (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res.*, **32**, e38.
- Goodier,J.L. *et al.* (2001) A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.*, **11**, 1677–1685.
- Martens,J.H. *et al.* (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.*, **24**, 800–812.
- Waterston,R.H., *et al.* (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520–562.
- Moore,R.A., *et al.* (1937) Production of tumors of the prostate of the white rat with 1:2-benzopyrene. *Am J Cancer.*, **30**, 731–741.
- Rugg-Gunn,P.J. *et al.* (2010) Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc. Natl Acad. Sci. USA*, **107**, 10783–10790.
- Yamazaki,M. *et al.* (2010) TARPs gamma-2 and gamma-7 are essential for AMPA receptor expression in the cerebellum. *Eur. J. Neurosci.*, **31**, 2204–2220.
- Xu,B. *et al.* (2012) The role of Pax2 in mouse prostate development. *Prostate*, **72**, 217–224.
- Ordway,J.M. *et al.* (2006) Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis*, **27**, 2409–2423.
- Massimi,I. *et al.* (2013) The HMGA1 protooncogene frequently deregulated in cancer is a transcriptional target of E2F1. *Mol. Carcinog.*, **52**, 526–534.
- Gaboli,M. *et al.* (2001) Mzf1 controls cell proliferation and tumorigenesis. *Genes Dev.*, **15**, 1625–1630.

36. Liu, Y. *et al.* (2008) Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem Cells*, **26**, 2475–2484.
37. Suzuki, M.M. *et al.* (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.*, **9**, 465–476.
38. Schubbert, S. *et al.* (2007) Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer*, **7**, 295–308.
39. Carnero, A. (2010) The PKB/AKT pathway in cancer. *Curr. Pharm. Des.*, **16**, 34–44.
40. Katoh, Y. *et al.* (2009) Hedgehog target genes: mechanisms of carcinogenesis induced by aberrant hedgehog signaling activation. *Curr. Mol. Med.*, **9**, 873–886.
41. Jurkowska, R.Z. *et al.* (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem*, **12**, 206–222.
42. Severson, P.L. *et al.* (2012) Agglomerates of aberrant DNA methylation are associated with toxicant-induced malignant transformation. *Epigenetics*, **7**, 1238–1248.
43. Widschwendter, M. *et al.* (2007) Epigenetic stem cell signature in cancer. *Nat. Genet.*, **39**, 157–158.
44. Fernandez, A.F. *et al.* (2012) A DNA methylation fingerprint of 1628 human samples. *Genome Res.*, **22**, 407–419.
45. Boyer, L.A. *et al.* (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*, **441**, 349–353.
46. Saunders, A. *et al.* (2013) Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. *Stem Cells*, **31**, 1227–1236.
47. Wang, M.L. *et al.* (2013) Targeting cancer stem cells: emerging role of Nanog transcription factor. *Onco. Targets. Ther.*, **6**, 1207–1220.
48. Luo, Y. *et al.* (2012) Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. *J R Soc Interface.*, [epub ahead of print].
49. Loh, Y.H. *et al.* (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.*, **38**, 431–440.
50. Varley, K.E. *et al.* (2013) Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.*, **23**, 555–567.

Received September 10, 2013; revised January 8, 2014;
accepted January 17, 2014

FIGURE LEGENDS

Figure S1. Verification of differentially methylated CpG islands by COBRA.

Genomic DNA from B[a]P-treated and control mice was treated with sodium bisulfite, and the CpG islands of interest were amplified with gene-specific primers, and subjected to COBRA. Digested fragments on the gel are indicative of methylated restriction sites within the CpG island. *In vitro* methylated mouse genomic DNA served as positive control (Pos). The symbols (+) and (-) show the presence and absence, respectively, of the restriction enzyme in the reaction mix. M= 100 bp ladder DNA marker.

Figure S2. Hypermethylated CpG islands associated with the *Hoxa* and *Hoxb* clusters.

