Epigenetic mechanisms in colorectal cancer

Joaquín Custodio Rojo | Tesi doctoral 2014

Epigenetic mechanisms in colorectal cancer

Memòria presentada per

Joaquín Custodio Rojo

per optar al grau de

Doctor en Biologia Cel·lular

Tesi realitzada sota la direcció del Dr. Miquel Àngel Peinado Morales
a l'Institut de Medicina Prediciva i Personalitzada del Càncer

Tesi adscrita al Departament de Biologia Cel·lular, de Fisiologia i
d'Immunologia de la Facultat de Medicina
Universitat Autònoma de Barcelona
Bienni 2013 – 2014

Director	Tutora	Doctorand
Miquel A. Peinado Morales	Rosa Miró Ametller	Joaquín Custodio Rojo

Barcelona, setembre de 2014

“La frase más excitante que se puede oír en ciencia, la que anuncia nuevos descubrimientos, no es ¡Eureka! (¡Lo encontré!) sino 'Es extraño...'.”

“La ciencia se construye a partir de aproximaciones que gradualmente se acercan a la verdad.”

—Isaac Asimov—

Als meus avis i iaios,

Als meus pares i germà

i a l'Anna,

Gràcies

INDEX

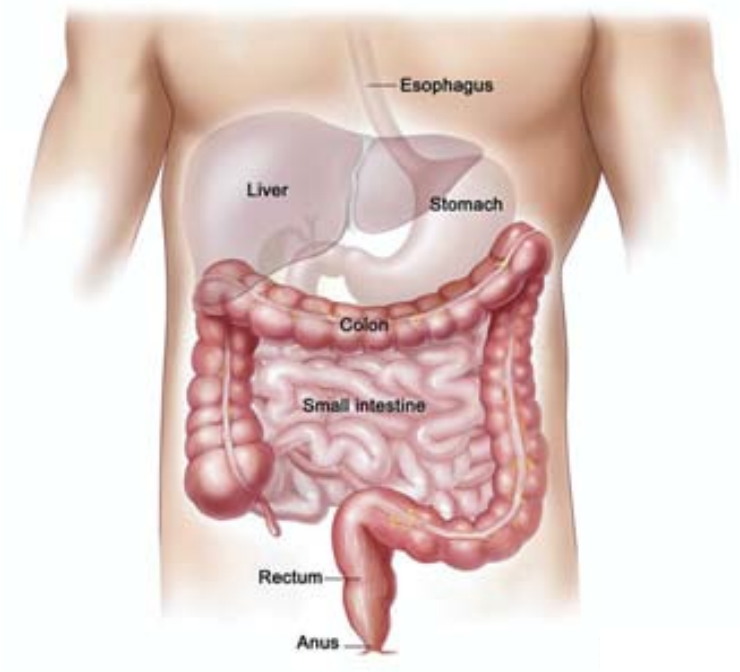
[illegible]

[illegible]

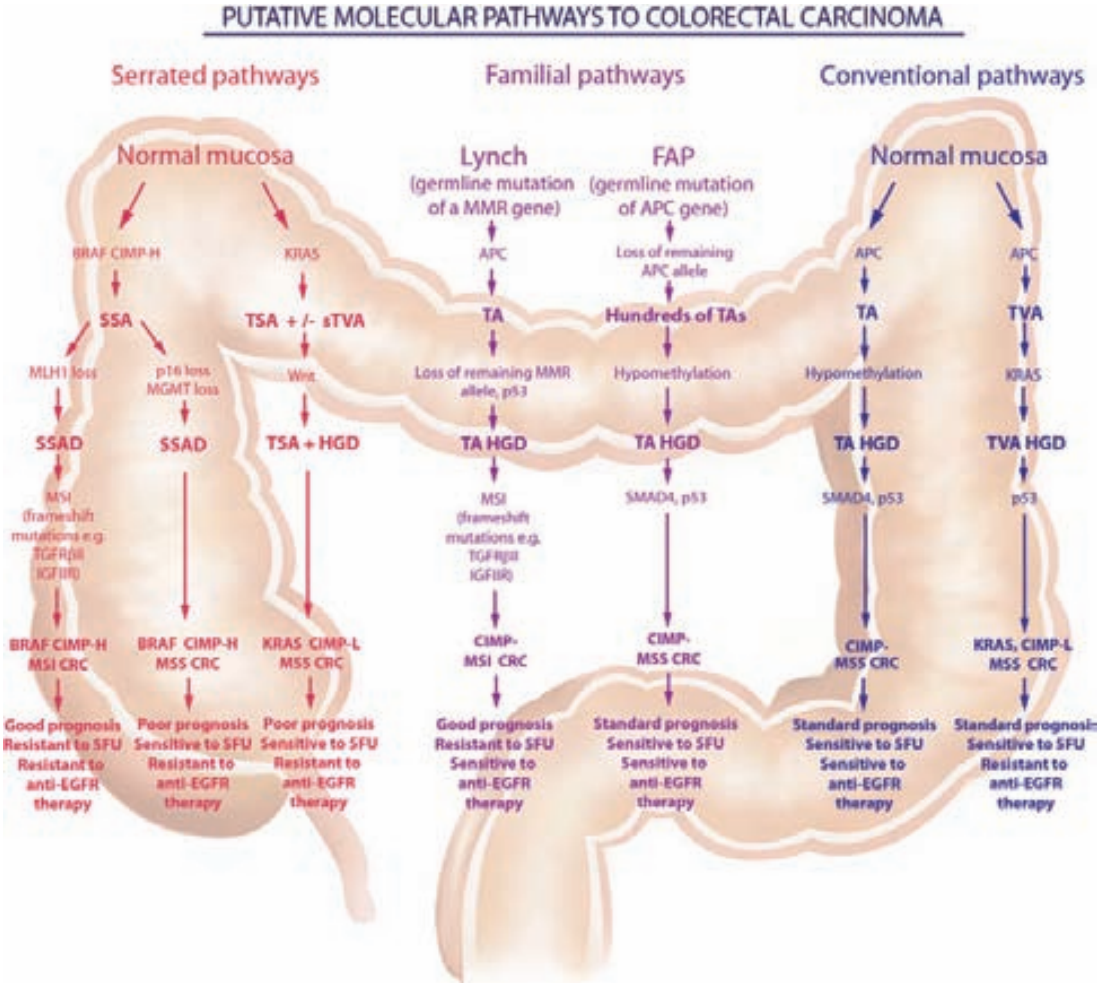
3C	chromosome conformation capture
5-AzadC	5-aza-2′-deoxycytidine
5-FU	5-Fluorouracil
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
AKR	aldo-Keto reductase
ATRA	all-trans retinoic acid
ChIP	chromatin immunoprecipitation
CNA	copy number alteration
COAD	colorectal adenocarcinoma
CpGi	CpG island
CRC	colorectal cancer
CRISPR	clustered regularly interspaced short palindromic repeats
CTCF	CCCTC-binding factor
DNMT	DNA methyltransferase
DZNep	3-Deazaneplanocin A hydrochloride
ecCEBPA	extra-coding CEBPA ncRNA
EGFR	epidermal growth factor receptor
EMT	epithelial to mesenchymal transition
eRNAs	enhancer RNA
FAP	familial adenomatous polyposis
FDA	Food and Drug Administration
FOBT	fecal occult blood test
GO	gene ontology
HDAC	histone deacetylase
HNPCC	hereditary non-polyposis colorectal cancer
IBD	inflammatory bowel disease
ISC	intestinal stem cells
Kb	Kilobase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LMR	low-methylated region
lncRNA	long non-coding RNA
miRNA	micro-RNA
ncRNA	non-coding RNA
bp	base pair
PRC	Polycomb repressive complex
PTMS	post-translational modifications
RA	retinoic acid
RAR	retinoic acid receptors
RARE	retinoic acid response elements
RXR	retinoic X receptors
TCGA	The Cancer Genome Atlas
TDG	thymine DNA glycosylase
TET	Ten-eleven translocation
TF	transcription factor
TFBS	transcription factor binding site
TMAs	tissue microarrays
TNM	tumor node metastasis
TrxG	Trithorax group
TSS	transcription start site
VEL	variant enhancer loci

[illegible]

[illegible]


Esophagus
Liver
Stomach
Colon
Small intestine
Rectum
Anus

[illegible]



?

?fa??2?f n?n?u?c ?n?s ?i ne??ece?a??n?o???r ??ah??cae?f ????ael ??mn?r ?ner ?????2. pPl?

?

?r ????r ????r ?1 ean?o?ni ?

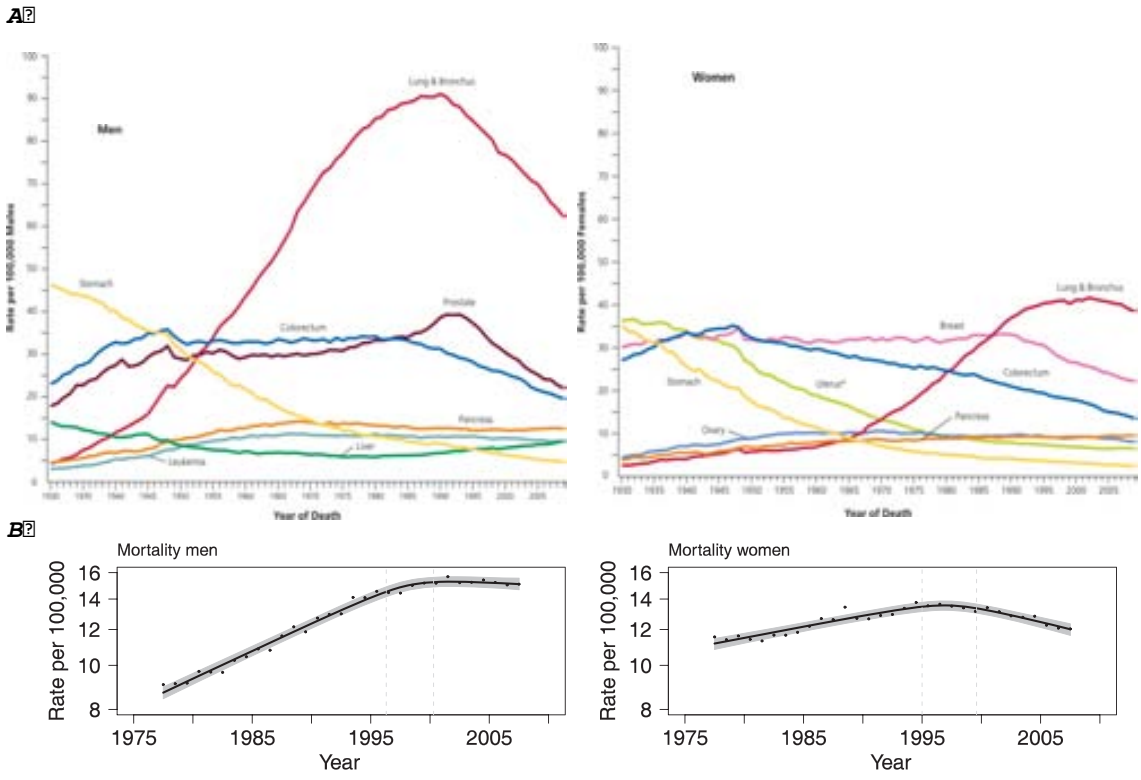
?id C??e?d???c??C?tl ???u ?ni C???al?? ?ui C?t?tes??c??ui Ce?des?enGiaIniae?en??
RiCd??xRRR0Rni0CI%??id C??e?d???c??C?tl ?en??ic??uile??iuuic???al?? ?
???en?RiCd?Rt????x?tcpd?C?F??iI?de?tcW?qóó/%?c??en??entC??uile??iuuic?
??c??C?tc?en??RiCd?W?Ren?q0LE?u tddi c?c?R???1?1??t?Icil???tc?L887??c??te?n?1?
???c??letu ?e????c?tc?C??1??i ?en??tc?t???c??W?u ?otcIa1???ei e?d?i ?L0L?u tddi c?
??1?1??t?Icil???1?Cs??C??s?L8Z8?0??

?id C??e?d???c??C?? ??el?u ?c??c??Riu ?c??duile??9a?ddsW?c??en??tc?t???c????tl?
ntIn?C?tc?en??tc?aleCt?dp????iaceCt?l?x??iae?eRizentC?l%?k?isd??F???icW?L88L%?W
?uicI?en?u ?en??ntIn?le?tc?t???c????i ??id C??e?d???c??C?tl ?tc??aleC?d?W?? ?R?
????d?c?W? ?le?Cc??aGi 1????c????iaen?Cc??aGi 1?,??c??en??d R?le?tc?t???c????tc?

[illegible]

	Birth to 39	40 to 59	60 to 69	70 and older
Male	0.08 (1 in 1212)	0.94 (1 in 106)	1.40 (1 in 71)	4.19 (1 in 24)
Female	0.08 (1 in 1236)	0.75 (1 in 134)	0.98 (1 in 102)	3.80 (1 in 26)

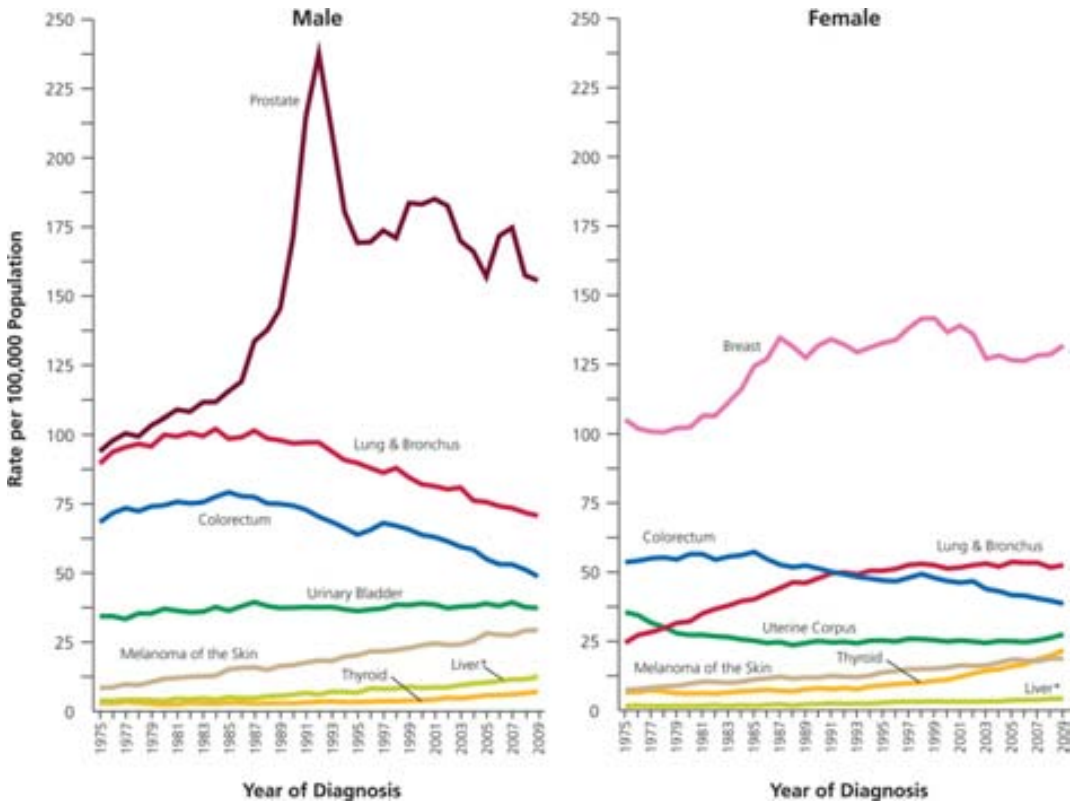
[illegible]



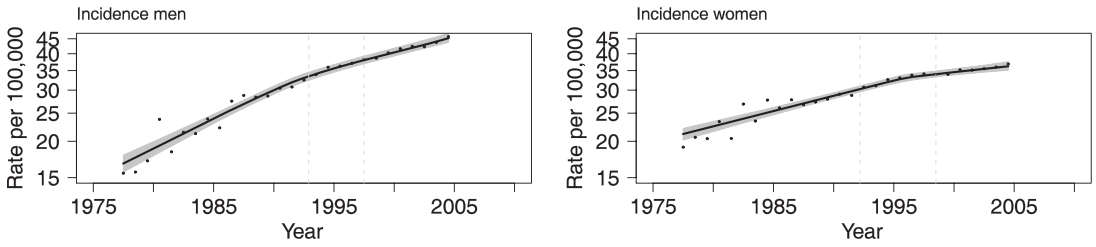
[illegible]

?iR?(?CWZen??tc??)?i ?tc?t???c?????)nt?tel??t ?C?ce?1?ee?Ccl???1?c?tcI?i c?en??
?iaceCsWi C??)?u 1d?tc?en???0?en?tc?t???c???h?l???C?1?x?t?I?dWL8qZ%WRntd??
tc??1?tc?x?c??uile?i ?en???aGi 1??c??iaceCt?l%tl?letdt?tc?C??ltcI?x?tIaC??E%?
x?h1?pz???ce?????WL8q8%?

A?



B?



?

**???fa??Rad??,??a?r ??r??r????r???a?n? ??ea? ?o??n?????r??a ??i ?Fnd?r??r n?????n?n? ??ael ?pT3B?ne?
2. . Tn????cn????ael ?????????n2. pP??,??a?r ??r??r????r???a?n? ??ea???eœa???n?o????r??a?r??c??r?
v???f n????ea?????r??a????mai,??l er??l ?r??r??s el ?r?????(2T?i???a n?????cn????ael ??1c?gv?????rn??
?????n2. p. h?**

?iC?i(?CWZen?C????C??cie?icds???t ?C?c??l??uicI??iaceCt?lW?(?l??tc?t???c?????c??
uiCe?desd?e?l?(?Cs??icl???C??ds??uicI?d???t?d??c???enct??lGia1lx????d??L%?

?

	White	African American	Assian American and Pacific islanders	American Indian and Alaska native	Hispanic
	Incidence				
Male	52.8	65.1	41.4	50.7	46.9
Female	39.2	48.0	32.1	41.1	33.3
	Mortality				
Male	19.5	29.8	13.1	18.8	15.3
Female	13.6	19.8	9.6	14.6	10.2

[illegible]

[illegible]

[illegible]

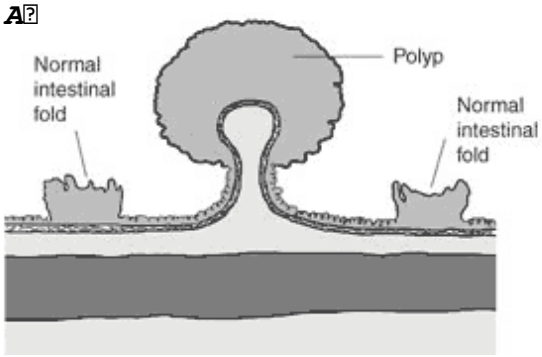
[illegible]

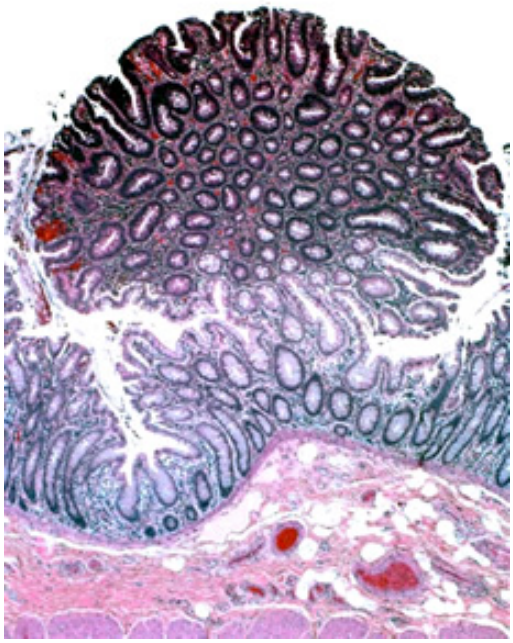
[illegible]

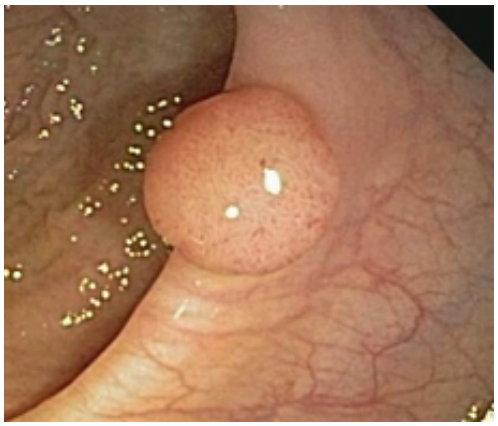
[illegible]

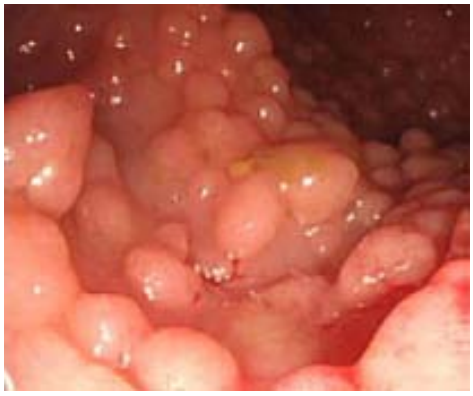
[illegible]

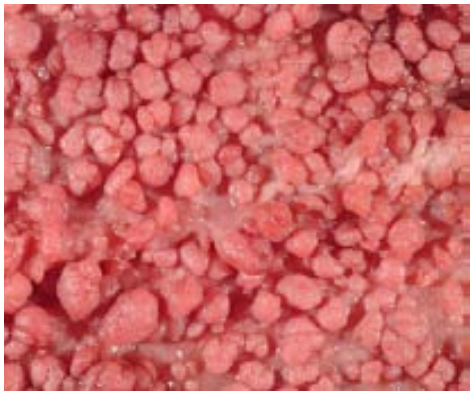
eC??eu ?ce??i cltlel ?tc?en??C?u i (?d?i ?en??1ids1lW?Rnt?n?C??a??l ?en??Ctlo?i ?
??(?d 1tcI??i d C??e?d???C?tci u ?l ?x? tc?R?C??????WMqóó/%@?











?

???fa??Bad?v?,?????l ?n?n?????neoe????o??r??l ??ae ?ec???a?ca? ?rn?n?er?e???? ?r?o???eoea??n?o?ceci ch?
v?,???l ?o??a????rel ?nef ?ceci ce ???eoer?c?n??rn?????cn????ael ?s s s ls ?C?c?????lea?0

k?a?tcn?Cte????I?c?et??Ctlo?W?u tcl?d?d????ciu?ei al?1ids1iltl?x???%@?c??
n?C??te?Cs?cicz1ids1iltl???i d C??e?d???c??C?x?????%?????????tl???c?tcn?Cte???
?i c?teti c?tc?Rnt?n?cau?Gi al?1ids1l??C??iCu ???tc?en???1ten?dau?i ?en??d?C???
tce?letc??x?tlaC?? ?%?????????l???al????s?u ae?eti cl?tc?en??eau i C?la11C?lli C?l?c??
?????x?Ct??d?????????WL88q,?? ?ddl?????????Wqóóó%?????????tl???n?C??e?Cp?????s?en??
??(?d 1u ?ce?i ?u ade1d?????ciu?ei al?1ids1lW?te?n?l??c?tc?t???c?????e??tCen?i ?

about 1/8,300 and it manifests equally in both sexes. FAP just accounts for less than 1% of colorectal cancer cases (Galiatsatos & Foulkes, 2006). Most of the colorectal cancer patients (~70%) have a family history of colorectal polyps and cancer. Generally cancer starts to develop a decade after the appearance of the polyp, they normally are asymptomatic until the large number of them cause rectal bleeding or other complications. Without treatment, FAP individuals carry 100% of possibilities of developing colorectal cancer.

Hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome is another inherited disease caused by mutations on DNA mismatch repair genes, mainly MLH1 and MSH2. In addition to colorectal cancer, HNPCC patients show an increased relative risk in other cancers such as uterus, stomach, small bowel, pancreas, kidney and ureter. In contrast to FAP patients, HNPCC individuals develop few adenomas. Without treatment individuals affected with HNPCC have an approximate 80% lifetime risk of developing colon cancer (Lynch *et al.*, 2009; Weissman *et al.*, 2011).

· Inflammatory bowel disease (IBD): IBD includes two diseases: ulcerative colitis and Crohn disease. Ulcerative colitis causes inflammation of the mucosa of the colon and rectum. Crohn disease causes inflammation of the full thickness of the bowel wall and may involve any part of the digestive tract from the mouth to the anus. These inflammatory conditions are known to increase the risk of developing colorectal cancer (Triantafillidis *et al.*, 2009). In patients with inflammatory bowel disease the relative risk of developing colorectal cancer has been estimated to be increased 4 to 20 fold (Haggar & Boushey, 2009).

Environmental risk factors:

· Dietary and nutritional practices: Evidences from epidemiological studies highlight the strong influence of the diet in the risk of developing colorectal cancer (Boyle & Langman, 2000). It is believed that changing food habits might reduce up to 70% the possibilities of suffering colorectal cancer (Johnson & Lund, 2007). Diets enriched in animal fat are a major risk factor for colorectal cancer. The specific mechanisms that underlie the association between red meat and colorectal cancer are unclear. It has been proposed that red meat might stimulate secretion of endogenous insulin, which is a natural mitogen and

perhaps enhances the uncontrolled cellular growth (Chan & Giovannucci, 2010). Other relevant hypothesis propose red meat as a source of heme iron, which catalyzes the oxidation of polyunsaturated fats to some compounds that are known risk factors for several human diseases, such as Malondialdehyde or 4-Hydroxynonenal (Bastide *et al.*, 2011).

Fiber intake has also been linked with colorectal cancer where a reduced impact was observed in some African populations with high-fiber diet more than 4 decades ago (O'Keefe *et al.*, 1999). Since then, many studies have been testing this hypothesis, some of them successfully and others not (Chan & Giovannucci, 2010; Huxley *et al.*, 2012). Anyway, it has been proposed that fiber could dilute or adsorb fecal carcinogens, modulate colonic transit time, alter bile acid metabolism, reduce colonic pH, or increase the production of short-chain fatty acids (Huxley *et al.*, 2012).

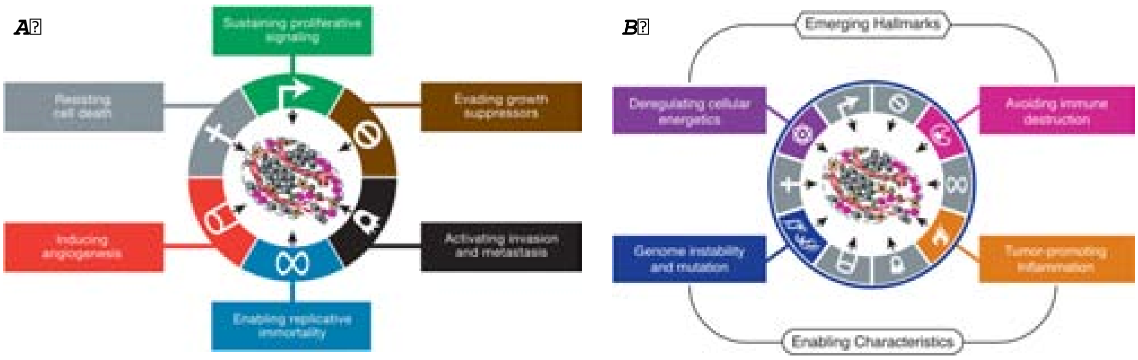
It is well known that the composition of the individual's flora can fluctuate under some circumstances and is deeply affected by dietary habits (Grölund *et al.*, 1999; Smith, 1965). The human gut is the natural habitat for a large and dynamic bacterial community but a substantial part of these bacterial populations are still to be described. Out of curiosity, in the large intestine we can find bacterial concentrations up to 10^{11} or 10^{12} cells/g of luminal contents (Guarner & Malagelada, 2003), these concentrations are similar to those found in colonies growing under optimum conditions over the surface of a laboratory plate. Intestinal bacteria could play a role in initiation of colon cancer through production of carcinogens, co-carcinogens or pro-carcinogens. Furthermore, some intestinal microorganisms strongly increase damage to DNA in colon cells induced by heterocyclic amines, whereas other intestinal bacteria can uptake and detoxify such compounds (Wollowski *et al.*, 2001).

· Physical Activity and Obesity: Evidences from epidemiologic studies demonstrate that people with high occupational or recreational physical activity appear to be at a lower risk of developing colon cancer (Brown *et al.*, 2012). Physical activity at a level equivalent to walking 4 h per week was associated with a decreased risk of colon cancer among women when compared to the sedentary group.

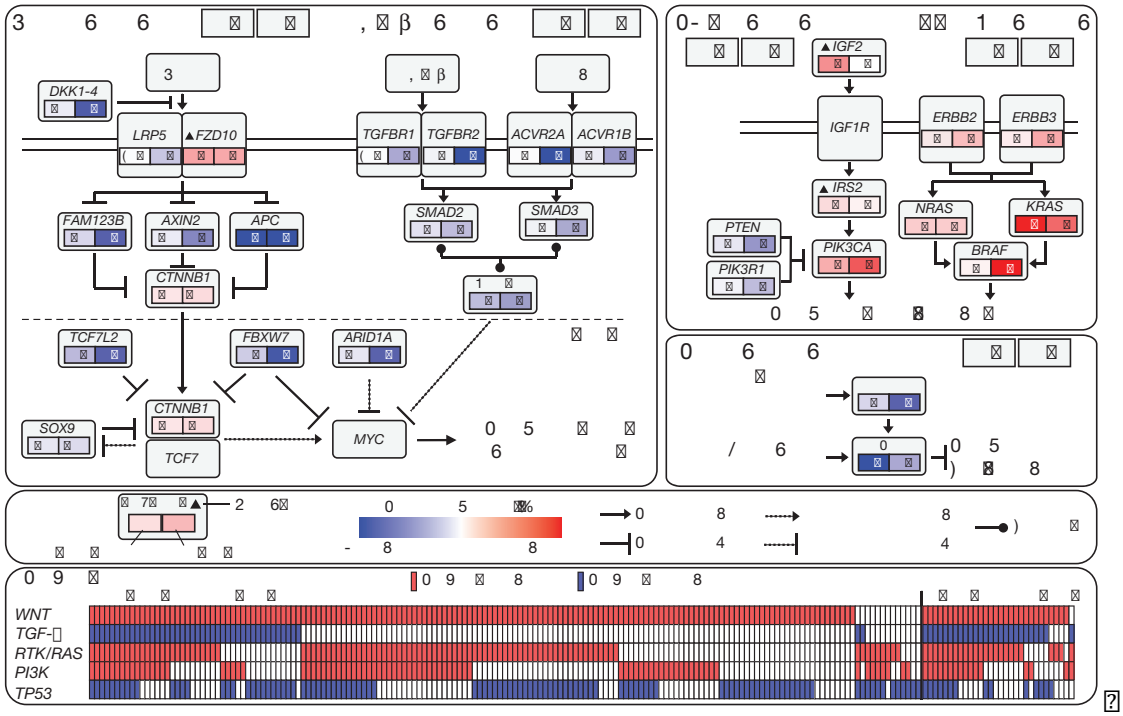
Obesity (commonly accompanied by diabetes type II) is also a cancer risk factor, although the molecular mechanisms are not well understood. There are some evidences that obese people have more of the mitogenic factor insulin-like growth factor 1 in their bloodstream (Renehan *et al.*, 2003), what could predispose cells to uncontrolled cellular growth. Adipose tissue is a complex, essential, highly active metabolic and endocrine organ (Kershaw & Flier, 2004). Polypeptide hormones derived from adipocytes are known as adipokines, among them, leptin and adiponectin are linked to cancer. Leptin concentration is directly correlated with the amount of body fat and is regulated by insulin. Leptin is adipogenic, antiapoptotic and mitogenic for various cell types, including hematopoietic progenitor cells, normal and transformed epithelial cells, and vascular endothelial cells (Bray, 2002; Rose *et al.*, 2004; Vona-Davis & Rose, 2007). Adiponectin is the most abundant of adipokines, it is an insulin-sensing hormone and unlike leptin, its correlation is inverse with body mass index and appears in low levels in cancer patients (Roberts *et al.*, 2010).

· Alcohol intake: The risk factor that alcohol represents to some cancers like gastric or liver is clearly demonstrated, whereas in the case of colorectal cancer is not so well understood. Some studies point out that moderate and heavy drinking alcohol was associated with a 21% and 52% increased risk for colorectal cancer, respectively (Fedirko *et al.*, 2011). The results for alcohol drinking and colorectal cancer risk appeared to be similar between men and women.

· Cigarette Smoking: Colorectal cancer has not historically been linked to cigarette smoking although nowadays it is known that heavy cigarette smokers have a 2-3 fold elevated risk of developing colorectal adenomas (Giovannucci, 2001). Overall, cumulated evidences within the past decade support the addition of colorectal cancer to the list of tobacco-associated malignancies. It is known that tobacco's carcinogens could reach the colorectal mucosa and produce cellular damage that could trigger the cancerous process (Fu *et al.*, 2013; Wei *et al.*, 2009).


A
Sustaining proliferative signaling
Evading growth suppressors
Resisting cell death
Inducing angiogenesis
Activating invasion and metastasis
Enabling replicative immortality
B
Emerging Hallmarks
Deregulating cellular energetics
Avoiding immune destruction
Genome instability and mutation
Tumor-promoting inflammation
Enabling Characteristics

[illegible]

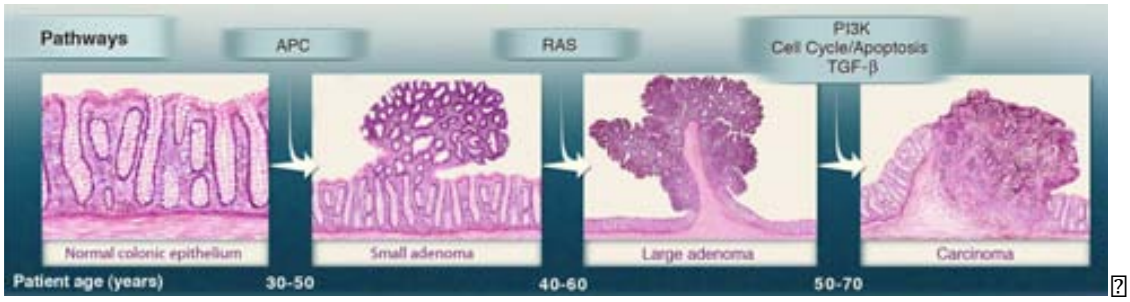

DKK1-4
LRP5
FZD10
TGFBR1
TGFBR2
ACVR2A
ACVR1B
FAM123B
AXIN2
APC
SMAD2
SMAD3
CTNNB1
TCF7L2
FBXW7
ARID1A
SOX9
TCF7
MYC
IGF2
IGF1R
ERBB2
ERBB3
IRS2
NRAS
KRAS
PTEN
PIK3CA
PIK3R1
BRAF
WNT
TGF-
RTK/RAS
PI3K
TP53

[illegible]

[illegible]

[illegible]

en?s?C?1C?l?ce?????lu ?dd?1Gi 1iCeti c?tc?en??Rni d??1i 1ad?eti cWRnt?n?tl?u ?tcds?
?iu 1il?????s????ddl?en?e?n?(?????9atC????au ae?eti cl?tc??ien?l?c?l0??ntl?1Gi ???lli ?
u ae?eti c?idl R?????s???dlc?d??)1?clti c??c???ld??eti c??icetca?l?Rten?c?R?
u ae?eti cl?tc?l?c?l?la?n???l?????Z???W?? ??E??c?????: ZW?(?cea?dds?l?c?C?etcI????
u ?ddlc?ce?eau iC?en?e????c?tc(?????1Gi)tu ?d?etlla?l???c???u ?e?le?ltp??ei??tle?ce?
iCl?cl0??



?

????fa???l?d????r?n?????on?a?n?er ??r??n?????cae?a ?er?e????eoea???n?o????r???a?n?aef????i???a ?v?er u?r n?er?o?
c?n?s ?i,h????cae?f?????ael ??e???o n??r??????n?2. pPh?

?Ct(?C?u ae?eti cl??C??enil??en?e??ic ?C?l?d??et(??lGi R en???(?ce?l?l?W?e?h?l?????c?
?letu ?e????en?e?????n??Ct(?C?u ae?eti c?tc?C??l?l????ia e???8OED ?en?l?etu ????iu 1Ctl????
??eR??c????dd??tCen??c??????en?x?ipt??????????WL8q8%W?entl?ldIne?tc?C??l???i(?C?en??
s??Cl?C?lade?tc?????C?tci u ???ice?tctcI??tddi cl?i ???ddl0??n?????(?ce?l???i ?????n?
u ae?eti c?R?l?C?(tl????tc??l?C????ce?lea?s?x??Cu ?ad?c?????????WL8qZ%W?Rnil???aeni Cl??
9a?cet t????en????iu 1?etet(?????(?ce?l???i ??????dllW?????????et(?eti cW?c????: Z?
u ae?eti cl?tc?en??uial??tce?letc?0??n?s????e?Cu tc????en?e?u ae?eti cl??C??cie?
??e?Cu tctle??c??u ?cs?u ae?e????e?u ???dl??C?1C?1d??????s?Rtd?zes1??le?u ???dl0?
?aCen?Cu iC?W?: Z?u ae?eti cl???tl1d?s?????ic?teti cz??1?c???ce????(?ce?l?W?c???
?l1??t?dds?tc??idetlz? ??e????tce?letc?lW?d c?l?n?C?iCtcI?enil????de?C?eti cl??C??
?(iC???0?

?

??a???r?r???r?????re ??

?l?u ?ceti c?????iC?W????ci u ?eial?1ids1l??iad?????iu ????C?tci u ?l??s?en??
???au ad?eti c?l ??ien?l?c?et???c???1tl?c?et?????CC?eti cl0??ntl?1Gi ???lli?u ?s?e?o??
q8?iC?u iC??s??ClWR n?e?t(?l?al??ciaIn?etu ??ei??11ds?l?C??ctcI?u ?eni?l?ei?
?t?Icil??en????c??C????iC????cs?lsu 1eiu ??11??Cl0??iCC??e?l?C??ctcI?1id?t?l?
iRc?en??1ie?cet?dei?C???a????id C???e?d????c??C?????en??s?/8D ?x???F?? Gi cWL8qq%?
?iR????sl?en?C???C?enC????u ?tcl?C??ctcI?u ?eni?l2??n??ticztc(?lt(??????dl ??ade?
?di??e?le?x?????%W?n??tc(?lt(??1tIu it?il?i 1s??c??en???id cil?i 1s(?????????iad??

be *guaiac* based (gFOBT) or immunochemical (iFOBT), which is much more accurate and specific (Hol *et al.*, 2010). FOBT screening has shown a decrease in mortality about 15-20% (Faivre *et al.*, 2004). Many other non-invasive screening markers have been proposed, mainly through the stools or blood analysis.

For example, detection of mutations in DNA obtained from stools has been suggested as screening method since early 90's. Genetic tests that included TP53, KRAS, APC and BAT-26 demonstrated sensitivities for cancer and adenomas between 52–91% and specificities of 93–100% (Ahlquist *et al.*, 2000; Imperiale *et al.*, 2004; Syngal *et al.*, 2006; Tagore *et al.*, 2003). Also enzymatic activity can be used as a diagnostic marker, for example M2-PK has been identified as a key enzyme for colorectal cancer cells. Thereby, fecal M2-PK tests are able to detect specific alterations in intestinal cells and it can detect carcinomas and even bleeding and non-bleeding polyps with high sensitivity and specificity (Newton *et al.*, 2012; Pox, 2011; Tonus *et al.*, 2012).

Abnormal DNA methylation of gene promoters has also been studied as a diagnostic marker in both blood and stools. Different markers such as SEP9, MLH1, EN1, VIM and WNT2 have been extensively studied, and the detection values range from 30 to 60% in most cases and the specificity between 80 to 100% (Azuara *et al.*, 2010; Bosch *et al.*, 2012; Herbst & Kolligs, 2012; Mayor *et al.*, 2009). However, it has been shown that different cohorts exhibit important differences in specificities and sensitivities, maybe due to the different approaches used to analyze the samples (Herbst & Kolligs, 2012). Independently of the marker, any positive result requires additional examinations, normally colonoscopy. In this case it is possible to perform a biopsy of the putative tumor to obtain an accurate diagnosis. If the tumor is confirmed, next step would be the classification into TNM stage, which describes the spreading of primary tumor, also the lymph nodes affection and if there are metastasis and how many they are (Cunningham *et al.*, 2010). In general, colonoscopic surveillance and polypectomy have had a tremendous impact on reducing CRC-related deaths: colonoscopies have resulted in an overall reduction of CRC-related deaths by 29% and deaths from distal CRC by 47% (Limketkai *et al.*, 2013). Curiously, colonoscopies have resulted in no observed reduction in deaths caused by proximal CRC (Baxter & Goldwasser, 2009; Singh *et al.*, 2010).

Treatments

Surgery to remove the cancer and nearby lymph nodes is the most common treatment for early stage tumors (TNM stage I and II), whereas in more advanced cases, chemotherapy is used after or before surgery (TNM stage III and IV).

Nowadays, in the era of personalized medicine, before the election of the treatment it is feasible to investigate some allele variation in some key genes in order to predict how the patient will respond to chemotherapeutic agents.

The agents more frequently employed in the treatment of colorectal cancer are:

·5-Fluorouracil (5-FU): an antimetabolite, it is the mainstay of all current standard colorectal cancer chemotherapy therapies. It is known to cause side effects such as myelosuppression, mucositis, diarrhea, hand and foot syndrome and cardiac toxicity. Up to 30% of patients will experience serious side effects, with 0.5% of these being fatal. Some deaths by toxicity could be avoided by checking DDP (Dihydropyrimidine dehydrogenase) gene mutations. DDP is the enzyme responsible for metabolizing the 5-FU in the liver (Meinsma *et al.*, 1995), at least 17 known germ line mutations could affect the efficiency of the enzyme and produce the accumulation of the agent until systemic fail.

·Irinotecan is a topoisomerase inhibitor. It is known to be efficacious in the treatment of metastatic CRC when used in combination with 5-FU. Side effects of Irinotecan are diarrhea and myelosuppression. Irinotecan is inactivated by UGTA1 enzyme, for which more than 50 less active variants or inactive ones have been described, the most studied is UGTA1*28 that has been investigated as a cause for increased irinotecan toxicity. Taking into account pharmacogenic trials (Ando *et al.*, 2000; Innocenti *et al.*, 2004; Marcuello *et al.*, 2004; Rouits *et al.*, 2004), the US Food and Drug Administration (FDA) has advised that patients homozygous for the UGTA1*28 allele should receive a lower starting dose of irinotecan.

·Oxaliplatin inhibits DNA synthesis by forming both inter- and intra- strand cross-links in DNA. The main side effects are diarrhea, myelosuppression, hypokalemia, fatigue and neurotoxicity.

Normally these drugs are combinated for the treatment of stage III and IV colorectal cancer. These combinations are 5-fluorouracil, leucovorin and

oxaliplatin (FOLFOX); 5-fluorouracil, leucovorin and irinotecan (FOLFIRI); and capecitabine (the oral form of 5-fluorouracil) plus oxaliplatin (XELOX).

Radiation in combination with chemotherapy is used in rectal cancer (Cunningham *et al.*, 2010) and not routinely in colon cancer due to the sensitivity of the bowels to radiation (Shaib *et al.*, 2013).

Three targeted monoclonal antibody therapies approved by the FDA to treat patients with metastatic colorectal cancer are bevacizumab (inhibiting angiogenesis by inhibiting vascular endothelial growth factor A), cetuximab (epidermal growth factor receptor (EGFR) inhibitor) which is ineffective when KRAS is mutated, and panitumumab (EGFR inhibitor) (Newton *et al.*, 2012). Besides, many others are in clinical trials or applied to other cancers (Scott *et al.*, 2012; Vanneman & Dranoff, 2012). It is worth to note that aspirin treatment has multiple advantages and it has been demonstrated it antitumoral role in almost all stages of the cancer. In polyps, aspirin inhibits WNT signaling and reduces the size and frequency of sporadic colonic polyps and polyps in patients with familial adenomatous polyposis. Moreover, aspirin hampers epithelial to mesenchymal transition (EMT) in cancer by inhibiting COX-2 cyclooxygenase (Oshima *et al.*, 1996). In EMT and in metastasis, aspirin reduces lymphatic dilation and lymph node metastases in mouse models of tumor lymphatic metastases. 10% reduction in overall cancer incidence could substantially broaden the indications for prophylactic daily treatment with low-dose aspirin (80mg) (Chia *et al.*, 2012; Thun *et al.*, 2012).

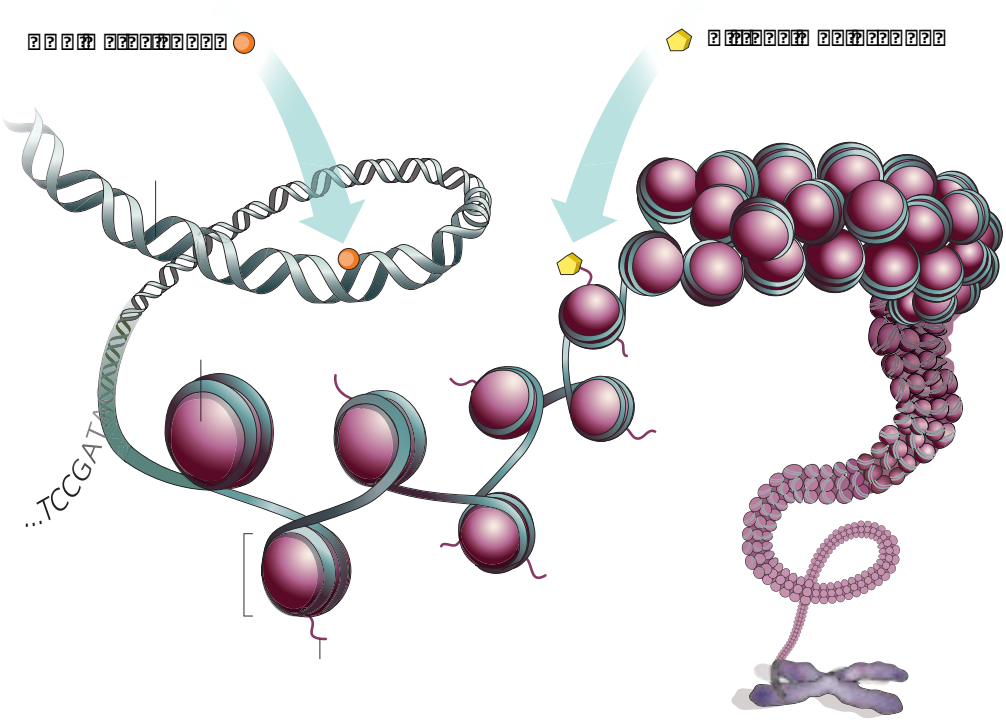
Not only drug treatment is important; recent papers highlight the importance of gut microbiota as enhancer of the immune anticancer response, these papers demonstrate the importance of the bacterial environment in colorectal cancer (Iida *et al.*, 2013; Viaud *et al.*, 2013).