Visualization of the NimbleGen array data for B[a]P-treated (T0, in blue and T2, in red) and control mice (C, in green). At this level of resolution, each peak corresponds to a CpG island. The methylation signal, obtained by the NimbleScan software, is plotted along the chromosome as a *P*-value on the y-axis (starting from 0 when *P*-value is 1). The *P*-value is derived from the Kolmogorov–Smirnov test comparing the log₂ enrichment ratios between MIRA and input within a 750 bp window.

Figure S3. Methylation profiling in repetitive DNA elements in tumors from B[a]P-

treated mice. (A) Bisulfite sequencing of LINE L1, IAP-LTR and SINE B1 elements.

Values indicate percentages of mCpGs. **(B)** Quantification of mCpGs in major repetitive DNA elements in tumors from B[a]P-treated mice *versus* control. Fisher's exact test was used to calculate the statistical significance of difference in mCpGs between tumors and controls.

Figure S4. Functional pathway analysis of aberrantly methylated genes in tumors from B[a]P-treated mice. Merging networks were generated by Ingenuity Pathway Analysis[®] (IPA[®]: v 9.0), using the relaxed list (T2 vs C). The convergence of genes in the *MEK/ERK*, *AKT* and *Shh* networks is noticeable. Red and green nodes represent hypermethylated and hypomethylated genes, respectively.

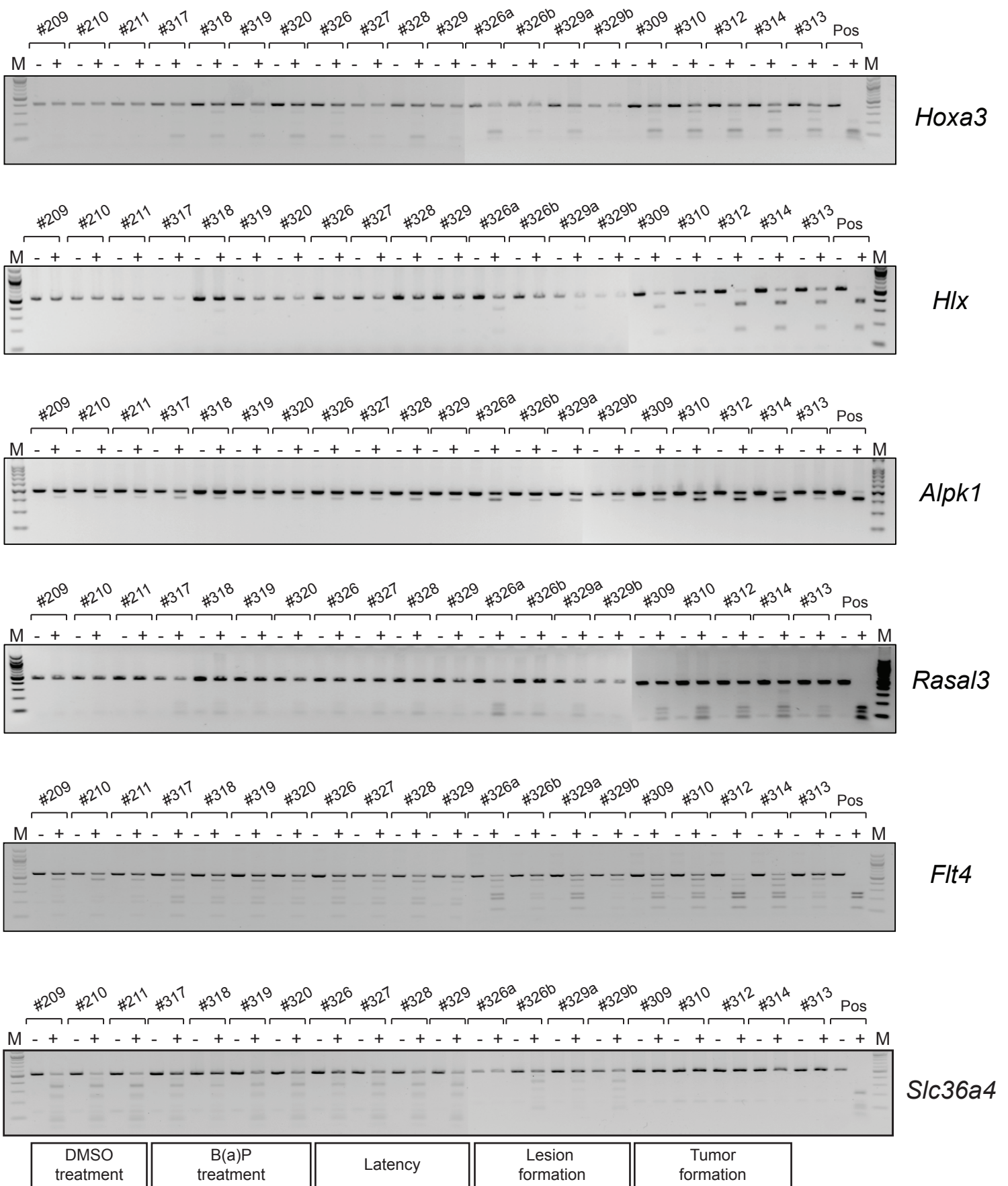