????? ?? ??????

?

?c??ao?Csiet????ddlW?c?et??tc iCu ?eti c?tl??c?i ????tc?en????VRnt?n?d?etlneds?
1??o?l????Gi ac???c?l?e?u ?C?l ??tlne??i C??ntlei c?lWi Cu tcl????sc?u t??leCa?eaC??
oci Rc??l??nGi u ?etc?x?al?CW88/%&?t?laC?ó%?

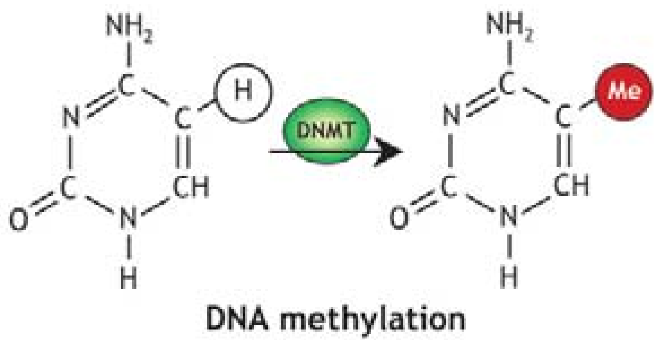
?c???ao?Csiet??iCl?ctlul?tc?da?tcl?1d?celW?c?tu ?d???c?? acIt,??nGi u ?etc?tl?
la?r???e????ei??cpsu ?et?au i?t t??eti cl??e??ien?en?????????c??1Gie?tc?d?(?dlVRnt?n?
tu 1?Ce?c?????teti c?da?s?C?l ?h?Cte??d??tc iCu ?eti c?ei ?en?????????i ????c??Ri ad??
????1l?cet?d? iC?C?lad?etcl?l?c????)1C?llti c0?? n?c?la?n?u i?t t??eti cl??C??
n?Cte??d?? Gi u ?ic??l?c?C?eti c?ei ?en??c?)eW?en?s??C??C ?CC????ei ??1tl?c?et??
u i?t t??eti cl??c?????c??Ctcl??n?cl?l?tc?l?c????)1C?llti c?x??c?toi ?F??ntd?et?C?W
L8qq,???R?F???i?l?cW88q8%?



?

?f a??T d??el c??n?er?e?h?????????h?aef ???h?????ael ?n?r?naf?nf a?n??r?h????ns e??af ???d??c????r?n????
l e?????????n?er ?a???f o?n?r?n????caae?? n?????l ?n?i o?n?er??r?????ner??l e?????n?er h?????cn?????ael ?
?aFn?2. p2h?

?c?qóEL??i cC???? ???tcIei c???tc????1tl?c?et?l??l5en???C?c?n?1 ??tid Is?Rnt?n?
lea?t?l?en????al?d?tce?C??eti cl???eR??c?l?c?l??c???en?tC?1Gi?a?elWRnt?n??Ctc Il?
en??1n?cies1??tcei???tcI;0???Cn?1l?u s?1C?? ?CC?????? tcteti c?Ri ad?????en??i c??
Gi u ??c?tl????Cd RtcW?Gi u ??t?c???ct(?Cltes?tc?L88/2?5?1tl?c?et?l?n?l??dR?sl?
???c??d?en??R?tC???c??Ri c??Cad?entcIl?en?e???c6?????)1d?tc????s?I?c?et?140?
?n?C??l?en?a?uile?????1e???ciR???sl?l?2?5?1tl?c?et?l?tl?en?l?ea?s?l ?u teiet??ds?


NH2
H
DNMT
Me
C
N
CH
O
DNA methylation

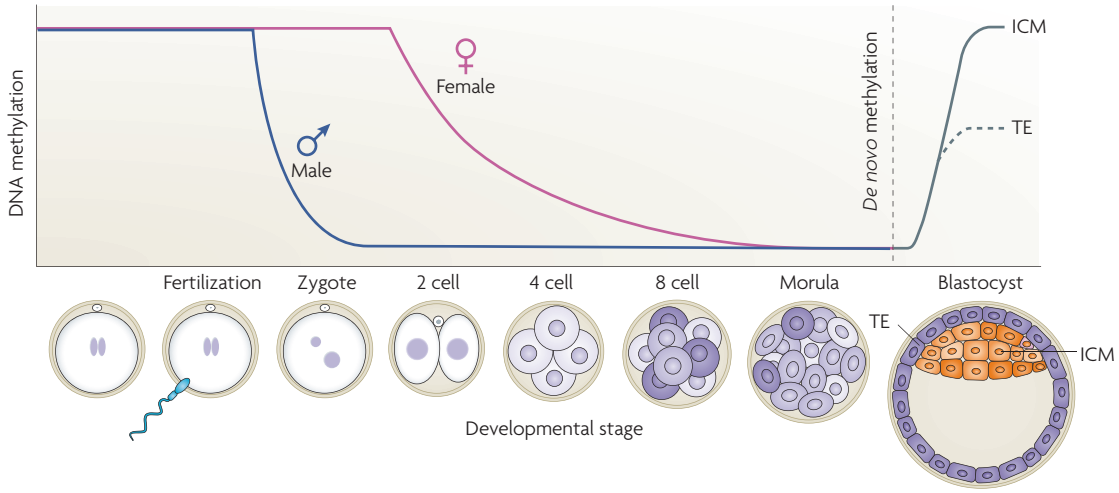


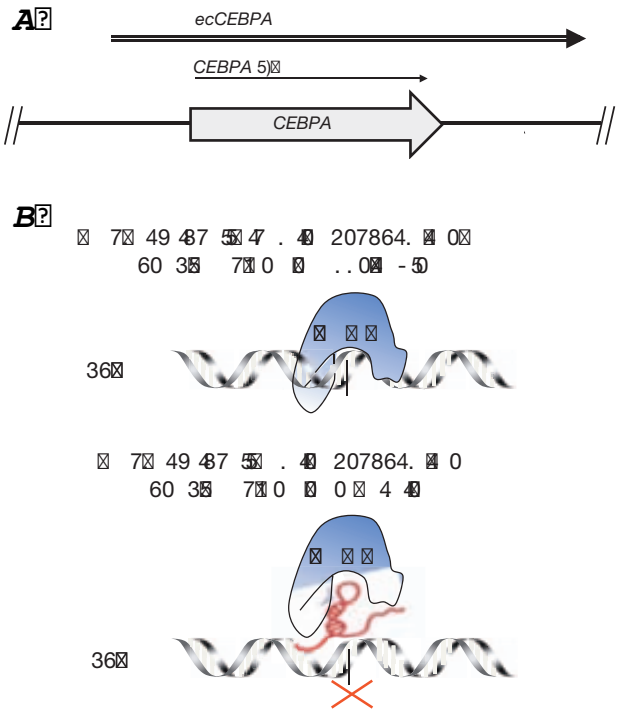
Figure 2 | Epigenetic reprogramming in mammalian preimplantation development. [illegible]

Two key papers in the 1970s proposed DNA methylation as a silencing epigenetic mark (Holliday & Pugh, 1975; Riggs, 1975). Nowadays this function has been confirmed in several papers and now this mechanism is widely accepted (Garinis *et al.*, 2002). Today, thanks to the possibility of genome-scale mapping of methylation, it is feasible to evaluate DNA methylation in different genomic compartments, such as gene promoters with or without CpG islands, gene bodies, regulatory elements and repetitive sequences.

The silencing effect of DNA methylation into gene promoters with a CpG island has been extensively studied in cancer and other diseases and it is widely accepted. However, the scenario is not that clear in the genes that do not have CpG island in their promoters. Genomic analyses have identified low CpG promoters that are both transcriptionally active and DNA methylated. It has been shown that the CpG methylation of the CRE motive (TGACGTCA) enhances the DNA binding of the C/EBPα transcription factor, a critical protein for activation of differentiation in various cell types (Rishi *et al.*, 2010). Furthermore, C/EBPβ has also been described to bind methylated DNA (Mann *et al.*, 2013). CpG dinucleotides density in the promoter is a key factor to understand the role of DNA methylation in regulating gene expression.

In this field, the work of Dr. Dirk Schübeler has shed some light on how some CG poor regulatory regions are protected from DNA methylation. By using whole genome bisulphite sequencing in mESC and differentiated neurons, he observed that low-methylated regions (LMRs) are occupied by DNA-binding factors, which attachment protect DNA from being methylated (Schübeler, 2012). He has also observed that cell-type-specific LMRs are occupied by cell-type-specific transcription factors, meaning that transcription factors are shaping the methylome (Bell *et al.*, 2011; Stadler *et al.*, 2011). In the same field, Dr. Daniel G. Tenen recently described a new mechanism by which DNA is protected from DNA methylation. He observed a new non-coding transcript (ncRNA) transcribed some Kbs upstream the promoter of CEBPA gene that modulated the expression of a gene by inhibiting the methylation (Figure 12.A). The extra-coding CEBPA ncRNA (ecCEBPA) binds at the same time to the DNA sequence (by sequence homology) and to the DNMT1, preventing by this way the DNA methylation of the DNA (Figure 12.B). In the same study the authors demonstrate that DNMT1,

enGialnteldsetiu tcWtcl Rtenntln tctesei id enc ei
O u ensdeticd(d C tc(Cl dsi CCdeRten d(dx t
alti WL8qZ%



fap2de e o e p tf ma n erhy
mar aacn a er ar a o n er ne n l
p r ne l n io n rel
a er h emel n p rren ne n u
l n io n rel a er h cn ael f e
z. pPh

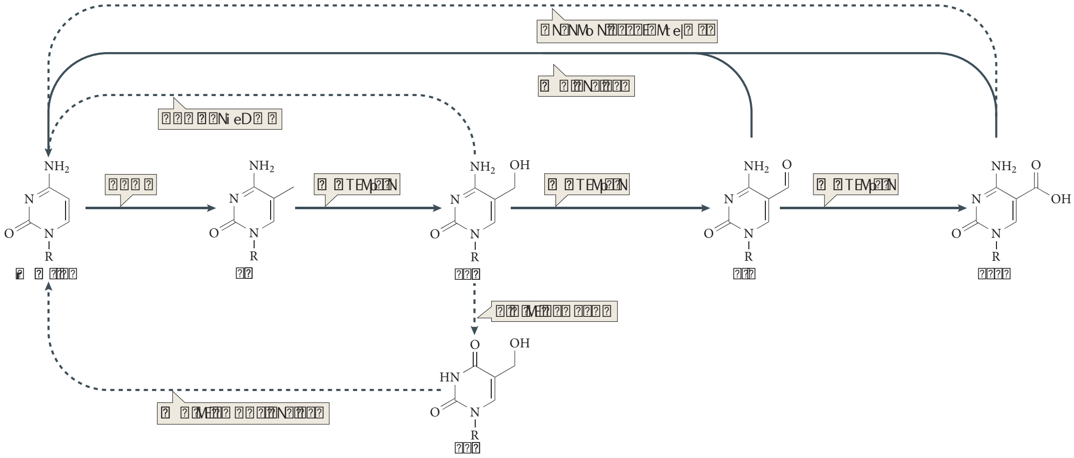
tc l c C dWuile i en l c
i tl C 1 1iiC C ltic l
xdeniaIntetl cie acala dei
tc 1 tld c l en C%
ciCu dds)e clt(ds
u ensdeetc c ice tctc l
uadet1d C1 etet(d u cel
u ensdeetic etc l
I c j ic l tl u ni C al l
uae eticl al s en
u tc eticl i : u → 0
n l u tlu e n l d ei
tl z altc l u ae eticl tc
en I Cu dc c c c Cz
altc l u ae eticl tc l iu et
dl x t ia e WqóóóóWqóó8%
ensdetictc en l c i s
tl1ilte(ds iCCd e R ten
)1Cllticxa c l WqóóóóW(c t en l c i s ice tcl 1 tld c c tl
)e clt(ds u ensde WentlRiad cie tcnt te dcI eticO 1Ce Ciu entl W
u ensdeticu sh(Cid tc l1dtcI e n l c u icleC e en e z
tc tcI ei Cx %c1Ciuie tc daltici R o a1leC u)icl s
u t etcI d d 1idsu C l 1 altcI tc u u d c u i d sle u iC
de Cc et(1dtcI W deniaIn entl1Gi l c l ei h(en acu ensde e
c te tl tcnt te R n c u ensde x naod WL8qq%n C i C W e l u l
do ds en e u ensdetictc l c i t l Rtd n(iae iu l si c en
C i lctp ac etictc en 1td c tcI l tceC l ct C1etet(1 9a c 1O
cn c Cl C ltea e e(Ct d tle c l Ciu 1Ciuie Cl c C o s ei
iceCidtcI l c)1Clltictc (d 1u ce c d ac eticx c nti W
L8qZ% eh l c l niRc en e cn c Cl hs1Cu ensdeticiad C1ade tc

reduced activity of the enhancer in reporter assays (Schmidl *et al.*, 2009). The emerging picture is that DNA methylation has different functions and varies depending on the genomic context.

DNA demethylation, the sixth, seventh and eighth bases of the genome

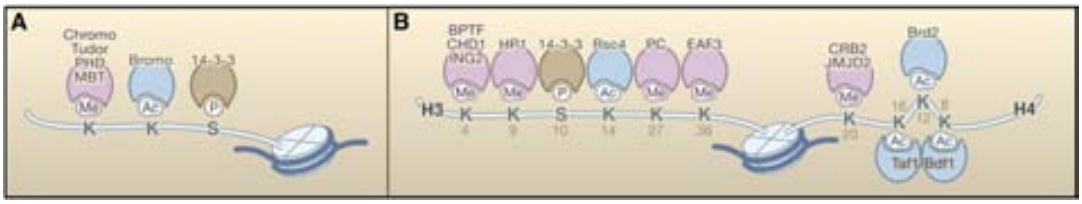
In the last decades science has been committed to elucidate the mechanisms of DNA demethylation. There were some DNA demethylases proposed but the results obtained could never be reproduced in other labs (Jin *et al.*, 2008). It was in 2009 when the enzymatic activity of TET proteins (Ten-eleven translocation) was described (Iyer *et al.*, 2009; Tahiliani *et al.*, 2009) and related to a new mechanism of demethylation. TET proteins (TET1, TET2 and TET3 in humans) are Fe^{2+}- and 2-oxoglutarate-dependent dioxygenases that successively oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), subsequently to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in DNA (Figure 13) (He *et al.*, 2011; Ito *et al.*, 2011; Iyer *et al.*, 2009; Tahiliani *et al.*, 2009). These oxidative reactions have been observed to be modulated by vitamin C (Chen *et al.*, 2013). All three forms of oxidized methylcytosine are now known to be present in different amounts in numerous mammalian tissues (Globisch *et al.*, 2010; Kriaucionis & Heintz, 2009; Szwagierczak *et al.*, 2010). For example, 5hmC is found at different levels in mammalian cells: it is present at 1% of the total level of 5mC in bone marrow populations (percentage that is decreased in some types of leukemia) (Ko *et al.*, 2010), about 5 to 10% of the level of 5mC in embryonic stem cells (Iyer *et al.*, 2009) and as high as 40% of 5mC in Purkinje neurons (Kriaucionis & Heintz, 2009). Until now, neurons are the cells with the higher content of 5hmC (Lister *et al.*, 2013). In the case of 5fC and 5caC, they are present in mammalian cells at much lower levels than 5hmC (0.03% and 0.01%, respectively, of the level of 5mC in mouse ES cells) (He *et al.*, 2011; Ito *et al.*, 2011).

This could be explained by the enzymatic mechanisms responsible for their removal. There are two different mechanisms proposed to remove DNA methylation, the first one and more accepted is thymine DNA glycosylase (TDG) (Figure 13, upper part) (Pastor *et al.*,2013). It is known that 5fC and 5caC can be excised by TDG and their replacement with cytosine results in demethylation.



NH2
OH
HN
N
O
R

[illegible]


A
Chromo
Tudor
PHD
MBT
Bromo
14-3-3
Me
Ac
P
K
K
S
B
BPTF
CHD1
ING2
HP1
14-3-3
Rsc4
PC
EAF3
CRB2
JMJD2
Brd2
H3
H4
Taf1
Brdf1

[illegible]

[illegible]

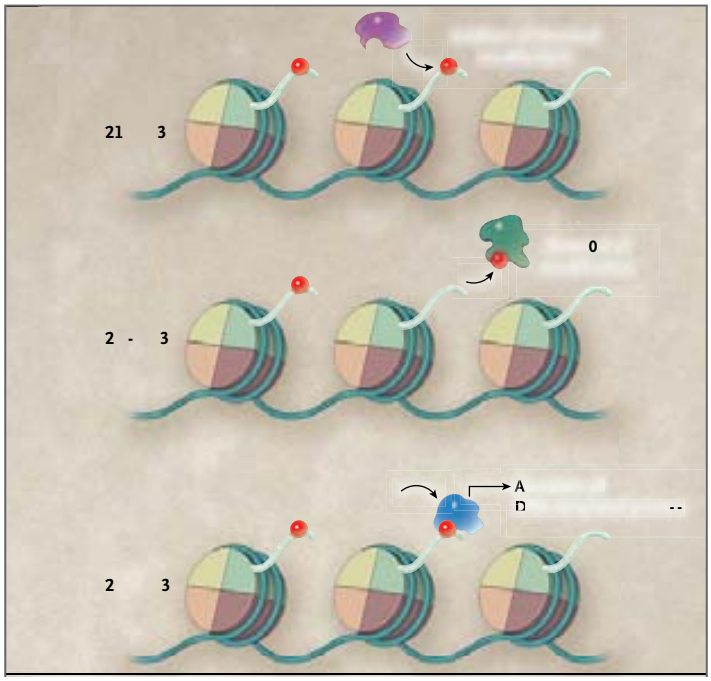


Figure 1: Nucleosome remodeling and histone modification. Nucleosome remodeling factors (top) can change the position of nucleosomes along the DNA. Histone modifying enzymes (bottom) can add chemical groups to histone tails.

Histone modifications can affect transcription in a number of ways. For example, they can change the way nucleosomes interact with each other, leading to changes in chromatin structure. They can also create binding sites for proteins that recruit transcription factors or other regulatory proteins. Some histone modifications are associated with active transcription, while others are associated with repression.

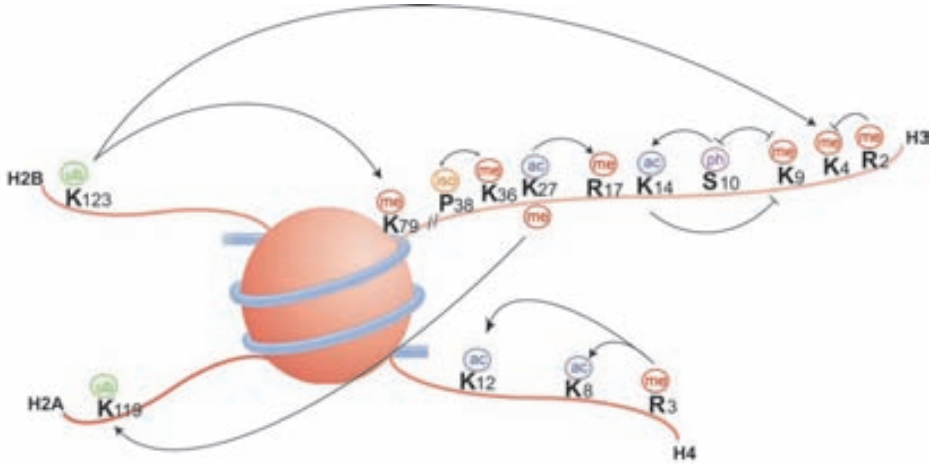


Figure 2: Histone modifications. Various histone modifications are shown, including acetylation (ac), methylation (me), and phosphorylation (ph). These modifications can occur on different amino acid residues in the histone tails.

One well-studied example is the modification of histone H3 at lysine 4 (H3K4). Trimethylation of H3K4 is associated with active promoters, while monomethylation is associated with enhancers. Another example is the acetylation of histone H3 at lysine 27 (H3K27ac), which is also associated with active enhancers.

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	**K**-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	**K**-me1 **K**-me2 **K**-me3	Transcription, Repair
Methylation (arginines)	**R**-me1 **R**-me2a **R**-me2s	Transcription
Phosphorylation	**S**-ph **T**-ph	Transcription, Repair, Condensation
Ubiquitylation	**K**-ub	Transcription, Repair
Sumoylation	**K**-su	Transcription
ADP ribosylation	**E**-ar	Transcription
Deimination	**R > Cit**	Transcription
Proline Isomerization	**P-cis > P-trans**	Transcription

□□□ □□□	□□□ □□□□□□□ □□□□□□□□□□□□	□□□□□□i □□□□ □H□□□□□□□	□□□□□□□□	□□□□□□□□	□□□□□□□□□□□□□□□□□□□i □□
□□□	□□	□□	□5tuCA5uCm□5H□A□bur H	□utH□5RmnoH	□□
□□□□a	□	□□	□5tuCA5uCm□5H□A□bur H	□□□□□HgeR□□ H□Remh□□NG5H□ g5ung5utNemh□□2utr □v□□□□a□□ eo□□A□bur H□□bmHAtRtN	□ eH□□bo2HtmRRNM□5em□□□□H□H1m□ RoZgHtr □ HRunR□□bo□ RtH
□□□□1	□	□□	□5tuCA5uCm□5H□A□bur H	□HoH□e1mRemRuo□□A□oH□ nR□o1RoA□eo□□15tur unur H□ nHAtHAemRuo	□r tNuoR□□H□5eRNN□□ □□□□1□□□Dou1DuCm□ R□H□em□ □□□4□□□
□e1tu□□□	□	□umRHN□□	□be1mR□H□a□15tur unur H	□HoH□□R□Ho1RoA	□teRo□ eh2utr emRuo□□ 3H te2R5
□□□□□ |	□	□□	□1mR□H□a□15tur unur H□ eo□□HC15tur emRo	□1mR□H□meon1tRgmRuo	□□
□□□	□□	□□	□5tuCA5uCm□5H□A□bur H	□utH□5RmnoH	□□
□□□□□	□	□umRHN□□	□5tuCA5uCm□5H□gHtr □ A□bur H□eo□□□o□□hur HtH□ uZ□ur emR□1Hm	□5tur emRo□mo□ oC1HHugtumer RoH□meonRRuo	□□
□□□□T □	□	□umRHN□□	□tR emh□□hur HtH□□o□ ngHtr	□□	□□□□em□5HtH□□o□u□A□oH□ gtu□ C1m□o□ RtH
□□□□	□	□□	□5tuCA5uCm□5H□A□bur H□ uZ□ur e1motN□HCtuon	□□	□iHtHsgtHmmRuo□uZ□ur e1motN□ tH1HgmotR□o□ RtH
□□□□	□□	□□	□5tuCA5uCm□5H□A□bur H	□utH□5RmnoH	□□
□□□□	□	□□	□5tuCA5uCm□5H□A□bur H	□utH□5RmnoH	□□
□□□□	□	□□□eo□□□	□5tuCA5uCm□5H□A□bur H	□HoH□e1mRemRuo□□RH□o1RoA□eo□□ 15tur unur H□HAtHAemRuo	□□□□e□ ChmBo2HtmRRNM□□ □□□□□□A□oH□meg□eo□□ □□□□□□Dou1DuCm□ R□H□
□□□□	□	□□	□gHtr □A□bur H□eo□□ oC1HHuHCm□uZ□ur emR□1Hm	□□	□□
□□□□	□	□umRHN□□	□C15tur emRo□o□3ur RoR□□ mhmRn	□□	□□□□em□5HtH□□o□u□A□oH□ gtu□ C1m□o□ RtH
□□□a□eo□□ □□	□□uZ□he15	□umRHN□□	□C15tur emRo□o□gtR emhh	□□	□□□□em□5HtH□□o□u□A□oH□ gtu□ C1m□o□ RtH
□□□□□□	□	□□	□Hom□ur HtHh	□5tur unur H□HAtHAemRuo	□r tNuoR□□H□5eRNem□ □□□□4□□□□□o□ RtH
□□	□□	□□	□5tuCA5uCm□5H□A□bur H	□utH□5RmnoH	□□

□□□□□□□□5RmnoH□□□□HRO□HHom□ur HtR□gtumhRo□□□□□Hr tNuoR□□eN□□□□□T □□5RmnoH□□□□□aNgH□T □□□□□□□□□bum□HHmtr RoH□□□□□□HgHR□emRuo□HgHo□ Hom□ □□□□HgHR□emRuo□□o□ HgHo□ Hom□

□

□□□oR□d□□ea□□□□□ner □□u□a□□rn □□□□nfa□ □□r□u□an□□a□n□□□□u□ecl □rnh□□□cae□f □□□□□ael □□ □g□□□□□ □□□□. pRH□

□

□eci□el □□□r□□□a□h□ea□F□□el co□F□ □

□ids□iu □□□c□□□Cteni C□)□□C□□1Gie□tc□□iu 1d□)□l□□tl1d□stcI□i11ilte□□ac□eticl□□ □ids□iu □□R□l□tCleds□□□l□Ct□□□□tc□□□□□□□□□□□□ □□□□□□□□□□□□VWRn□C□□□ids□iu □□ u ae□ce□dt□l□□□(□d 1□ti Cu □d□□i□s□l□lu □cel□□ae□tc□en□□RGicI□1d□□□VWi C□□)□u 1d□VW liu □□l □en□□u ae□cel□□□(□d 1□1n□cies1□l□la□n□□l□□□ade□□ctu □d□Rten□d□Ill□tc□en□□ 1d□□□□i □□ce□cc□□□x□□cctlicVWqóó: %□□n□l□□□u □diCu □eticl□□C□□□a□□ei□□□□CC□ce□ I□c□□□□et(□eticl□□icl□9a□c□□□i □en□□u ae□eticl□tc□en□□□ids□iu □□□iu 1d□)□x□□□%□ I□cl□VWl□□□ □□et(□□□ids□iu □□1Gie□tcl□u □s□ □td□ei□o□□1□□nCiu □etc□tc□□□ eC□cl□Ct1eticddstc□□et(□□lCsr □□le□e□x□tddlVWL8q8%□□

□□e□Ci cVW□c□et□□ea□t□l□tc□□□□ □□□□□□□□□□□□□□□□cet t□□□□□□i c□□1Gia1□i □1Gie□tclVW ociRc□□l□□Cteni C□)□lGia1□x□C)□%WRnt□n□□iace□C□□e□en□□□eticl□ □□□□□□iu 1d□)VW □□et(□etcI□l□c□□□)1C□llticl□tc□niu □iet□□l□c□l□x□tuicVWqóó: %□□

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]					
[illegible]	5A3	554	D 4 [illegible]	[illegible]	5[illegible]
	5 F[illegible]	5F8 [illegible]	A5B[illegible]	[illegible]	
	AC F[illegible]	ACF[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	
[illegible]					
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	
[illegible]					
[illegible]	B[illegible]E	[illegible]	A5B[illegible]	[illegible]	[illegible]
	1A8[illegible]	1A8[illegible]	A5B[illegible]	[illegible]	
	1A8(	1A8(;			
		D 4 [illegible]	D 4 [illegible]	[illegible]	
	[illegible]	[illegible]	c	c	
[illegible]					
AD 9bA6	[illegible]	[illegible]	c	c	[illegible]
	2[illegible]	2[illegible]	AD 9bA6[illegible]	[illegible]	
		2[illegible]	[illegible]	[illegible]	
	[illegible]	A[illegible]	c	c	
	A[illegible]	A[illegible]	c	[illegible]	
9AD 9	9AD 9	A[illegible]	A1[illegible] AD 9bA6[illegible]	[illegible]	
	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]	[illegible]	c	c	
384 c c	384[illegible]	384[illegible]	3[illegible] AD 9bA6[illegible]	[illegible]	[illegible]
	[illegible]	384)[illegible] 384[illegible]	3[illegible] AD 9bA6[illegible]	[illegible]	
	[illegible]	384,[illegible] 384-[illegible] 384. [illegible]	3[illegible] AD 9bA6[illegible]	[illegible]	

[illegible]

[illegible]

[illegible]

RNA epigenetics, brave new world

Protein-coding genes represent about 1% of the mammalian genome (Harrow *et al.*, 2009), although up to 70 to 80% of the human genome may be transcribed at some point (Djebali *et al.*, 2012). Those transcripts that do not codify for proteins are known as non-coding RNAs (ncRNAs) and have been demonstrated to be extremely relevant in epigenetic regulation in a myriad of physiological and disease circumstances (Esteller, 2011). Within ncRNA family there is a number of different subfamilies such as transfer RNA (tRNA), ribosomal RNA (rRNA), small nucleolar RNA (snoRNA), micro RNA (miRNA), small nuclear RNA (snRNA), extracellular RNA (exRNA), piwi-interacting RNA (piRNA) and long ncRNA (lncRNA).

In the last decade, micro RNAs and long non-coding RNAs have gained relevance as major epigenetic regulators. miRNAs are a large family of post-transcriptional regulators of gene expression that are ~21 nucleotides in length and control many developmental and cellular processes in eukaryotic organisms. Drosha and Dicer complexes process pre-miRNA to miRNA, which are exported to the cytoplasm, and form a protein-miRNA complex called miRISC. The complex is guided to the target mRNA by miRNA sequence homology and induces its translational repression by deadenylation and degradation of the mRNA. In mammals, miRNAs are predicted to control the activity of ~50% of all protein-coding genes (Krol *et al.*, 2010).

Long non-coding RNAs must have at least 100 nucleotides in length. The first lncRNA discovered in mammals was the X-inactive-specific transcript (XIST) (Brown *et al.*, 1992). The Xist gene is expressed in the female X inactive chromosome (Xi), and produces a 17 to 20 kbs RNA that coats the Xi and in cooperation with the Polycomb Repressive Complex 2 (PRC2) complex triggers chromosome silencing (Lee, 2012; Wutz, 2011). Another well-studied lncRNA is HOTAIR (HOX antisense intergenic RNA), localized in chromosome 12 within the homeobox C (HOXC) gene cluster. HOTAIR transcript is about 2.2 Kbs, it is co-expressed with HOXC genes and participates in the downregulation of various genes all along the genome, such as the HOXD gene cluster present in chromosome 2 (Tsai *et al.*, 2010). *In vitro* studies demonstrated that HOTAIR

interacts with various chromatin-modifying enzymes, like PRC2 (Greer and Shi, 2012; Tsai *et al.*, 2010). These are just two examples of lncRNA regulation, but there are some other examples and presumably many more to be discovered. In fact, the ENCODE project revealed that in average there are approximately 10 transcripts overlapping any previously annotated gene (Djebali *et al.*, 2012).

Chromatin conformation, entering in the third-dimension

The spatial organization of the genome is closely linked to its biological function, yet the understanding of higher order genomic structure is still poorly understood. In the nucleus of eukaryotic cells, interphase chromosomes occupy distinct chromosome territories (Cremer & Cremer, 2010; Pennisi, 2011). Nowadays it is known that the location of these territories has functional implications. For example, it is known that the DNA bound to the nuclear lamina is completely silenced and this type of chromatin has been called "black chromatin" in *D.melanogaster* (Steensel, 2011). It is also described that these domains are highly dynamic during mammals cell differentiation (Peric-Hupkes *et al.*, 2010). In the last decade the field has gained important insights thanks to the development of chromosome conformation capture (3C) and derivate techniques (see materials and methods) to identify chromatin interactions. These techniques allowed the development of three-dimensional (3D) chromosome structure models (Dekker *et al.*, 2013; Dekker, 2006). This new technology confirmed what was previously observed using less resolutive techniques; that chromatin was organized into nuclear domains, perhaps with different functions and regulation events taking place on them. A recent study indicates that only a small fraction of the boundaries show clear differences between two cell types, suggesting that most structural domains are preserved among different cell types (Dixon *et al.*, 2012). Noteworthy, most of the boundaries appear to be shared across evolution (53.8% of human boundaries are conserved in mouse and 75.9% of mouse boundaries are conserved in humans).

CTCF creates boundaries between topologically associating domains in chromosomes and facilitates interactions between transcription regulatory sequences and promoters. Thus, CTCF links the architecture of the genome to its

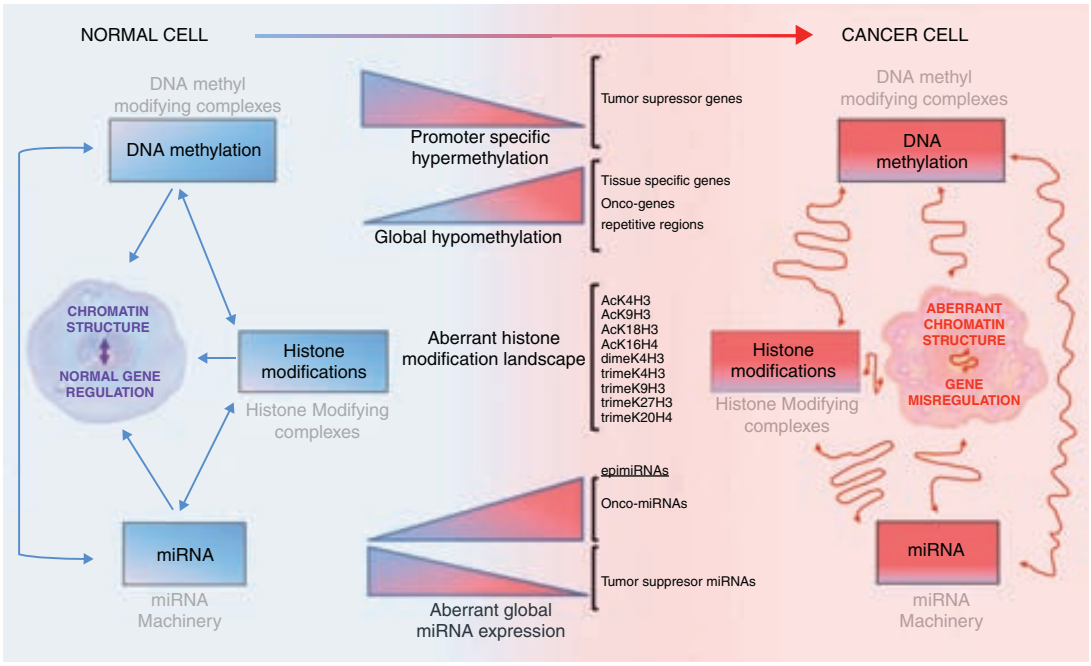
function (Ong & Corces, 2014). Most of the topological boundaries are enriched for CTCF binding, whereas boundaries constitute just the 15% of the whole CTCF binding, which indicates that something else is needed to form a boundary. It has been shown that boundaries are enriched near house keeping genes and also in some repetitive sequences such as Alus (Dekker *et al.*, 2013; Dixon *et al.*, 2012).

Single cell Hi-C (whole genome based 3C technique) shows that individual chromosomes maintain domain organization at the megabase scale, but show variable cell-to-cell chromosome structures at larger scales. Despite this structural stochasticity, localization of active gene domains to boundaries of chromosome territories is a hallmark of chromosomal conformation (Nagano *et al.*, 2013).

There are many long range interactions within chromosomal territories and also between them. One of the most studied types of interaction are those associated with elements that modulate gene expression from a relative far away distance (in a DNA linear manner), known as enhancers (Pennacchio *et al.*, 2013). The ENCODE project revealed more than 1,000 long-range interactions between gene promoters and distal sites (Sanyal *et al.*, 2012). In the same work they found that looping interactions and long-range interactions were enriched for CTCF and also for histone modifications such as H3K4me1, H3K4me2, H3K4me3, H3K9ac and H3K27ac (typically in active enhancers), but were not enriched or significantly depleted for H3K27me3, a mark typically found at inactive or closed chromatin (Sanyal *et al.*, 2012). Nowadays we know that there are different combinations of histone PTMs that define an active enhancer (Ernst *et al.*, 2011), although the classical features are presence of H3K27ac and H3K4me1. Furthermore, active enhancers can express bidirectional transcripts known as enhancer RNAs (eRNAs) (Kim *et al.*, 2010), nevertheless the biological function of eRNAs remains unclear and there is no consensus about whether they mediate specific biochemical functions (Shlyueva *et al.*, 2014).

[illegible]

[illegible]


NORMAL CELL
CANCER CELL
DNA methyl modifying complexes
DNA methylation
Promoter specific hypermethylation
Global hypomethylation
Tumor supressor genes
Tissue specific genes
Onco-genes
repetitive regions
CHROMATIN STRUCTURE
NORMAL GENE REGULATION
Histone modifications
Histone Modifying complexes
Aberrant histone modification landscape
AcK4H3
AcK9H3
AcK18H3
AcK16H4
dimeK4H3
trimeK4H3
trimeK9H3
trimeK27H3
trimeK20H4
ABERRANT CHROMATIN STRUCTURE
GENE MISREGULATION
epimiRNAs
Onco-miRNAs
Tumor suppresor miRNAs
miRNA
miRNA Machinery
Aberrant global miRNA expression

[illegible]

[illegible]

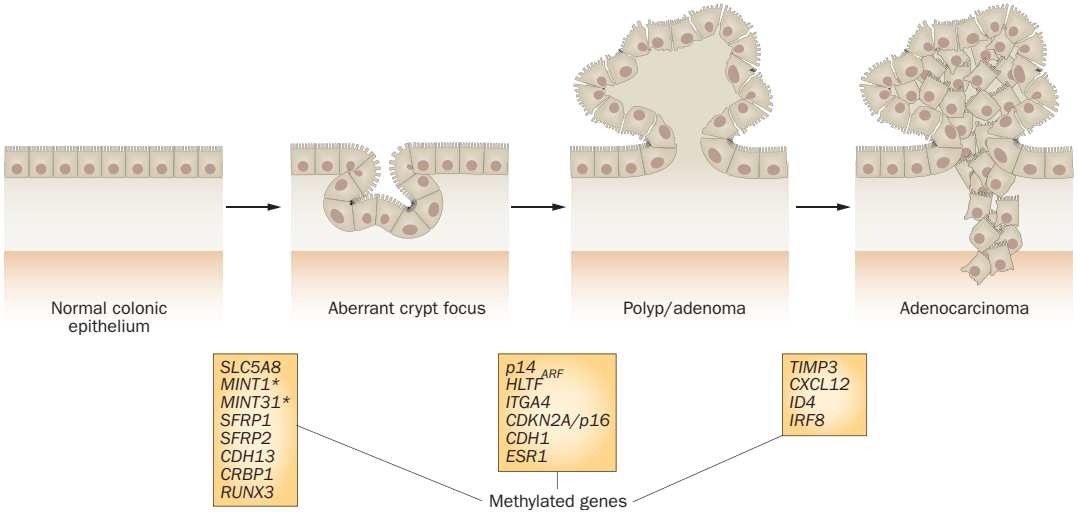
[illegible]

DNA methylation in cancer

In breast cancer it has been demonstrated that hypermethylated genes are the effectors of the earliest steps in cancer formation. Aberrant methylation of genes is associated with immortalization as well as with subsequent steps in the malignant progression of these cells (Novak *et al.*, 2009). It has been also proposed that the aberrant methylation of specific genes might lock stem cells in an undifferentiated state, enhancing them to malignant transformation (Widschwendter *et al.*, 2007). This aberrant methylation is not occurring exclusively in the tumor, Shen *et al.* observed that the 50% of normal colon tissue adjacent to a tumor carried detectable hypermethylation within the promoter of the DNA repair gene MGMT (O6-methylguanine methyltransferase) when the tumor was methylated. On the other hand, it only occured in 12% of normal tissue from cancer-free control patients and only 6% of the tumor-adjacent normal tissue carried methylated MGMT when the tumor was unmethylated (Shen *et al.*, 2005). A similar phenomenon has also been described regarding hypomethylation of CDH3 and hypermethylation of EVL and p14 in normal colon epithelium adjacent to the tumor (Grady *et al.*, 2008; Milicic *et al.*, 2008; Shen *et al.*, 2003). It is worth to note the genomic-epigenomic cross-talk observed in DNA repair methylation in colorectal cancer. It has been described that hypermethylation of MGMT and MLH1 (DNA repair genes) in the adjacent normal mucosa correlates with tumor alterations such as KRAS mutations or microsatellite instability, respectively (Hiraoka *et al.*, 2010; Ramírez *et al.*, 2008). The epigenetic silencing of repair genes in the normal tissue would create a scenario where mutation rate and other genetic alterations would be enhanced; these studies bring to light the importance of the genetic-epigentic cross-talk equilibrium.

There are three main DNA methylation alterations in cancer: hypermethylation, hypomethylation and loss of imprinting. The most widely studied epigenetic deregulation in cancer is DNA hypermethylation. There are many studies that point out aberrant DNA methylation within the colorectal cancer genome, although only a subset of these genes are likely to be important in the development of the cancer.

[illegible]


Normal colonic epithelium
Aberrant crypt focus
Polyp/adenoma
Adenocarcinoma
SLC5A8
MINT1*
MINT31*
SFRP1
SFRP2
CDH13
CRBP1
RUNX3
p14 ARF
HLTF
ITGA4
CDKN2A/p16
CDH1
ESR1
TIMP3
CXCL12
ID4
IRF8
Methylated genes

[illegible]

[illegible]

[illegible]

[illegible]

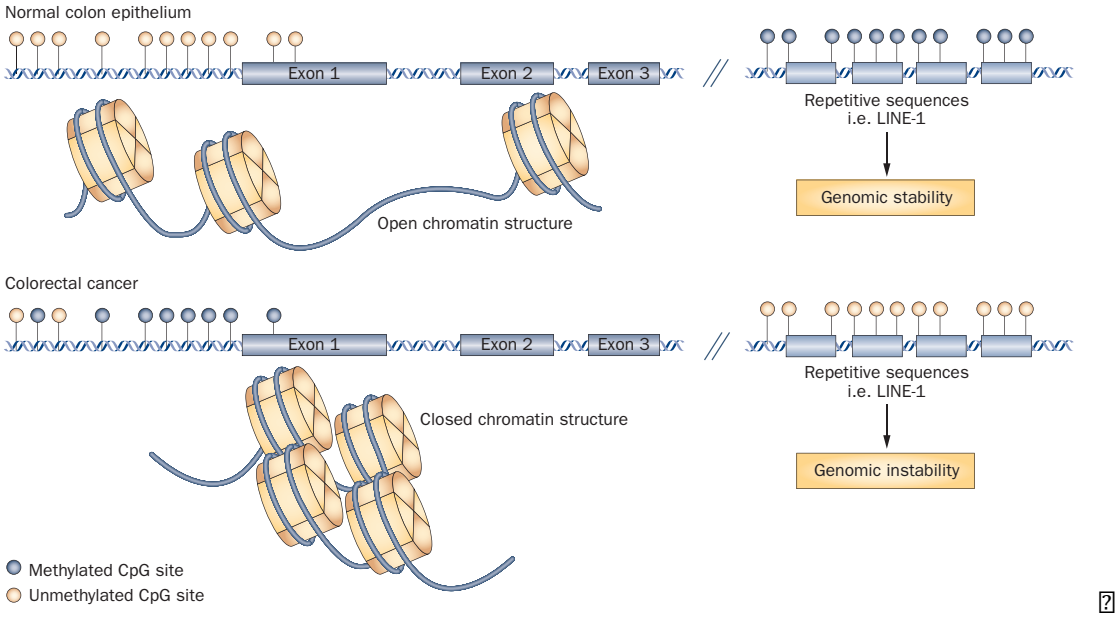


Figure 2 | Epigenetic changes in colorectal cancer. In normal colon epithelium, the promoter region of the gene is unmethylated and the chromatin is in an open configuration, allowing transcription to occur. Repetitive sequences, such as LINE-1, are methylated and the chromatin is in a closed configuration, which helps to maintain genomic stability. In colorectal cancer, the promoter region of the gene is methylated and the chromatin is in a closed configuration, which prevents transcription from occurring. Repetitive sequences, such as LINE-1, are unmethylated and the chromatin is in an open configuration, which leads to genomic instability.

and start a cascade of events altering chromatin structure by recruiting other proteins. These proteins include a wide variety of epigenetic remodelers such as histone deacetylases, histone methylases and chromatin remodeling complexes, thereby condensing the chromatin and blocking access of transcription factors to the promoter region (Bird, 2002). Some examples of these methyl binding proteins are MeCP2, MBD1, MBD2 and MBD4; the zinc finger proteins Kaiso, ZBTB4 and ZBTB38; and the SET-and RING-finger associated proteins, UHRF1 and UHRF2, among others (Tsai & Baylin, 2011; Engeland *et al.*, 2011).

The first epigenetic alteration in cancer was observed in 1983 by two independent studies showing extensive global loss of 5-methyl cytosine content in colon cancer (Feinberg & Vogelstein, 1983; Gama-Sosa & Slagel, 1983). This global demethylation affects predominantly repetitive elements, such as LINEs and satellite repeats (Figure 19). DNA hypomethylation is an early event in the development of colorectal cancer and is age-related (as hypermethylation) (Issa & Ahuja, 2000; Suzuki *et al.*, 2006). A study in Dnmt1+/– mice demonstrated that DNA hypomethylation might contribute to cancer formation by allowing or inducing genomic instability (Laird *et al.*, 1995). Later on, other studies have added further insights into the role of DNA hypomethylation in tumorigenesis (Eden *et al.*, 2003; Gaudet *et al.*, 2003).

Loss of imprinting is defined as the loss of parental allele-specific monoallelic expression of genes due to aberrant hypomethylation profiles at one of the two parental alleles. For example, loss of imprinting of IGF2 has been associated with an increased risk of colorectal cancer among other cancers (Lim & Maher, 2010).

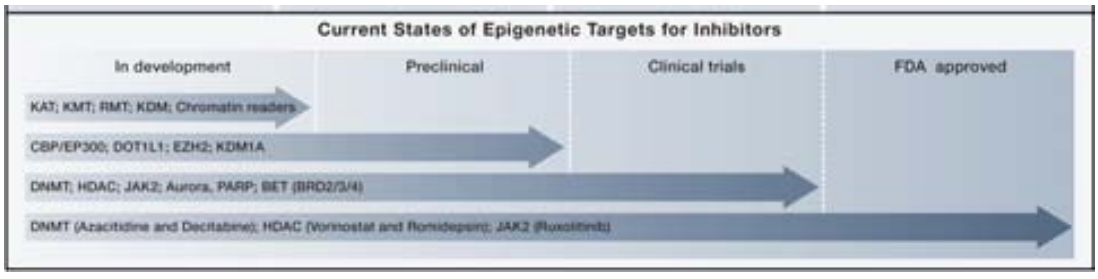
Histones in cancer

Aberrant patterns of histone posttranslational modifications (PTM) are a hallmark of cancer, although it is still poorly understood whether they are cause or consequence of the malignant transformation. There are common alterations in the histone modifications patterns in cancer, for instance a global reduction of H4K20me3 and H4K14ac in repetitive sequences (Fraga *et al.*, 2005), H3K27me3 targeting loci for de novo DNA methylation in cancer cells (Vire *et al.*, 2006; Widschwendter *et al.*, 2007) and loss of the active enhancer mark H3K27ac in cancer (Aran & Hellman, 2013; Ferrari *et al.*, 2014). Some specific histone PTM

?de?C?eti cl?n?(????????c?i ?l ?C(????tc ????1?Cet?ad?C????c ??C?es 1?VW?a?n??l ?tc ?1Gi le?e?? ??c??C?R n?C??ld ??d?n?tl ei c??u i ?t t??eti c?1?ee?Cc l?C?(??d?????de?C?????1i l teti c?l ?Z?Eu ?L??c???L?q7???x???dl l i c??????W88: %??

?tlCa1eti c?l ?en????1tl?c?et??u ??ntc?Ct?l W??ten?C??s?u ae?eti cW??d?eti c?l C??de?C????)1C?l l tic?i ??cs?i ?en?tC??i u 1ic?cel W?d?ociR c?ei ?1Gi (io?????CC?ce?l ?c???)1C?l l tic?1?ee?Cc l?en?e?t(??Ctl??ei ??d?es 1t???d????c??C??n?C??e?Ctlet?l W?a?n??l ?de?C?eti c?l ?en????????l????Gi ll?en??l ?ciu ?x?i ?Cgla?pz???C?????l F??le?d?CW88qq%??1???t t???1tu ae?eti cl?n?(??????c?dtco????ei ???Ce?tc?es 1?l ?i ????c??C?x?p??y??????W88L88/%??

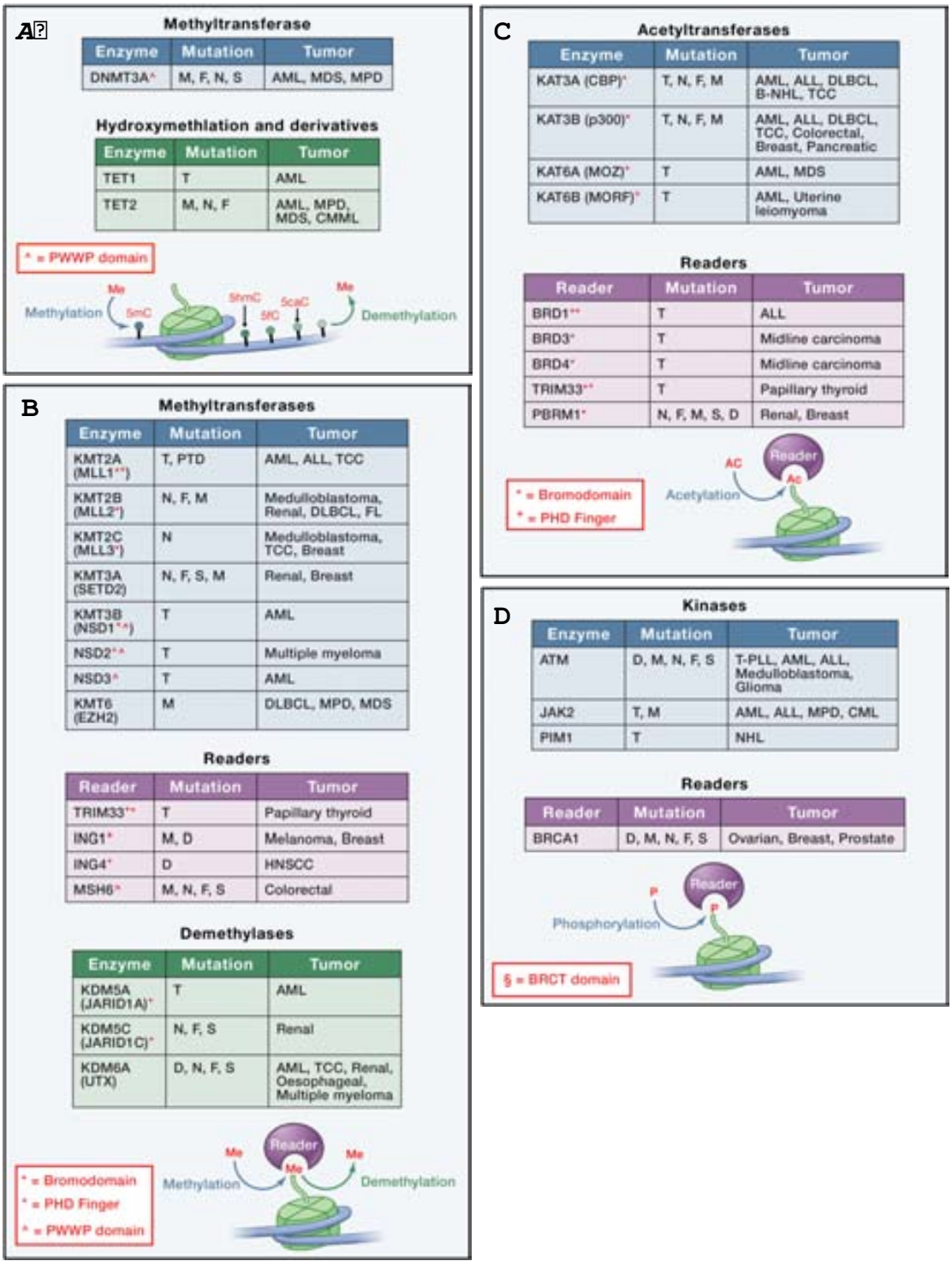
??c?et??d?l ti cl?tc??nGi u ?etc?u i ?t t?Cl???c??l d ??d??de?C?eti cl?tc?en????1tl?c?et??d?c?l ??1???ci e?i cds?tu 1ds??????al?et(??Gi d??i C?en?l ??1Gi e?tcl?tc????c??C??ae??di ?1Gi (t????1ie?cet?d?e?Cl?el?i C?en?C?1?aet??tce?C(?ceti c0????cau ??C?i ?tcnt?tei Cl ?n?(???dC????s??????c????(?d 1?????l ?tcle??nGi u ?etc?C?l ad?ei Cl ?x?tlaC?1L8%?k??Rl i c?F??i ap?Ct??l W88qL%??



?

??f a??2. ad??????????l ????c??n n????c ae?? ??ea??c???r ni????af ????u?ce cl ?r n??r ??n????f aa?r n?n?nf ?e??u?a?ef ??c???r n?????a?c?? ?h?????cn????ael ??s er? ??ef g?a???m2. p2l?

?iu ??? ?en???CaI l???1t?e????tc?tlaC??L8?xe?Cl ?etcI????? ?l W?????l??c?????L%W n?(???dC????s??????c??11Gi (????s?en??????0??ntl?la???l l?l ?en??C?d?)? ?en??tce?C?le?en?e??1tl?c?et???CaI l?n?(??1Gi ?a?????tc?en??t?d??? (?C?en??d?le?????????0??iR?(?CW?en??C??des?tl ?en?e?en??t?d??? ??1tl?c?et???CaI ??tl?i (?Cs?tl ?letd?tc?tel?tc ?c?s??c??u i C??lea?t?l ??C??c ????????ei ??(it???CaI?l t???? ??el?tc?en??ciCu ?d?etl a?0??i C?tcle?c????????l?tcnt?tei Cl W?Rnt?n?h?(??(?Cs?tteed???cpsu ??1???t t?tesW??i ad?ah?(??l?(?C??1t???? ??el?C?cI tcI?Giu ??etla W?c?al????c??(iu tetcI?ei ??ic??u ?CGiR?ei)t?tesW?C?(?Cl t?d????C?t????CCnsenu t???c??c?aGi dIt??d?ei)t?tesW??1?c?tcI?i c?en???CaI?x?tca?t?F???d??tW88/%??c?tlaC?1Lq??C?1C?d??e????en?1u ?tc?u ae?eti cl?? ??etcI??1tl?c?et???cpsu ?l?tc???c??CW?c??1?Cn?1l??CaI?e?Cl?e??d??1Gi e?tcl?i C?? ?ceteau i C?en?C?1t?1C??


A
Methyltransferase
Enzyme | Mutation | Tumor
DNMT3A^ | M, F, N, S | AML, MDS, MPD
Hydroxymethlation and derivatives
Enzyme | Mutation | Tumor
TET1 | T | AML
TET2 | M, N, F | AML, MPD, MDS, CMML
^ = PWWP domain
Me
Methylation
5mC
5hmC
5fC
5caC
Me
Demethylation
B
Methyltransferases
Enzyme | Mutation | Tumor
KMT2A (MLL1*^) | T, PTD | AML, ALL, TCC
KMT2B (MLL2*) | N, F, M | Medulloblastoma, Renal, DLBCL, FL
KMT2C (MLL3*) | N | Medulloblastoma, TCC, Breast
KMT3A (SETD2) | N, F, S, M | Renal, Breast
KMT3B (NSD1*^) | T | AML
NSD2*^ | T | Multiple myeloma
NSD3^ | T | AML
KMT6 (EZH2) | M | DLBCL, MPD, MDS
Readers
Reader | Mutation | Tumor
TRIM33*^ | T | Papillary thyroid
ING1* | M, D | Melanoma, Breast
ING4* | D | HNSCC
MSH6* | M, N, F, S | Colorectal
Demethylases
Enzyme | Mutation | Tumor
KDM5A (JARID1A)* | T | AML
KDM5C (JARID1C)* | N, F, S | Renal
KDM6A (UTX) | D, N, F, S | AML, TCC, Renal, Oesophageal, Multiple myeloma
* = Bromodomain
* = PHD Finger
^ = PWWP domain
Me
Reader
Me
Me
Methylation
Demethylation
C
Acetyltransferases
Enzyme | Mutation | Tumor
KAT3A (CBP)* | T, N, F, M | AML, ALL, DLBCL, B-NHL, TCC
KAT3B (p300)* | T, N, F, M | AML, ALL, DLBCL, TCC, Colorectal, Breast, Pancreatic
KAT6A (MOZ)* | T | AML, MDS
KAT6B (MORF)* | T | AML, Uterine leiomyoma
Readers
Reader | Mutation | Tumor
BRD1** | T | ALL
BRD3* | T | Midline carcinoma
BRD4* | T | Midline carcinoma
TRIM33** | T | Papillary thyroid
PBRM1* | N, F, M, S, D | Renal, Breast
* = Bromodomain
* = PHD Finger
AC
Reader
Ac
Acetylation
D
Kinases
Enzyme | Mutation | Tumor
ATM | D, M, N, F, S | T-PLL, AML, ALL, Medulloblastoma, Glioma
JAK2 | T, M | AML, ALL, MPD, CML
PIM1 | T | NHL
Readers
Reader | Mutation | Tumor
BRCA1 | D, M, N, F, S | Ovarian, Breast, Prostate
P
Reader
P
Phosphorylation
§ = BRCT domain

[illegible]

[illegible]

ALDO-KETO REDUCTASES IN CANCER

Aldo-Keto reductases (AKRs) are multifunctional enzymes that catalyze the reduction of several endogenous and xenobiotic aldehydes and ketones (Penning & Drury, 2007a). The AKR superfamily comprises 15 families containing over 150 members, which are further classified into subfamilies (Salabei *et al.*, 2011). These enzymes utilize a wide range of substrates, among them steroid hormones, prostaglandins, retinals, lipids and sugars. Most of them prefer the use of NADPH over NADH as the reducing cofactor (Barski *et al.*, 2008; Penning & Drury, 2007b). The AKRs are about 35–39kDA in weight and generally cytosolic and monomeric, although some, such as AKR7 can exist as a dimer and associated to membrane (Kelly *et al.*, 2002; Kozma *et al.*, 2002). Some of AKR proteins do not have or have a very poor enzymatic activity for typical substrates (such as MVDP, AKR1C12-C13, Kvβ), and some of them have been shown to have structural roles, like Rho in the crystalline (Fujii *et al.*, 1990; Weng *et al.*, 2006).

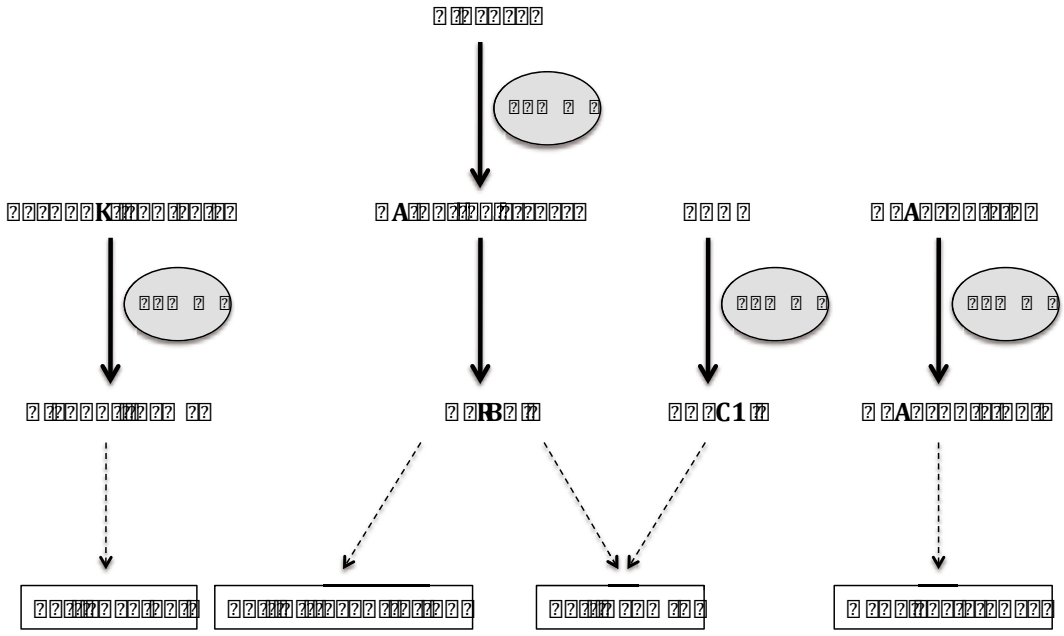
Mammalian AKRs have been classified into 3 families: AKR1, AKR6 and AKR7. Among them, the AKR1 family is the largest and has been further divided into subfamilies, from A to E. To date, 13 human AKRs have been described, and they participate in xenobiotic detoxification, biosynthesis and metabolism. Increasing evidence suggests the involvement of human AKR proteins in cancer development, progression and treatment. Some of them demonstrate multiple functional features in addition to being a carbonyl groups reductase. In this section we are going to review the most studied AKRs in cancer: AKR1B1 and AKR1B10.

AKR1B1 in human diseases

The most widely expressed member of the AKR1B subfamily is AKR1B1. Historically, it is the one that has received most of the attention, because of its role in diabetic complications and development discovered many years ago (Pastel *et al.*, 2012). AKR1B1 is considered a causative factor for diabetic complications by converting glucose to sorbitol under hyperglycemia. Several studies have shown that inhibition of AKR1B1 could prevent, delay or even reverse tissue injury induced by hyperglycemia (Gabbay, 2004). Moreover,

□□□□q□q□lniRl□ntln□11□□t□t□tes□tc□□□□a□tcl□len□C□□Ca□t□d□lnsltidlt□□dd□a□leC□e□l□tc□da□tcl□□□(□c□□□□Ids□□eti c□□c□□1Gi□a□el□x□□□%□1C□□aCliClW□dc1n□et□□d□□ns□□l□xla□□tct□□l□u t□d□□ns□□W□qW□zc□1neni9atci c□□□c□□q/zo□ei□leGi c□%W□Giu□et□□□d□□ns□□l□x□□C□i)s□□cp□d□□ns□□□□c□□cteGi□□cp□d□□ns□□□%W□□Giu□et□□o□ei c□l□xcteGi□□□ei1n□cic□%W-W□zacl□eaC□e□□□o□ei c□l□x□□□ei c□□□c□□□□□%W□t□□□C□isdxu□ensdcsi)□d□□c□□h□)□c□□tic□%□□c□□1)t□tp□□□1nil1nid1t□l□x□□□□□□□□□□W□L88ó,□□Ct(□le□(□□□□□□□□□W□L88E%□□□□□q□q□u□s□□□e□□l□act(□Cl□d□1Gie□□et(□□□cpsu□□□l□tcle□□□C□icsd□□

□□□□q□q□tl□a□t9ateialds□□)1C□ll□□□tc□hau□c□etlla□lW□c□□te□n□l□□□□c□□□l□Ct□□□□□a1C□lad□e□□□tc□u□cs□□□c□□C□2□□d□□□□CW□leGi□seiu□lW□ldi□d□leiu□lW□d(□CW□□□C(t□□dW□li1n□l□□dW□h□□□□c□□c□□oW□ot□c□sW□z□□dd□□□ae□□d□ao□ut□W□z□□dd□□□ae□□d□ao□ut□W□nGict□□d□ao□ut□W□dsu1niu□W□u□d□ciu□,□□c□□□iRcC□lad□e□□□tc□1Gile□e□□□c□□Cx□□tc□F□□□eC□lnW□L8qL%□□



□□fa□□22□d□□aece □□□er□e□□r□□□cn□s□i□s□□a□□□□□p□p□□□□ruecu□□□□□□r□□r□□r□□a□e□□n□□□□nfl eal□

□n□C□□□□C□□□(□Cial□u□□n□ctlul□1Gi1il□□□ei□□)1d□tc□Rns□□□□q□q□tl□□□(□ce□l□ial□iC□en□□eau iC□□iC□□)□u1d□W□e□l□bciRc□en□e□□□□q□q□□□c□tc□a□□□□CaIC□□ltle□c□□□□stC□□□a□tcI□en□□□□C□icsd□lGia1□□□liu□□□ceteau iC□□CaIlW□a□n□□l□□i)iCa□t□tc□x□□t□□tc□□□□□□W□L8qL,□□tc□F□□□ccteIW□L88U%□□□aCen□Cu iC□W□□□□q□q□1Giuie□l□□□dd□laC(t(□d□□s□□□duutc□etcI□□seiei)t□□□C□icsd□□iu1iac□l□□□l□□□l□Ct□□□□□□iC□W□□□□q□q□1Gi□a□□l□ns1□Clds□□ut□□en□ed□□□□l□ei□)t□□□et(□□leC□llW□

[illegible]

[illegible]

[illegible]

[illegible]

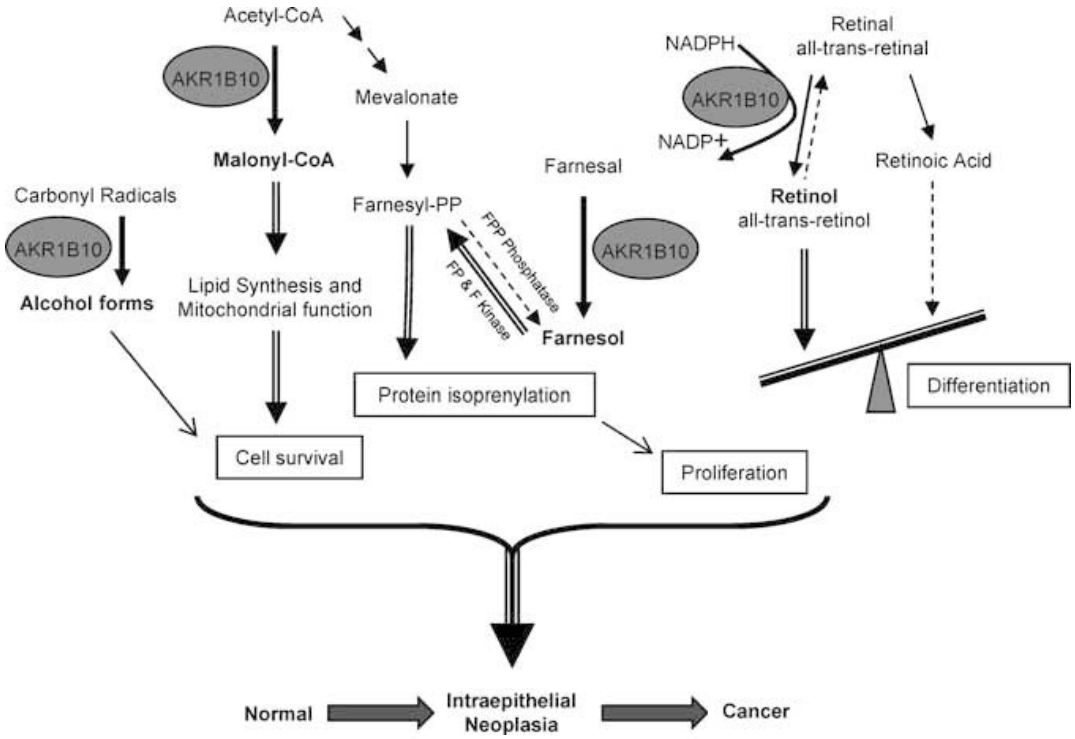
[illegible]

efficiently glucose to sorbitol (as AKR1B1 does). AKR1B10 has similar reducing efficiency with AKR1B1 in a wide range of substrates, such as lipid peroxides (Aldini *et al.*, 2007; Ravindranath *et al.*, 2009), cigarette and environmental pro-carcinogen polycyclic aromatic hydrocarbons (PAHs) (Quinn *et al.*, 2008) and cytostatic anticancer agents such as daunorubicin (Heibein *et al.*, 2012; Plebuch *et al.*, 2007). However, AKR1B10 has a special affinity to reduce All-trans-retinaldehyde (and other retinaldehydes) to retinol, about 50-fold higher than AKR1B1, what makes AKR1B10 a crucial enzyme in the retinoic acid biosynthesis pathway in some tissues (Ruiz *et al.*, 2009). It has been recently observed the ability of AKR1B10 to do protein prenylation, which is a lipid modification, involving the covalent addition of either farnesyl (15-carbon) or geranyl (20-carbon) isoprenoids to C-terminal cysteines of the target protein (Chung *et al.*, 2012). This modification allows the proteins to be located in cellular membranes. Proteins that undergo prenylation include Ras and Ras-related GTP-binding proteins (G proteins) and kinases, which are involved in cell growth, differentiation, maintenance of the cellular cytoskeleton and vesicle trafficking (Casey, 1994; Gibbs *et al.*, 1996; Zhang & Casey, 1996).

In contrast with AKR1B1, AKR1B10 shows restricted tissue distribution, predominately expressed in the small intestine, colon, stomach and adrenal gland (Cao, 1998; Hyndman & Flynn, 1998). But this restricted expression changes drastically in different cancer types where it is significantly upregulated. For example, in B-cell and T-cell acute leukemia as well as in chronic leukemia (Laffin & Petrash, 2012), liver carcinoma (Schmitz *et al.*, 2011), lung adenocarcinoma, squamous carcinoma and smoking-related Non-small-cell Lung Cancer (NSCLC) (Kang *et al.*, 2011; Laffin & Petrash, 2012), pancreas carcinoma (Chung *et al.*, 2012), breast carcinoma (Ma *et al.*, 2012) and oral cancers (Nagaraj *et al.*, 2006). The fact that AKR1B10 is upregulated in so many cancers is not a coincidence and actually the literature is plenty of examples about the benefits that overexpression of AKR1B10 have for cancer cells (Figure 23). As it occurs with AKR1B1, AKR1B10 can reduce damaging carbonyl radicals and anti-tumor drugs (doxorubicin and others) promoting cancer cell survival.

Also, AKR1B10 blocks ubiquitin-dependent degradation of the Acetyl-CoA carboxylase-α (ACCA) and thus enhances fatty acid/lipid synthesis, which affects

u tei ni c Ct d ac eti cWi) t et(le eal c cd1t 1Gi) t eti c i dl c (cea dds cn c tcI dd aC(t(dx F i WL8qq%@



fa 2P d caece rn a nau aeo e p p. r a eri o m ar i o r oc l n eo l r n r r n r el e n e a n reo r n a rue u l r n r a a re r h cn ael f r mn2. p2h

i C en c ó: D i nau c 1c C et c Cl CCs et(etc I l c u ae eti cW2c en 1Gi e tc C9atC2l 1C csd eti c i C tel et(tesWRnt n tl nt(sa (C) 1C llti c I q q8x nac I WL8qL%@s 1C csd etc I c 1Cn1l en C 1Gi e tc lW q q8 l 1Gi u i etc I c C 1Gi d C eti c@tc dsW en tc C l d(di q q8) 1C llti c tc c i 1d let dl C 1a del tc c cpsu et 1 a eti c l C etc dei C etc idW i c i u te ce C l tc C etc it t tW c nt e tc tl c dc I Gi u t C cet eti c ei aci 11il dda d C 1Gi d C eti c xatp WL88ó, CCbl WL8qL%@e d Ri Cen ei c ie en c R Gi d i q q8 C ceds 1Gi 1i l Ws n lnt c i db la l l 1: Ze eCl e@n s i l C(en e en et(eti c i 1: Zs Ct u s tc cn c en) 1C llti c i q q8 tc il z 1c u cc CWen e Ri ad) 1d tc en a 1C lad eti c i q q8 tc c Cl Rten 1: Z i (C) 1C llti c xC9a ceds u ae e%@i R(CW en C l i (C) 1C llti c i q q8 tc c Cl Rten l C Rten iae 1: Z) 1C llti c xn lnt WL8qZ%@W i 1: Z l ac ie) dalt(ds C 1 1i clt d i C en C lad eti c i

AKR1B10 and the implications of the protein in such crucial pathways makes the regulation much more complicated.

However, gathering all the data about AKR1B10 in cancer, appears to be obvious the advantages of upregulating this gene, whereas in biology there is always a case that breaks the rule, and in this case is colorectal cancer, where AKR1B10 is downregulated in most of the patients (Laffin & Petrash, 2012; Ohashi *et al.*, 2013). We are interested in this dual role of AKR1B10 in cancer, and mainly in the special relation between AKR1B1 genes and colorectal cancer.

RETINOIC ACID PATHWAY

Vitamin A and its natural or synthetic derivatives (retinoids) have been shown to act as cancer chemopreventive agents (Niles, 2004). The relationship between Vitamin A and cancer was firstly described by Wolback and Howe in 1925; they observed higher proliferation rate in vitamin A deficient epithelial cells (Wolbach & Howe, 1925). Later on it was observed that Vitamin A deficiency elevated the spontaneous and chemically induced tumors in animals (Lasnitzki, 1955). Underpinning these reports, it was also seen that pharmacological concentration of Vitamin A added to the diet decreased the incidence of chemically induced tumors in animals (Bollag, 1972; Harisiadis *et al.*, 1978).

Vitamin A or its derivatives (retinoids) cannot be synthesized *de novo* by any animal species and are only obtained through diet mainly in the form of retinol, retinyl ester or β-carotene (depicted as a carrot in Figure 24), the latter two would be converted into retinol in the intestinal cells. Vitamin A or retinoic acid pathway (head and tail molecules of the pathway) begins with the absorption of the vitamin, which is modified through a number of reductive reactions to the bioactive compound, all-trans retinoic acid (ATRA). In fact, diet contains very little all-trans retinoic acid and intestinal cells are mainly exposed to retinol and retinal.

Retinoids are a family of numerous compounds (over 4000) chemically related to vitamin A. The basic structure of retinoid molecules is composed by a cyclic end group, a polyene side chain and a polar end group (Figure 24). Although many retinoids can be found physiologically, just few of them are known to be bioactives *per se*. Retinoids biological active forms can be summarized under three main structures: retinol and retinoic acid (or RA).

· Retinol has six biologically active forms (all-trans, 11-cis, 13-cis, 13-di-cis, 9-cis and 11, 13-di-cis retinol), being retinol all-trans the predominant physiological form (Baron *et al.*, 2005).

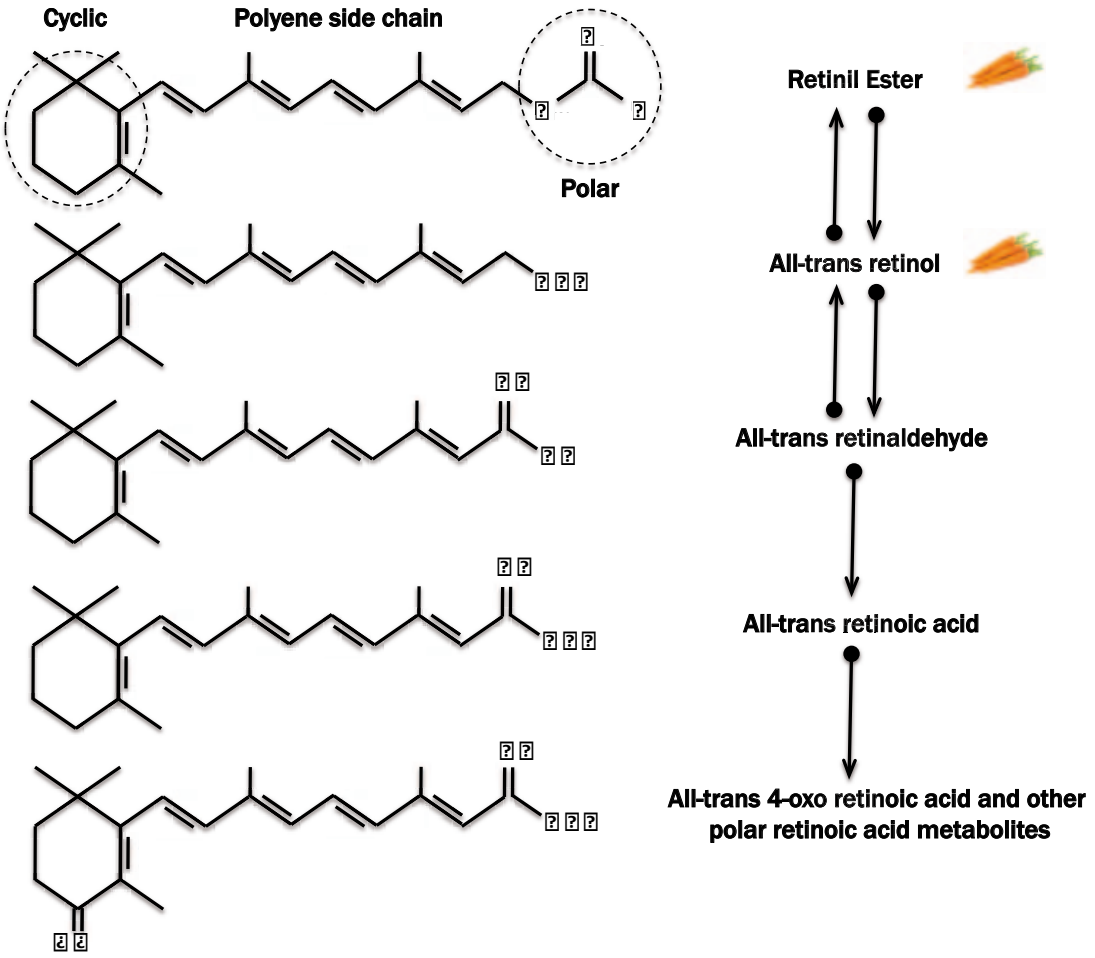
· Retinoic acid (RA) active forms include all-trans retinoic acid (ATRA), 9-cis retinoic acid, 11-cis retinaldehyde, 3,4-didehydro retinoic acid, and less commonly found 14-hydroxy-4, 14-retro retinol, 4-oxo retinoic acid, and 4-oxo retinol (Baron *et al.*, 2005).

It is known that retinol is acting in a non-retinoic acid receptors way, whereas RA and oxo-RA work activating the classical pathway. The starting metabolite can be different, but all of them converge to retinol, which is reversibly oxidized by retinol dehydrogenases producing retinal. Subsequently retinal could be irreversibly oxidized to all-trans retinoic acid (ATRA) by retinal dehydrogenases. Retinoic acid is oxidized to more polar metabolites (such as 4-oxo-RA, which can also activate RARs) by cytochrome P450 family members such as CYP26A1, -B1, and -C1 (Tang & Gudas, 2011) (Figure 24).

Retinoic acid exerts its functions through two distinct nuclear receptors, retinoic acid receptors (RAR) and retinoic X receptors (RXR). Both of them have three different subtypes: α, β and γ. Each of the receptors possess different affinities for their ligands, but once bound RAR and RXR form heterodimers and activate downstream factors by binding to the retinoic acid response elements (RARE), which are normally located in the gene promoter (Bushue & Wan, 2010; Tang & Gudas, 2011). More than 500 genes have been reported to be regulated by retinoic acid (Theodosiou *et al.*, 2010).

Retinoic acid receptor pathway controls many important and diverse functions such as cell proliferation, cell differentiation, neural and vision functions, immune processes and the proper establishment of the body plan during early development.

In the last years retinol has been proposed to inhibit the growth of colorectal tumor cells by an independent RA-receptor mechanism (Park *et al.*, 2005). Although it is unclear how this inhibition is achieved, several labs have found out different antitumor pathways activated by retinol treatment such as the inhibition of β-Catenin (Dillard & Lane, 2007) or the loss of Phosphatidylinositol 3-Kinase Activity (Park *et al.*, 2008). In summary, retinol, ATRA and oxo-RA (and close derivates) are the main bioactive retinoids with antitumorigenic activity described.



?

?????f a???2R?d???a?r ??c?o?l ?n???e?o?h? ?e?h????a?n?r?e??????c?t?n?s ?i ??r ??h ????l ???o?m?f ?n?f a? h?????aaen?
?r ?????n? ?n????c?a? ?r ???e?h??? ????el cefr ?ir?h?????n?h?

?1t??u ti d l t??d???c????ctu ?d?u i ??d?l ea?t?l ?h?(?????u i cl eC?e???en?e?(te?u tc???
??1Ct(?eti c?l ??tC??ed s??i CC?d?e???Rten??h tI n?C?Cdl o?l ???(?d 1tc I???c??C?x?ac?F?
?ie?cW?L88L,?? ?d?????????W?qó78%??aCen?Cu i C?W?a??ei ?en????cetz1Gi d ?C?et(?W?LGi z
?t ?C?cet?eti c??c??1Gi z?1i 1ei et??? ??el ?i ?C?etci t?l W?en?l ???C????tc I ?al ?????l ?
?n?u i 1C?(?cet(???I ?c el x??c?dt?????W?L8q8,??C?tc?F ?? s?d?C lotW?L8q8%?ei ?eC??e?
1C?u ?dI c?ce??c??u ?dI c?ce?d?1ti cl x?d??c l ?????W?qóó/,???(tc?o?????W?qóó:,?
??d1?Cc?????W?qóóE,??i cI?????W?qó7/,??aCt?????W?L888%??c???(?c?ei ?1C?(?ce???
l??i c??eau i C??(?ce?x?i cI?????W?qóó8,??bc?C?????W?L88q,??aei ?????W?qóó/,?
??1ei Ctci?????W?qóóZ%??n?l ??C?etci t?z??l ???eC??eu ?cel?n?(?????c??11d????ei ?
?du i le??(?Cs?tl la?O?

?n??tu 1i Ce?c???l ?C?etci t?l tc???c??C?1C?(?ceti c??c?eC?eu ?ce?n?l ?C????t(?????
d e?i ??ee?ceti cW?Rn?C??l ?en????C?lad?eti c?i ??c?i l?ci al?C?etci t?l 1?enR?s?tl ?
letddC?ie?R?dd?ac??Clei i?O??n?C???C??iu ???da?l???ia e???C?lad?eti c?l ?liu ??l ?en??
C?etci t????t??1?enR?s?x????1?enR?s%?l ?c?l W?i C??)?u 1d?W?????S?Rnt?n??tl 1d?sl ?

hypermethylation in some cancers (Hayashi *et al.*, 2001; Narayan *et al.*, 2003; Wang *et al.*, 2003); ALDH1A2 which is downregulated in prostate cancer (Kim *et al.*, 2005); CYP26A1, which is found highly expressed in breast cancer (Osanai *et al.*, 2010) and LRAT which expression is also decreased in some cancers (Mongan & Gudas, 2007).

To date, no attempt has been made to fully characterize the epigenetic regulation of the retinoic acid pathway in cancer. In this study, we perform a comprehensive analysis of genetic and epigenetic alterations in this pathway and we evaluate the prognostic and diagnostic applications of gene expression and DNA methylation markers in human colorectal cancer.

OBJECTIVES

AKR1B1 enzymes have an important role in multiple biological processes, such as inflammation, differentiation and sugar metabolism. AKR1B1 genes are included in the retinoic acid pathway; these enzymes reduce retinal to retinol, which finally will end up in retinoic acid. We hypothesize that retinoic acid pathway genes, especially AKR1B1 subfamily, display an altered expression pattern through the specific DNA hypermethylation during colorectal tumorigenesis. This work is focused on understanding the epigenetic alterations of retinoic acid pathway genes in colorectal cancer and the applications that those changes may have in terms of diagnostic, prognostic and therapeutic strategies.

The general objective of this thesis is to bring light into the mechanisms regulating AKR1B10 and AKR1B15 genes and what is their role in colorectal cancer.

To address this goal we propose the following specific objectives:

· To elucidate the epigenetic mechanisms regulating the expression of AKR1B10 and AKR1B15 genes in normal colon and in cancer cells.

· To determine the functional implications of AKR1B10 and AKR1B15 deregulation in colorectal cancer.

· To evaluate the utility of AKR1B1 CpG island hypermethylation as a non-invasive screening biomarker for colorectal cancer.

· To characterize the gene expression and changes in the DNA methylation profiles of retinoic acid pathway genes in normal colon and in tumor samples.

· To evaluate the use of molecular alterations of the retinoic acid pathway genes in the management of colorectal cancer patients (prognosis and therapeutic strategy).

SAMPLES

Patients' samples

Normal and tumor tissues were available from 25 colorectal cancer patients. DNA methylation and mRNA expression of the AKR1B1, AKR1B10 and AKR1B15 genes were analyzed. A set of 16 patients samples were collected from the Institut Català d'Oncologia (ICO, Barcelona, Spain) and 9 from the Hospital Universitari Germans Trias i Pujol (Badalona, Spain).

The mean age of these patients at the time of collection was 69 years, in a range between 48 and 86 years (Table 7). Both genders were well represented, with 12 women and 13 men (Table 7). Tumors were evaluated according to the tumor node metastasis (TNM) staging system (Hu *et al.*, 2011). Samples were fresh-frozen at -80ºC within the two hours of removal.

Patient Code	Age	Gender	TNM stage
CR010	49	man	II
CR012	65	man	III
CR013	86	woman	III
CR015	69	man	III
CR016	56	woman	III
CR017	76	woman	III
CR018	75	woman	III
CR019	72	woman	III
CR020	63	man	III
54081201	73	woman	III
09-4	80	man	III
54081057	70	man	III
54081060	59	woman	IV
54081064	48	man	III
09-63	63	woman	III
09-160	84	woman	IV
09-454	80	woman	III
09-736	67	man	III
09-817	73	man	IV
09-813	51	woman	II
09-437	64	man	III
09-530	68	man	IV
09-12	73	man	IV
54080980	73	woman	IV
09-697	81	man	IV

Table 7 | Relevant clinical data for the twenty five patients examined, the two sources of samples are represented in different colors, lime for Hospital Universitari Germans Trias i Pujol and green for Institut Català d'Oncologia.

Human stool samples and associated carcinoma tissues

A total of 143 stool samples were collected at the Institut Català d'Oncologia (Barcelona, Spain); 45 of them from patients with colorectal carcinomas, 35 from patients with colorectal adenomas, 5 from patients with colonic inflammatory bowel disease (IBD) which could be either ulcerative colitis or Crohn's disease, finally 58 control subjects who underwent surgery for diverticular disease or an endoscopy that revealed no lesions. Stools DNA obtained from healthy individuals paired by age and sex were used as control group. For 7 carcinoma patients (Carcinoma 27, 28, 29, 33, 34, 38 and 41) we also had the DNA from the tumor, what enabled us to compare among the DNA methylation of the tumor and the stool of the same patient.

[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]
L751	[illegible]	[illegible]	[illegible]
L752	[illegible]	[illegible]	[illegible]
LL397	[illegible]	[illegible]	[illegible]
LL349	[illegible]	[illegible]	[illegible]
LL392	[illegible]	[illegible]	[illegible]
LL352	[illegible]	[illegible]	[illegible]
LL393	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

[illegible]

ii ii7 u o Monu s 3 Uj S s T- H 3l n7da o Rts nl di 3 Ni n07 n t t M tNl 6 I u t3a 7nl 7d t t3t 7 d s ts Mns a ii its 6 i 8I 1

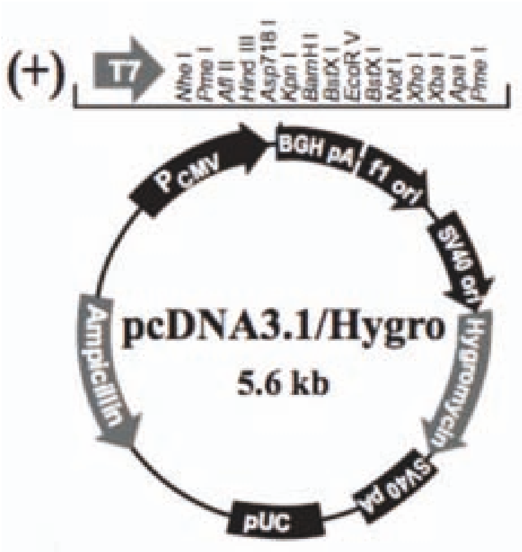
Cell Line	Description	Medium
HCT116	Human colon carcinoma	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
Sw480	Human colon adenocarcinoma,Dukes' type B	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
SKO	Human colon adenocarcinoma	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
Caco-2	Human colon adenocarcinoma	DMEM/F12/100mM Pyr/200mM GLN/20% FBS
HT29	Human colon adenocarcinoma	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
LoVo	Human colon adenocarcinoma, Dukes' type C, grade IV	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
DLD1	Human colon adenocarcinoma, Dukes' type C	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
DKO	HCT116 deficient in two major DNA methyltransferases, DNMT1 and DNMT3b	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
HepG2	Hepato celular carcinoma	DMEM/F12/100mM Pyr/200mM GLN/10% FBS
CT26	Mouse colon carcinoma	DMEM/100mM Pyr/200mM GLN/10% FBS

i q x ii Hil Hm m HPme nl p s ml ii Pn p Hp Pn 1 Pp pl H cf E Hi s gp 8 PP i He R cnE s He cnHu l R E h iHl e Pn R E l i s uPn nHe f ii ns H l p m ns e ®, nuPl ns n | mip | | Uf

ii7 u o Ni3No inu 4) - h s 7t3v s 3 nu d 77 M 7R s n1 no 3a s HTR ts j T 1 H Ni3No i 7r 71 niinu ts M o nl l s 3tns 7R ii7 u o 3ovd7ts tb U 3tl 7 d o u r RN7ts M ovd7ts B) 1HMg 6 s mt3onM s R oi7 R R I R s 3a s tiN3 G 3n 4 ni 1 ii7 u o Ni3No ts 7 s n s 3t tn3t 7 s a r d otn t iiv no ns 3 l ts 3tns 6 nNs M R H) 9) I 1

no a wd otl s 3R ii7 u o nNs 3 N7ts M 3a a 1 n v3nl 3 o 6 ovd s iN dn7t3tm ii7 u o w iN I R a s ii7 u o di 3 7 ddondot 3 no 3a 7d t wd otl s 3 1

[illegible]



[illegible]

Plasmid	Resistance	Digestion enzymes	Description
AKR1B10-pCMV-HA	No	NotI / XhoI	Transient Transfection
AKR1B15-pIRES-hrGFP-1a	No	NotI / XhoI	Transient Transfection
AKR1B10-pcDNA3.1 (+) Hygro	Hygromycin	NotI / XhoI	Stable Transfection
AKR1B15-pcDNA3.1 (+) Hygro	Hygromycin	NotI / XhoI	Stable Transfection
pcDNA3.1 (+) Hygro	Hygromycin	Not Digested	Empty vector

[illegible]

Vector and insert were ligated using T4 DNA Ligase (New England Biolabs, Beverly, MA, USA) in proportion 1:3 respectively. The reaction took place at 16ºC during 4 hours.

Vector amplification

All vectors were transformed in DH5-α *Escherichia coli* strain by heat-shock procedure. Transformed bacteria were seeded into LB agar, at 50 µg/ml of ampicillin. Plates were cultured overnight at 37ºC.

Individual clones were selected and amplified in liquid LB at 100 µg/ml of ampicillin, in a volume of 5ml or 250ml according to the experiment.

For small amounts of plasmids, and starting from 5ml of grown LB media, we used the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, St Lois, MO, USA), eluting in 100 µl of pure grade water. In larger scale experiments, we started with 250ml of grown LB media and purified the vector with the GenElute™ HP Plasmid Maxiprep Kit (Sigma-Aldrich, St Lois, MO, USA), eluting the purification in a final volume of 3ml of pure grade water.

CELL CULTURE

DNA transfections

All transfections were carried out in 6-well plates, seeding 150,000 cells per well at day 0. Cells were transfected with Lipofectamine® 2000 (Invitrogen, Carlsbad, CA, USA). Transfection (day 1) was performed when cells were 70% confluent, as recommended by the manufacturer. Ten µl of Lipofectamine® 2000 were diluted in 480 µl of Opti-MEM® (Invitrogen, Carlsbad, CA, USA) and incubated for 5 minutes at room temperature. Then we diluted 3 µg of the plasmid in 495 µl of Opti-MEM® and mixed together with the Lipofectamine® 2000/Opti-MEM® solution, incubated for 10 minutes at room temperature. After, we added 1 ml of the mix to each well in a drop wise manner. Cells were incubated at 37ºC for 5h, after this time the transfection media was removed and substituted by fresh medium.

In the case of vectors carrying resistance for hygromycin, selection of transfected HCT116 cells was done by adding 100 μM of hygromycin B (Invitrogen, Carlsbad, CA, USA) to the media for 10 days (changing the media every two days); afterwards the antibiotic concentration was decreased to 50 μM. Untransfected HCT116 cells were used as a control and treated with hygromycin B in the same conditions until death.

Lentiviral infection

Infectious lentivirus particles were generated and packed using the 293T cell line. We used a second-generation system with co-transfection of three plasmids: (1) the lentiviral vector pLX302 that contains the gene of interest (AKR1B10 or AKR1B15) and LTRs (Long Terminal Repeats), (2) the pCMV-VSVG plasmid for pseudo-typing with VSVG, allowing the lentivirus to infect a broad range of cells, and (3) the psPAX2 plasmid with *gag*, *pol*, and *rev* genes (Naldini *et al.*, 1996). As a transfection and infection control we used a GFP lentiviral plasmid instead of AKR1B10 or AKR1B15. The virus generation and the infection process were performed following the standard protocol (Tiscornia *et al.*, 2006).