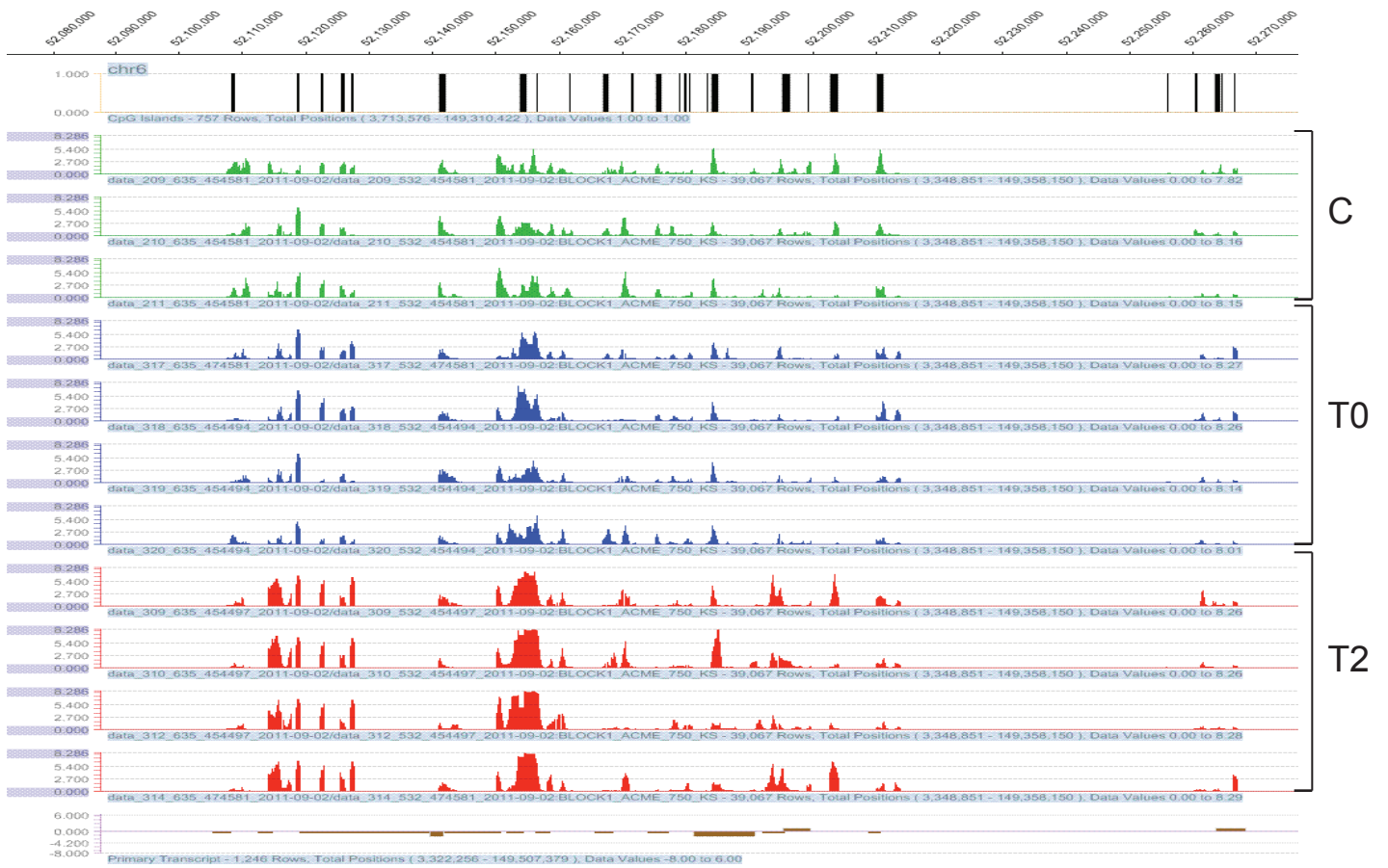
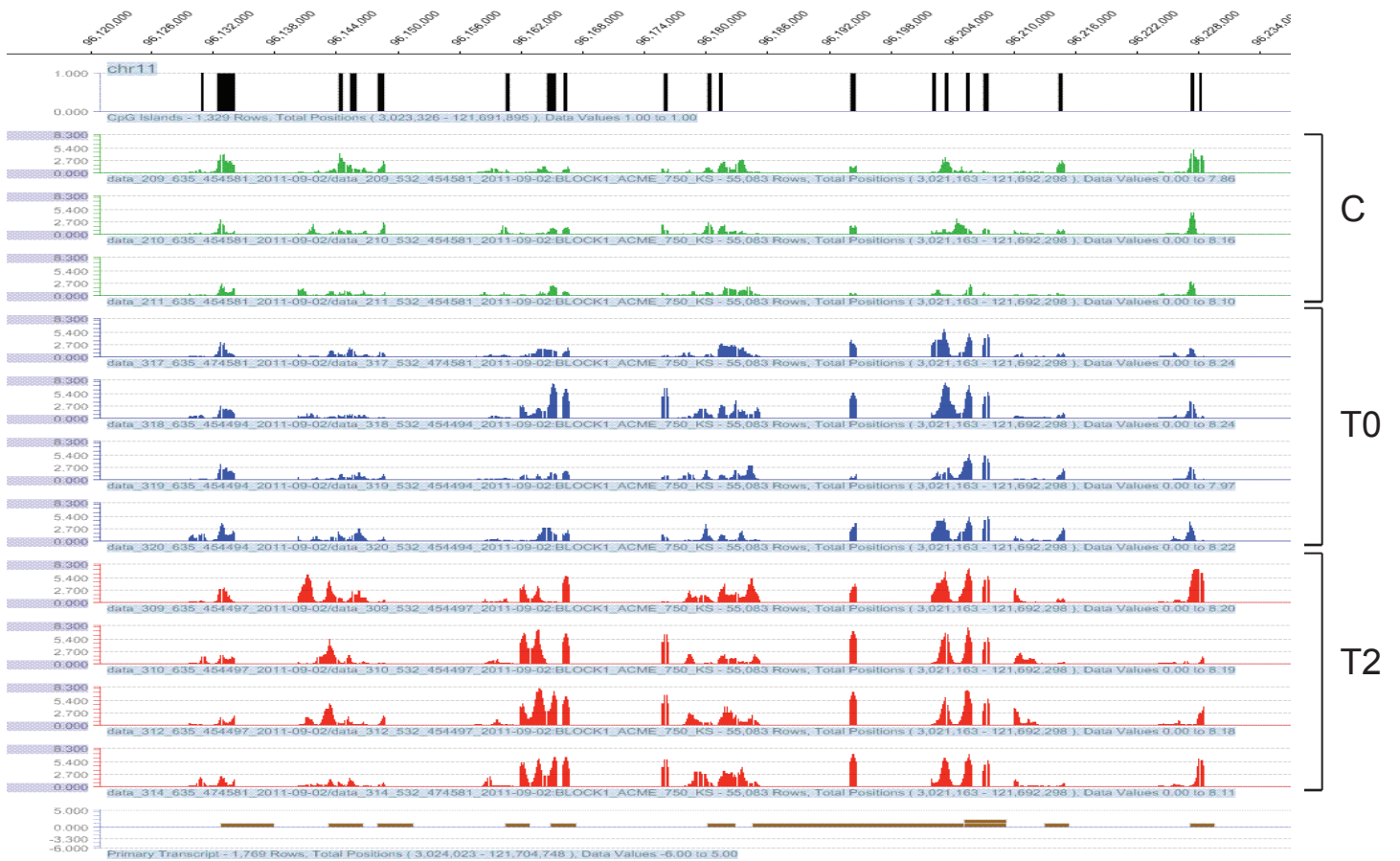