Dznep treatment

In order to know the relevance of the EZH2 protein in the inactivation of the enhancer, we treated HCT116 cells with the EZH2 inhibitor 3-Deazaneplanocin A hydrochloride (DZNep) (Sigma-Aldrich, St Lois, MO, USA). Cells were treated for 72 hours with 5μM of DZNep, replacing the medium every 24h.

All-trans retinal treatment

HCT116, Sw480 and HT29 colorectal cancer cell lines were treated with All-trans retinal (Sigma-Aldrich, St Lois, MO, USA) for 24h at 0.5 and 2μM.

5-aza-2'-deoxycytidine treatment

To study the importance of the methylation in the regulation of the genes studied, we treated cells with the demethylating drug 5-aza-2'-deoxycytidine (5-AzadC) (Sigma-Aldrich, St Lois, MO, USA). 5-AzadC is a nucleoside analogue that is converted into nucleotide and incorporated into the DNA. There, 5-AzadC can

trap DNMTs by forming covalent complexes (Sheikhnejad *et al.*, 1999), resulting in global DNA demethylation in a replication-dependent manner.

We seeded a total of $1.5x10^6$ HCT116 cells (10^6 in the case of control cells) in a 10 cm^2 plate and treated for 48 hours with 0.5 μM of 5-AzadC, changing the medium every 24 hours. After 48 hours we removed the 5-AzadC medium and let the cells recover with normal medium for 24 hours. Then cells were collected.

To verify the success of the treatment we analyzed the re-expression of the silenced genes EN1 and INHBB, by qPCR and the demethylation of their CpG island associated promoter by bisulfite sequencing (primers are in the Annex II and III).

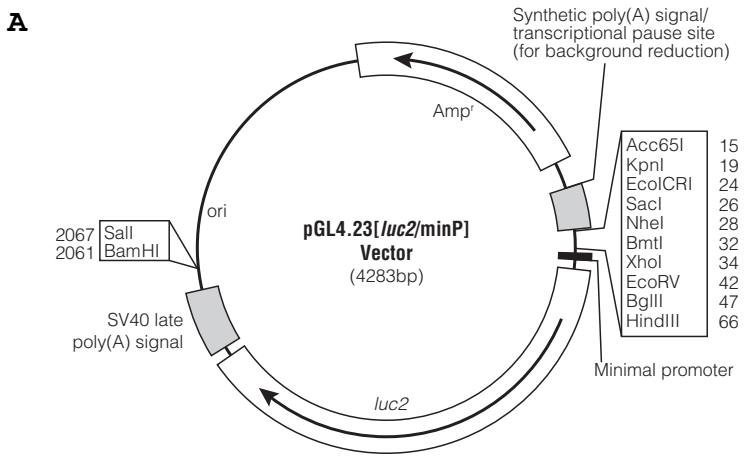
Luciferase enhancer assays

To test the enhancer capability of some sequences of the AKR1B1 locus we used the pGL4.23[*luc2*/minP] vector (Promega, Madison, WI, USA), which is specially designed to test luciferase enhancer activity of the fragments inserted. We first amplified the selected sequences using the Phusion® High-Fidelity DNA Polymerase (Promega, Madison, WI, USA) to avoid any PCR error. Then PCR amplicons were inserted in the multicloning site (MCS) of the pGL4.23 vector (Figure 26A).

pGL4.23 vector is mainly composed by a MCS, a minimal promoter and the firefly luciferase gene (Figure 26B). The different constructs and the empty vector were transfected in different wells, each of them in quintuplicate (in 5 wells) and in at least three independent experiments. The assay was done using the Dual-Luciferase Reporter Assay kit (Promega, Madison, Wi, USA) in accordance to the manufacturer's instructions.

Luciferase activity was assayed using 10 μl of the cell supernatant and the Dual-luciferase reporter plasmid system (Promega, Madison, WI, USA). Luciferase measurements of the different fragments were normalized against Renilla luciferase activity (cotransfected with the pGL4.23) to determine the relative luciferase activity and the fold activation. Luciferase/Renilla signal was analyzed with a Centro LB960 luminometer (Berthold technologies, Bad Wildbad, Germany). Empty pGL4.23 vector was included as control. Next, we detected

3a to1N t o7 3mt3t7ts ii7u t3a t o s 3m ot s 37ts 3t 73U ts d s s 3 wd otl s 371



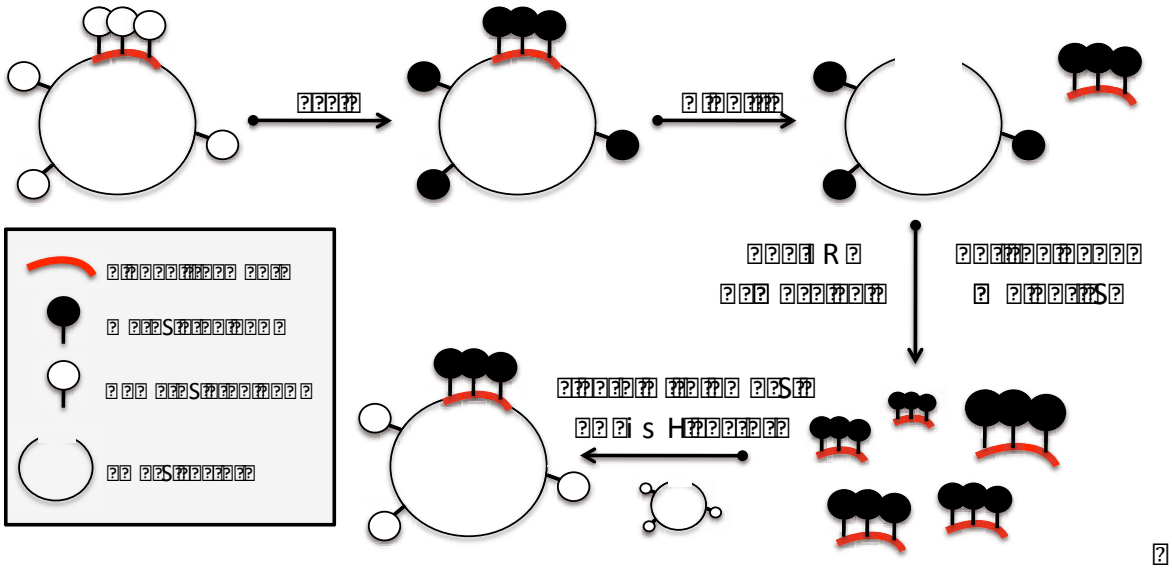
P Hm â8 x U . faOG ye Rn j u ls m e ' U p p e s P n nl . faO Hp Rn l plH cf

n 3 73 a nu l 3a vi 3tns 37 3a s a s o d tit3v n 7 i 3 o M s 3Ru ddit m ot 3tns n 3a 73s o l 3an 6tMNo HjI1 3o s 7 3 3a d G1HU m 3no u t3a 3a ts 7o3 o M s 3Nsl 3a vi 3 no l 3a vi 3 1Rts no o3n 3ol ts u a 3a o t o s 7 ns l 3a vi 3tns nNi 3 3a s a s o 3mt3v1

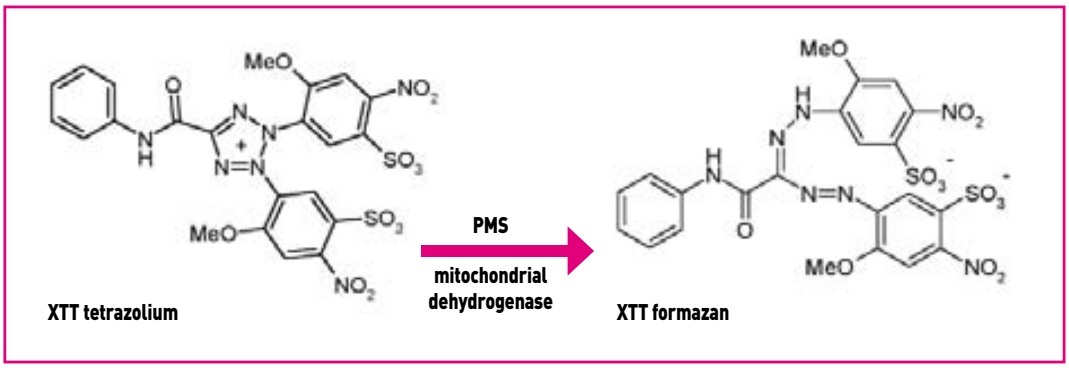
to73 l 3a vi 3 3a d G1HU º o M s 3 m 3noR N7ts M 3a 77 l 3a vi3o s 7 o7 6 Rd7u at aR R IR 3ou o7u tM 73 u t3a ts bvl 3n o 3a l 3a vi3 o M s 31 a tM 73tns u 7oNs ts 3n 9- M on7 M iRu a o 3a 7to s u 7 N3 s dNot t N7ts M 3a N i n7dts i s i s Bd r t3 6 a o vB M iR po sR ol s vI 7 l ii l nNs 3m 3a dNot t don N3u 7N7 3n 77 77 3a l 3a vi3tns 73 3N7 h 3a o M s 3R v 3u n ts d s s 3 tM 73tns 7Ru t3a l s l 1 n3a s bvl 7 o t7n7 atbnl o7 3a 3 N3 3a 7 hNs 1 nu m oR l t7 in r u a s 3a 7t3 u t3ats 3a 3 oM 3 7 l 3a vi 3 Ru a ti l u tii N3 n3a R l 3a vi 3

□

□s □□Nsl □3avi□3□□1□□□3a □□□77□□l □3avi□3tns □u □7□□noo□□3□3a □□□tM□73tns □u t3a□3a □□
□l □□□l N73□□□□□□in□r □□1□□s □Nsl □3avi□3□□□□□□□□u □7□□i7n□□□□□□□□□7□□□dn7t3tm□□
□ns 3oni□no□□l □□1□


□□□□□
□ □□□□□□□□
□□□□□□□□□□□□
□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□
□□□□

□



□

□P□Hm□□a 9x□□P□□i□□□□iip□m□□H□lBs n s□□□□□ts l□□□□ o l□mhpsiH□i□□□sne □r□n□p□ilf□

□a□o□ no□R□3a□□□ns□□s 3o□3tns□n □3a□□□v□□t7□dondno3tns□i□3n□3a□□sNl □□o□n □ l □3□□nit□□iiv□ □□3tm□□ □□ii71□ □a□□ □77□v□ u□7□ d□o nol □□□ □□□no□tsM□ 3n□ 3a□□ l □sN □□3No□oO7□ts 73oN□3tns 71□

□

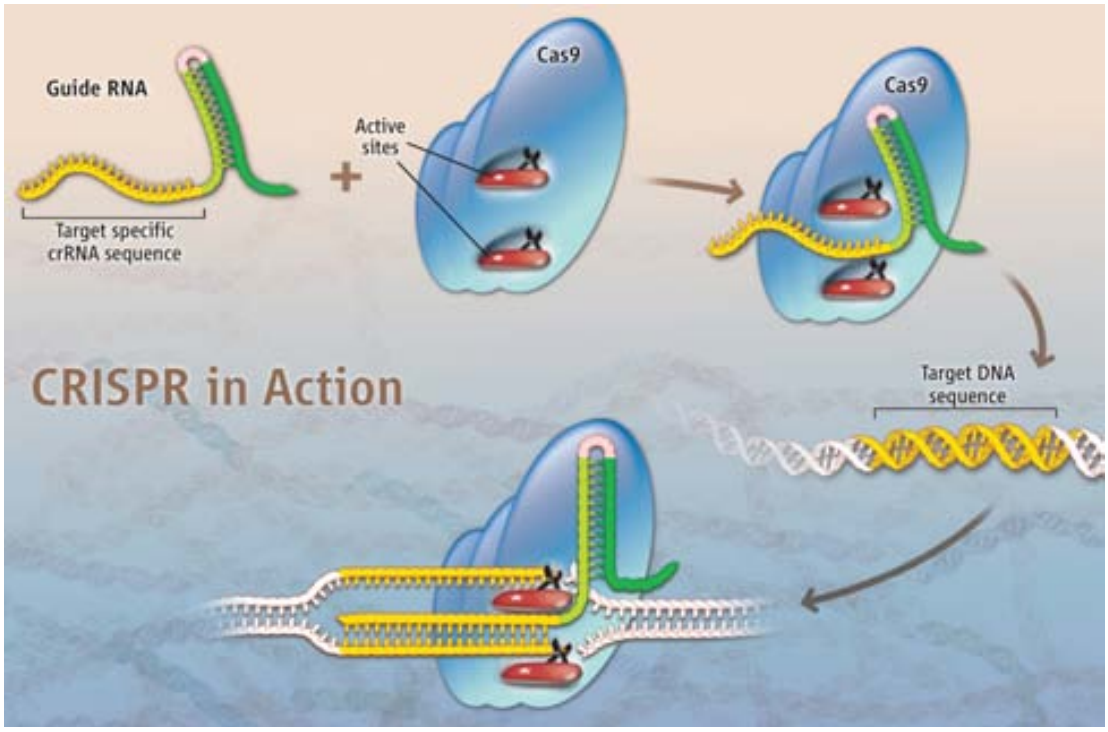
□□nse □□□□□HHn□v□□□□□□□□

□a□□iN73□o□□□□□MNi□oiv□□s 3□o7d□□□□□□ano3□□□its□onl t□□□d□□37□6□□□□□□I□□vd□□□□□ 7v73□l □t7□□Noo□s 3iv□3a□□l n73□□nl l nsiv□N7□□□□□□B□Nt□□□□□s□ns N□i□□7□□ 3□□as ninMv□no□M□snl □□□s Mts□□otsM□6□ □it□□□□□□□RH) 9UI1□□a□o□□□o□□3u n□□t73ts□3□ □nl dns□s 37□3n□3at7□7v73□l K□□□MNt□□□□□□□□□s□□□s□□s□ns N□i□□7□Rts□3at7□□□7□□3a□□ □□□□□□□□77n□t□3□□□s N□i□□7□R□□□781□□a□□MNt□□□□□□□□t7□□□□nl □ts□3tns□n □3a□□ □s□nM□snN7□□□□3□ot□i□□o□□□□□s□□3o□□o□□□ 6□□3tm□3ts M□□s□□3o□s 7B□□3tm□3ts M□ □□□□□□□□□□I□ts 3n□□□7ts M□□□atl □ot□□MNt□□□□□□□□6M□□□I□3o□s 7□otd31□□a□□M□□□□ □nl □ts□7□3a□□3□oM□3ts M□7d□□t t□t3v□n □3a□□□o□□□□□ut3a□3a□□7□□ ni□ts M□dond□o3t□7□n □ 3a□□3o□□o□□□□ts 3n□□□7ts M□□3o□s 7□otd31□□ a□s□3a□□M□□□□□s□□3a□□□□781□o□□□wdo□77□□□ ts□3a□□□□iiR□3a□□M□snl t□□3□oM□3□7□hN□s□□□□□s□□□□□l n□t t□□□no□d□ol □s□s3iv□ □t7oNd3□□□6□□s st7tRH) 9UI1□□□□□□□dlotl □o7□7□hN□s□□7□□o□3n□□3□□□ts□□s s□w□□1□

□a□□M□□□g□□781□nl di□w□t7□b□□oNt3□□□3n□3a□□3□oM□3□7□hN□s□□□□v□3a□□□□7□Bd□tots M□ anl ninMv□□□□3u□□s□3a□□M□□□□□7□hN□s□□□□s□□□3a□□□nl di□l □s3□3n□3a□□3□oM□3□ 7□hN□s□□□ts□3a□□M□snl t□□□□□1□□no□7N□□□77 Ni□□ts□ts M□n □□□781R3a□□M□snl t□□3□oM□3□ 7□hN□s□□□l N73□□i7n□□ns3□ts□3a□□□noo□□3□□on3n7d□□□o□□□e□□□s3□□ n3t□6□□□ I□ 7□hN□s□□□tl l □□t□3□iv□ niinu ts M□3a□□3□oM□3□7□hN□s□□1□a□□□ts□ts M□n □3a□□ M□□□g□□781□nl di□w□n□□itb□7□3a□□□781□3n□3a□□M□snl t□□3□oM□3□7□hN□s□□□7n□3a□3□ 3a□□u ti□B3vd□□□781□□s□□N3□□n3a□73o□s□7□h □□□□□□□N7ts M□□□□nN□i□□□3o□s□□□o□□r□ 6□□□I□6□tMNo□□H8I1□□□□□□□□□s□□□□o□d□to□□□3aonNMa□ns□□n□3u n□M□s□o□i□o□d□to□

[illegible]

[illegible]


Guide RNA
Target specific
crRNA sequence
Active
sites
Cas9
Cas9
CRISPR in Action
Target DNA
sequence

[illegible]

[illegible]

[illegible]

[illegible]

DNA METHYLATION ANALYSIS

Genomic DNA extraction

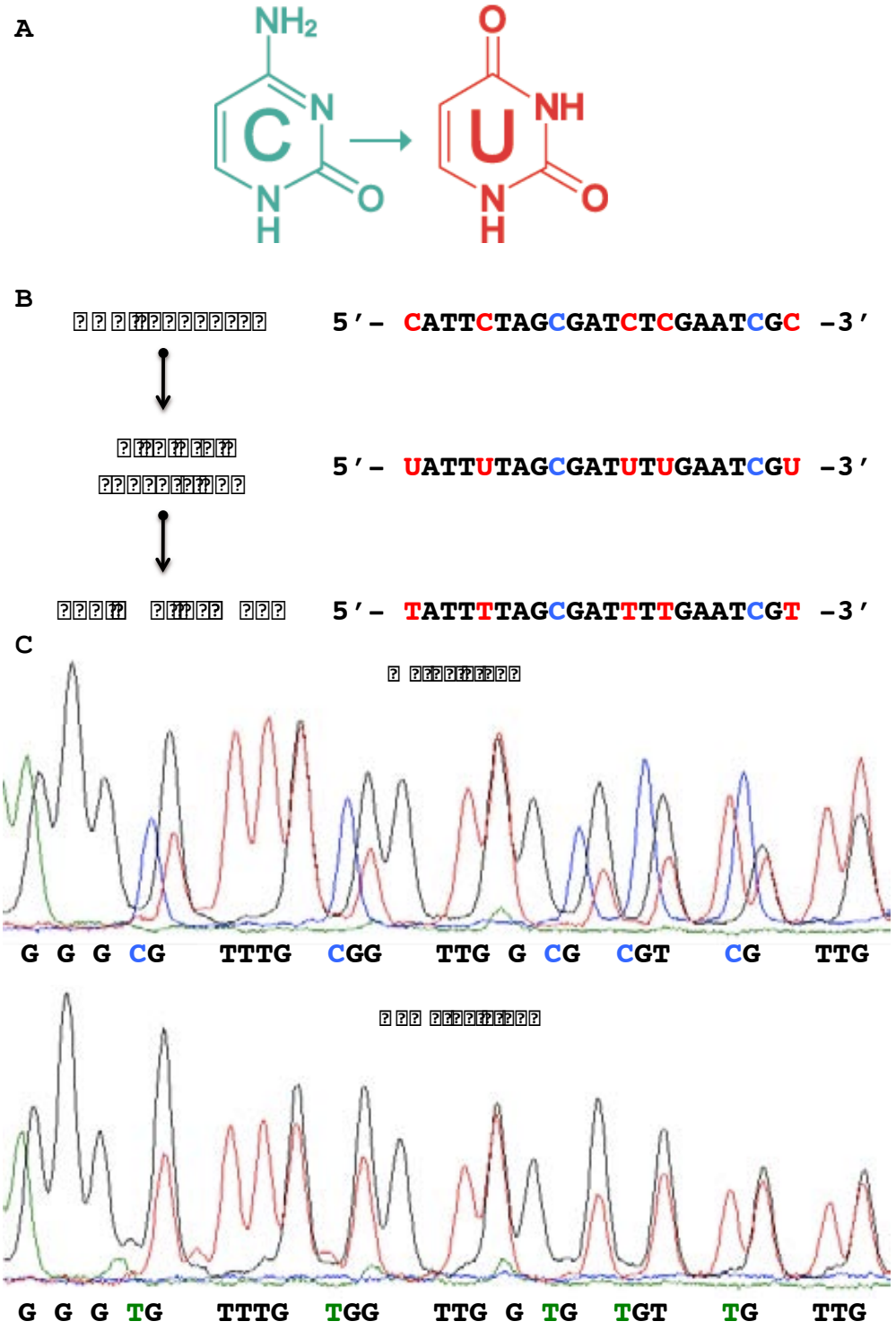
Genomic DNA extractions were done using the PureLink™ Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, USA), according to manufacturer's instructions. DNA was eluted using elution buffer and quantified using the NanoDrop (Thermo Scientific, Rockford, IL, USA).

Bisulfite treatment

There are different methods to study DNA methylation at specific genomic loci (Laird, 2010; Lister & Ecker, 2009). They range from methylation-sensitive restriction enzymes to Methylated DNA immunoprecipitation (MeDIP) (Down *et al.*, 2008) and bisulfite modification of DNA (Frommer *et al.*, 1992).

Sodium bisulfite modification followed by sequencing is currently the gold standard in DNA methylation analysis as it provides single base resolution and the highest coverage (Lister & Ecker, 2009). Sodium bisulfite treatment deaminates unmethylated cytosines (C) to produce uracil (Figure 30A), while methylated cytosines (mC) remain unaltered. Uracils are read as thymines by DNA polymerase. In that way, umethylated cytosines will appear as thymines in sequencing of PCR amplicons obtained after sodium bisulfite treatment of DNA. Alternatively, methylated cytosines will be preserved as cytosines in the electropherogram (Figure 30B and 30C). By comparing the modified DNA with the original sequence, the methylation state of the original DNA can be inferred.

Bisulfite conversion was performed using the EZ DNA Methylation™ kit (ZymoResearch, Orange, CA, USA), using 250 ng of DNA as starting material and eluting in 50 µl of elution buffer.


A
B
5′- CATTCTAGCGATCTCGAATCGC -3′
5′- UATTUTAGCGATUTUGAATCGU -3′
5′- TATTTTAGCGATTTTGAATCGT -3′
C
G G G CG TTTG CGG TTG G CG CGT CG TTG
G G G TG TTTG TGG TTG G TG TGT TG TTG

[illegible]

[illegible]

[illegible]

that this degradation affects between 84–96% of the total DNA (Grunau *et al.*, 2001).

Because of the poor quality of the DNA recovered from bisulfite treatment, nested PCRs were performed to improve the performance. Two µl of bisulfite converted DNA were used in the first PCR (external). The second PCR (internal) was performed using one µl of a dilution of the first one. Dilution ranged from 1:1 to 1:20. Bisulfite primers are listed in Annex II.

Infinium 450K methylation arrays

To obtain a genome-wide profile of DNA methylation in HCT116 control cells and treated with 5-AzadC, we performed 450K Infinium methylation array, which has a good coverage of the whole genome, specially CpG islands (97% are covered) and other interesting regions, such as shore CpGi islands and promoters.

One µg of DNA was quantified using the Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA), and transformed using the EZ DNA Methylation ™ kit (ZymoResearch, Orange, CA, USA). Biological duplicates of each experiment were performed. Methylation probes used in this retinoic acid pathway experiments are listed in the Annex VI.

RNA EXPRESSION ANALYSIS

RNA extraction

Total RNA extractions were done using TRIzol® (Invitrogen, Carlsbad, CA, USA), according to manufacturer's instructions and following the single-step RNA isolation method (Chomczynski & Sacchi, 2006). All processes were carried out on ice and centrifugations steps at 4ºC in order to preserve RNA integrity.

Cultured cells were trypsinized and washed with cold PBS 1X twice before RNA extraction. Then cells were resuspended using one ml of TRIzol® per 2.10^6 cells. In the case of tissue samples, mechanical homogenization with the razor blade was crucial to obtain high recovery yields. At that point samples could be stored at -80ºC for at least one year. During the extraction protocol, we applied a recombinant DNAse I treatment (Applied Biosystems, Foster City, CA, USA) to

avoid DNA contamination mainly from those samples coming from transfection experiments.

RNAs were quantified in the NanoDrop (Thermo Scientific, Rockford, IL, USA) and were run in a 1% agarose gel with ethidium bromide staining in order to assess the integrity of the RNA.

We also performed RNA expression microarrays and RNA-seq of HCT116 control cells and treated with 5-AzadC. Each experiment was performed in biological duplicates. For these experiments, total RNA was extracted with the miRNeasy® Mini kit (Qiagen, Venlo, Netherlands) which allows the recovery of all sizes RNA. The quality of these RNAs was checked using Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and a minimum score of 9 was required to be included in the experiment. Two µg of RNA were used for each RNA-seq experiment. For RNA expression microarray analysis, we used Agilent SurePrint G3 8x60K array (Agilent Technologies, Santa Clara, CA, USA) and the data obtained were normalized and analyzed statistically using Limma (R package) (Smyth *et al.*, 2005).

Reverse transcriptase PCR

For each reaction we used 500 ng per sample that were reverse transcribed by M-MLV (Invitrogen, Carlsbad, CA, USA), in the presence of pd(N)6 Random Hexamers (GE Healthcare, Piscataway, NJ, USA) and RNasin® ribonuclease inhibitor (Promega, Madison, WI, USA). The reaction was performed according to manufacturer's recommendations in a final volume of 25 µl. Two negative controls were added, one without the M-MLV enzyme and the other without the RNA.

Quantitative Real-Time PCR

The expression levels were analyzed using the quantitative Real-Time PCR technique in the LightCycler® 480 (Roche Diagnostics Corporation, IN, Indiana, USA) platform with Fast Start DNA Master SYBR®Green I mix (Roche Diagnostics Corporation, IN, Indiana, USA) in a final volume of 10 µl.

Samples were diluted 4 to 20 fold (except for 18S control that was diluted 1000 fold), depending on the expression levels of each gene. Every sample was

analyzed in triplicate. When possible, forward and reverse primers were designed in different exons to avoid genomic DNA interference. Amplified products were run in an agarose 2% gel and sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). We used 3 to 6 housekeeping genes (PUM1, MRPL19, PSMC4, 18S, B2M (β2-microglobulin) and PPIA) in order to normalize the expression levels. Efficiency was calculated for each PCR reaction as described (Schefe *et al.*, 2006). Standard deviation was calculated on PCR triplicates. Expression primers are listed in Annex III.

CHIP AND CHIP-SEQ EXPERIMENTS

We used chromatin immunoprecipitation (ChIP) against some post-translational histone modifications, histone variants and transcription factors (TF) (summarized in Table 11), in order to characterize the epigenetic landscape of the genes of interest.

In this study we used two different applications of ChIP: the locus specific ChIP (analyzed by quantitative PCR, ChIP-qPCR) and the genome-wide ChIP (analyzed by next generation sequencing, ChIP-seq). Both applications were performed using the Millipore EZ-Magna ChIP™ A/G kit (Billerica, Massachusets, USA) according to the manufacturer's instructions.

First we collected cells by trypsinization (as explained previously). Then we proceeded to the cross-linking step (1% formaldehyde, 37ºC, 15 min) followed by a 5 min inactivation step with 125mM glycine. $20x10^6$ cells in 1mL of cell lysis buffer were sonicated using Bioruptor sonicator (Diagenode, Liège, Belgium) with 60 s 'on' and 60 s 'off' for 34 min.

Fragmented chromatin was aliquoted in 50 µl (10^6 cells), which is the recommended starting material for the EZ-Magna ChIP™ A/G kit.

In this study we carried out whole genome ChIP-seq in HCT116 control cells and HCT116 treated with 5-AzadC (as described previously). Besides, ChIP-qPCR was performed for specific regions in different cell lines and also to validate ChIP-seq data.

[illegible]

Antibody	Source	Catalogue ID	Ab/IP	ChIP-seq Control cells [1]	ChIP-seq Treated cells[2]
H3	abcam	ab1791	5 µg/IP	1	1
H3K4me3	Millipore	NG171579	0.6 µg/IP	1	1
H3K4me	Millipore	07-436	3 µg/IP	3	2
H3K27me3	Millipore	07-449	5 µg/IP	1	1
H3K27ac	Millipore	05-1334	3 µg/IP	2	2
H3K9me3	Diagenode	pAB-060-050	1 µg/IP	1	1
H3K9me2	Diagenode	pAB-056-050	3 µg/IP	1	1
H3K36me3	abcam	ab9050	5 µg/IP	1	1
H4K20me3	abcam	ab9051	5 µg/IP	5	9
H2A.Z	abcam	ab4174	5 µg/IP	2	5
mH2A	Dr. Marcus Buschbeck lab		5 µg/IP	4	4
RNApol II	Millipore	05-623	1 µg/IP	3	4
RNApol III	Santa Cruz	sc-21754	2 µg/IP	5	9
CTCF	Millipore	17-10044	5 µg/IP	2	5
IgG	Jackson Immunoresearch	011-000-120	5 µg/IP		

[illegible]

[illegible]

[illegible]

PROTEIN ANALYSIS

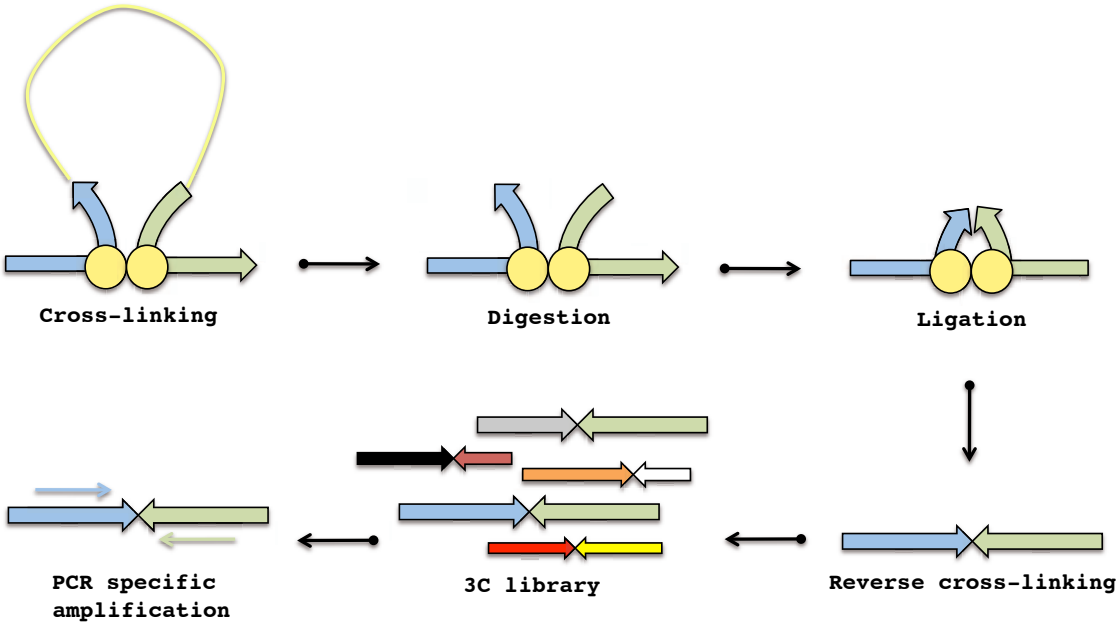
Western blot

Tissues were snap-frozen in liquid nitrogen, then smashed and homogenized in lysis buffer, containing 10 mM Tris-HCl (pH 7.5), 1 mM MgCl2, 1 mM EGTA, 10% glycerol, 0.5% 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesul-fonate, 1 mM β -mercaptoethanol, and 0.1 mM phenylmethyl sulfonyl fluoride. Extracts were vortexed for 30 minutes at 4°C and, after centrifuging for 20 minutes at 15,000 g, the supernatants were stored at -20°C. Before doing the western blot experiment, extracts were boiled for 5 minutes in Laemmli sample buffer and equal amounts of protein (20-30 μg) were separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred electrophoretically to a immobilon-P membrane (Millipore, Billerica, Massachusets, USA). The membrane was blocked in TBS 0.5% plus 5% nonfat dried milk for 90 minutes at room temperature and probed with a primary antibody AKR1B1 (1:3000; Dr. Jaume Farrés Lab) or AKR1B10 (1:3000; Sigma-Aldrich) at 4°C overnight. The membrane was washed thrice for 10 minutes in TBS 0.5% and incubated with anti-rabbit secondary antibody (1:3000; Dako, Agilent Technologies, Santa Clara, CA, USA) for 90 minutes at 4ºC. Targeted proteins were visualized using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Rockford, Illinois, USA). We used GAPDH (Abcam, Cambridge, Massachusetts, USA) as a loading control.

Immunohistochemistry

A tissue microarray with 196 tumors from patients diagnosed of advanced colorectal cancer and homogenously treated with combinations of 5-FU and irinotecan was constructed. Sections were stained using a fully automated staining system for immunohistochemistry (Ventana Medical Systems, Tucson, Arizona, USA). The primary antibody used was anti-AKR1B10 (Sigma-Aldrich) and anti-AKR1B1 (Dr. Jaume Farrés Lab) at dilution 1/100. Immunostaining of HepG2, HCT116 and SW480 cell lines were used as a controls. The staining was semiquantitative evaluated as diffuse positive, focal and negative.

[illegible]


Cross-linking
Digestion
Ligation
PCR specific amplification
3C library
Reverse cross-linking

[illegible]

[illegible]

ligation product is the result of a unique interaction, the quantification and comparison of different interactions was done by quantitative PCR using specific primers for the given interaction. The quantification of these ligation products in a given population of cells leads to relative frequency maps where physical interactions can be inferred. The specific primers used in the AKR1B1 locus are described in the Annex VIII table.

BIOINFORMATICS

The Cancer Genome Atlas project

Most of the genomic data used in this project were downloaded from The Cancer Genome Atlas (TCGA) project. This huge project attempts to get a better knowledge about the most common cancers (17 cancers in the last update, 09/10/2013), by applying genome-wide techniques that provide information on gene expression, DNA sequence and DNA methylation to large series of cases. We have used the TCGA database to investigate DNA methylation, gene expression and gene mutation in the regions or genes of interest, using publicly available bioinformatic tools or developing our own scripts.

TCGA: Gene expression

Colon cancers expression array data was downloaded from The Cancer Genome Atlas database (tcga-data.nci.nih.gov).

Gene expression data from the Agilent G4502A microarray (Agilent Technologies, Santa Clara, CA, USA) was available for 155 colon cancer patient samples. There were 12 tumor-matched normal tissues and 7 normal unmatched tissues available that were also included in the analysis.

Expression results were obtained using the TCGA dataset on Oncomine database (www.oncomine.com) and using in-house made R scripts. TCGA scripts were constructed with R (http://www.r-project.org/) using the Bioconductor open source R package (http://www.bioconductor.org) (Gentleman *et al.*, 2004). Expression probes are depicted in Annex IX.

Cancer Type	Code	Normal Samples	Tumor samples
Bladder Urothelial Carcinoma	BLCA	78	13
Brain Lower Grade Glioma	LGG	2	140
Breast invasive carcinoma	BRCA	91	548
Colon adenocarcinoma	COAD	38	254
Glioblastoma multiforme	GBM	1	112
Kidney renal papillary cell carcinoma	KIRC	160	283
Kidney renal clear cell carcinoma	KIRP	44	81
Liver hepatocellular carcinoma	LIHC	49	98
Lung adenocarcinoma	LUAD	3	220
Lung squamous cell carcinoma	LUSC	40	149
Pancreatic adenocarcinoma	PAAD	6	24
Prostate adenocarcinoma	PRAD	49	153
Rectum adenocarcinoma	READ	7	95
Stomach adenocarcinoma	STAD	2	69
Thyroid carcinoma	THCA	28	230
Uterine Corpus Endometrioid Carcinoma	UCEC	25	334

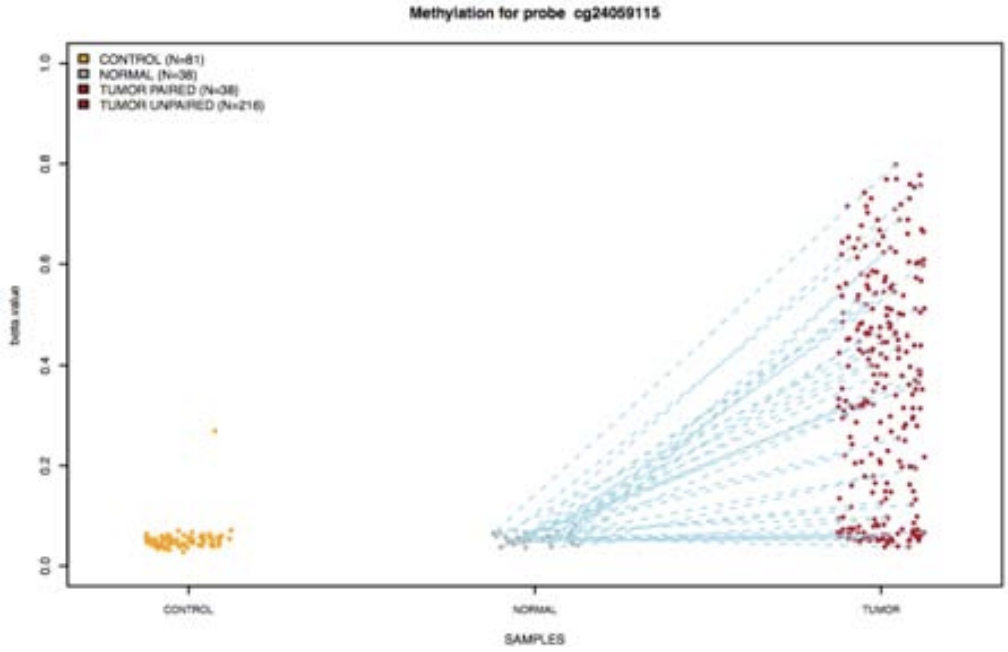


Figure 32 | Example of the output generated by the script developed to visualize DNA methylation data from the TCGA.

The second method is the log2 ratio of the intensities of methylated probe versus unmethylated probe (Irizarry *et al.*, 2008). We have used the beta value to represent DNA methylation data as it offers a biological meaningful information.

For DKO DNA methylation we used already published data from Dr. Peter Jones, stored at Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo), under the code GSM896399 (De Carvalho *et al.*, 2012).

TCGA: Gene mutation and copy number alterations

We also used the TCGA data to investigate genetic alterations in our genes of interest. We used genetic TCGA data through the cBioportal webtool (http://www.cbioportal.org) (Gao *et al.*, 2013).

Other bioinformatic tools

We have used DAVID (http://david.abcc.ncifcrf.gov) webtool for Gene ontology (GO) and KEGG pathways analysis (Kyoto Encyclopedia of Genes and Genomes) (Huang *et al.*, 2009).

In order to define the RA pathway genes expression in normal colon mucosa we used the Refexa standarized database (http://sbmdb.genome.rcast.u-

tokyo.ac.jp/refexa). High and moderate expression was assessed using RefExa criteria.

For the composition of the survival test, we used 4 gene expression datasets of human colorectal cancers (GEO: GSE14333, GSE12945, GSE17536, GSE17537)(Bailey *et al.*, 2012; Jorissen *et al.*, 2009; Staub *et al.*, 2009) with follow-up information available to analyze the survival rate depending on gene expression. Using PrognoScan algorithm (Mizuno *et al.*, 2009) patients were divided into two groups depending on gene expression levels. High or low expression groups were made using the minimum p-value approach. Survival curves were constructed using the Kaplan-Meier method.

DNA METHYLATION PROFILE OF AKR1B GENES IN COLORECTAL CANCER

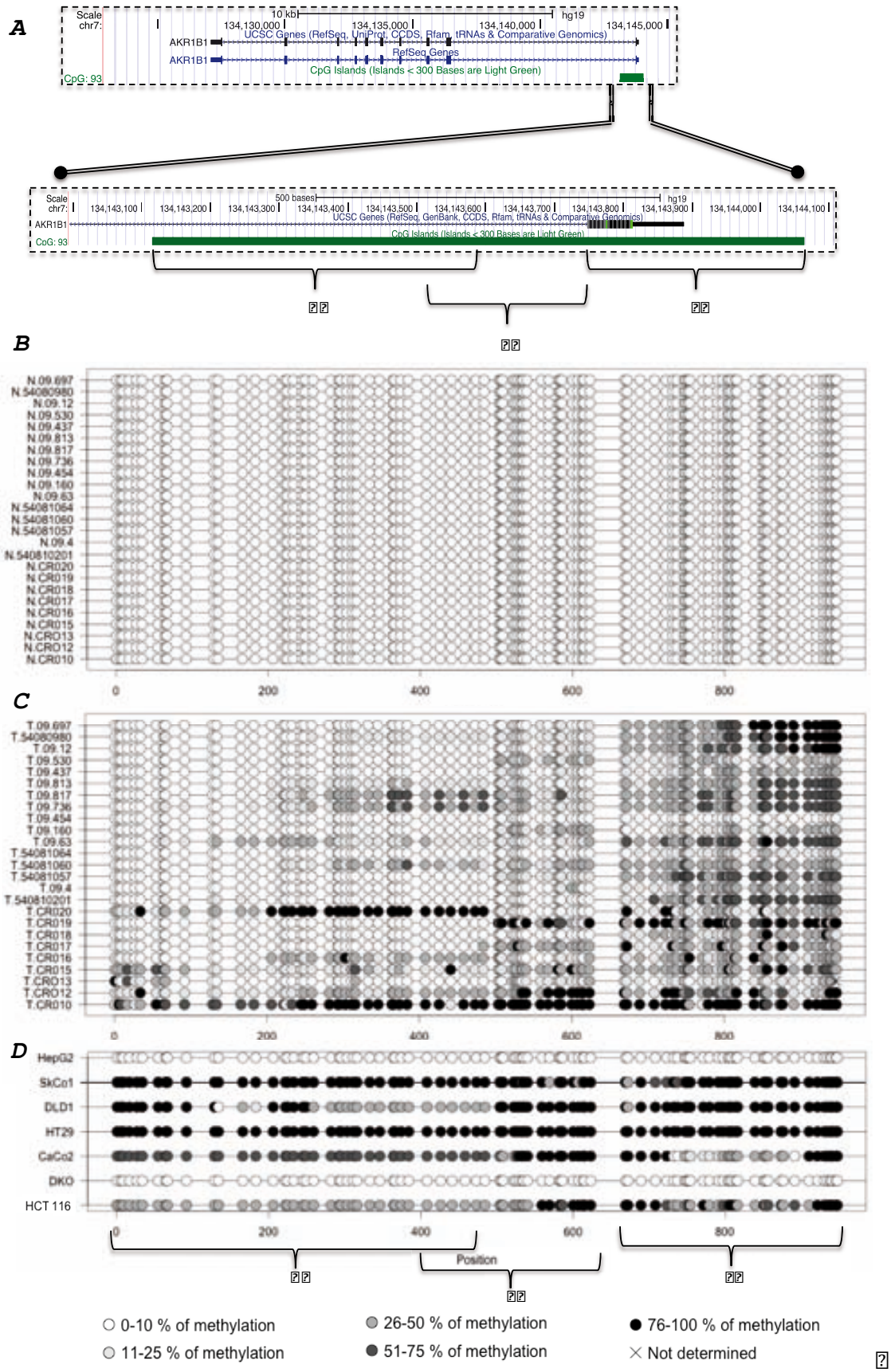
DNA methylation of AKR1B1 by bisulfite sequencing

Silencing of multiple cancer-related genes is associated with *de novo* DNA methylation of promoter CpG islands. In this study we have focused on AKR1B1 gene, which has a CpG island (CpGi) associated to its promoter (Figure 33A). We analyzed the DNA methylation status of AKR1B1 CpGi island in a panel of 7 cancer cell lines and in 25 colon tumor samples, as well as in their associated normal mucosa by bisulfite sequencing (see materials and methods for more detail).

To cover the 93 CpG dinucleotides that comprise the CpG island, we needed to perform 3 independent bisulfite PCRs, named A, B and C (Figure 33A). The results obtained from the three amplicons within the AKR1B1 CpG island have provided an overview of the methylation landscape of this gene in colorectal cancer (Figure 33).

All 25 normal colon mucosa samples displayed a fully unmethylated pattern all along the CpG island (Figure 33B). On the contrary, almost all tumors analyzed (24 out of 25) exhibited hypermethylation. It is important to note that hypermethylation is more prone to occur in the C segment of the CpG island (Figure 33C), where the AKR1B1 transcription start site (TSS) is located (Figure 33A). Clonal analysis of tumor samples with partial DNA methylation showed the coexistence of densely methylated with poorly methylated molecules in all cases, confirming the presence of cell populations with heterogeneous DNA methylation profiles.

Most of the colorectal cancer cell lines presented a high grade of DNA methylation in the AKR1B1 CpG island (HCT116, CaCo2, HT29, DLD1 and SkCo1), except for DKO (deficient for DNA methylation) (Rhee *et al.*, 2002) which exhibited a fully unmethylated pattern. Hepatocellular carcinoma cell line HepG2 was also fully unmethylated. In the cell lines analyzed, there was no enrichment in hypermethylation regarding the C amplicon. In most cases the whole CpG island displayed a homogeneous DNA methylation profile.