Figure S1



Hoxa cluster



Hoxb cluster

Figure S2

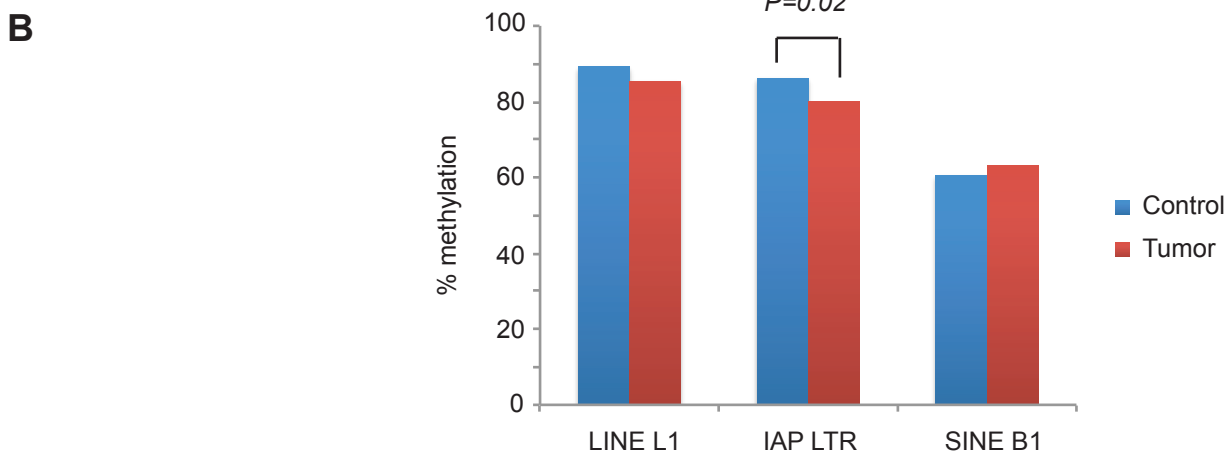
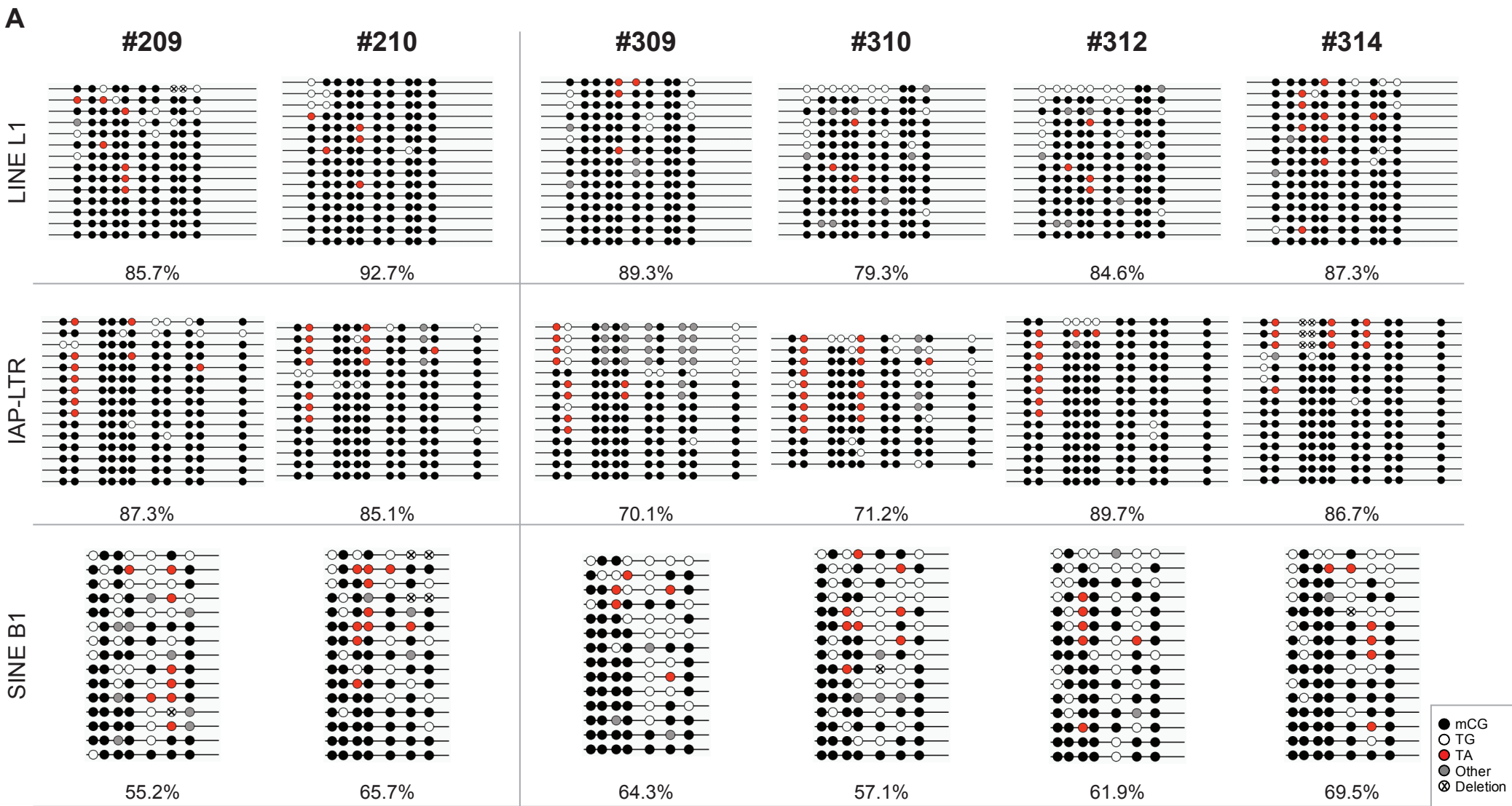


Figure S3

Table S1. Primers used for qRT-PCR

Gene symbol	Forward	Reverse	Product size (bp)
<i>Gata6</i>	5'-TTGCTCCGGTAACAGCAGTG	5'-GTGGTCGCTTGTGTAGAAGGA	105
<i>Mapk11</i>	5'-TCAGTGCGGCCGAAGCCTTG	5'-CGTGCCTCCTTGGCCTCAA	110
<i>Wnt4</i>	5'-AGACGTGCGAGAACTCAAAG	5'-GGAACTGGTATTGGCACTCCT	126
<i>Mapk3</i>	5'-TCCGCCATGAGAATGTTATAGGC	5'-GGTGGTGTGATAAGCAGATTGG	248
<i>Fzd3</i>	5'-ATGGCTGTGAGCTGGATTGTC	5'-GGCACATCCTCAAGGTTATAGGT	109
<i>Nanog</i>	5'-TCTTCCTGGTCCCCACAGTTT	5'-GCAAGAATAGTTCTCGGGATGAA	100
<i>Foxd3</i>	5'-GACCCCGAACAAGCCCAAG	5'-GAAAACGGTTGCTGATGAACTC	137
<i>Dnmt1</i>	5'-CCTAGTTCCTGGCTACGAGGAGAA	5'-TCTCTCTCCTCTGCAGCCGACTCA	137
<i>Dnmt3a</i>	5'-GCCGAATTGTGTCTTGGTGGATGACA	5'-CCTGGTGGAAATGCACTGCAGAAGGA	147
<i>Dnmt3b</i>	5'-TTCAGTGACCAGTCCTCAGACACGAA	5'-TCAGAAGGCTGGAGACCTCCCTCTT	145
<i>Gapdh</i>	5'-AGGTCGGTGTGAACGGATTTG	5'-TGTAGACCATGTAGTTGAGGTCA	123