A
Scale
chr7:
10 kb
hg19
134,130,000
134,135,000
134,140,000
134,145,000
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparative Genomics)
AKR1B1
RefSeq Genes
AKR1B1
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 93
500 bases
134,143,100
134,143,200
134,143,300
134,143,400
134,143,500
134,143,600
134,143,700
134,143,800
134,143,900
134,144,000
134,144,100
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
CpG Islands (Islands < 300 Bases are Light Green)
CoG: 93
B
C
D
HepG2
SkCo1
DLD1
HT29
CaCo2
DKO
HCT 116
0
200
400
600
800
Position
0-10 % of methylation
11-25 % of methylation
26-50 % of methylation
51-75 % of methylation
76-100 % of methylation
Not determined

DNA methylation of AKR1B10 by bisulfite sequencing

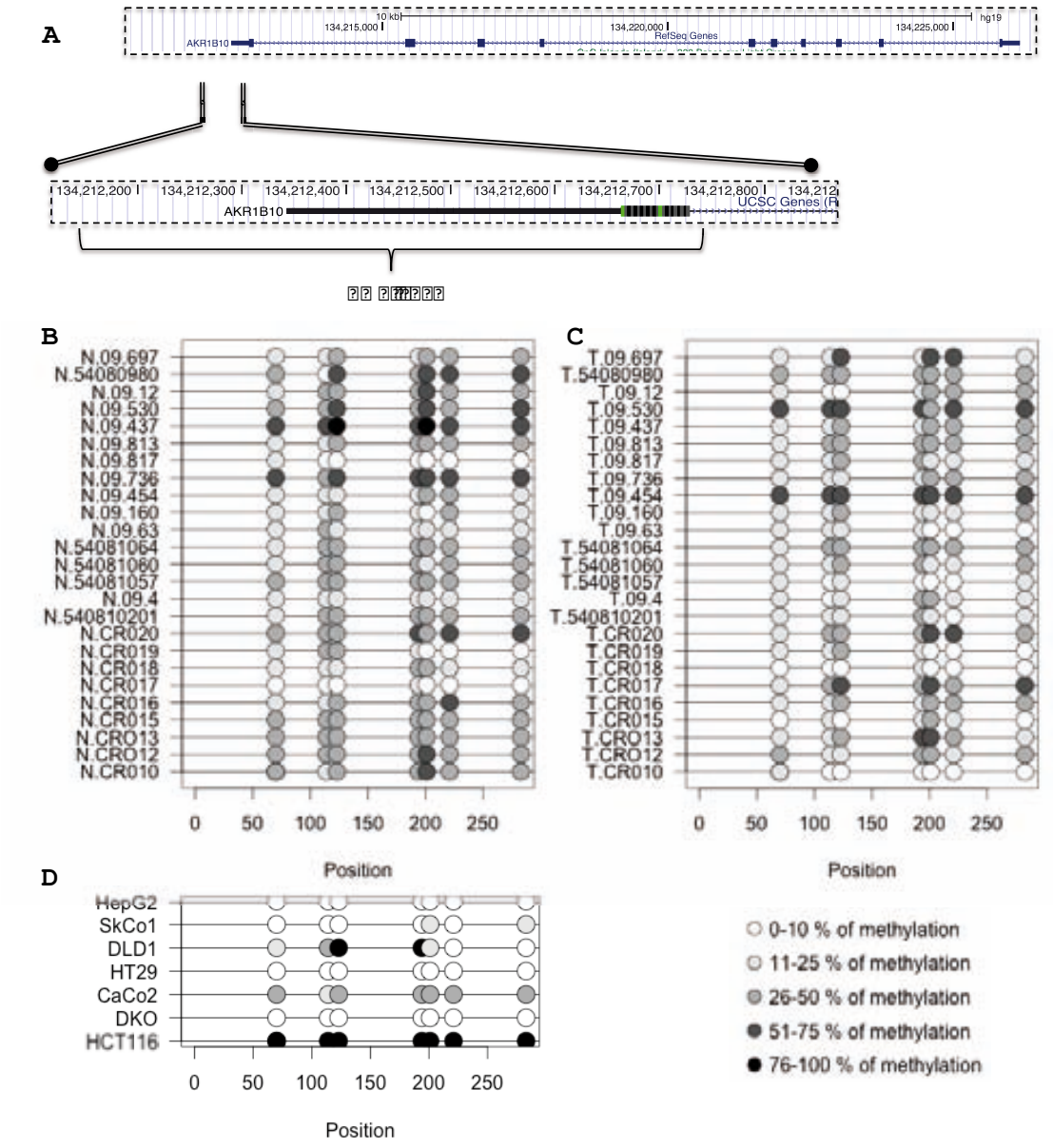
The DNA methylation level of the promoter region of the AKR1B10 gene was analyzed although it did not have an associated CpG island. The analyzed amplicon was composed by seven CpG dinucleotides (depicted in Figure 34A).

We have studied AKR1B10 promoter DNA methylation in the 25 colorectal tumors and in the associated normal mucosa. Seven cancer cell lines were also analyzed.

The DNA methylation pattern of AKR1B10 was heterogeneous among the different normal samples. For example, samples N.09.817 and N.CR017 were almost unmethylated, whereas N.09.437 was heavily methylated. In contrast to AKR1B1 promoter, AKR1B10 promoter tended to be methylated in normal colorectal tissue (Figure 34B).

DNA methylation of AKR1B10 in colorectal tumor samples showed a great variability among the different samples. Some of them were almost unmethylated such as T.CR010, while T.09.454 and T.09.530 were almost fully methylated (Figure 34C).

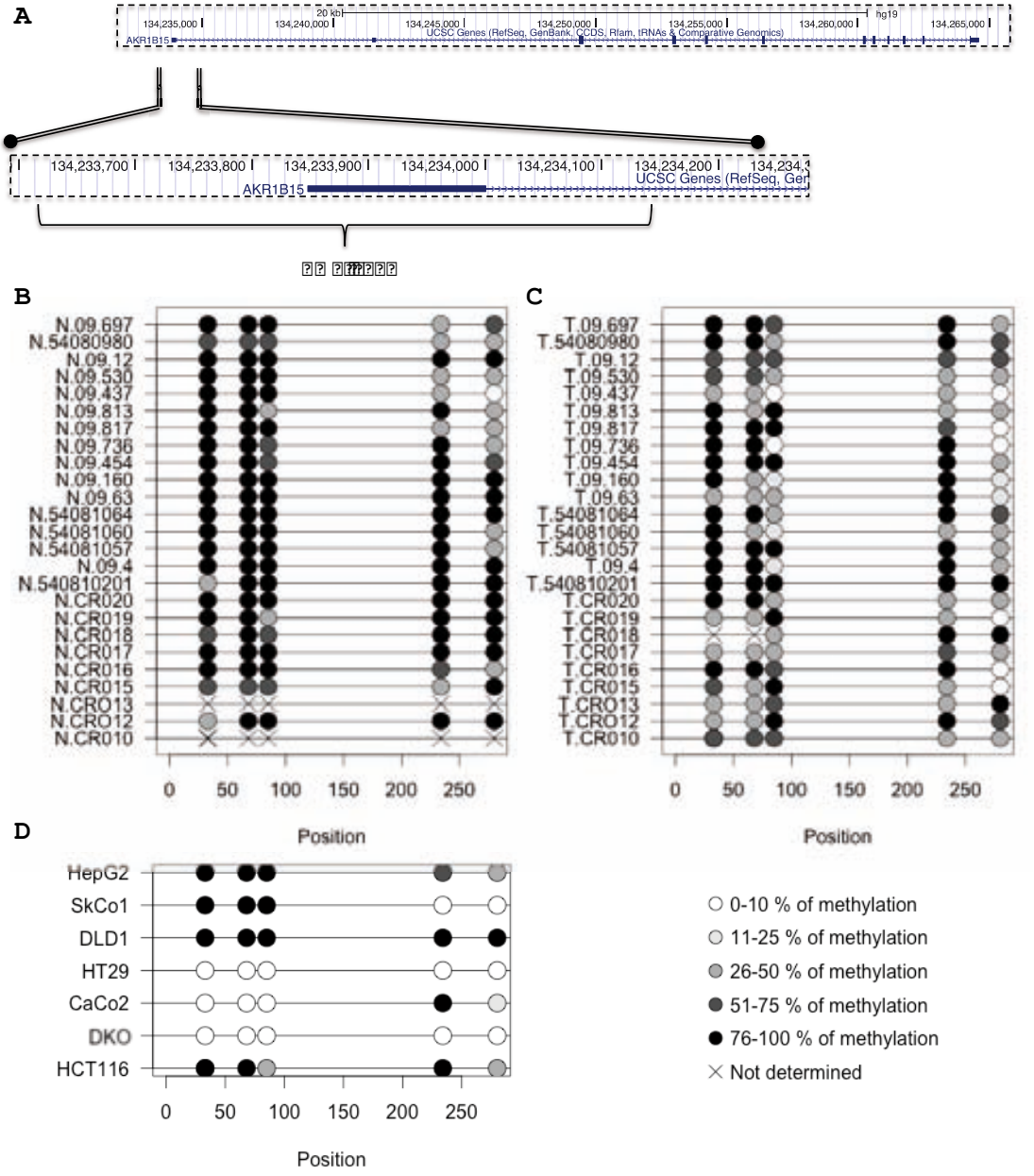
As it occurred in normal colon mucosa, AKR1B10 promoter tended to be methylated in colorectal tumor samples, but slightly less methylated than in the normal tissue. Regarding cell lines methylation, HT29, DKO and HepG2 displayed a completely unmethylated profile, while SKCo1, DLD1, CaCo2 and HCT116 presented different degrees of DNA methylation (Figure 34D).



ahG. g3 C Cd A Dm l mA A 1 mhyemA A yaD N eDmAyla 3 C Cd N l f Dl mAyl lay Ap2AmhN DmDmAN amy yA A 3 N l laNmhyN eDh 3 N l f Dl mAh C Cd Ae A Dh A DDA yh

N l f Dl mA h C C2 f yaD y t a A A

R ((t elaR c n9m e gR cga nch 9gRe Rot t 9oc h gR uc g9gR ga ge9 9 ((W v l e h), ((KA e ((WR c 9 S acl h m e g9 f lo g gR t gRal ge9 9 gR W S e o l9ge c9 gR Sn9t 9g n v e on FW KA



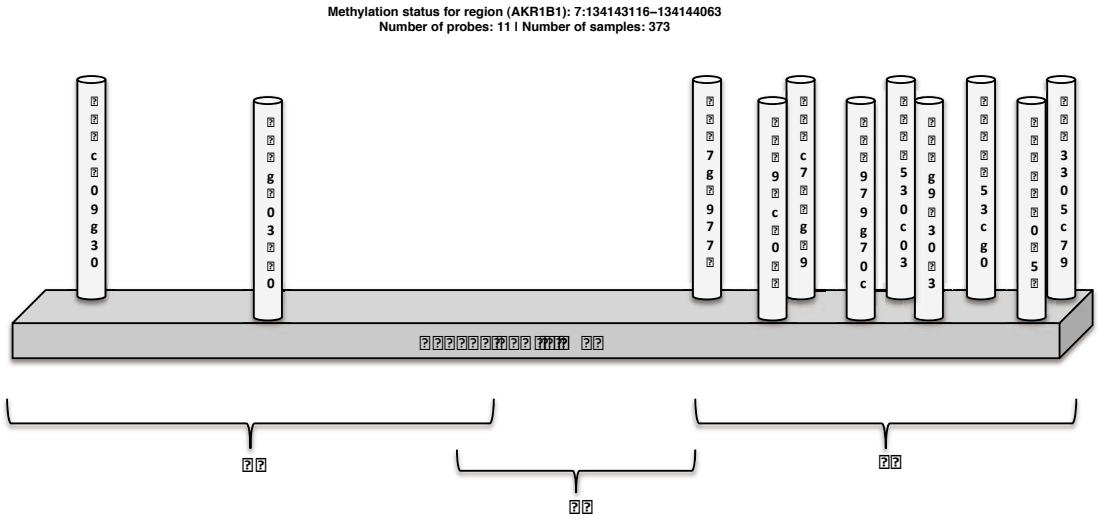
ah G2 g 3 C C2 A Dm l mA A l mhh yem A A yaDl N eDmA yla h 3
C C2 N lf Dl mAyl lay Ap2 AmhN DmDmA Na myy A Am3l N l laN mhy N eDyh
3 N lf Dl mAm C C2 A e A Dm A hDDA yh

c eg ec cR9m e e on FW (W Sn9t 9g n g c g9 Re Rla t gRal g e 9nt l 9l9 t Sl c ecSl ac enRaS9t gRal ge9 e 9l9 got 9n t Sl chte el nla g9 mR gR SS e (, Sn9t 9g nA n e gR ll le c S l he) 5 m n 9t Sl g la o t gRal g h mRe 9) l S) l u9(h y((z m n f ne la t gRal g A

C C N lf Dl mA ay A

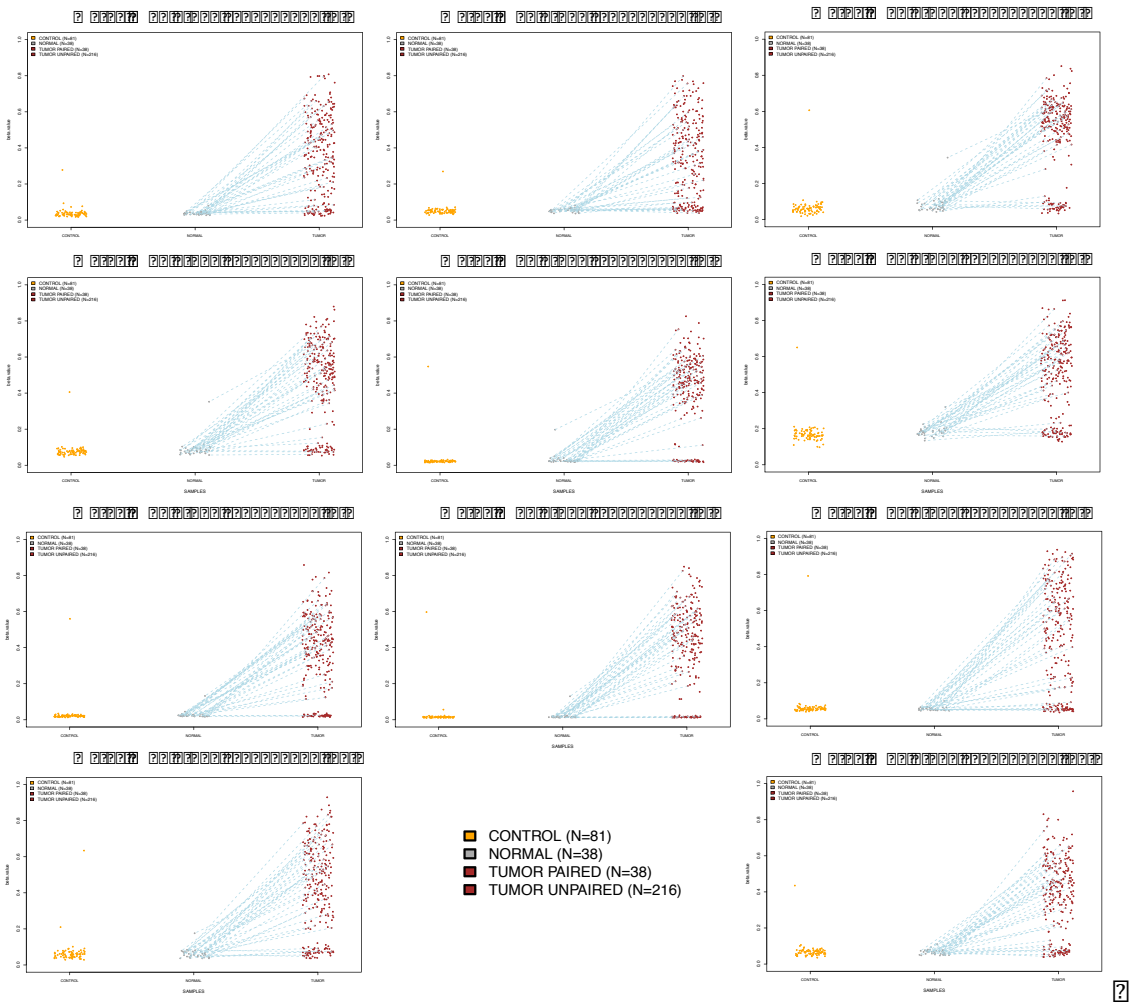
9n ng9 9 ant 9on Sn fe9oc 9 c nf ge9 ch m oc e S g 9R9ng9 9l9n g l 9 n e 9t S ge gc v K t gRal ge9 g c g m9t R n 9t g l c v K vt g ne lc t gR9 cKA R 9R9ng 9 cecgc 9)W3 9l9 got 9n c t Sl c F2 t g R 9nt l 9l9 t o9c cA R f cgo e gR 9l9 t gRal ge9 g oce c ne Sg vc t g ne lc t gR9 cKA

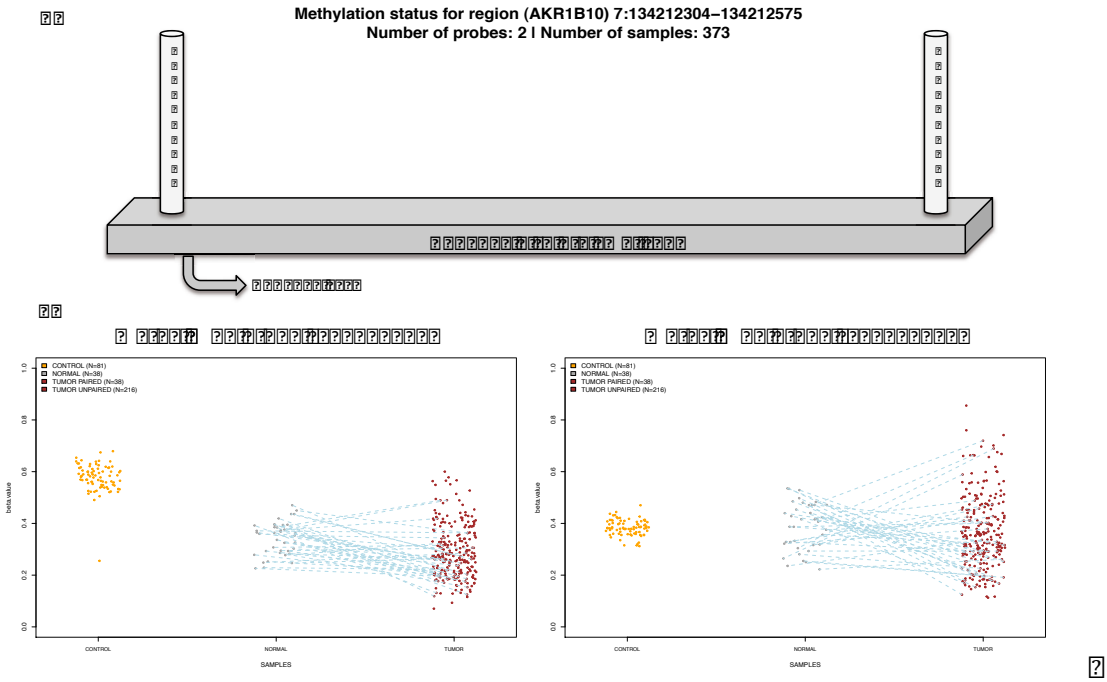
lar gR t gRal ge9 e gR S ecl cc9 e g megR gR ((Sn9t 9g nhgR c t n e9 Sn fe9ocla cgo e a ecol g A Rc n e9 ec 9f n a ((Sn9 c e gR llot e 3W, t gRal ge9 nn ac v e on Fz KA



ah G5g CC DaN A 2d N lf Dl mAe hm yDm l A C C e yDA Dm l mA m G yaD N e D mAy R A 3y Dym e l h

9 ant e 9on Sn fe9oc 9 c nf ge9 ch ((S ecl m c 9t Sl g la o t gRal g e 9nt l 9l9 llc ka S nt gRal g e 9l9n g l n e 9t v e on F7 KA


CONTROL (N=81)
NORMAL (N=38)
TUMOR PAIRED (N=38)
TUMOR UNPAIRED (N=216)



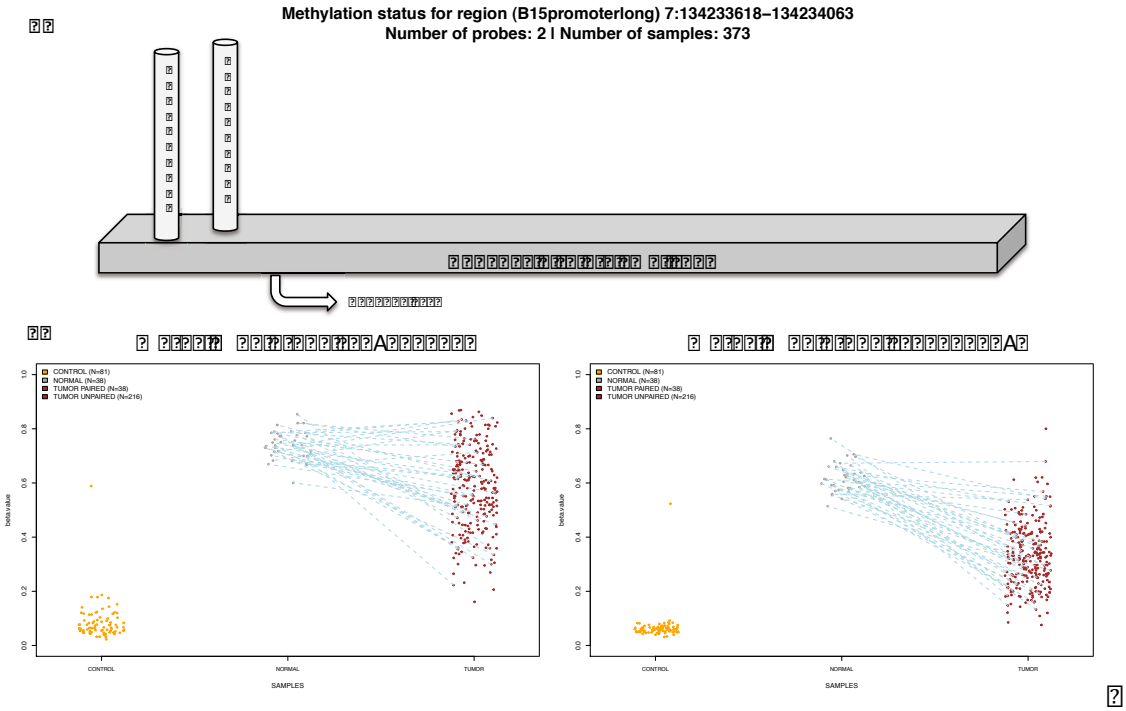
ah G4 g 3 lo m Da N A 2d N l f Dl mA eh m y Dm l A C Cd eh mN ml hl 3 mDA Am h AmN N l f Dl mA h ya Dy mh lo m eh m y mA l A A C Cd eh mN ml hl

ce gR t gRal ge9 g m 9c nf me f ne ge9 e 9nt l c t Sl c v), yW b t gRal ge9 K f t 9n e got 9n c t Sl c v e on F2 KA cSeg gRec R g n9 ega m 9ol 9g9 c nf l n g a e t gRal ge9 one gR n9oc Sn9 ccA R c g t g R megR9on Sn f e9oc ecol g n colgcA

C C2 N l f Dl mA ay A l

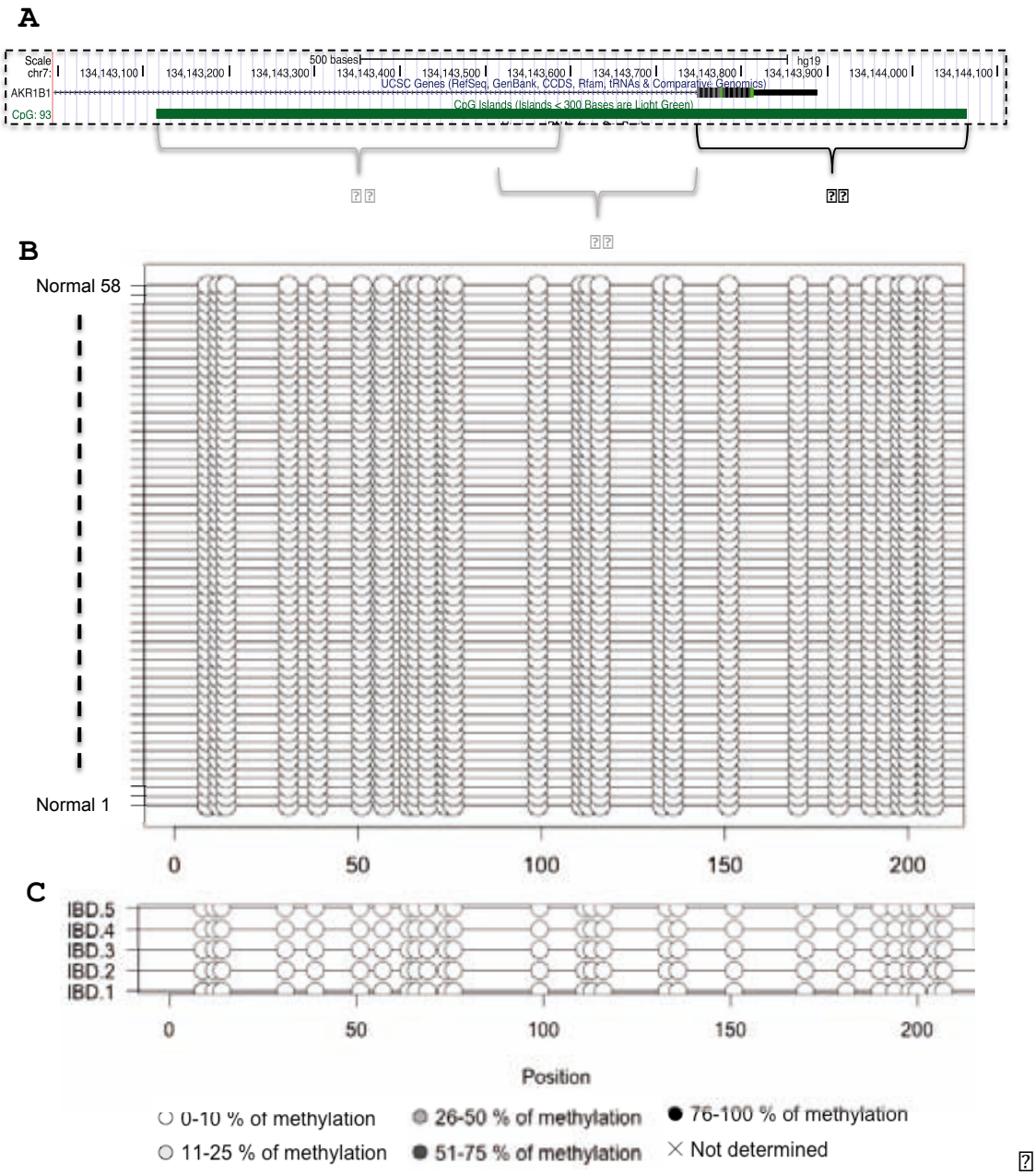
R n m n gm9 t gRal ge9 Sn9 c megRe gR Sn9pet l Sn9t 9g n9 gR ((W A R cego ge9 gR gm9 Sn9 c gR e e ge9 gR n Se g e on F5 A

((W Sn9t 9g n pRe eg Re R9nt 9 n g f lo c9 t gRal ge9 e gR 9nt l 9l9 t o9c v S e 9 gR Sn9 9 ce n KmRel got 9n c t Sl c ecSl a t gRal ge9 gR g m c t 9n 9n cc t f n S e 9 gR c t Sl Sn9 e on F5 KA



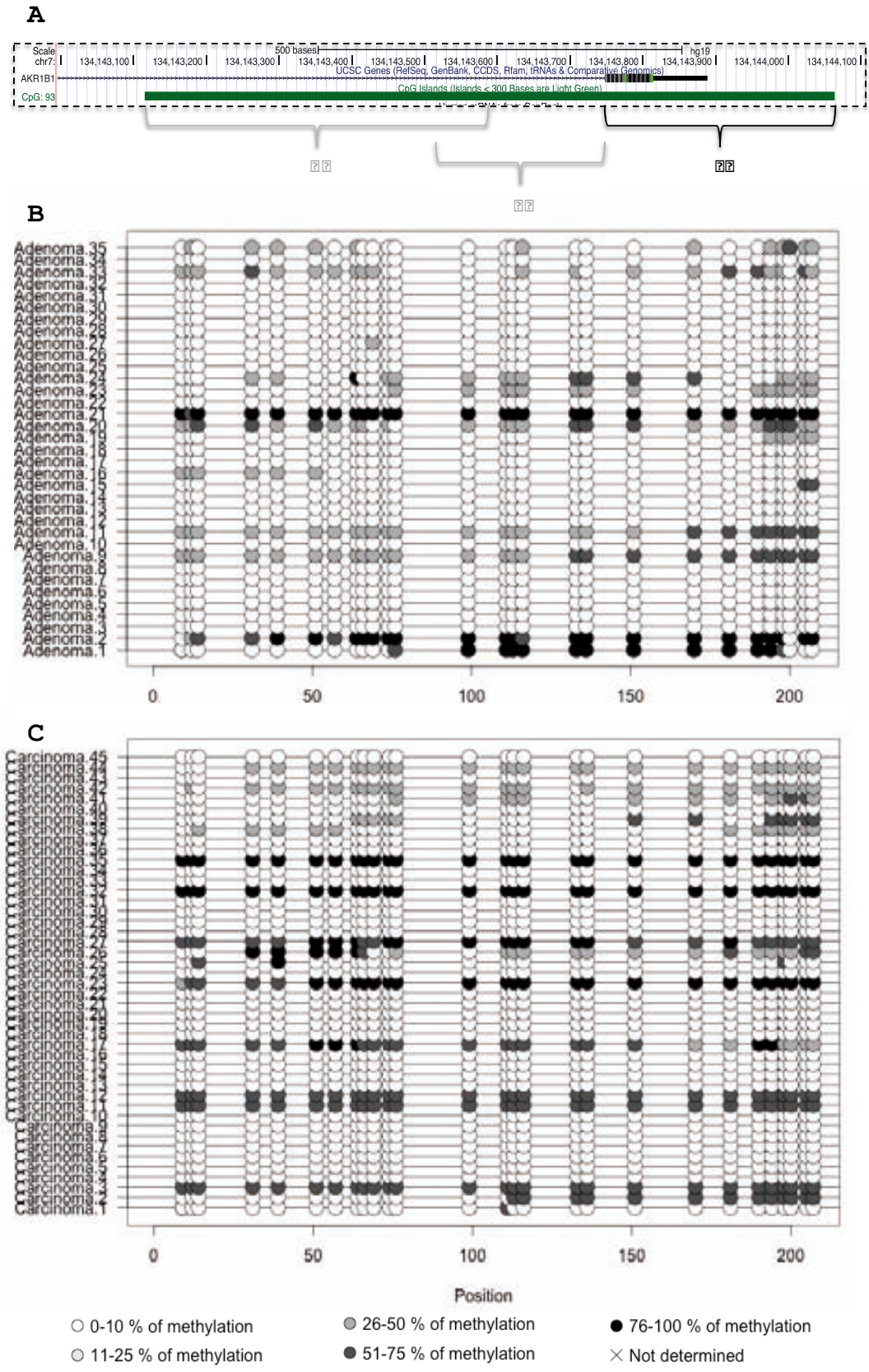
C C N l f D l mA A y l m D

t gRal ge9 e9t nu nc 9n 9 ye f cef e 9cec 9 9l9n g l n Sn onc9n ce9 c R f pg cef la tgo e A e n gS lc R f oc e gg t Sg g9 et Sn9f omn g Sn9g9 9lc e le e l Sn ge h lgR9o R 9 e eg e9t nu nc R f cg le cR A n m f lo g gR t gRal ge9 9 gR c((S ecl c m e 9cge t nu n 9n 9l9n g l ne e 9R9ng 9t S9c a 3F tg99l c t Sl c c g ne lc t gR9 cK lar gR t gRal ge9 9 gR c t g megRe gR ((S ecl oc gRec ec gR 9 gR g ecSl ac t 9n h io g S RaS nt gRal ge9 v e on FF KA R n colgc9 g e h9t gR cg99l m n 9t S n meg gR gR t g R got 9n gecco mR g n c f e l lA ll gR cg99lc n9t R lgRa 9 9nc lar pRe eg olla o t gRal g S gg n v e on 3, KA Rec 9e e c megR gR g g R g m R f 9g9 c nf t gRal ge9 9 gRec S ecl e a9 gR 9nt l 9l9 t o9c c t Sl c lar A gR c 9 e l t t g9na 9m l ec c t Sl ch m e 9g9 c nf t gRal ge9 h c gR SS megR R lgRa 9 9nc v e on 3, KA



ah. dg3AmN DnlmAmlNeDmAo AlCCeyDAh3 N lf DlmAhyaDymlmDyNeDyhmN DfmAmhyb3 N lf DlmAhyaDymlmDyNeDyhmN ADN N lmfmo Dyy3e lAlyb

nle cg99l ctt Slchm9t S ge gcmegR 9t hmm9olgg g t gRalge9 e F9og9 FWtt Slchm Re Rh Snc gc g ge9 hg9 F7A 3 b ve on3(KhmRe RactctdngmRgm9o e gRcg99l ctt Slc9 nle 9t S ge gcNz9og9 3Wtt Slchm Re Rh Snc gc FWAWb g ge9 ng ve on3(KA



ah. C g 3 AmN Dn l mA m l N e DmA o A l C C e yDA h 3 N l f Dl mA h ya Dy mG2 ylmDy N e Dy hmN AmN e l Alyh 3 N l f Dl mA h ya Dy m 2 ylmDy N e Dy hmN h AmN e l Alyh

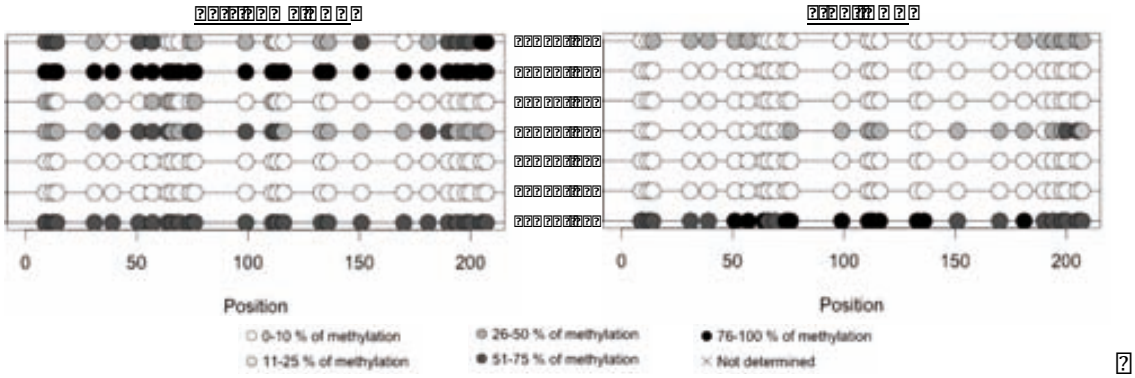


Figure 1

Abstract

lgR9oRgRcmcolgcmcoSS9ngRSSle((t gRal ge9 mcm e 9cge t nu n 9n e 9t e ge9 megR9gRnchongRnmcgo e c t ocg S n9nt g9mollamccccgRme 9nt gef ccc9gRecmg nt e ge9 A

Introduction

9 ce ne gRgSnet na got 9nc n9t l9cla n l g gaccoc g g9cRn 9t t 9 t 9l ol n ce gonc v o Rnl S ge d R l nm), ()Km S9cgol g gR g gR t gRal ge9 S gg n c 9o e 9l9 n 9olcet el nla n Sn9o e 9gRn n gaScAR n 9n m lar gR t gRal ge9 Sn9el c e g9g l9 (z got 9ngaS c f el l gRn9o RgR S9ngA

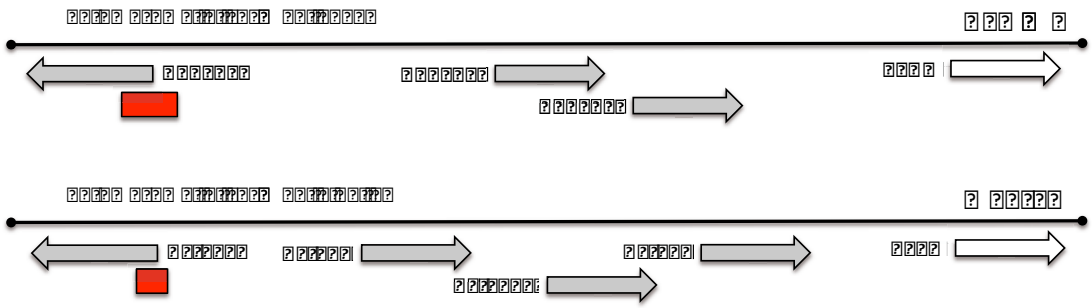
TCGA Cancer type	% AKR1B1 hypermetylation
Prostate adenocarcinoma	90.8
Colon adenocarcinoma	81.9
Rectum adenocarcinoma	81.1
Breast invasive carcinoma	77.2
Liver hepatocellular carcinoma	76.5
Stomach adenocarcinoma	49.3
Lung adenocarcinoma	36.6
Uterine corpus endometrioid carcinoma	26.9
Lung squamous cell carcinoma	8.7
Glioblastoma multiforme	4.5
Kidney renal clear cell carcinoma	4.2
Bladder urothelial carcinoma	3.8
Thyroid carcinoma	3.5
Pancreatic adenocarcinoma	3.3
Brain lower grade glioma	0.7
Kidney renal papillary cell carcinoma	0

DCG Al mye yo fe hN lf DlmA A C Ce yDA AC5 hAl fe y maN A Ah

lh 9g9 la9l9 h log lc9n got cg9t Rgot 9ncSnc g Re R n g c9 RaSnt gRal ge9 9 ((hmRe RS9e gc9ogg Rn l f 9 gRc lg n ge9 e nc9 gR e cgef gn gv l (FKA

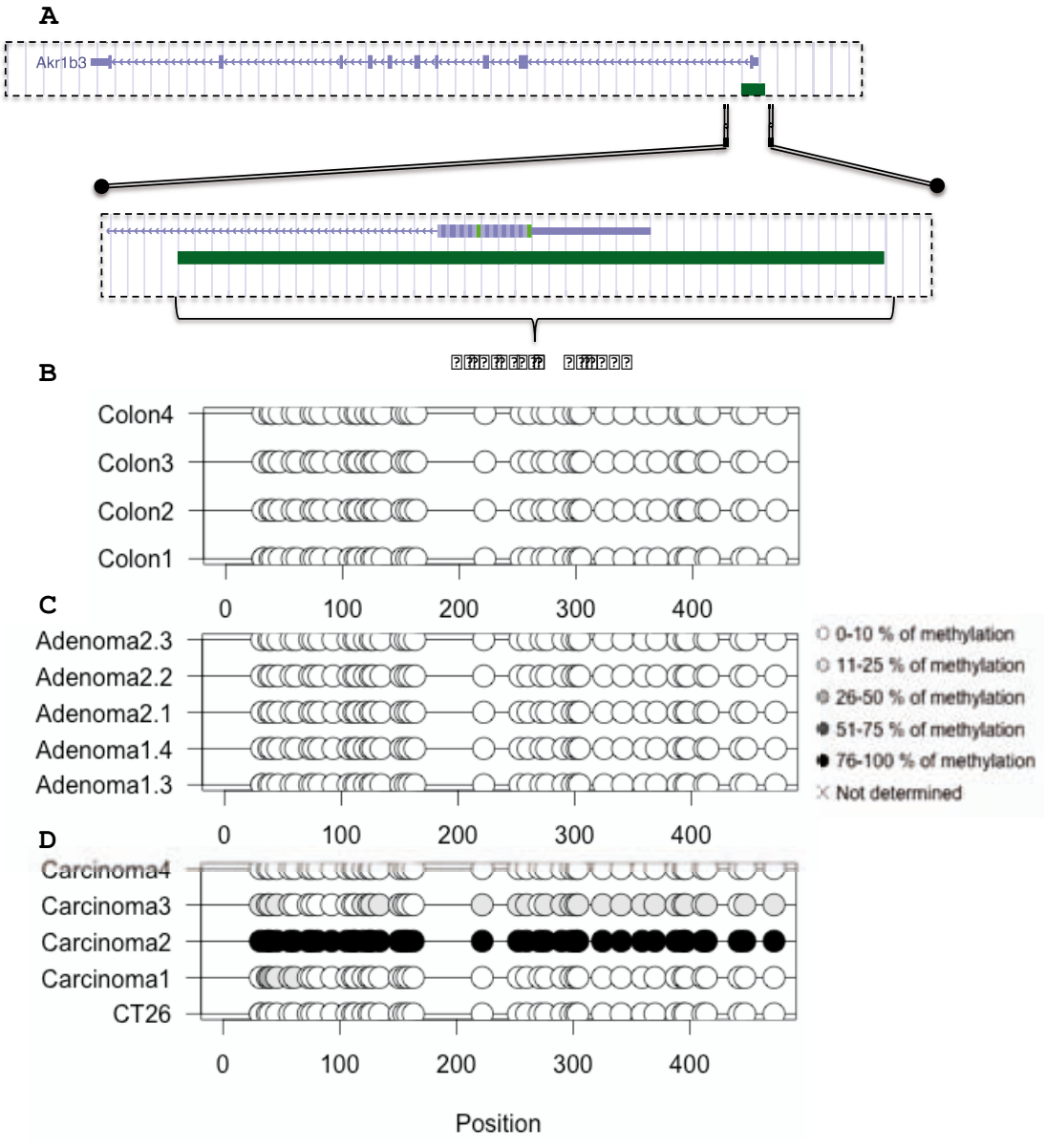
N lf DlmA AN

R ecgne oge9 l9 ge9 9 (t ela megRe gR 9t ec Sngel a9 c nf gm Rot t 9oc A gR c9 Rot ch gR n n n gRn c mRel e t e gR n n 9onA cSeg gRec e n hgR n n n Sl ga 9 tet el nege cN gR 9 oc 9 g e cSocg9 S ecl e gR mR9l h e9 e 9gR cS e ch c n tet el nla9ne g g f t 9n hgR e R9ne c n gR t t SVe on 3FKA



ah G AmN Dm lmA m C A yA aN A N may h e yDA y h h eh y Al y hyt a hyh

oc9gR9 c nf ge9 9 gRec e9 gRn9o R f 9loge9 h m m9 n a gR RaS nt gRal ge9 SR 9t 9 m9ol lc9 9 on e t 9oc A 9 f lo g eg m e g9 lar gR t gRal ge9 9 gR S ecl cc9 e g meg R v ((t 9oc 9ngR9l9 K e 9nt l 9l9 h 9t n e 9t t e t Sl cA R t Sle9 tgo e a ecol g t i o e ac Se g e e on 33A ecol g Snet nc t i o c n h g e gR p A



ah . g 3 ah A A Dm l mA A l mhh ye mA N eD mAy yla f ya D y t a A Ah b 3 N l f D l mA yl lay A AmhN D nDmA N a my y1 3 2 AmN y N eD y A A 3 h p5 DDA N may mDmA A h DDA 3h

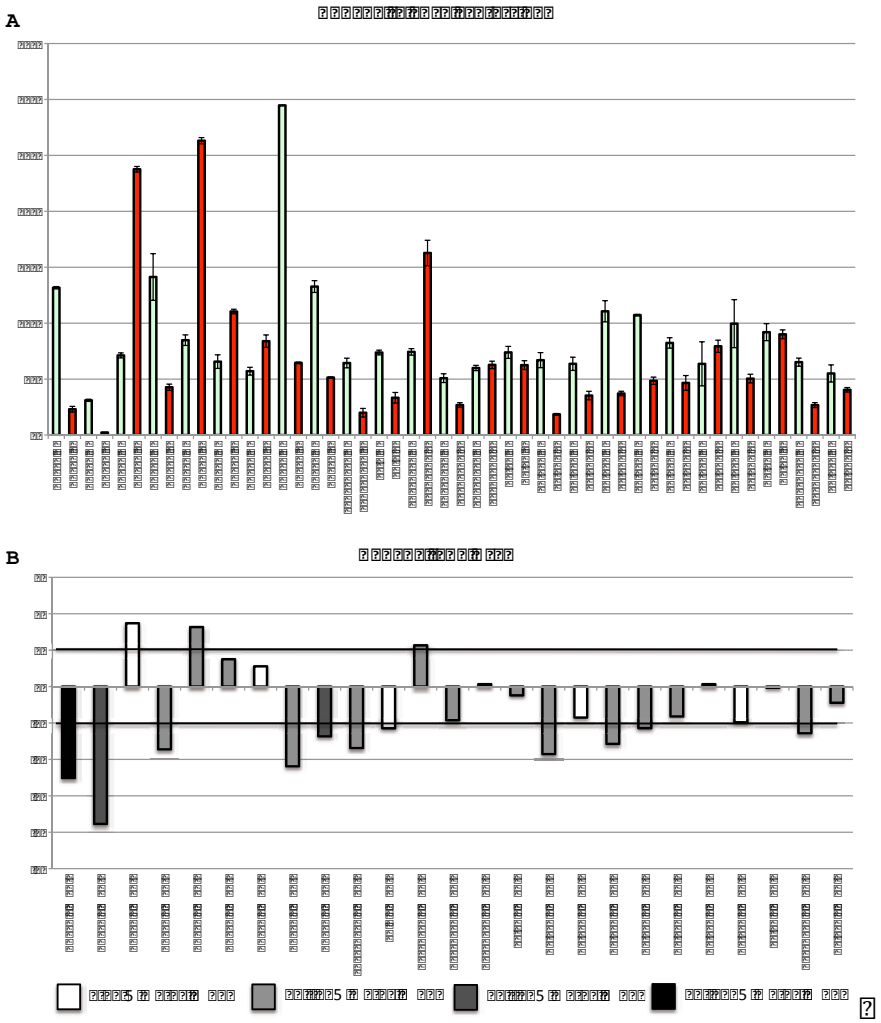
lar on 9nt l 9l9 c t Sl ch f 9t c 9on n e 9t cA t Sl c n n Sn9f e a gR 9n g9na 9 o n gll v k n l9 K

brief description of the characteristics of the samples is presented in Table 14 in the materials and methods chapter.

As occured in humans, the CpG island was not methylated in normal colon samples (Figure 44B). There was also a complete unmethylated pattern in the adenomas studied (Figure 44C), unlike the human counterpart that were frequently hypermethylated (Figure 33).

In the case of the CT26 carcinoma cell line and "carcinoma4" sample, we could not detect any DNA methylation in the region analyzed; whereas in "carcinoma1" and "carcinoma3" we detected partial methylation of a few CpG sites; and finally "carcinoma2" was fully methylated (Figure 44D).

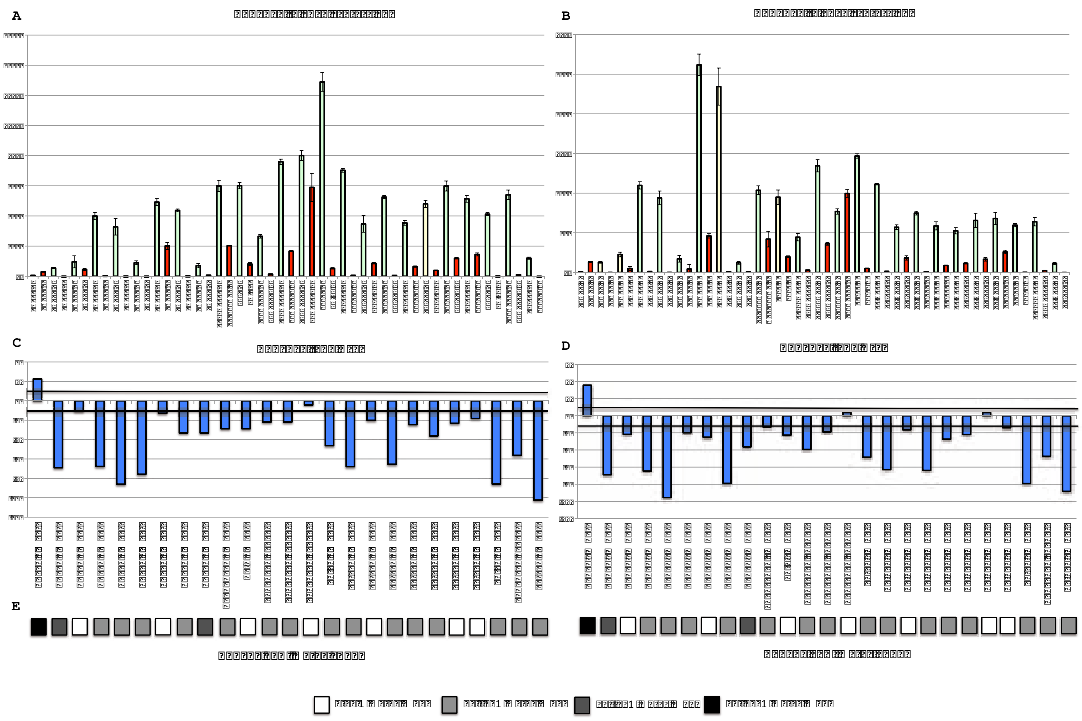
These results confirm that the hypermethylation of AKR1B1 CpG island in colorectal cancer is also observed in its murine ortholog *Akr1b3*, however it seems to occur less frequently than in human, at least in regard to the murine cancer model used.


A
B

□□ue □□gRec□gRn□cR9l□□e g9□□□9o gh□□□(□(□n□c□□9m n□□ol□g□□□e □((□9og□9□□) W□ v33b K□S□ge□ gc□□ □□oSn□□ol□g□□□e □F□9og□9□□) W□v( ) b K□S□ge□ gcA□□c□□□mR9l□h□gR□c□□ n□colgc□□9 □ent □gR□g□gR□□c□t Sl□c□gR□g□□ecSl□a□Re□R□n□□□□□t □gRal□ge9 □f□lo□c□□n□□ □lc9□gR□□9 □c□gR□g□□n□□t 9n□□□9m n□□ol□g□□N□9 □gR□□□9 gn□nal□S99nla□t □gRal□g□□□ c□t Sl□c□□n□□□□n□la□□□□□g□□□9n□□f □ □oSn□□ol□g□□□v□□, (F□S□ge□ gK□v□e□on□□3W□KA□ □R□□n□l□ge f□□□pSn□cce9 □f□lo□c□9□□□□(□(□9n□□□□R□c□t Sl□□n□□□□Se□g□□□e □□e□on□□ 3W□A□□ Re□□□e □□e□on□□3W□□gR□□n□Sn□c□ g□□□f□lo□c□□9mn□cS9 □□g9□gR□□l9□) □ 9nt □□l got 9n□n□ge9 □□□gm□□ □S□en□□□c□t Sl□cA□

N □□□□ne h□yy□mA□□A□D y□y□m□□□□C□Cd□A□□□□□C□C2□A□□mDmh□□l□D□□A□□h□ y□N e D□y□

□ □□9 gn□cg□meg R□□□□(□(h□gR□□□pSn□cce9 □9□□□□□(□(, □□ □□□□□(□(W□□□ □c□m□c□ t 9n□□□□□Sla□□□□□g□□□□one □□gR□□□□ □□n9oc□Sn9□□ccA□5) b □S□n□□□ g□9□□gR□□□□c□c□ v) F□9og□9□□) W□□pRe□eg□□□□9m n□□ol□ge9 □9□□□□□(□(, □□pSn□cce9 h□mR□n□□c□socg□ 3b □v(□9og□9□□) W□9□gR□t □n□n□□oSn□□ol□g□□□v□e□on□□3z □□□ □□□KA□



□□□ah□□ 5g□N □□□□ne h□yy□mA□□□□□C□Cd□A□□□□C□C2□A□□□p2□e □□h□□□y□N e D□y□A□D K□□b□□A□□3□ h□D□l □ □□ne h□yy□mA□h□e □o □□h□□AmhN □D□mDmA□y□N e D□y□h□□mDmh□□□A□D□□l □h□□A□□A□1aN mhy□N e D□y□ □A□h□□b□□□A□□□3□□m□p□r□c□3R□o □□h□□A□□□l□ □□i □□a□y□N □□A□□mo Ah□□a D□l □mA□□A□□e my□l□ □□mA□y□ a e h□□a D□l □mA□m□l □□□ne h□yy□mA□□A□1aN mh□□mN e □h□□□o □□□□A□mhN □D□mDmA□l□yya □b□□yaN N □h□K□y□l□□□ □e□□y□D□A□□□□C□C□□□□N □l□f D□l □mA□yl□l ay□mh□□□□□y□N e D□b□

The scenario is very similar for AKR1B15, which was downregulated in 88% of the samples (22 out of 25) and there is just 1 case out of 25 where the gene was upregulated (Figure 46B and D). The behavior of AKR1B10 and AKR1B15 was almost identical in all samples analyzed, which may indicate a co-regulation of both genes.

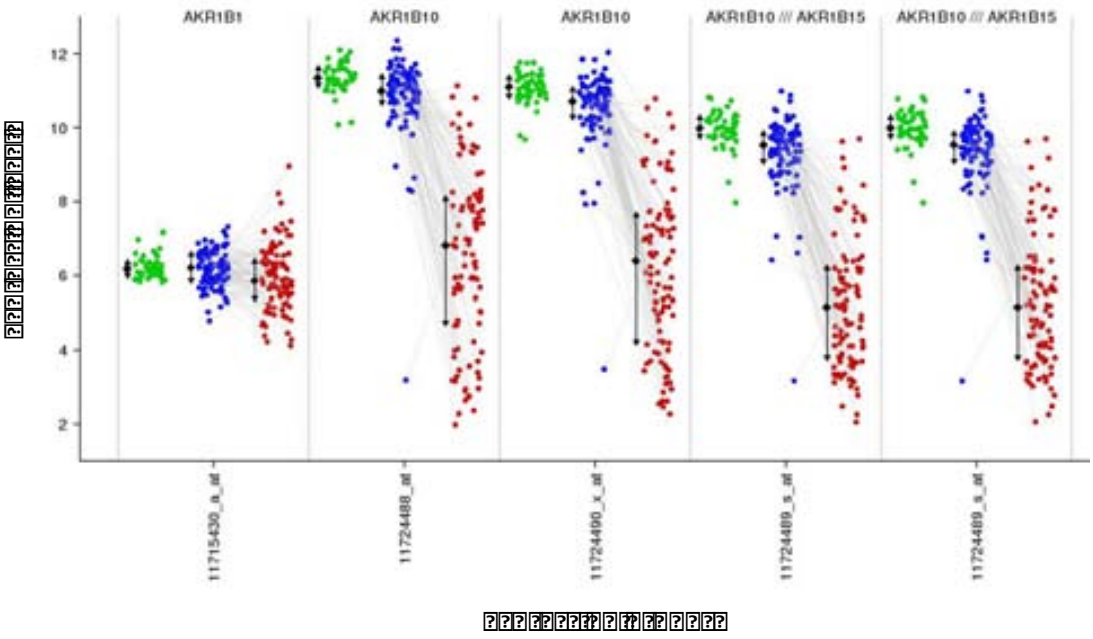
As we have previously reported, promoter DNA methylation of AKR1B10 and AKR1B15 is almost unaltered during cancer development despite both genes display a dramatic downregulation. However, we could observe a link between the DNA methylation of AKR1B1 CpG island and the expression of AKR1B10 and AKR1B15 (Figure 46E): hypermethylation of AKR1B1 CpG island was associated with strong downregulation of both genes (Figure 46C and D). An exception is CR010 patient, in which, despite the low expression of both genes in the tumor tissue, there was an apparent upregulation explained by the lack of expression of AKR1B10 and AKR1B15 in the normal colon.

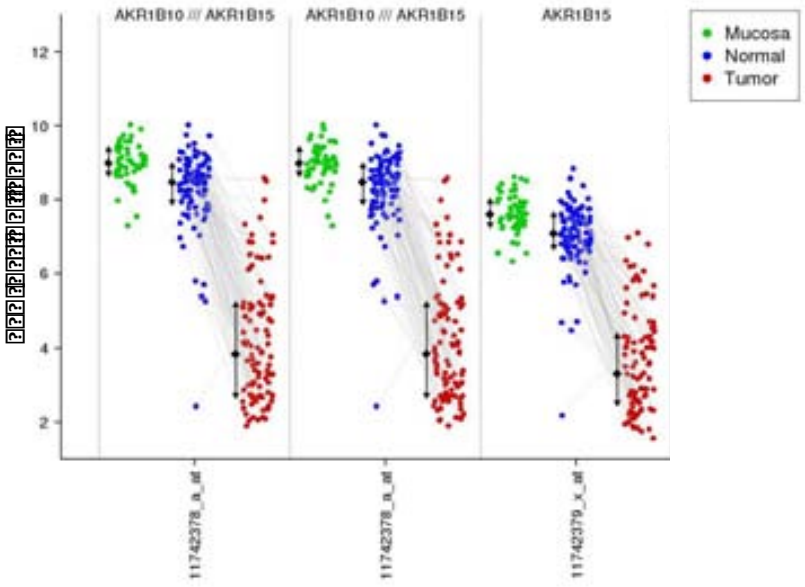
mRNA expression of AKR1B family in a second set of samples

Next, we extended our study by applying microarray expression analysis to a cohort of 100 patients from the Institut Català d'Oncologia (ICO, Barcelona), with the corresponding normal and tumor sample for each patient. Additionally, to discard possible sample cross-contamination and in order to identify possible disease-related features in the normal tissues of patients, this analysis included the colonic tissues from a control cohort comprising 100 healthy individuals. We then analyzed the expression of AKR1B1, AKR1B10 and AKR1B15 in the expression microarrays.

Although AKR1B1 displayed a modest tendency to be downregulated in colorectal cancer, this was not a general fact and many tumors did not change the expression levels and in some of them were even upregulated (Figure 47).

On the contrary, AKR1B10 and AKR1B15 showed high expression values in normal colon that suffered a dramatic downregulation in almost all tumor samples analyzed (Figure 47).



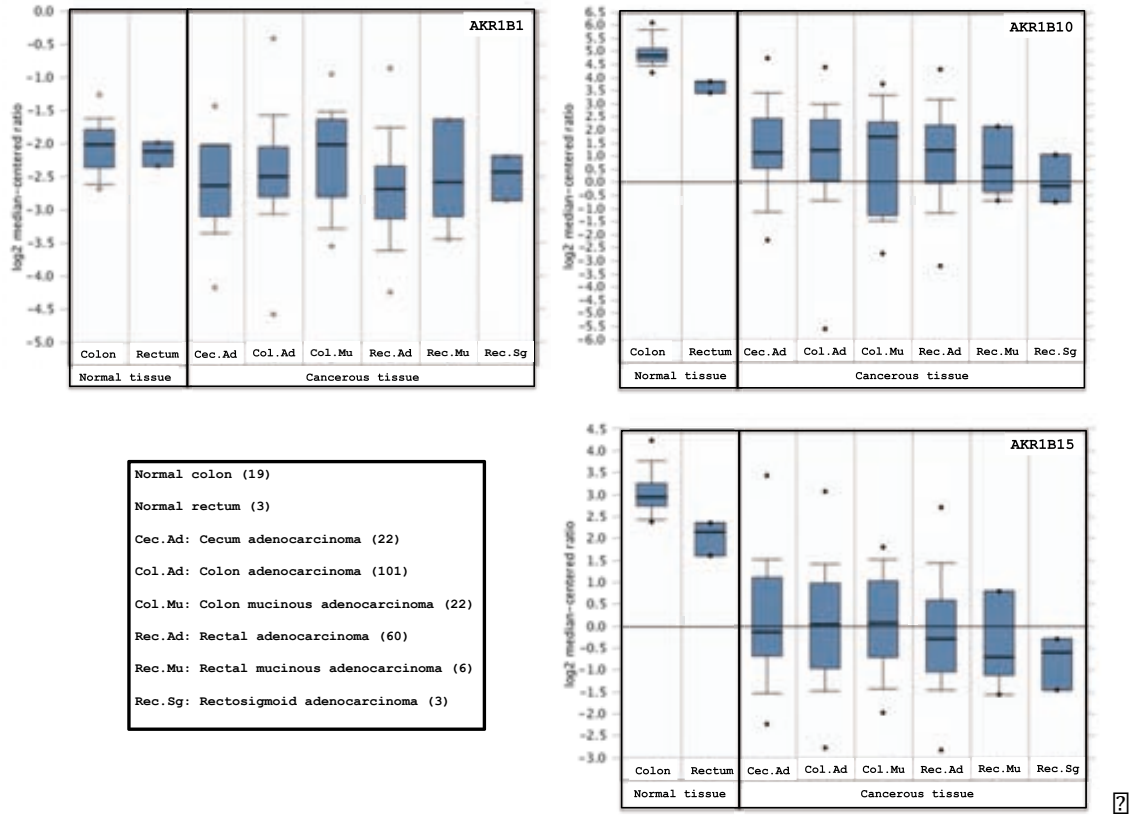


ah 0g C CR C Cd A C C2 A ne h yy mA A G h Al y ly m y N e D yb A h AB AnhN D f N a ny hnN D f mAnhyR a B AnhN D mDmA N a ny yym l mf laN nh A Ah B mDmA laN nh y N e D b

g ec m9ngR g9 9g gR g e ll c t Sl c n9t R lgRa 9 9nch ((, W pRe eg Re R f lc9 t pSn cce9 mR n c e 9nt l t g R t o 9c hc9t c t Sl c ln a ecSl a 19m pSn cce9 f lc e gR n 9 got 9n gecco cA

N ne h yy mA m C N D A y N e D y

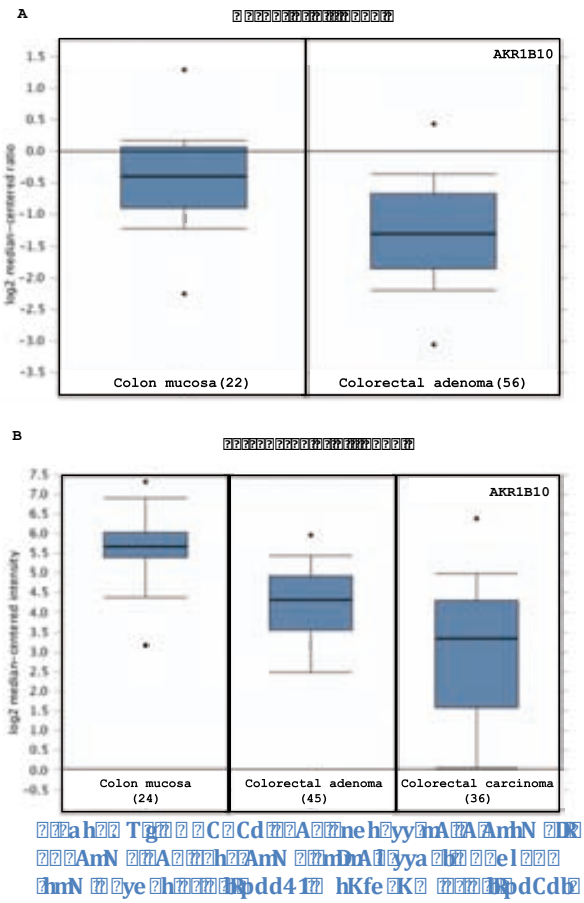
9 tgn gR 9on 9 loce9 ch m e g9 lar gR pSn cce9 9 ((h ((, ((W e gRen c g9 t Sl c 9nn cS9 e g9 gR ln a t ge9 Sn9s gv e on 32KA



ah 4g ne h yy mA m C A N D A h Al Amh N D A mDmh l D laN mh l yya y r mN l y3 aN hy Ae h Al y y A l 1 Aa N h m y N e D y e h mma eh

c Sn fe9ocla9 c nf h ((m c 9m n ol g e ct ll co c g9 n e 9t c mR n c ((, ((W m n Sla 9m n ol g e t 9cg 9l9n g1 n ct Sl c v e on 32KA R n n n e n c e pSn cce9 e 9nt l3gacco h n n e ((cA 9n p t Sl ((, ((W n cle Rgla t 9n pSn cc e gR 1 n e g cge n gR n gR e gR n got A gR ct t le h h got 3got 9nc R f 9gR c e gla l cc pSn cc gR 9l9 got 9nc v e on 32KA lgR9o R gR n n n c9t e n c e ((pSn cce9 got 9n 9 ge9 h g a c t S9ng g3g9 g u e g9 9o g gR ecS nega e gR ot n 9 c t Sl c 9n R n9oSA

((2

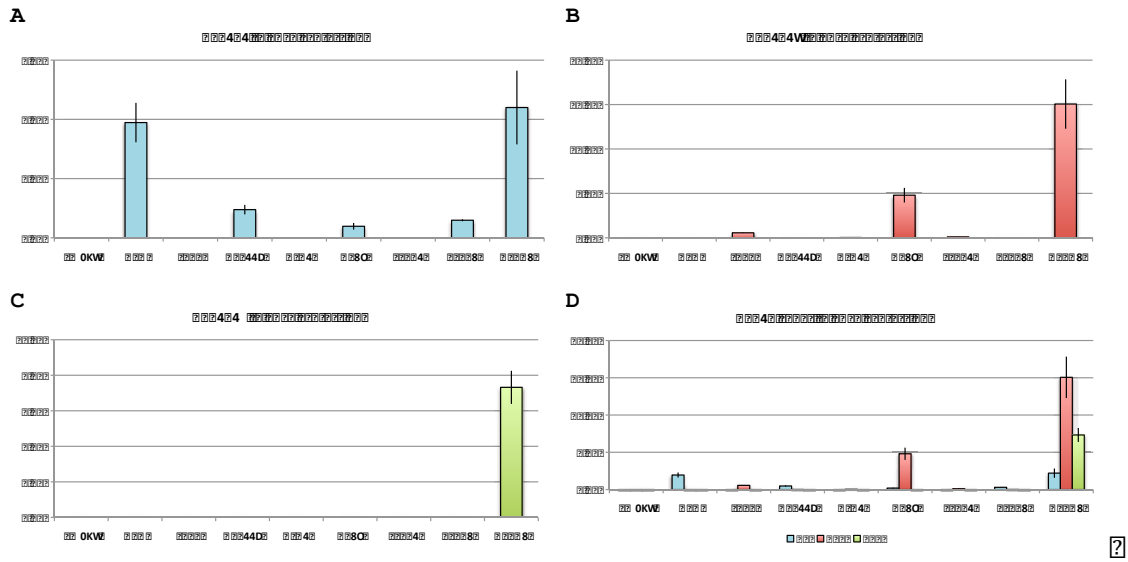


ce gR pSncce9 g9 gm9e S g cgoecvcSn lm e gel 9m nolge9 9 ((, e 9l9ngl 9t cv eon 35K mRe R e eg c gR nla e e 9 gRecglgnge9 S9e gc 9og gR et S9ng 9 ((, e gR got 9n SSn f l9St gA

N ne h yy mA m C N A y A h DDA y

R pgcg S mcg9 lar gR t pSncce9 9 (ce S l9 e Rg9l9ngl n le c 9 RS g9llolnne 9t ll le v S) KA

et elnlag9 mRg mlnaa9cnf e nchl9l9ngl n lle c cR9m R gn9 9oc pSncce9 lf lc nne v eon W KA 9mf ngR c ne9R cmR mg lu9og ((, v eon W Kh mRe Rect 9n g la pSncc e 5h Re Rla e S) S99nla9n 9g pSncc e gR9gRn lle cA ((W ecf ma l g t Sle a i e 9l9ngl lle cv eon W KD go llakm so cg g g gR pSncce9 9 ((We S) RS g9n e 9t ll le A S) pSncc Re R f lc9 ll ((c v eon W K m cg u c S9cegef 9 gn9l9n ((, pSncce9 e bngR n pSnet gcA



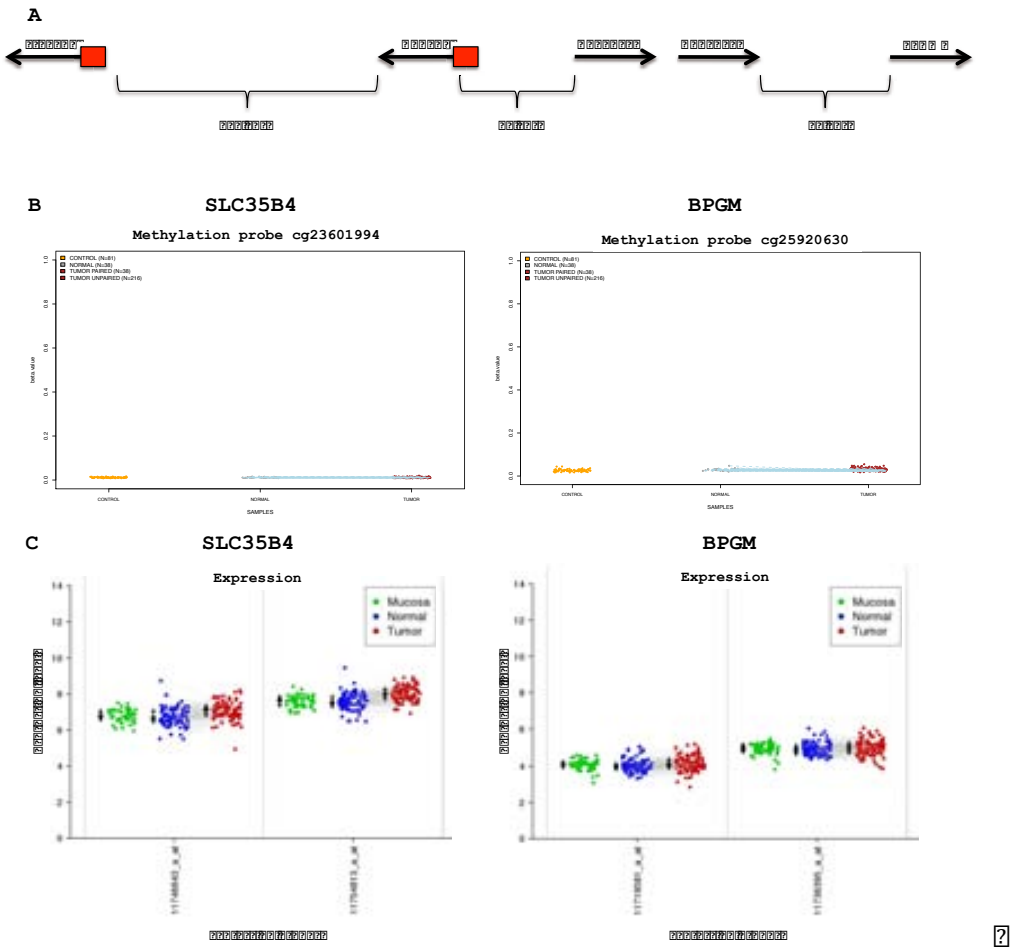
ah2dgN nehyymAmC Cr3R C Cdr3A C C2r3 Ay A e A Dm aN AmDmh D A hDDA yA A e plh

neh yymA m A yA h C Dm ay

m g g9 lo e g mR gR n gR 9m n ol ge9 9(c ac19 oc cS ea e9nh9 gR 9 gn nal eg ec 9 c i o 9 l9 n Se ge cel e e ge i olgeSl c e gR h e9 h mRe R ac 9t t 9 S R 9t 9 e 9l9n g1 n v ne9l m), , zD a9n m), , 5KA9 n cc gRec i o cge9 m lar gR pSn cce9 t gRal ge9 9 FW3 h mRe R n e R9n ch19 g 9og( ) W z W u ch n cS gef 1alh n9t gR n cg (3 e on W(KA

c n 1 g e e on W(hgR pSn cce9 9 FW3 m c 9g lg n e 9l9n g1 nA FW3 pSn cce9 ecSl a e g oSn ol ge9 e 9l9n g1 n mRel n t e o 1lg n A t gRal ge9 m c 9t Sl g1la c g e 9gR Sn9t 9g nc egR n e 9nt 1 n9oc gccoW v e on W(KA

9 loce9 hgR S 9m n ol ge9 9((, (W ac 9 e g9 gR (c19 och oc gR n ac 9g 9m n ol ge9 9 gR e R9ne c 9n RaS nt gRal ge9 A Rec t Sle c gR g 9m n ol ge9 9(c ecS ea e 9g 9 c i o 9 l9n Se ge cel e e A



ah2C g3 AmN Dm lmA A y lA y m C Dm2a y A l A mh A y A G2. b yt a h y A l e yD A yym l lm2e hmN ml hyb 3 mDmh l D N l f Dl lmA l m G2. A A Ayb3 neh yy mA l lm ml A y A D f mAmhy rh A3 mDmh l Dla N mh h3 A 1 uAmhN DqN a my yym l r Da 3b

m2neh yy mA m C A y

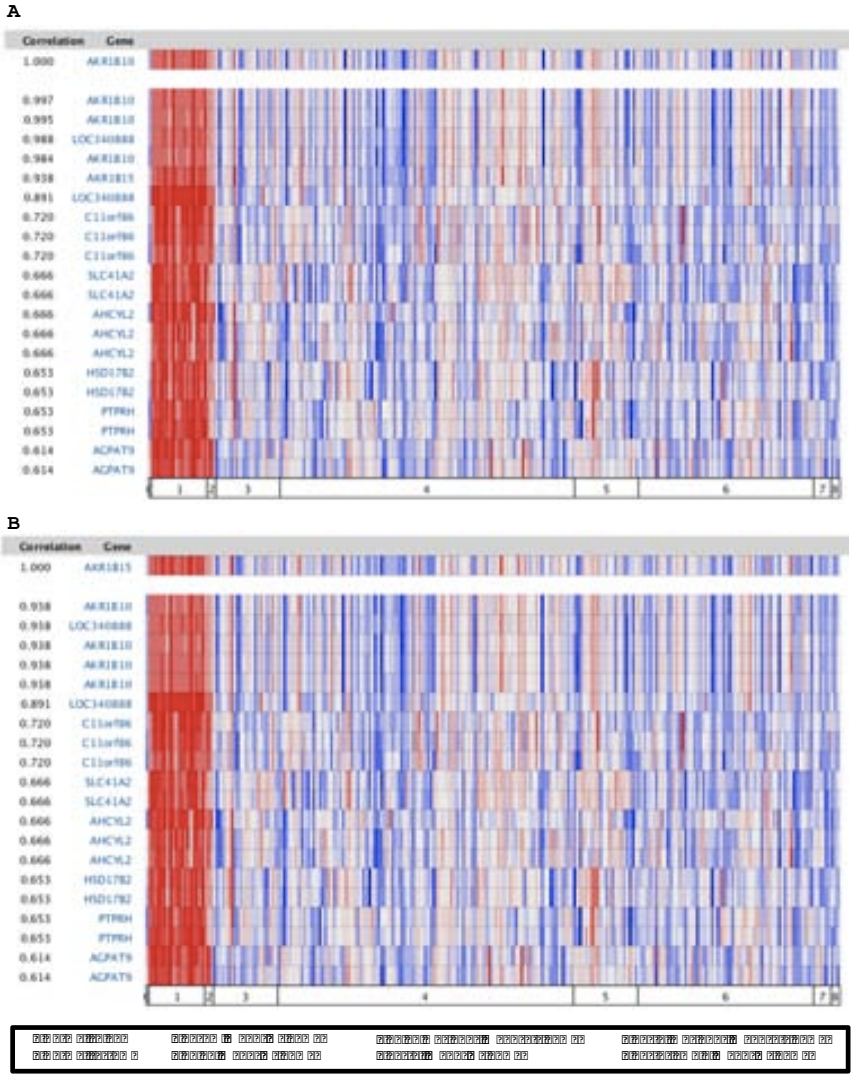
9m g2c l n gR gg R m nn g 9m n ol ge9 9 onne e 9l9n g l n2ac t e la ge ((, (W2 og2m m9 n e gRec cel e g ct olg 9ocla 9gR ch mR g2m9ol e e g 9t t 9 n ol g9ma t R ect 9n 9gR cA 9 n cc gR g2 lo cge9 m oc gm9e S g 9l9n g2l c gc9 c t Sl ch gR pSn cce9 g 9 g e m9t gR 9t e g c S9ng2l vc t g ne2lc t gR9 cK gR pSn cce9 g 9 gR [, , got 9nc lar gg R A c Se g e e on WJ h ((, (W ecSl a f ma Re R 9nn l ge9 h lt 9cg e ge l f lo ch e 9gR v, A5F2K C v, A5FzK g v e on WJ KA R c n colgc co cg 9y n ol ge9 9 ((, (W g c m9ngR g9 9g gR g ((m c 9g 9nn l g megR ((, egR n ((W mRe R9 ant c e S g

??colgc?

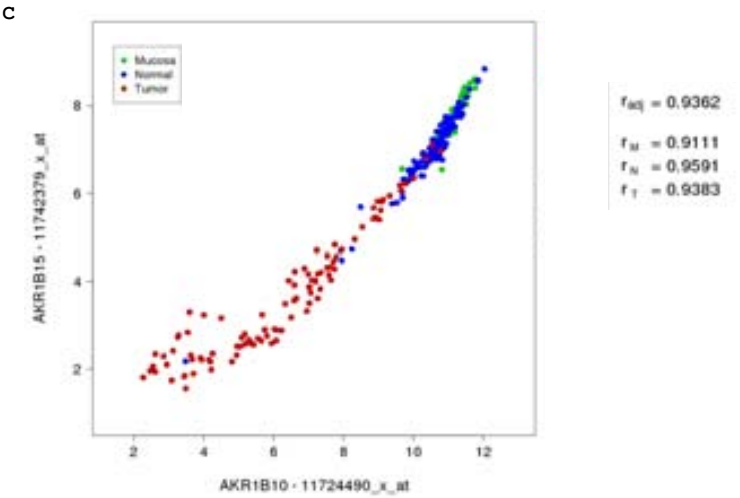
?

n??ol?ge9 ?9????(?(?f?ncoc????(?(, ?? ????(?(W?e ? 9nt ?l?? ??got 9n?

?9l9n??g?l?gacco?A?



C



?

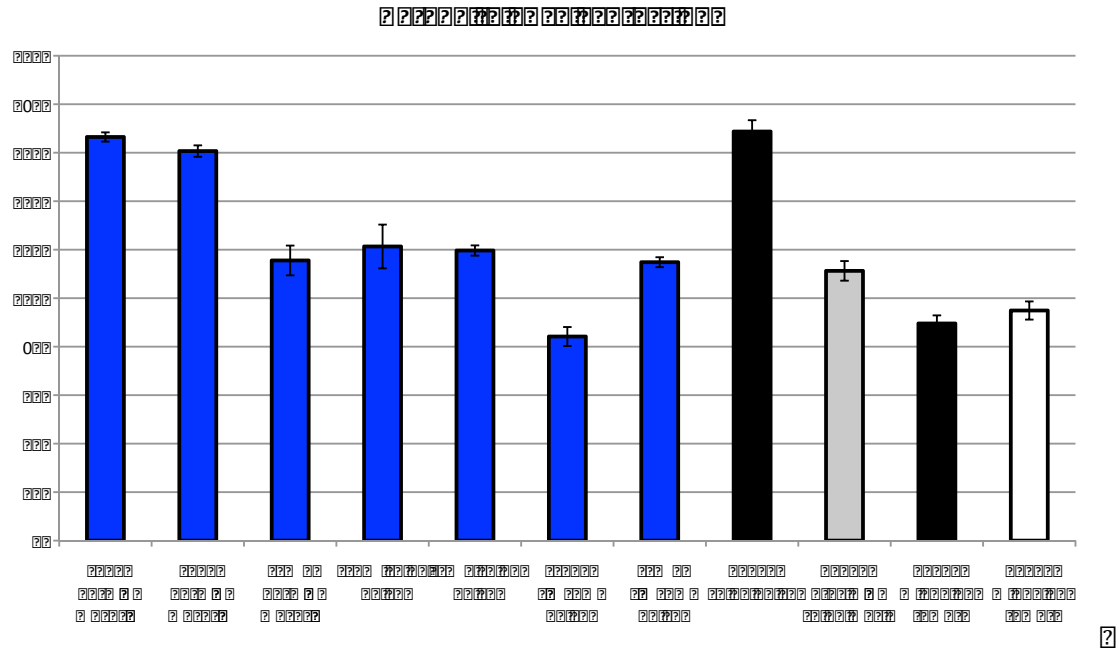
??ah??2p?g??mhh?D?l?mA?m???neh?yy?mA?Di ?Dy???lo ??A????C?Cd?r?3??A?????C?C2?r?3?o ?l????A?y?
?A?a?????A?????????mDmh??l?D?neh?yy?mA???l?br?3??msh??aD?l?mA?m????C?Cd?A????C?C2?ay?A??
????l?y?lyR?A??mDmA?N a?my??hmN ????D?f??mAmhyR?A?AmhN ?DN a?my??hmN ??mDmh??l?De?l??Aly??A??
?A????N ?l?????laN mhl?

( ))?

neh yy mA m C N Df A N may mDmA

c c ne Sn fe9oclahgR cgno gon 9 (l9 oc ac S nge lla 9 c nf gm Rot t 9oc A ((n 9ngR9l9 c 9gR R f S acl mR n c gR n ac 9t 9oc 9ngR9l9 u 9m 9n Rot ((, A lar gR pSn cce9 9 gR e n gt 9oc 9nt l 9l9 c t Sl chS9laSc n e 9t A

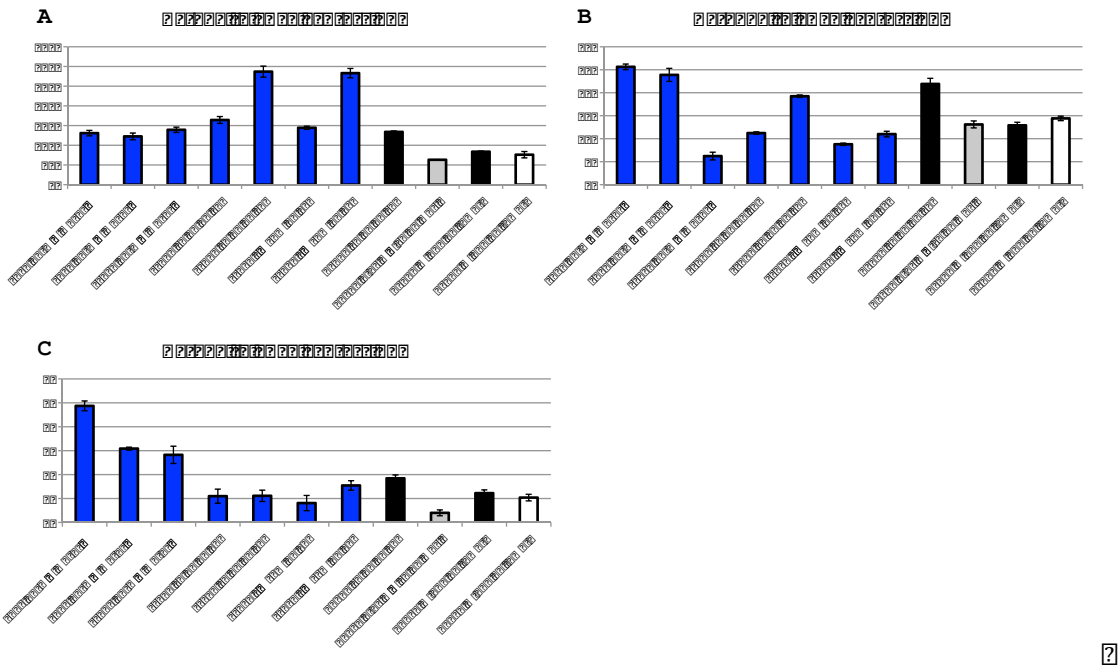
gR c 9 gR n m c lt 9cg 9 R gm gR pSn cce9 9 gR e n gt Sl che S gla9 gR t gRal ge9 cg gocv e on WFKtct el nla g9 mR g9 onn e gR Rot ((mR n t gRal ge9 9 gR Sn9t 9g n n e9 e 9gt cel e A



ah 2G g h Dl i neh yy mA A h Al N may y N e Dyb h mDmh A l y l N l f Dl mA yl lay m l e yDA b D 6 aDf N l f Dl R h f 6 e hl Df N l f Dl Ro f 6 aAN l f Dl A Da 6 aA hN A b

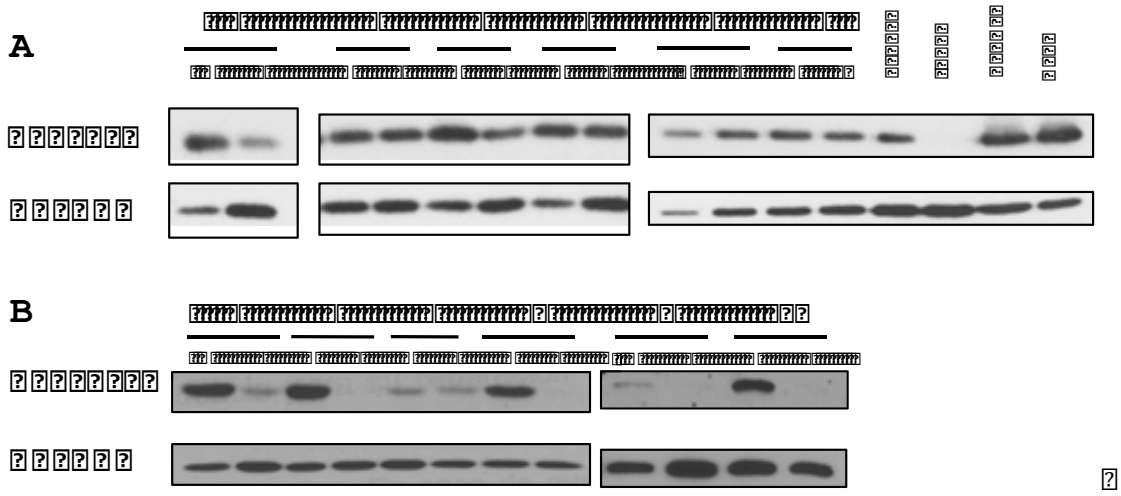
ec Re Rla pSn cc e 9nt l og gaccoch R9m f n eg ec cle Rgla 9m n ol g e n e 9t ch e 9 gn cg meg R S9laSc mR n gR n ec 9 cel e f f e n c pSn cce9 e c t Sl c F5) F5F v e on W3 KA

n e h cS e lla 9n v e on W3 KA R t gRal ge9 cg goc 9 S acl 9 c 9g SS n g9 gg R pSn cce9 9 h pRe eg t 9 cg g a9 9m n ol ge9 e nh cS e lla 9n v e on W3 KA R t gRal ge9 cg goc 9 S acl 9 c 9g SS n g9 gg R pSn cce9 9 h A

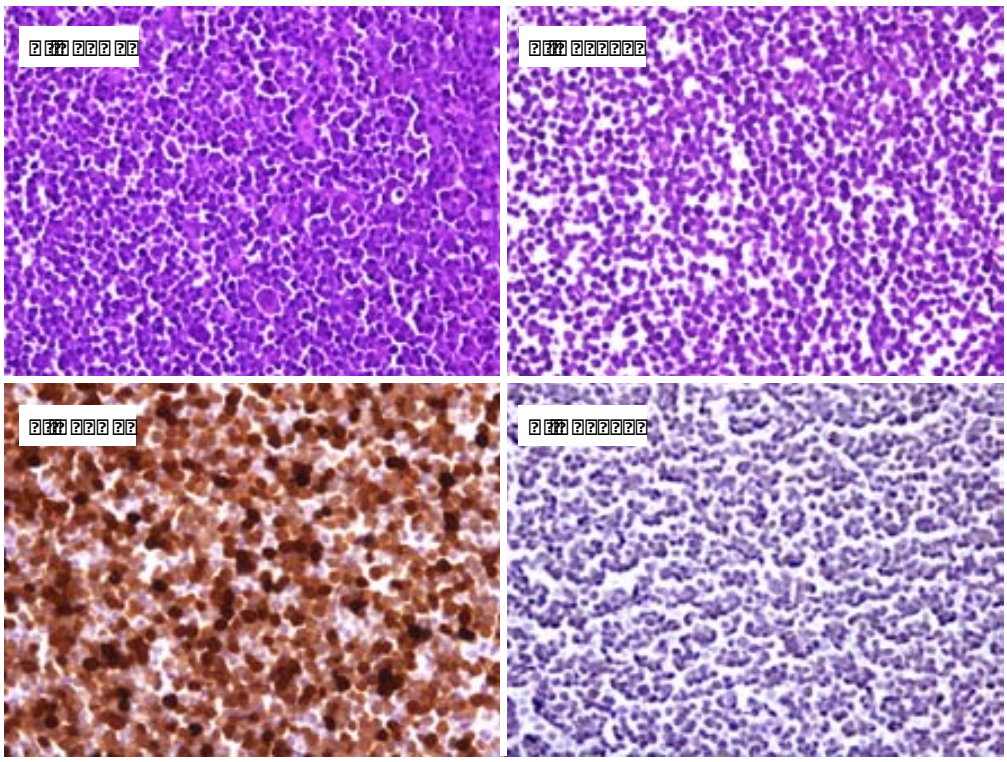


ah2. gA hDl i neh yymA A hAlN may yN eDyh h nDmh A lyl aym l e yDA h D6 aDf N l f Dl mA yl lay m l e yDA h D6 aDf N l f Dl Rh f6e hl Df N l f Dl Rb l6aAN l f Dl A Da6aA l hN A h

9 loce9 ht 9oc (c ecSl a e n g n c9 9m n ol ge9 e 9l9 nht cteg9 onc e Rot c A 9m f nhgR t gRal ge9 tg goc9 S acl 9c 9gtR9m l n cc9e ge9 megR gR pSn cce9 9 gR t 9oc A

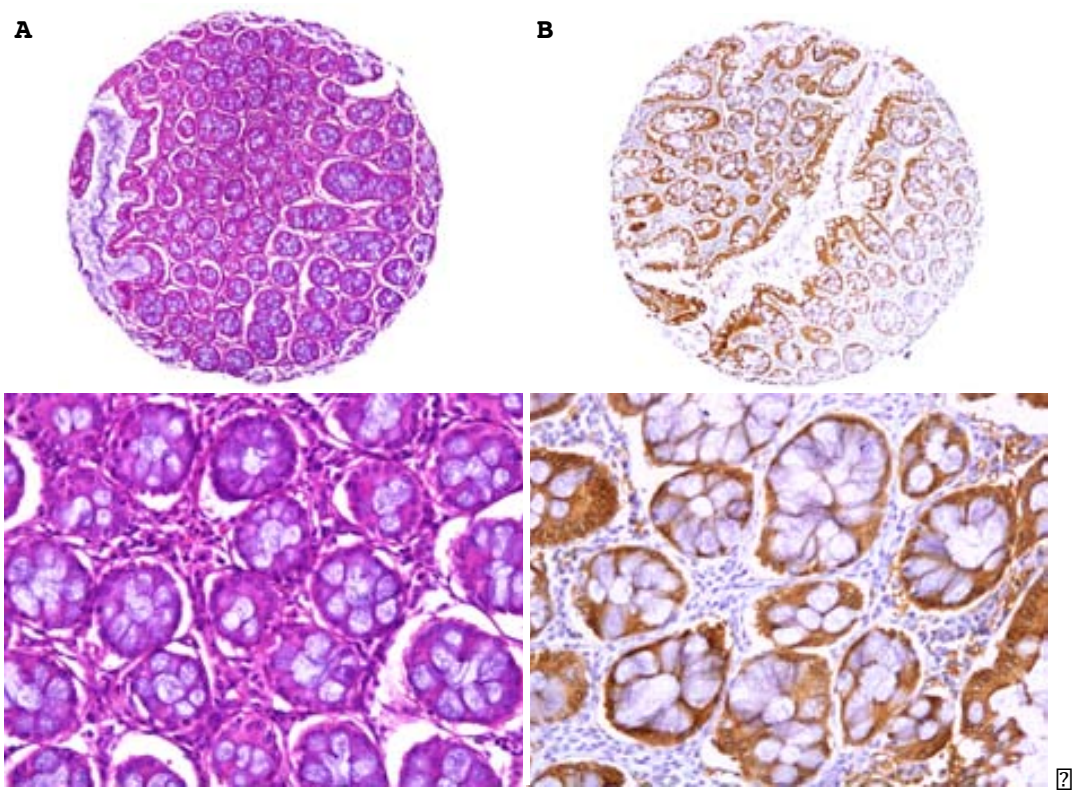

A
B

9ce cg e e e 9n n g9 R u gR i o lega 9 gR c t Sl vc t g ne lc t gR9 cKA

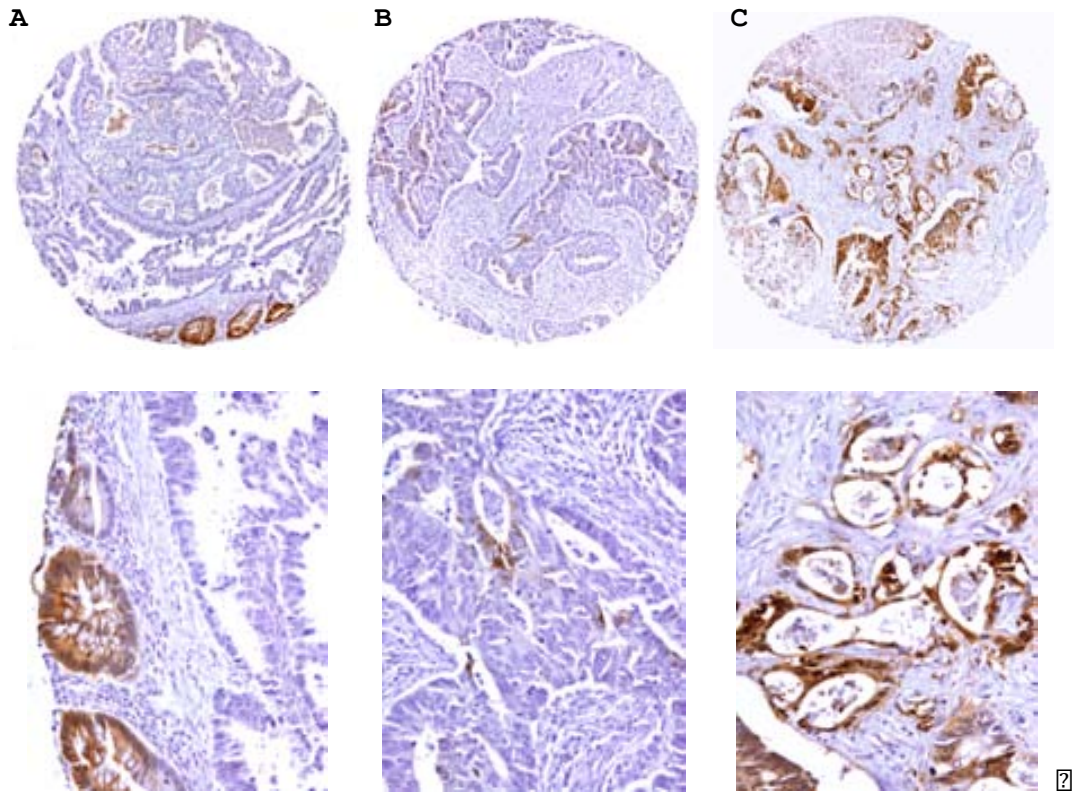


ah 25g A 3 N lmf DA A myA yl A A r 3m e p A CC5R ye l i Dh A 3 myh e hy Al 1 N N aAm yl m N yl hf o C Cd Al m f h

e on W7 m 9 c nf 9nt l 9l9 t o 9c a megR gR ((, ge 9 aA c Sn fe 9ocla 9 c nf a 9gR n t gR9 ch gR 9nt l 9l9 t o 9c ecSl a Re R R9t 9 9oc pSn cce9 9 ((, ce gRec g R 9l9 alm 9 c nf gR g l 9nt l 9l9 t o 9c 9 gn9lc m n S9cegef 9n ((, pSn cce9 A gR 9 gn nahz) AWb 9 gR got 9nc m n 9t Sl g la gef v e on W2 KHFFA5b R e oc pSn cce9 v e on W2 K FAzb Sn c g S9cegef pSn cce9 9 ((, v e on W2 KA gR ne llgR c g m 9 ant gR g5zA3b 9 gR 9l9 got 9nc lar R f cel 9n tcgn9 la 9m n ol g gR pSn cce9 9 ((, AR Re R ot n9 ((, gef got 9nc Sn lo c a cg gecge l lace c g9 e f cge g 9nn l ge9 c megR n l f g le e g A