29	chr10	92623450	92623801	3	4	1.28 +	92622620	92624620	Cdk17	NM_146239										
30	chr8	94984419	94984848	3	4	1.28														
31	chr4	100449306	100449671	3	4	1.26 +	100448283	100450283	Cachd1	NM_198037										
32	chr8	86521752	86522288	3	4	1.26 +	86520570	86522570	Samd1	NM_001081415										
33	chr11	104303386	104303915	3	4	1.25 -	104302605	104304605	17000811L1Ri	NM_001081045										
34	chr12	112615753	112616197	3	4	1.24 -	112614929	112616929	Cdc42bbp	NM_183016										
35	chr13	58229736	58230399	3	4	1.24 -	58228917	58230917	Hnrmpa0	NM_029872										
36	chr16	77014053	77014412	3	4	1.24 +	77013313	77015313	Usp25	NM_013918										
37	chr17	31992630	31992983	3	4	1.24 -	31991737	31993737	Sik1	NM_010831										
38	chr9	51856357	51856811	3	4	1.24 +	51855045	51857045	Rdx	NM_00110461+	51856254	51895843	Rdx						NM_001104616	
39	chr1	79855007	79855468	3	4	1.23 -	79854240	79856240	Serpine2	NM_009255										
40	chr2	3391942	3392381	3	4	1.22 -	3391258	3393258	Suv39h2	NM_022724										
41	chr5	34630959	34631308	3	4	1.22 -	34629973	34631973	Zfyve28	NM_001015039										
42	chr6	72047583	72048022	3	4	1.22 +	72046606	72048606	St3gal5	NM_011375										
43	chr11	4961494	4961748	3	4	1.21					4955133	4964330	Gas2l1						NM_030228	
44	chr19	41921986	41922455	3	4	1.21 -	41921622	41923622	Frat2	NM_177603										
45	chr3	121996168	121996637	3	4	1.21 -	121995621	121997621	Mir760	NR_030439				121995502	121997502	Mir760			NR_030439	
46	chr5	50449903	50450362	3	4	1.21 -	50449235	50451235	Gpr125	NM_133911										
47	chr8	119681885	119682234	3	4	1.21 +	119681034	119683034	Gan	NM_001081151										
48	chr1	91864838	91865192	3	4	1.2					91850252	91910152	Asb18						NM_139152	
49	chr6	90938042	90938471	3	4	1.2														
50	chr10	81007664	81008071	3	4	1.19 -	81006791	81008791	Gna11	NM_010301										
51	chr11	54679903	54680477	3	4	1.19 +	54678939	54680939	Hint1	NM_008248										
52	chr17	24625367	24625916	3	4	1.19 +	24624727	24626727	Caskin1	NM_027937										
53	chr4	128561354	128561613	3	4	1.19 +	128560383	128562383	Trim62	NM_178110										
54	chr7	144505319	144505985	3	4	1.19 -	144505128	144507128	Ebf3	NM_010096										
55	chr12	109513150	109513584	3	4	1.18 -	109512627	109514627	Ccdc85c	NM_001159910										
56	chr17	81127316	81127665	3	4	1.18 -	81126433	81128433	Map4k3	NM_001081357										
57	chr19	6364087	6364551	3	4	1.18 +	6362689	6364689	Sf1	NM_011750										
58	chr4	151235330	151235686	3	4	1.18 -	151234866	151236866	Camta1	NM_001166021										
59	chr6	92041715	92042089	3	4	1.18 +	92040411	92042411	Nr2c2	NM_011630										
60	chr1	134942579	134943008	3	4	1.17 +	134941588	134943588	Pik3c2b	NM_001099276										
61	chr4	21615970	21616419	3	4	1.17														
62	chr6	56747170	56747638	3	4	1.17 -	56746807	56748807	Kbtbd2	NM_145958										
63	chr7	71083454	71083916	3	4	1.17 -	71082801	71084801	Klf13	NM_021366										
64	chr10	60294100	60294364	3	4	1.16 -	60293329	60295329	Unc5b	NM_029770										
65	chr1	194681117	194681810	3	4	1.15 -	194680946	194682946	Sertad4	NM_198247										
66	chr14	65880563	65881032	3	4	1.15 -	65880300	65882300	Fzd3	NM_021458										
67	chr17	24686991	24687355	3	4	1.15 +	24685894	24687894	Pkd1	NM_013630										
68	chr2	158235529	158235788	3	4	1.15 +	158234588	158236588	Ralgapb	NM_177658										
69	chr7	70589788	70590232	3	4	1.15 +	70588657	70590657	Otud7a	NM_130880										
70	chr7	88851441	88851810	3	4	1.15 -	88850811	88852811	Homer2	NM_001164086										
71	chr17	53706359	53706752	3	4	1.14 +	53705295	53707295	Kat2b	NM_020005										
72	chr8	108160392	108160830	3	4	1.14 +	108159437	108161437	Ctcf	NM_181322										
73	chr11	32121975	32122424	3	4	1.13 -	32121293	32123293	Rhbdf1	NM_010117										
74	chr15	85036261	85036623	3	4	1.13 +	85035437	85037437	Fbln1	NM_010180										
75	chr5	111846327	111846796	3	4	1.13 +	111846185	111848185	Mn1	NM_001081235										
76	chr9	75458703	75459062	3	4	1.13 -	75458132	75460132	Tmod2	NM_016711										
77	chr9	95459855	95460235	3	4	1.13 +	95459235	95461235	Paqr9	NM_198414										
78	chr1	36569111	36569660	3	4	1.12 +	36567720	36569720	Cnmm3	NM_053186										
79	chr4	136833416	136833788	3	4	1.12 +	136832549	136834549	Wnt4	NM_009523										
80	chr5	37086978	37087332	3	4	1.12 -	37086208	37088208	D5Erttd579e	NM_001081232										
81	chr13	60277311	60278065	3	4	1.11					60276765	60277896	Gas1						NM_008086	
82	chr16	94511477	94511846	3	4	1.11														
83	chr11	69153940	69154289	3	4	1.1 +	69153270	69155270	A030009H04R	NR_027827										
84	chr12	70284760	70285215	3	4	1.1 +	70284144	70286144	Mgat2	NM_146035	70283720	70285054	Rpl36al						NM_025589	
85	chr13	15555811	15556250	3	4	1.09 +	15554555	15556555	Gli3	NM_008130										
86	chr8	97056618	97057087	3	4	1.09 +	97055927	97057927	Cpne2	NM_153507										
87	chr1	6204910	6205269	3	4	1.08 +	6203742	6205742	Rb1cc1	NM_009826										
88	chr1	91476844	91477378	3	4	1.08					91352385	91790857	Agap1						NM_178119	