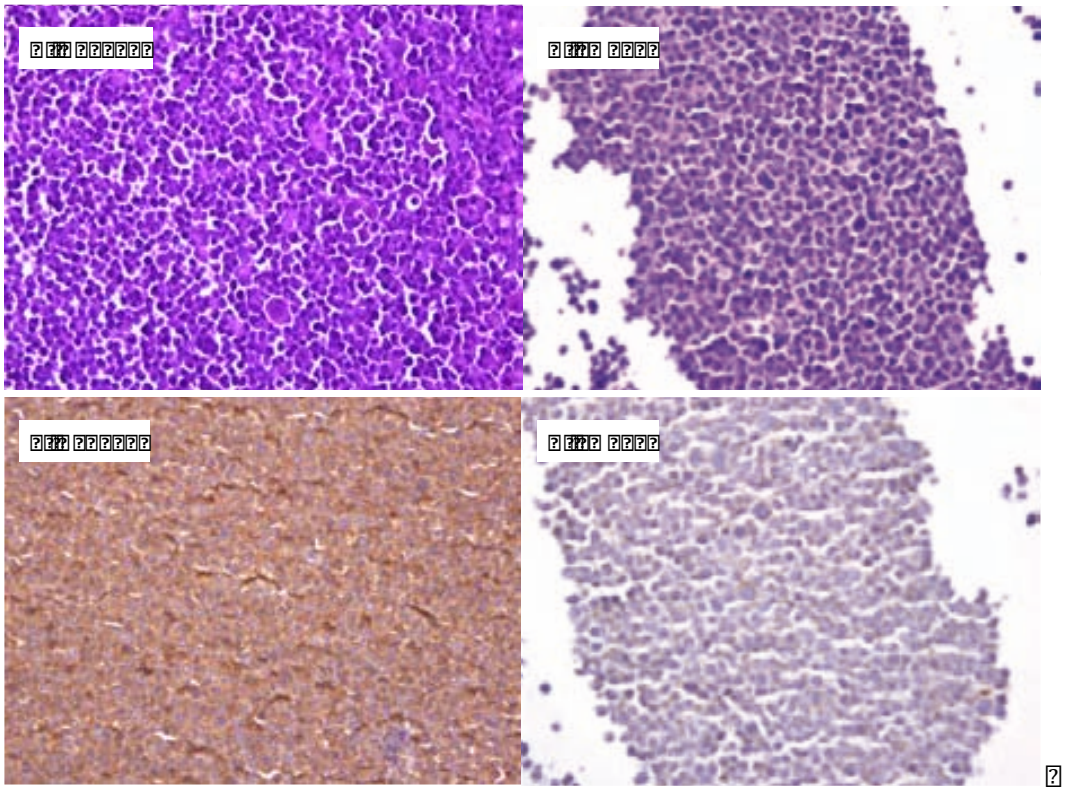
ah 20 g mhN DmDmA N a my yl A A A 3 f o l Al m f Al C Cdb



ah 24 g n N e Dy m l h h mDho l m ymhl h Al mDmA laN my mh C Cd yl A A 3 A l i y N e Do l ymN AmhN DmDmA yya A ml lmN m h N Al 3 ay yl A A 3e my l i laN mhb

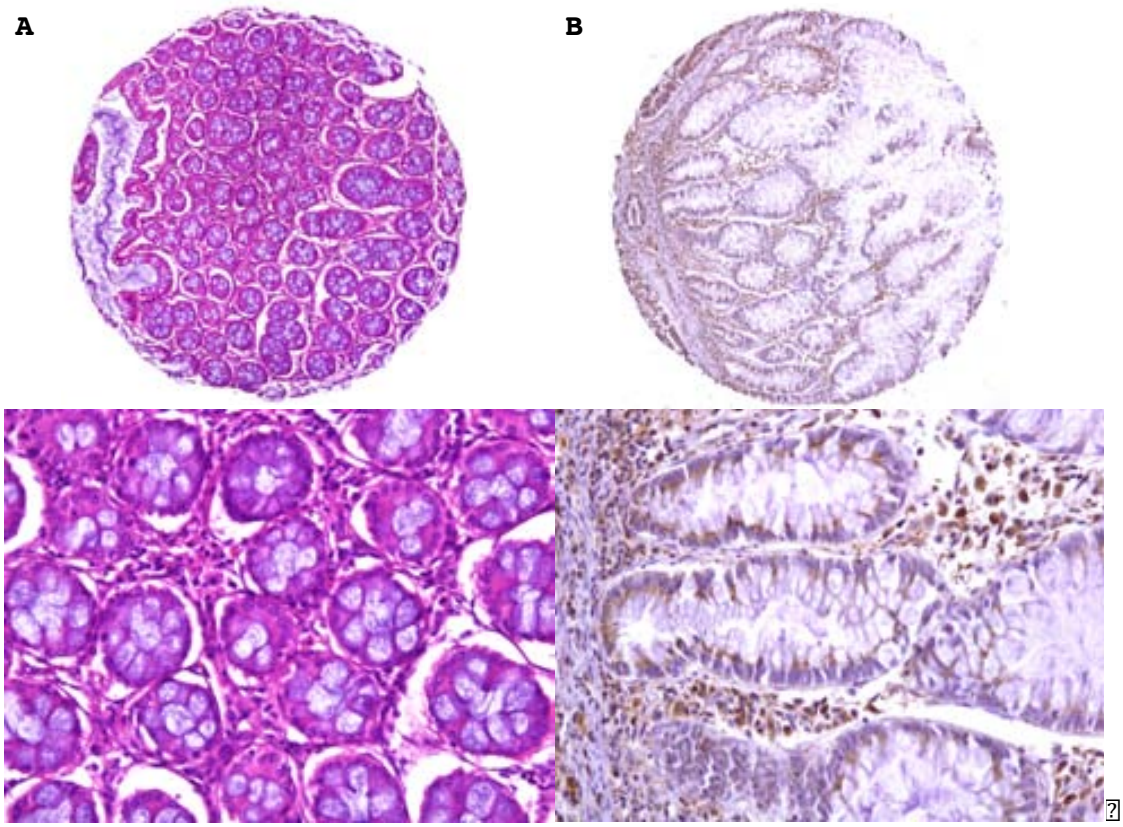
C C A mDnh l D A h1yya N hmhhfy

ce gR c t 9ll ge9 9 ch m lc9 R u gR Sn9g e pSn cce9 9 ((A gRec c m oc ((z c S9cegef 9 gn9l 9n pSn cce9 m32, c gef v e on W5KA

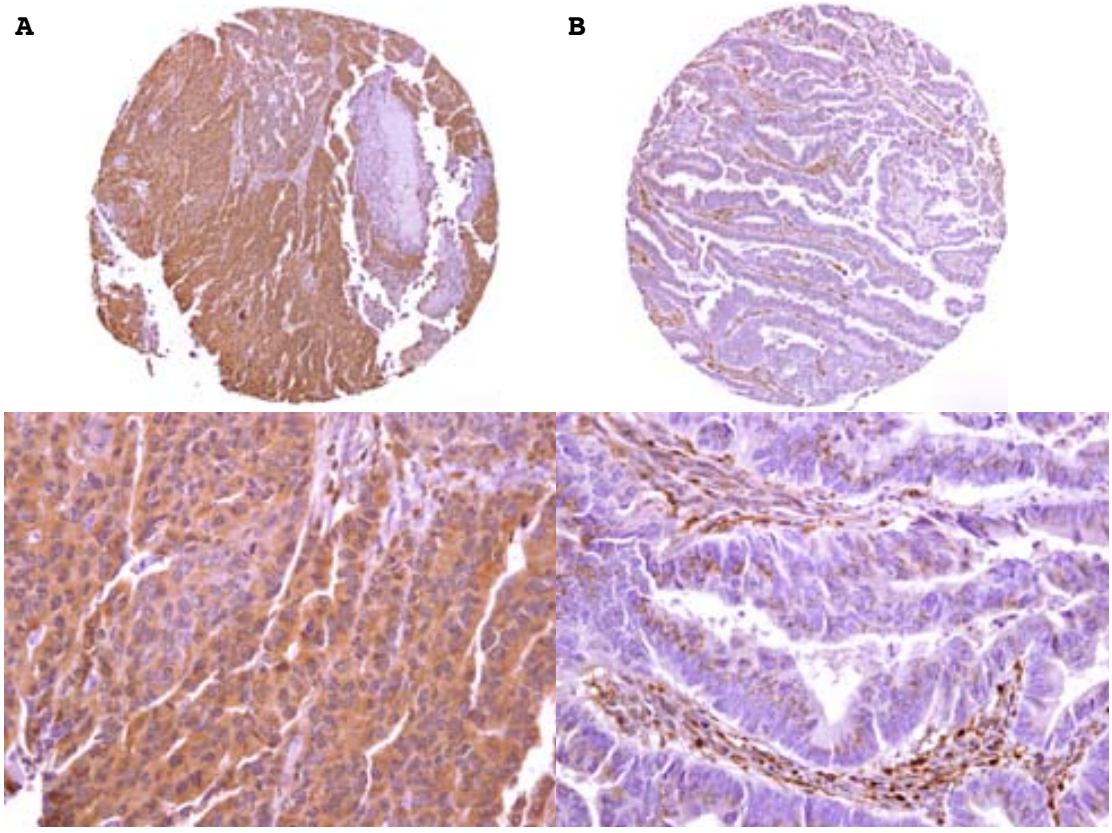


ah 2T g 3 A 3 yl A A mh CC5 A o . 4dRh ye l i Dfb 3 A 3 y l N N aAm yl m N ylhf o C C Al m fb

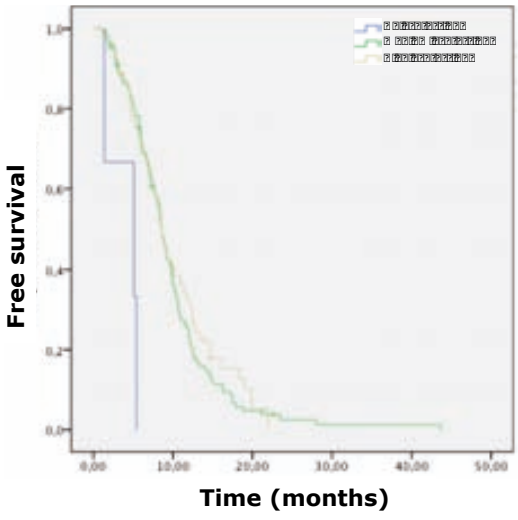
pgh m R u 9n gR pSn cce9 9 ((e gR 9nt l 9l9 v e on z, KA c pS g h ((n c pSn cc e gR ag9Sl ct 9 9nt l 9l9 llcA n e gR n colgch m 9 c nf gR g ll 9nt l 9l9 t o9c 9 gn9lc m n S9cegef 9n ((pSn cce9 A gR c 9 gR got 9nch52b n g e gR pSn cce9 9 ((v e on z (KA socg) b 9 gR t 9cg gR pSn cce9 9 ((v e on z (KA

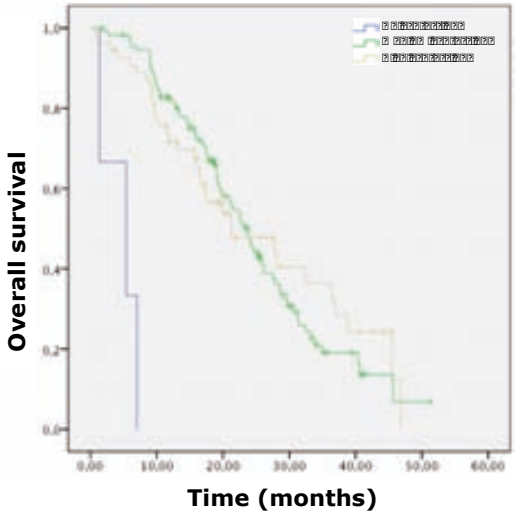


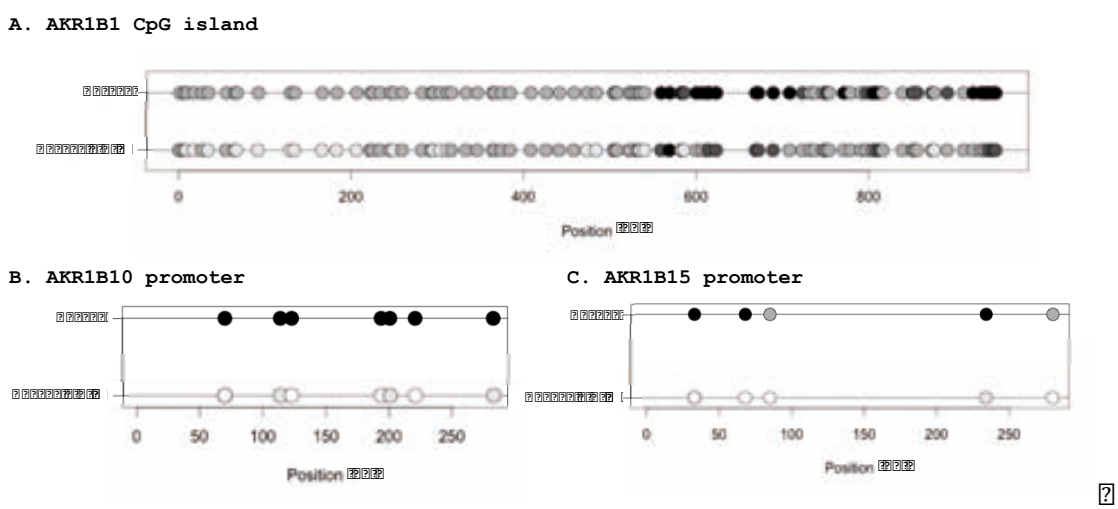
ah5dg3mhN DmDmAN amy yl AA3f o l Al mf Al C Cb



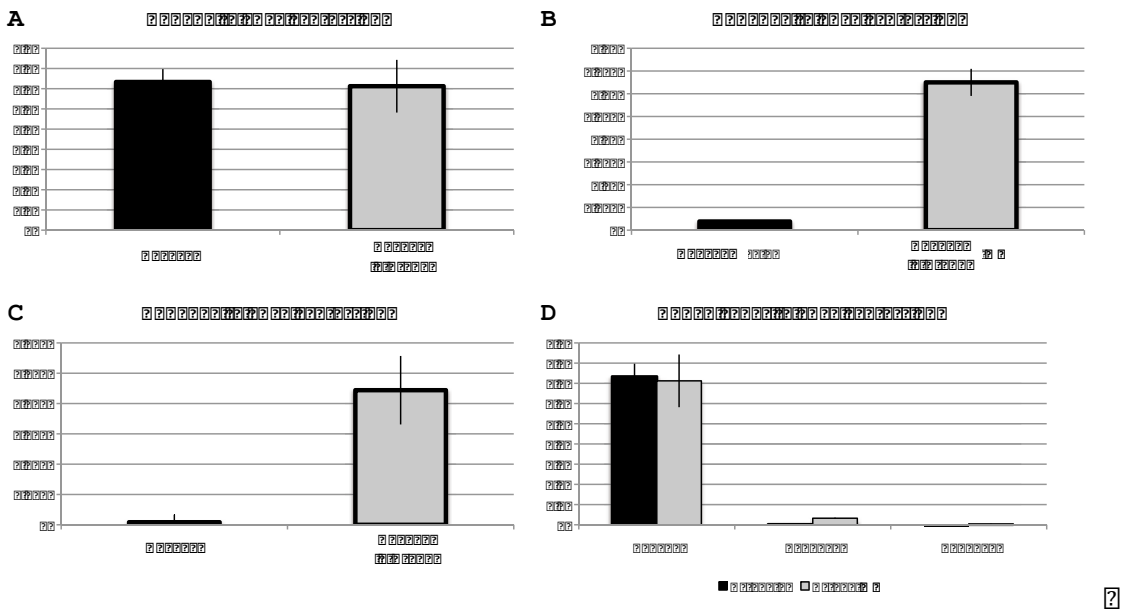
ah5Cgmn eDyhhhmDho lmymhl hAlmDmAlaN mhymhh C Cyl AAb
3emyl l laN mhb 3A l laN mhb


Free survival
Time (months)


Overall survival
Time (months)


A. AKR1B1 CpG island
0
200
400
600
800
Position
B. AKR1B10 promoter
0
50
100
150
200
250
Position
C. AKR1B15 promoter
0
50
100
150
200
250
Position

colgc



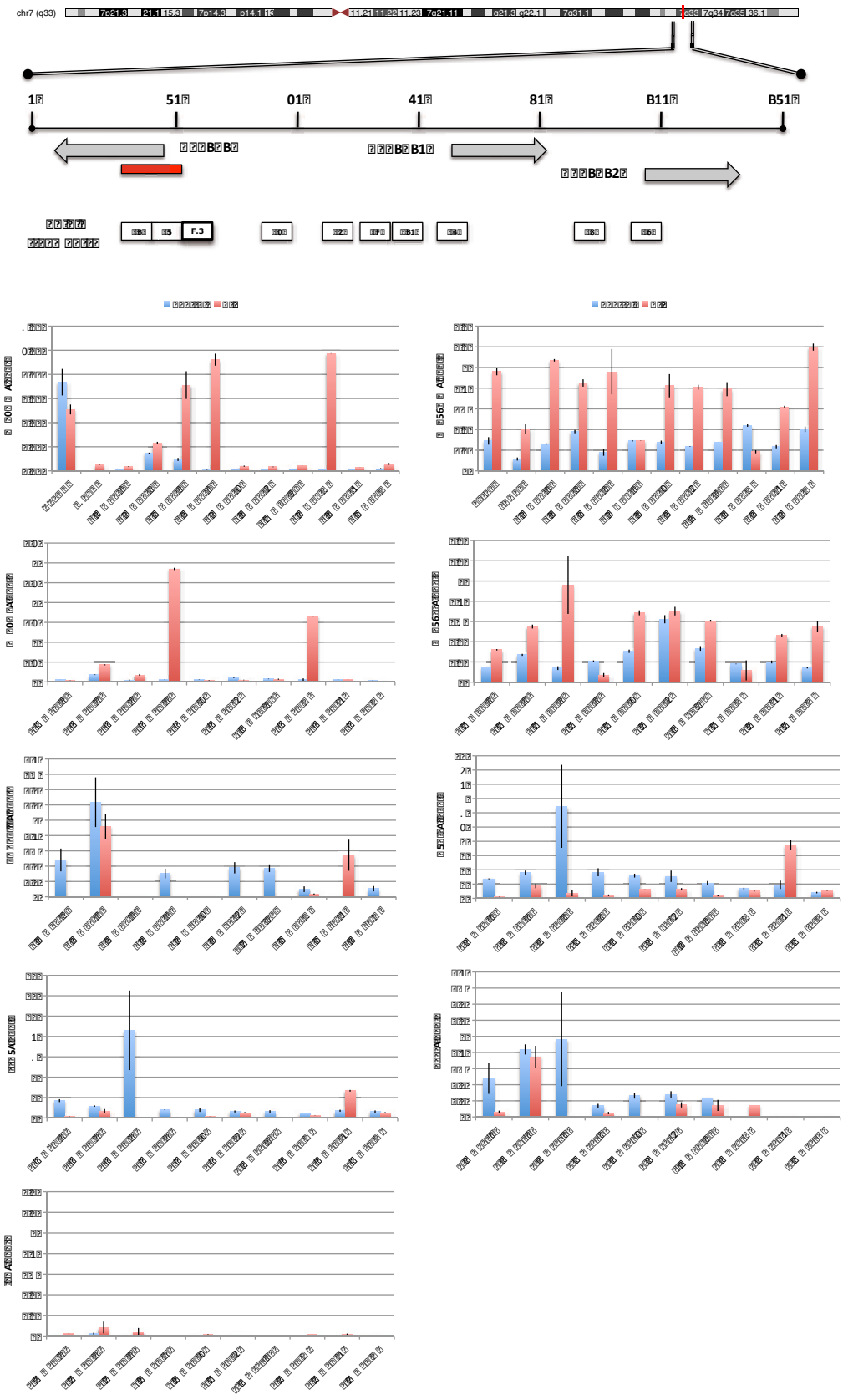
ah 5. g C CK C Cd A C C2h Dl i A neh yy mA A CC5 mAlhmD A 1h l o l 2s K h

hmN l A e hm D m l C Dm ay

ecg9 t 9 e ge 9 c v F 3t (h F 3t Fh F) 7 h F) 7t FKh Recg9 f ne g) A 9gR n e e e Sn9g e c v S9l h) K m n lar a Rn9t ge et t o 9Sn eSegge9 v R K e (z (z gn g megR Wy r 3t g ne lc t gR9 cK v e on z WKA

R n colgc S9e g 9og ct ll ne Rt g 9 F 3t F e gR ((Sn9t 9g ne e ((z A gn g llc m 9ol 9 c nf l9 l e 9 gR t nu e gR 9t e n e 9 e lo e gR ((, Sn9t 9g nh mRe R ac 9 cecg g megR gR n y pSn cce9 9 gR cel g A et el nlah m 9ol 9g g g F 3t (t nu e gR ((z llch lgR9o R oS9 gR gn gt gh h t gc 3 z cR9m ne Rt g 9n F 3t (SucA R gef R n t nu F) 7 m c e g n gn gt gh ll l9 gR l9 oc cS e lla e gR ((Sn9t 9g nA Rec n colg R c cS e l n l f oc eg ec cc9 e g meg R gR n gf ge9 9 gR R nc eg mell ec occ g nA R n c F) 7t F l f lc m n t 9 n g la ne R e gR (l9 oc e ((z llch Wy r gn gt g l nla e o l9 l e 9 gRec Recg9 t 9 e ge9 n9cc gR gen n e9 A Rec g m c Sn fe9ocla9 c nf e gR l e 9gR n cel g n e9 c h gef g meg R Wy r gn gt g g 9 ne o r , , 2 KA

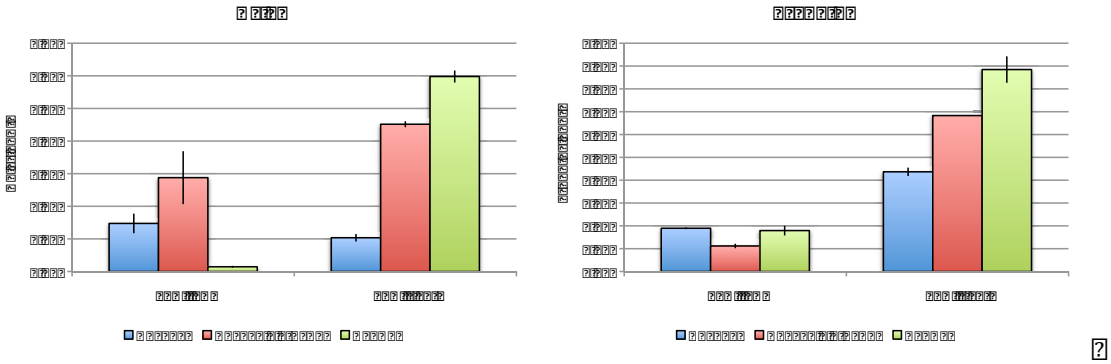
(F)



ah52gmDyme emDRRpbRpRAylmANh ypp0NGRpp0RG. NGAGp0AmAlhmDA2sKlhl CC5DybiDa yhehyAl A lmAhmNlA N NaAmehedlmAiDayhyomoAyylAhN AlhlmAmihAealhhmhhyA lyl AhilmAb

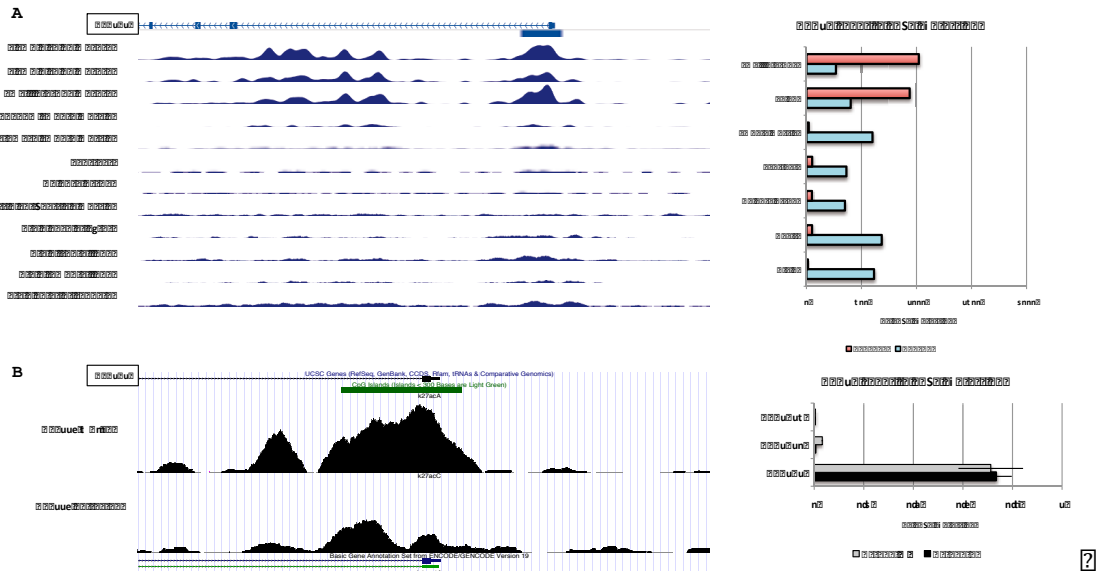
S9 gn gt gR n m c 19cc9 S9l e t 9cg9 gR n e9 h p Sg ((v n t g) K ((WSn9t 9g nh mR n gR n m c e A9n)) A hgR n m c n t ge 19cc9 ce l e n t g F e gn g g llch mRel h t g 2v ((WSn9t 9g nK e e g cegaA g n cge lah oS9 no gn gt g gh gn c neSge9 l n gef ge9 9 ((, ((W m n 9t S e aR 19cc9 l9 gR gen 9 och p Sg 9n m t g) hmR n gR S u ac lt 9cg o lg n A

lc9 lar gR e e 9 y l y1 gn c neSge9 g9nc 9 gR ((S ecl e ((zh ((z gn g meg R Wy r S) v e on zk A



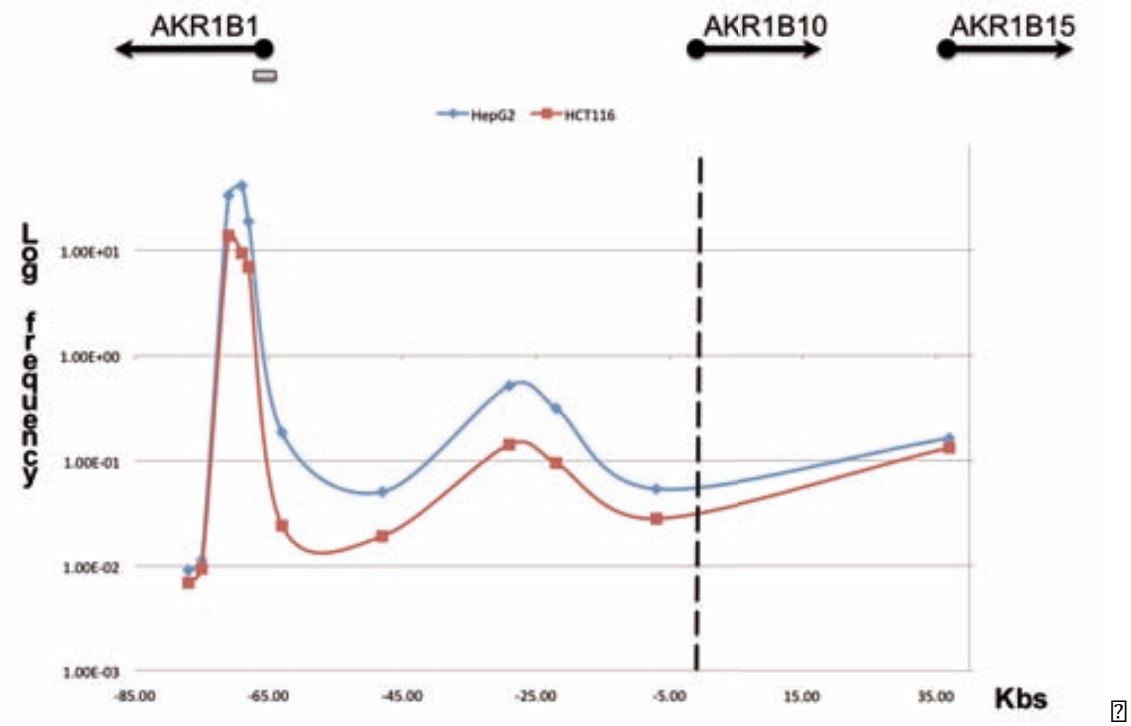
R e e 9 9gR g9 gR ((S acl ac no e l 9 ngR pSn cce9 9 ((, h c eg melll cR9m e pg pSnet gcA n m 9 c nf gR g gR e e 9 9gR m cle n c mR mn gn g meg R Wy r cle n c gR pSn cce9 9 ((, A S) pSn cc ((, R gR Re R nf lo c9 e e h cS e lla e gR me Rg ce 9 gR S eA

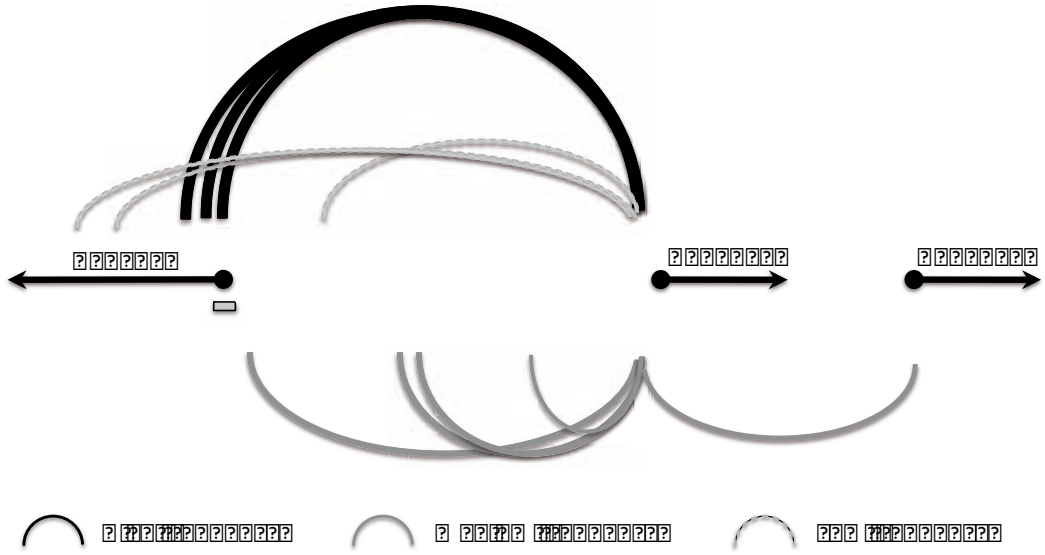
G p0 A AmhN Dl yya y A A2s K lh l CC5 A h DDDA oc 9on R yc i g e So le la f el l R yc i g g9 lar e g el gR a t e c9 gR gef R n Recg9 t nu F) 7 e e n g 9nt l gacco c e o gn g Wy r gn g ((z llcA cn l g e e on z 7 hgR n ac S u9 F) 7 e gR ((S acl so cge gR9c gacco c mR n ((, acpSn cc N9l9 t l e g cge A



ah50g3DDi Dym G p0 h e l A h Al AmhN DDlyya y rDDl e A D3 A l
neh yy mA m C C A C Cd A l my l yya y rh l e A D3 3 G p0 A mAlhmDD CC5
mDmh l DDA h DDDA A l h l o 2s K rDDl e A D3 A neh yy mA m C C R
C Cd A C C2 A mAlhmDDA CC5 2s K 1h l DDy rh l e A D3

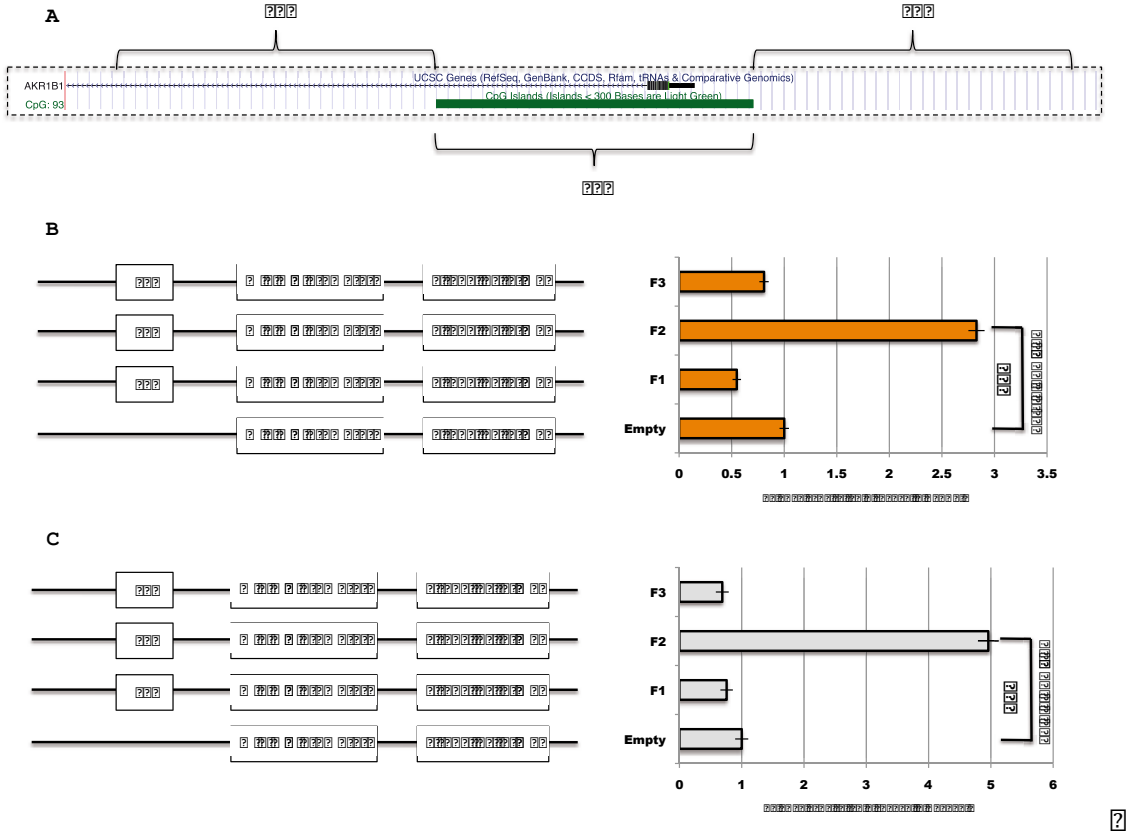
t 9 cgn g 19cc9 gR F) 7 t nu e gR ((Sn9t 9g n h e9 e
((z 9l9n g l n l le 9t S n meg R gR 9nt l 9l9 t o 9c
v e on z7KA g n cge lah gR F) 7 S u ec n 9f n g n Wy r
gn gt g mR g 9 gR y pSn cce9 9 ((, ((WA 9 gn cgh
gR pSn cce9 9 ((SS nc g9 o g a gR Se ge R cA
R c h colgc oc g9 RaS9gR cer Sog gef R n gef ega 9 gR ((
Sn9t 9g n h e9 9 ((, ((W c h ol ge9 A gR n nuc m n
lc9 lar oce R yc i g v p KA


AKR1B1
AKR1B10
AKR1B15
HepG2
HCT116
Log frequency
1.00E+01
1.00E+00
1.00E-01
1.00E-02
1.00E-03
-85.00
-65.00
-45.00
-25.00
-5.00
15.00
35.00
Kbs

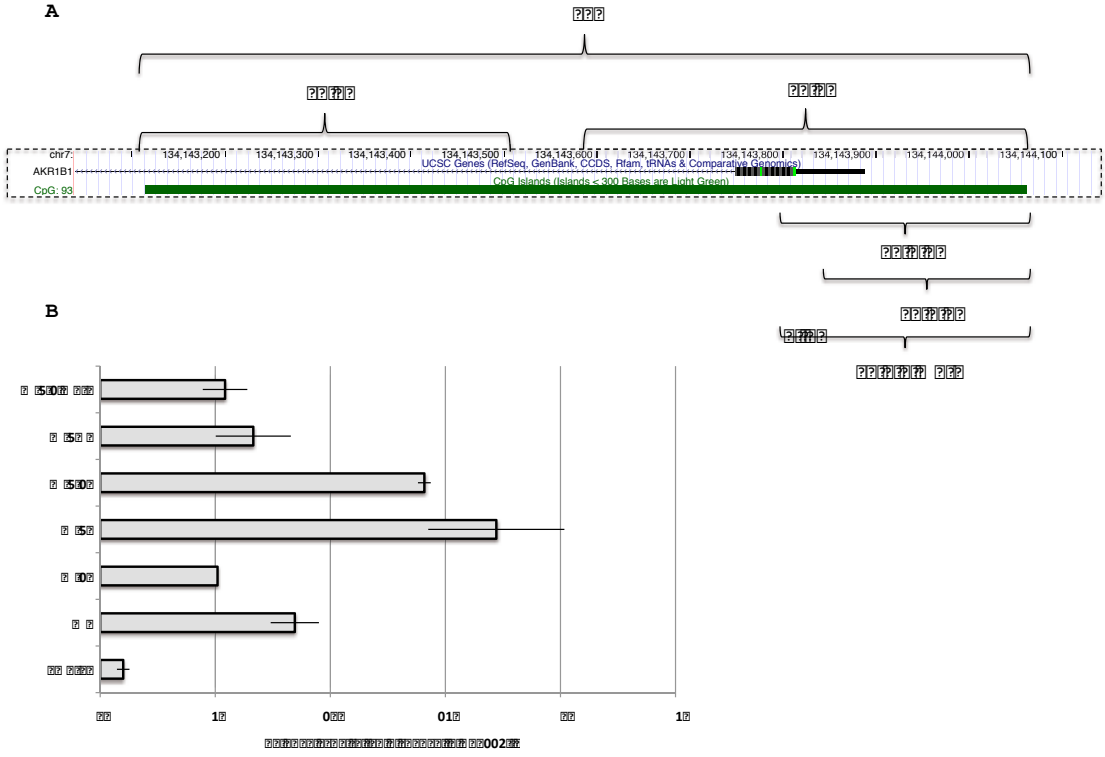


ahT Tg hAlh ta Af mAl lem Alyo A C Cm ayh

R gg gR g gR 9 9nt ge9 mc lt 9cg gR ct e S gla9 gR pSn cce9 9 co cgc gR g gR e gn ge9 ec t e ge e S gla9 gR pSn cce9 cg g he e ge gR g gR n t ocg 9gR n t R ect ge c 9 I9 cmeg RA RaS9gR cer gR g gRe ct R ect ac n l g g9 Se ge R c9 onne e gR ((S eA ongR nt 9n hgR F g S9e g 9og gR 19c n l ge9 cReS gm ((S acl gR ((, Sn9t 9g nh he 9n e f t 9n gR leu la h9l 9 gR S acl g9 n9nu c R nh cSn9S9c 9f A


A
AKR1B1
CpG: 93
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
CpG Islands (Islands < 300 Bases are Light Green)
B
F3
F2
F1
Empty
0
0.5
1
1.5
2
2.5
3
3.5
C
F3
F2
F1
Empty
0
1
2
3
4
5
6

ongRnt 9nh m m g g9 loe gg mRe?R Sn9ge 9n Sn9ge c m?n? n?cS9 ce?l?9n?gR? R n?n?elega9? ((S acl A?ongRnt 9nh?m?encg? efe gR?S acl e?gm9?S?ngch? t h?t ? gc? A(? A v?e?on?7 (K? cge gR?c h?t ? gc?megR?loe?n?c cc?a?v?e?on?7 (K?m?n?ler gR?g? n?t ? g? A?ec f ? t 9n?S9m?n?ob c R n?gR? gR?mR9l h?t ? g? h? R9m?f?n?h?t ? g? A(m?c?l?cc?gef?gR? h?t ? g? h?ogctgell?l?nla?ge c? R n?nA?pg?m?ct?l g gR?h?t ? g? A?h?mRe?R?m?c? ?lar oce ?e?n? g? gn? c?neSge9 g9nc?e e ?tcg?c?S n?e?g9nch? ?t 9 ?c9t e?g?chgm9? 9?gR?t ?m?n?c?l g g9 g?cg ?9n?cS ?eae R n? gef ega?megRe gR? ((S acl N?y I y1A



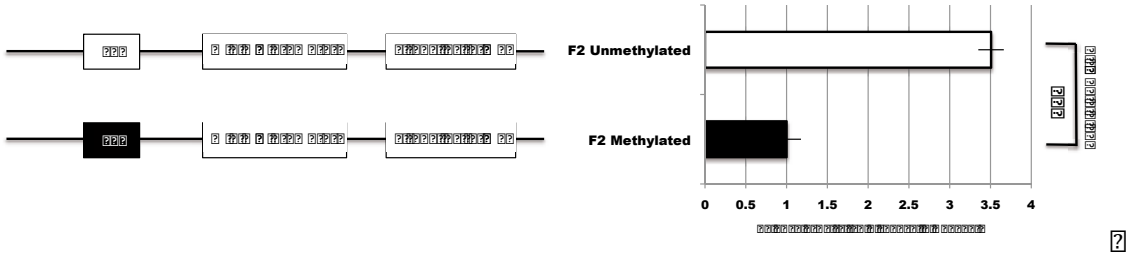
ah?DCg?A A h? l f m h?Al h?N ?Aly o A C?C e yD?A b?3 AmN e my ?mA m i Da l h?N ?Alb?3 D? l f A A h? l f mN e h o l N elf mAylha l b

y I y1 e g9?9f?nl?SSe c?io c g9?9 ?ant ?gR?en? S?nge?eS ge9 ?m f?lo g gR? R n? gef ega?9?tcR9ng?n?h?t ? gc?e ?lo e gR? v) A(K? p?lo e ?gR? v) A) K? ?h e llah? 9gR?n?9 cgno g e mRe?R?gR e e ceg?9 y I y1 m c t og g g9?n?t 9f gR? n?9 ege9 t 9ge v) A(t ogK v e?on?7 (K

llll9 cgno gc e lo e gR e g g tR9m gR Re R cg f lc9 lo e n c gef ega lmRel l ge9 9n t og ge9 9 gR ceg n colg e n c R n gef ega v e on 7 (KA R c cc ac S9e g 9og gR t S9ng 9 gR c gm9 9n gR R n gef ega 9 gR ((S acl A

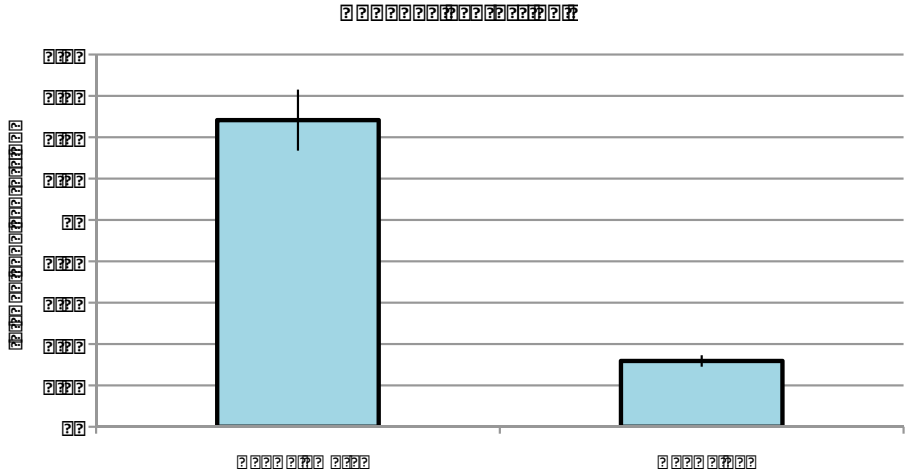
l f Dl mAs A A h a h y yy f y

g nm n ch m m g g9 R u R9m t gRal ge9 g gR R n gef ega 9 ((S ek c9on g e got 9nc e e g co R l9c n l ge9 cReSA 9 f lo g gR n9l 9 t gRal ge9 m n gn c g f g9n) I S 3A Fh 9gR t gRal g o t gRal g vc t g ne t gR9 cK e ((z ll le m n S n9nt R n lo e n c cc aA

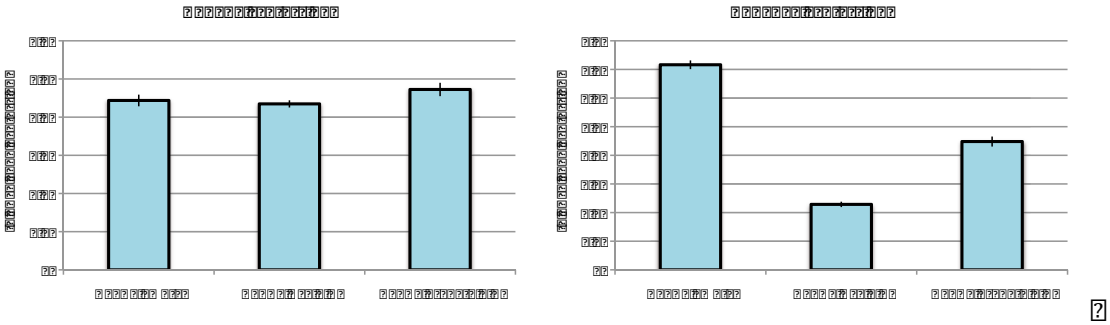


ah Op gA A h yy f y h ya Dy lh Ay l A p a AN l f Dl ro l mn3 A N l f Dl r D mn3

t gRal g) I S 3A F ce e gla oSn ol g vS f lo q, A (K gR pSn cce9 9 gR lo e n c 9oga FAW9l c 9t S n megR gR t gRal g) I S 3A F f g9n v e on 7) KA R c g coSS9ng gR R n gef ega 9 ((S e t 9 cgn g egc t 9 ol ge9 a t gRal ge9 A



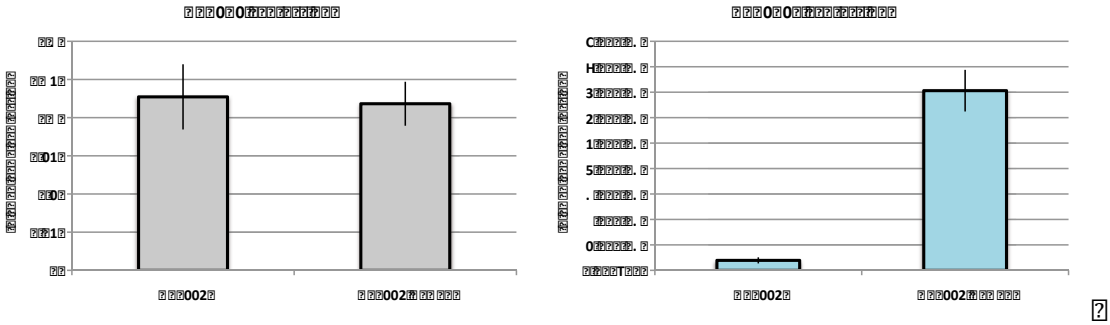
TFBS: ATTGGTTG ATT*****



ahd0. gC C A C d neh yy mA A e p N l f R e p s e p c s| h

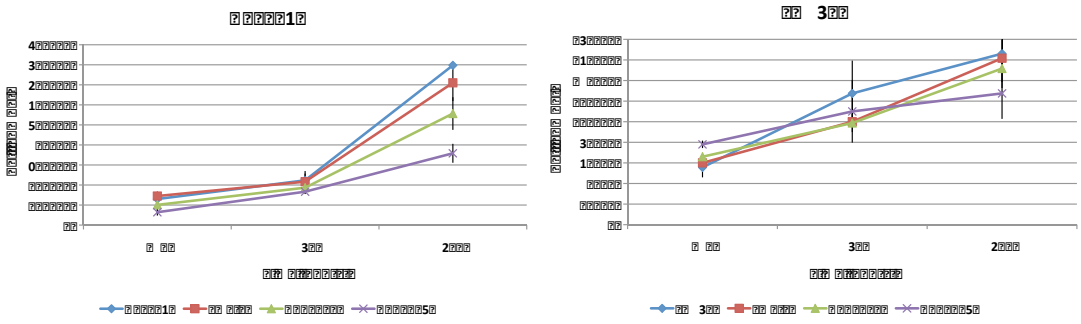
g n cge lahgR 9cc9 9gR gn c neSge9 g9nc Sn9o gR 9m n ol ge9 9 ((, ogRe 9g gR pSn cce9 9 ((hlgR9oRgR e e 9 9gR c ec e gR Sn9t 9gn9 gRec v e on 73KA pS net gc o nSe m colgc mR gm9 c nf toce gR lo e n c 9 cgno gcD gR gg gR e e 9 gR c Sn9g e c ac no e l9 ngR R nn9l9 gR ((S acl AongR n R pS net gc n 9 ent gR c 9 y I y1 e e g9 gR c i o e R pS net g A cS ee c i o c n9 g e p A

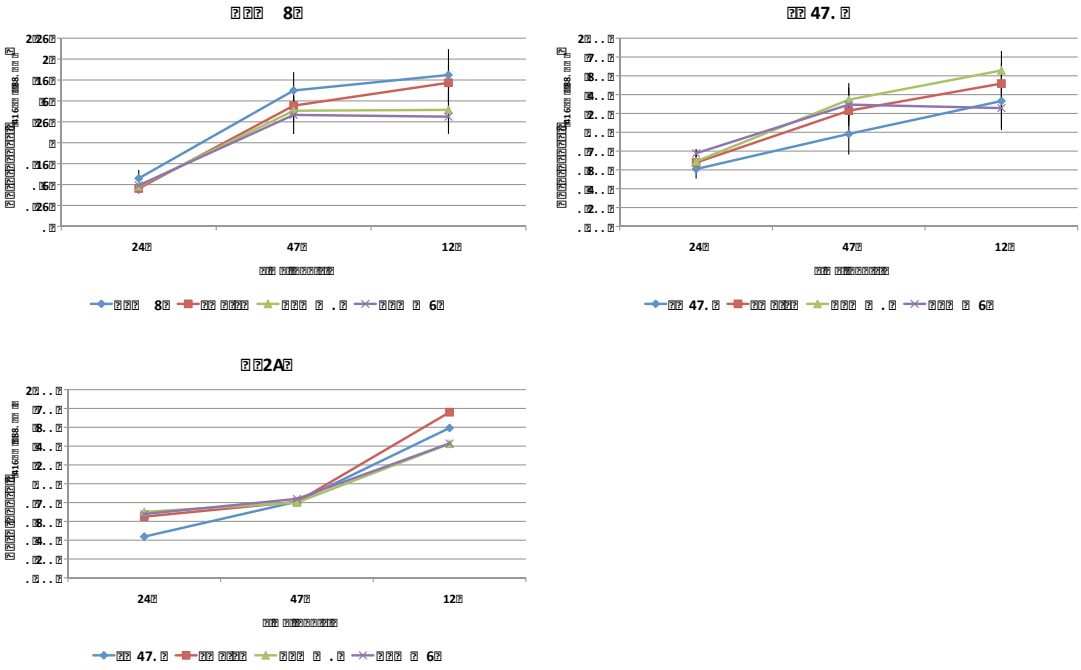
A



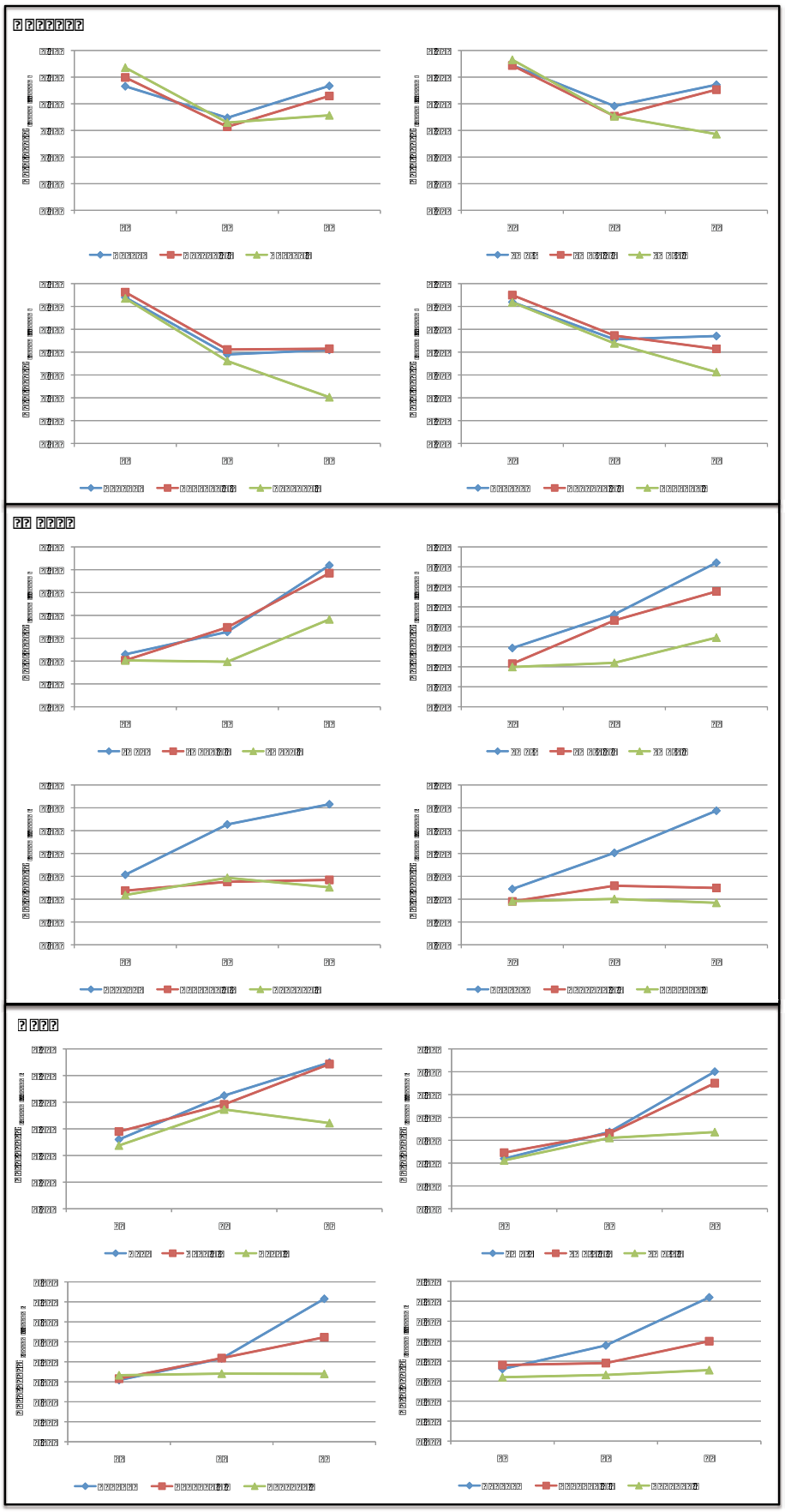
ah02 C C A C Cd neh yy mA A D y y m CC5 Dy1h o l el

g n cge lah m 9 c nf co cg ge1 e n c e gR pSn cce9 9 ((, mR g 9ol e e g gR g gR R n m c gef g o g9 gR 19cc9) n Sn cce9 t R ect A gR 9gR n R hgR pSn cce9 9 ((e 9g R g n gR gn gt gA ongR n R pSnet gc n 9 cc cc gR R c e F) 7 e e g n gR gn gt gA


3
3
3
4



et el nla g9 mR g 9 omn e gR Sn9le n ge9 cc ah gR cc a t 9 cgn g l9m n c9n ve e ge l cc llol n gef egaK e gR gn c g le c S e lla e W e on 7 7KA pgh m n m g g9 R u gR llol n gef ega9 gR llc mR m ge l m c g9 gR t eot A t9t e m ge l gn gt g gVt g ne l t gR9 cK megR cc a g) 3h32 7) R9onc e ((zh m32, 5 mel gaS gn c g megR gR ((, h ((W9n gR t Sga f g9n c 9 gn9l v e on 72KA ge l m9ol n o g9 n ge 9l a ((, ((Wh mRe R ac u 9m g9 R f geSn9le n gef n9l e 9l9n g l n mllcA g n ge l 9c c9 B It l llc m n Sla g eg Sn9le n g D R9m f n h g, AVB It l9 n ge lh9 gn9l llc m n t 9cg o g og llc gn c g megR ((, 9n ((Wm n 9g efe e one gR gn gt g gA llgR c Sn9le n ge9 cc ac g h co cg gey got 9n gef ega9 ((, ((We 9l9n g l nh f ngR l cc gRec g ac t o R t 9n l n mR m ge l g9 gR t eot A



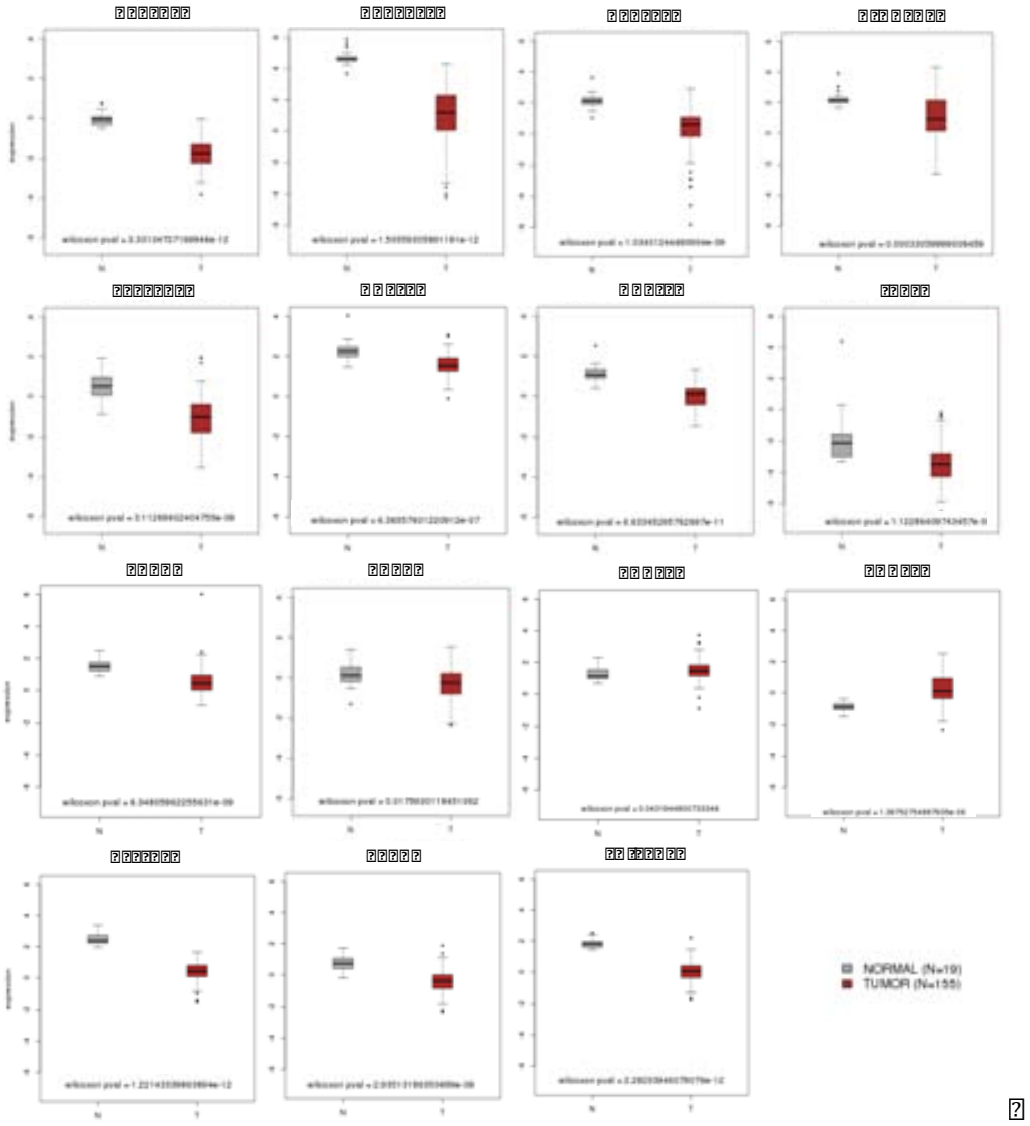
ah04g lA DlhlN Al A CC5Ro . 4d A pTRhAy lmhAmlo l C Cd A
C C2h hAl mA Alh l mAymhl A Dr) cN Do h lyl A h Al DlDA yh

Retinol-oxidizing	Retinal-reducing	Retinal-oxidizing	RA-oxidizing	Nuclear TF	Other functions
ADH1A	AKR1B1	ALDH1A1	CYP26A1	RARA	ABCG5
ADH1B	AKR1B10	ALDH1A2	CYP26B1	RARB	ABCG8
ADH1C	AKR1C1	ALDH1A3	CYP2E1	RARG	AWAT1
ADH4	AKR1C2	ALDH8A1		RXRA	AWAT2
ADH5	AKR1C3			RXRB	BCO2
ADH6	AKR1C4			RXRG	BCMO1
ADH7	SDR16C5				CRABP1
DHRS3	HSD17B6				CRABP2
DHRS4					DGAT1
RDH5					DGAT2
RDH8					LRAT
RDH10					RBP1
RDH11					RBP4
RDH12					RBP7
RDH13					RETSAT
RDH14					SULT1A1

z (o n g c e S g 9 g 9 n c v h h h W 9 g R n c e f 9 l f e g R S g R m a v (h h (h (K W g f c n R e R l e R g e l 3 A

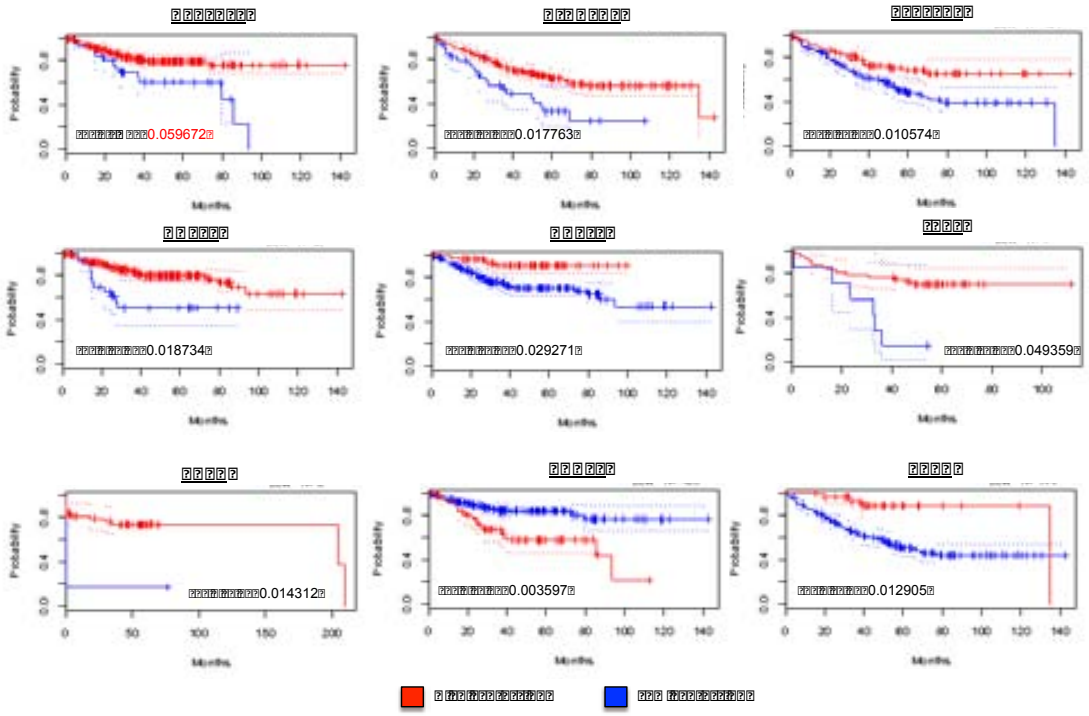
D h n e h y y m A m A y i m D A h l A m e e l o f y y m l o l y a h i l D A m D m h l D A h e l A l y

9 g e c e R g c e g 9 g R l e e t S l e g e 9 c 9 l g n g e 9 c e g R S g R m a e 9 l 9 n g l n h m l a r c g g c g e l c c 9 e g e 9 c 9 p S n c c e 9 f l c 9 g R z c e f 9 l f e S g R m a e 9 l 9 A g m n 9 g e m 9 t g R 9 l 9 n g l n g c g g R g e l o (5 9 n t l 9 l 9 (W 9 l 9 n g l g o t 9 n t S l c A



a h 0 T g A n e h y y m A D i D y m l h l A m e e l o f A y y A A l D D h A m D m h l D A h C T A m h N D m D m A y N e D y r A h f 3 A C 2 2 m D m A l a N m h y r A h 3 o h a y

on n colgc l nla cR9m gR g gR 9f n ll gn c neSge9 l Sn9 al 9 gR S gRm a m c lg n one 9l9n g l got 9ne ceA g9g l 9 W c 9og9)zK m n lg n Dt 9cg9 gR t v(FK m n 9m n ol g v ((h (Fh ((h) z (h (h 3h h h h (h pSn cce9 Sn9 c oc e gRec lacec n lecg e p A 9 n cc mR gR n n ge 9e e S gRm a lg n pSn cce9 9ol g confef l9og9t e 9l9n g l nh m lar F e n g pSn cce9 g c gc9 Rot 9l9n g l n gR g e lo confef l e 9nt ge9 A egR gR c g g m 9 cgno g Sl y e n confef l onf c 9n R lg n S gRm a vc t g ne lc l gR9cK g n cge lah m 9o gR g 9f n ll confef l 9n ec c h confef l m c ce ea gla 9m n e gR9c S ge gc gR g ecSl a lg n pSn cce9 Sn gg n 9n c9t 9 gR cv e on 2, KA



ah 4d g e DAs h yahi i Dah i y mh e l o f A yao mm e i Da I adhal2 r n el mh C Cd3h h hyd m ne h yy mA A Da dmDho ne h yy mAb ml DA yh D l mA A Al hi D

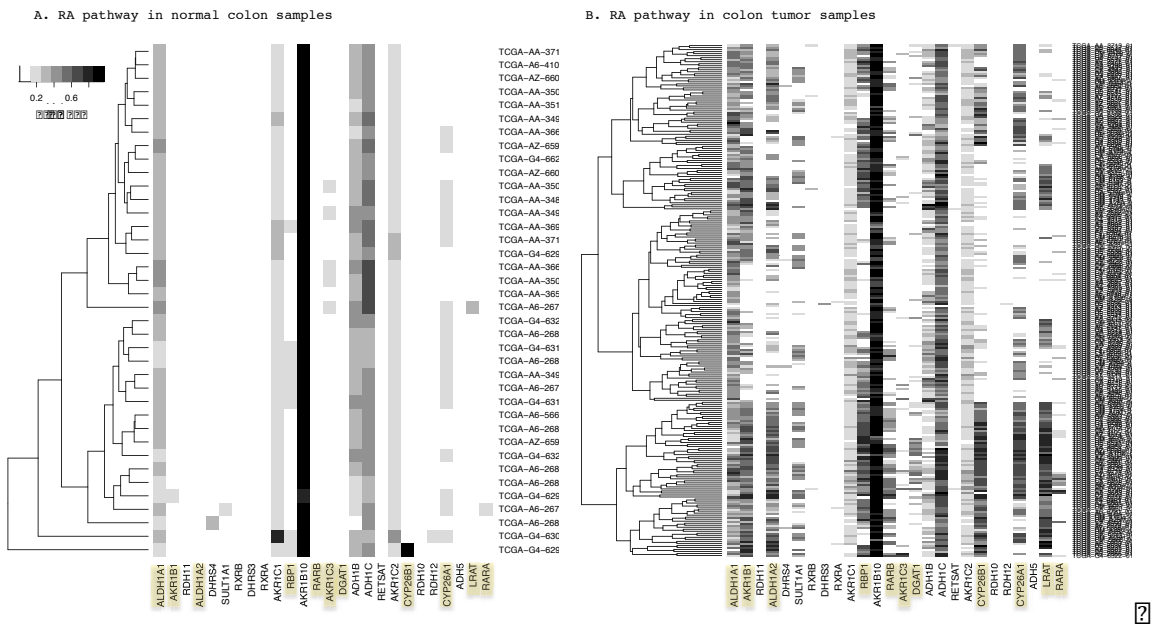
((, h ( ) h) z (h (h 3h h n c gR g pRe eg pSn cce9 9m n ol ge9 one gR 9l9n g l n n Sn9 cch 9m l f lc 9 gR c c n cc9 e g meg R m9nc t confef l h g v e on 2, KA gR 9 gn nah (, ac gR g g c g9 oSn ol g e

9l9n g1 nh e gRc c hgR oSn ol ge9 ec en gla e u g9S99n conf f l v e on 2, KA R c n colgc e e g gR g gR lg n ge9 9 S gRm a c ac 19c la n l g meg R gR conf f l 9 gR S ge g gR n ol ge9 9 c9t u a c t a R f SSle ge9 c Sn9 9cge g9n9 S99n conf f l 9n S ge gc meg R 9l9n g1 nc A

hmN ml h N l f D l mA e hm DA m l e l o f A AmhN D A m Dmh l D laN mh1 yya

R n c t 9cg9 gR S gRm a u a c n 9m n ol g e 9l9n g1 nh gR n ac 9 l n t R ect 9og R9m eg9 onc A n m lar gR n9l9 t gRal ge9 e gR cel e 9 gR n ge 9e e S gRm a c A

lar gR t gRal ge9 9 gR 9l9 t gRal ge9 g e 9nt l 9l9 h 9t S n meg R gR 9l9 got 9n gacco A 9t S ne 9nt l e cg got 9n t gRal ge9 h (c v (h (Fh (h ( ) h) z (h) z (h (h h h (Kmn Ra S nt gRal g e 9l9 nv e on 2 (KA



ah 4C g N l f D l mA l N ey m e e l o f A y A mDmh l D A hh l o h m l A mhN l y b l f D l mA e hm D y m AmhN D mDmA y N e D y r 3 A laN mh y N e D y r 3o f fe hN l f D l A y D l l b

9n t 9cg9 gR t h gRec ec gR ncg tgo a n S9nge eg c Ra S nt gRal ge9 e 9l9 nA g n cge la h 9 e n e t gRal ge9 9t S ne t g c h m c9 c nf v g 9g tR9m KmR g Re Rle Rgc gR S n f l 9

gRcc lg n ge9 e nla cg c9 got 9ne cecA llot e 3W t gRal ge9 Sn9 c c e gRcc lacecn e p A

c c ne 9f hgR n m n [F c 9m n ol g e S gRm a v e on 75K 2 9 gR t vz(AWFb K lc9 pRe eg ce e g Sn9t 9g n RaS nt gRal ge9 e gR got 9n c t Sl v l [W e on2 2(KA 9 ce n c RaS nt gRal g gR9c c ce e gla RaS nt gRal g v9t S n meg R 9nt l 9l9 S f lo q, A WK9n ge t 9n gR j W b 9 gR c cA

	ALDH1A1	AKR1B1	ALDH1A2	RBP1	RARB	AKR1C3	DGAT1	CYP26B1	CYP26A1	LRAT	RARA
Norm mean	0.290	0.019	0.080	0.094	0.036	0.084	0.014	0.091	0.128	0.060	0.052
Norm sd	0.063	0.019	0.025	0.045	0.013	0.041	0.002	0.133	0.042	0.048	0.021
Tum mean	0.420	0.397	0.343	0.385	0.188	0.069	0.081	0.323	0.359	0.330	0.084
Tum sd	0.150	0.234	0.230	0.257	0.211	0.092	0.137	0.270	0.249	0.299	0.106
Wilcoxon pval	3.17E-08	7.91E-15	4.37E-09	2.30E-07	0.130	7.64E-08	8.33E-06	9.63E-07	0.000	0.002	0.363
adj pval	2.06E-07	2.06E-13	3.78E-08	9.96E-07	0.198	3.97E-07	2.71E-05	3.58E-06	0.000	0.005	0.429
Hypermethylation cases (%)	45.669	79.921	71.654	66.142	45.669	7.480	39.370	57.874	61.811	55.512	27.559

D C2 g l yl D ya N N hf m l h Al D N l f D l A y A AmhN D A laN mh mDmA l yya h

y mi hf m A o A y h D l o l e l o f f A D y y m

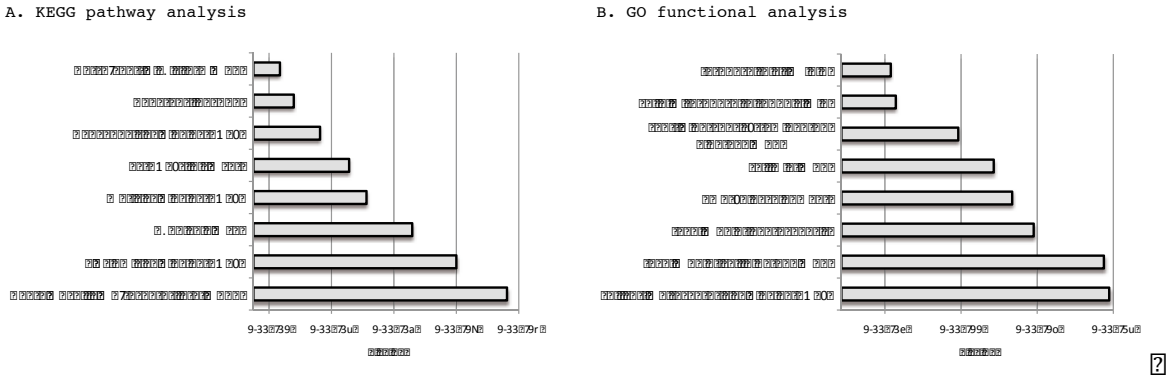
N l f D l mA msh aD l mA

pgh m pSl9n t gRal ge9 9nn l ge9 t 9 gR e n g c 9 ce n R n Ag R c Re Rle Rg e ot n9 cgo e c gR g 9y pSn cc 9n 9yt gRal g c n o ge9 lla n l g S nR Sc e g n ge SRace lla e gR o l oc cgno gon v uol u9 d lt ch), (FD 9og f h), ,) D n c nd l e ut 9n h , , 7KA

m9 n mRe R9gR n c n 9yt gRal g megR gR RaS nt gRal g S gRm a c g9 lo e g mR gR n gR n ac 9t t 9 S gRm a 9n e9l9 e l o ge9 cS ea lla lg n gRn9o R gR c RaS nt gRal ge9 Sn9 cc cA 9n n g 9t SlcR eghm lar 9yt gRal ge9 9 R9 9 gR (S gRm a c RaS nt gRal g e 9l9n g l n meg R gR n cg9 c n Sn c g e gR llot e 3W mn a vt g ne lc i gR9 cKA 9yt gRal g c m n l ar meg R gR o ge9 l 9g ge9 g 99l vRggSNI f e A e m 9f I K 9n v g9l9 aK S gRm a v a9g9 a l9S e 9 c 9t ck lacecA

g n cge lahe 9gR o ge9 l S gRm a lacec n 9 c nf n e Rt g 9 n ge 9e e n l g o ge9 cha o R c f l9St g l Sn9 cc ch on9 l e g n ge9 ch e n ge9 h l t e n ge9 n ge 9e e

feg?t e ??ce? ?le ?S?gRm?ac3v?e?on??2) KA?9??gR?nh?gR?c???g??coSS9ng???cS??e?a??
Sn9??cc9?9???9yt ?gRal?ge9 ?e f9lfe ???e???n? g?n?ge 9e???e??n?l?g????? ?cA?



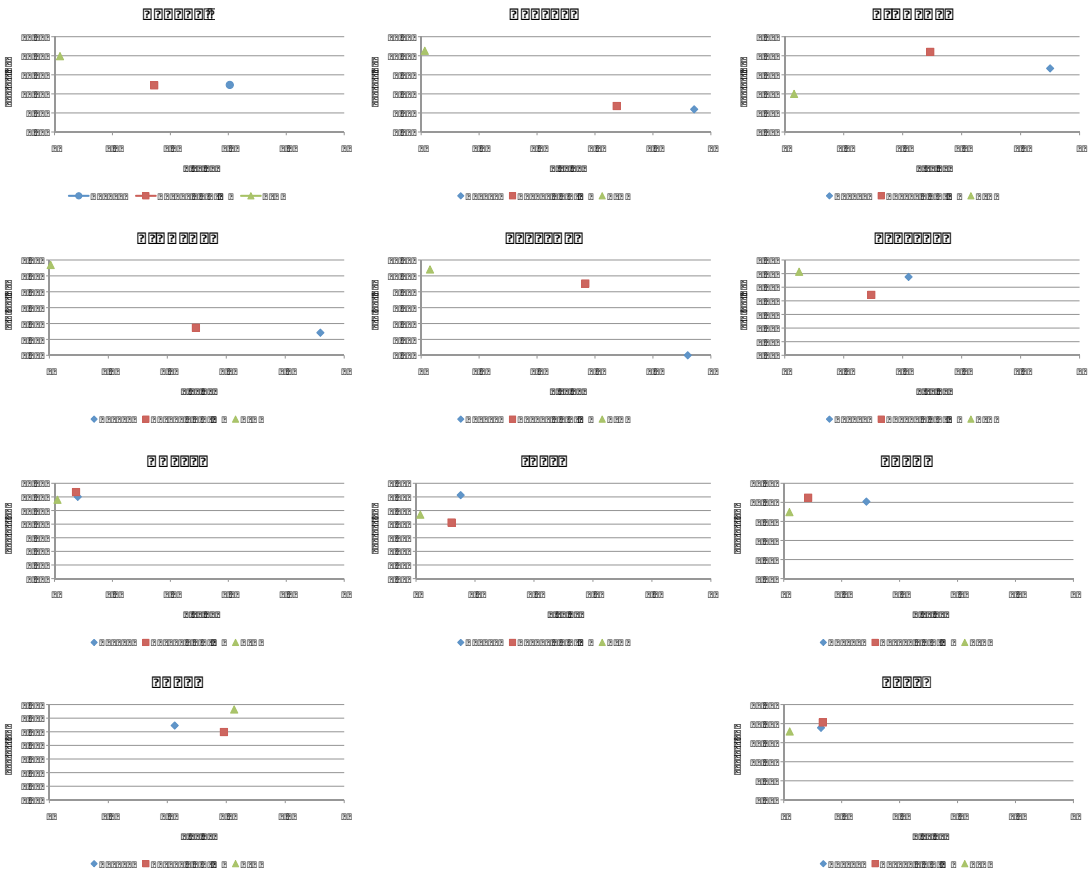
?

??ah??4p?g??msN ?l?f D?l ?mA??A?Dfy?y?m?l?????l?Am?????????e?l?o ?f??fe?hN ?l?f D?l?????A?yb?? ???D?f ?y??A?????Al??msN ?l?f D?l????e?hm??y?o ?h?y?D?l????1m?e?h?mhN ?????r?3??A????r?3??A?Dfy?yh?

?

??N ?l?f D?l ?mA?m?l????h?l?Am?????????e?l?o ?f y?e?hl??D?f ?h?ylmh????l????D?h????
??A???ne h?yy?mA?

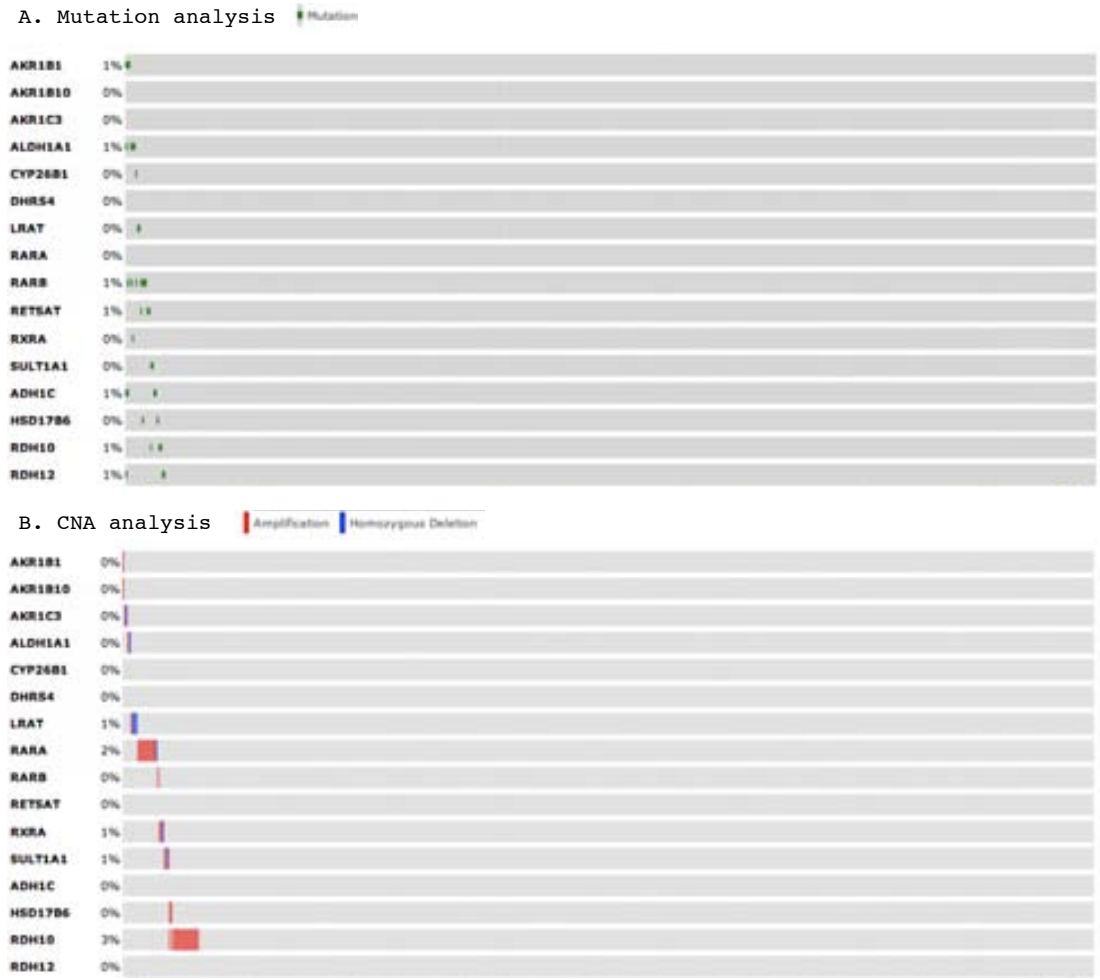
?g?R?c????? ??l??nla????t 9 cgn?g???gR?g??Se?? ?ge??lla?cel? ?????? ?c???? ????
n???gef?g????a?gn???ge ??gR????llc?megR?gR????t ?gRal?ge ????? g?Wy?r?????v?Ro? ??
?????h?), (, KA?9???g?????g?g?n?Se?gon????9og?gR??n9l?9?????t ?gRal?ge9 ?e ?gR??
cel? ?e ??9??gR??n?ge 9e????e??S?gRm?al?m??? ?lar????? ???pSn?cce9 ?v???el? g?
?on???ne g??F?2pz, ???nn?ak?? ??????t ?gRal?ge9 ?v?llot e ??3W,?K?Sn9?el?c?e ?
????((z?????((z?gn??g???megR?Wy?r????? ????????((z????llc?vu 9?u?9og??9n?
????(?? ??????F?K?mRe?R?R?f??l?na?9m??f?lc?9?????t ?gRal?ge9 A?
????((z??ecSl?a???R??fa?t ?gRal?ge9 ?v??g??f?lo?T, AWKe ?f?????S?gRm?a??? ?c?
v????(?(?h?????(?Fh?????(?(?h?????(?)?h????)z?(K?mRe?R?m?n??S?nge?lla?
??t ?gRal?g??? oS9 ? Wy?r???? gn??gt ? gA? ???? ??llc? cR9m??? ?? ???S?n?
??t ?gRal?ge9 ?9??gR??c?t ??Sn9t 9g?nc?v?e?on??2FKA?9on?9??gR??RaS?nt ?gRal?g???
?? ?c?v????(?(?h?????(?Fh?????(?) ?? ?????)z?(K??ecSl?a???? ?e ?n??c??9??
?pSn?cce9 ??cc9?e?g???megR?gR??19cc?9????????t ?gRal?ge9 ?v?e?on??2FKA?????(?(?
cR9m??? 9??cc9?e?ge9 ???gm?? ?????t ?gRal?ge9 ?? ???? ???pSn?cce9 h??og?gRec?
?? ???9? 9g?R?f???Se??cc9?e?g???? ??gR???n9l?9??t ?gRal?ge9 ?e ?gR9c??ue ??9??
?? ?c?ac? 9g??1??nA?? ?c?megR?e g?nt ??e?g??9n?19m?1?f?lc?9?????t ?gRal?ge9 ?e ?
gR??Sn? g?1????((z????lle ?v????)z?(h?????(h?????h?????h????(?? ??????K?
?pRe?eg???t e et ?l??R ??c?9n?9 ???g???le ?????pSn?cce9 A?



AmN ehm DA m e l o f A mDmh l D A h

9 g e ce Rgc e g9gR e f 9lf t g9gR S gRm a c e 9l9n g l n m ongR n e g nn9 g gR vgRn9o R e9S9ng l m g99lK 9n t og ge9 n i o e c 9Sa ot n lg n ge9 v Ke 9l9n g l got 9nc v e on 23KA ge 9e e S gRm a ec S99nla g a 9t e lg n ge9 cA R 9f n ll S gRm a t og ge9 ac l9mh n Sn c ge 3A b v) 39og9 W2WK9 gR c cD 9 9 gR c n f l Re R h i o a9 t og ge9 c v e on 23KA 9cg9 gR t n t ecc c t og ge9 c v g 9gcR9m KA

gR 9gR n R h lg n ge9 c hec oS g9 7A b v3z 9og9 W2WKhmegR 9gR e c l9cc c S e 9 gR v e on 23KA g n cge lah (, pRe eg h i o g 9Sa ot n e c f t Sle ge9 c v) A b KhgR g SS nc g9 pSl e gR oSn ol ge9 9 c nf e gR pSn cce9 lacec v e on 23KA



ahe4. gheAlmNalmAyr3AmefAaN hDhlmAyr3lADhe elo f AAAmDhlDAhyN eDyb3Cdnehyy mAiyl mefAaN hDhlmAb

EPIGENETIC REGULATION OF AKR1B1 GENE FAMILY IN COLORECTAL CANCER

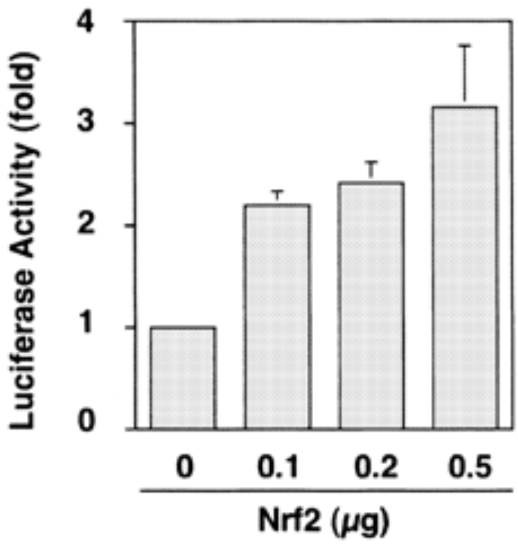
Epigenetic regulation of AKR1B1

The DNA methylation pattern in the genome is basically bimodal, with the large majority of CpG sites methylated at high levels (>85%) and CpG islands largely unmethylated (<10%) (Bergman & Cedar, 2013). This apparently stable DNA methylation scenario is modified in a number of diseases and it has been extensively studied in cancer. DNA methylation alterations in cancer are mainly due to global hypomethylation and focal hypermethylation (Bergman & Cedar, 2013). Many gene promoter hypermethylations have been described in cancers, and in most of them a clear association with a decrease in gene expression has been reported (Table 16).

We started with the observation that AKR1B1 promoter CpG island was hypermethylated in 24 out of 25 cases of colorectal cancer patients. The high incidence of this alteration was confirmed using the TCGA database, where we detected hypermethylation in about 85% of the colorectal cancer cases analyzed (Figures 33 and 37). The interesting fact is that AKR1B1 changes in DNA methylation do not appear to affect its expression, at least in a consistent manner. On one hand, we know that the gene promoter is highly hypermethylated in cancer, but we observed a very faint downregulation (2% of the cases) in the expression of AKR1B1 in colorectal cancer, at both mRNA (results from three different sources: our own data, TCGA and ICO data) and protein level (western blot and TMAs performed in our laboratory). On the other hand, we have checked the expression of this gene after the 5-AzadC treatment and, despite that the CpG island is hypomethylated, gene expression does not change at all. Gathering all these data we could confirm that the expression of AKR1B1 gene is not affected by the DNA methylation status of its promoter.

Gene	Protein	Function*	Effect of loss of function
APC	Adenomatous polyposis coli	Wnt signaling pathway inhibition	Increased Wnt/β-catenin signaling
MLH1	MutL homolog 1	DNA mismatch repair	Microsatellite instability
MGMT	O-6-methylguanine-DNA methyltransferase	Repair of alkylation DNA damage	Increased G>A mutation frequency
RASSF1A	Ras association domain family 1 (isoform A)	Negative RAS effector, proapoptotic, microtubule stabilization	Increased RAS/RAF/MAP kinase signaling, death-receptor-dependent apoptosis
SLC5A8	Sodium solute symporter family 5 member 8	Sodium and short chain fatty acid transporter, suppresses colony formation	Not known
RUNX3	Runt-related transcription factor 3	Transcription factor	Decreased TGF-β/BMP signaling
MINT1‡	Methylated in tumor locus 1	NA	NA
MINT31‡	Methylated in tumor locus 31	NA	NA
SFRP1	Secreted frizzled-related protein 1	Wnt antagonist	Increased Wnt/β-catenin signaling
SFRP2	Secreted frizzled-related protein 2	Wnt antagonist	Increased Wnt/β-catenin signaling
CDH1	E-cadherin	Calcium dependent cell–cell adhesion glycoprotein	Loss of cell adhesion, possible increased Wnt/β-catenin signaling
CDH13	Cadherin 13	Selective cell recognition and adhesion, antiapoptotic	Increased PI3K/Akt/mTOR signaling, MAPK signaling
CRABP1	Retinol-binding protein 1	Carrier protein for transport of retinol, promotes apoptosis	Not known
CDKN2A/p16	Cyclin-dependent kinase inhibitor 2A	Regulates cell cycle G1 progression	Increased cell proliferation
HLTF	Helicase-like transcription factor	dsDNA translocase, fork remodeling activity, ubiquitin ligase	Impaired DNA repair
CDKN2A (P14, ARF)	p14(ARF)	Inhibits E3 ubiquitin ligase	Decreased p53 stabilization and activation
ESR1	Estrogen receptor 1	Ligand-activated transcription factor	Loss of estrogen receptor signaling
TIMP3	Tissue inhibitor of metalloproteinase 3	Inhibition of MMPs and ADAMs	Increased EGFR signaling, TNF signaling
CXCL12	Chemokine (CXC motif) ligand 12	Alpha chemokine	Increased tumor cell metastases
ID4	Inhibitor of DNA binding 4	Transcription factor	Not known
IRF8	Interferon regulatory factor 8	Transcription factor	Interferon signaling
THBS1/TSP1	Thrombospondin 1	Cell-to-cell and cell-to-matrix adhesive glycoprotein	Decreased TGF-β1 signaling
DAPK	Death associated protein kinase	Induction of cell death	Interferon gamma signaling, TNF alpha signaling, Fas/APO1 signaling
VIM	Vimentin	Stablizing cytoskeleton	No known biological effect
SEPT9	Septin 9	GTPase, formation of filaments	Impaired cytokinesis and loss of cell cycle control

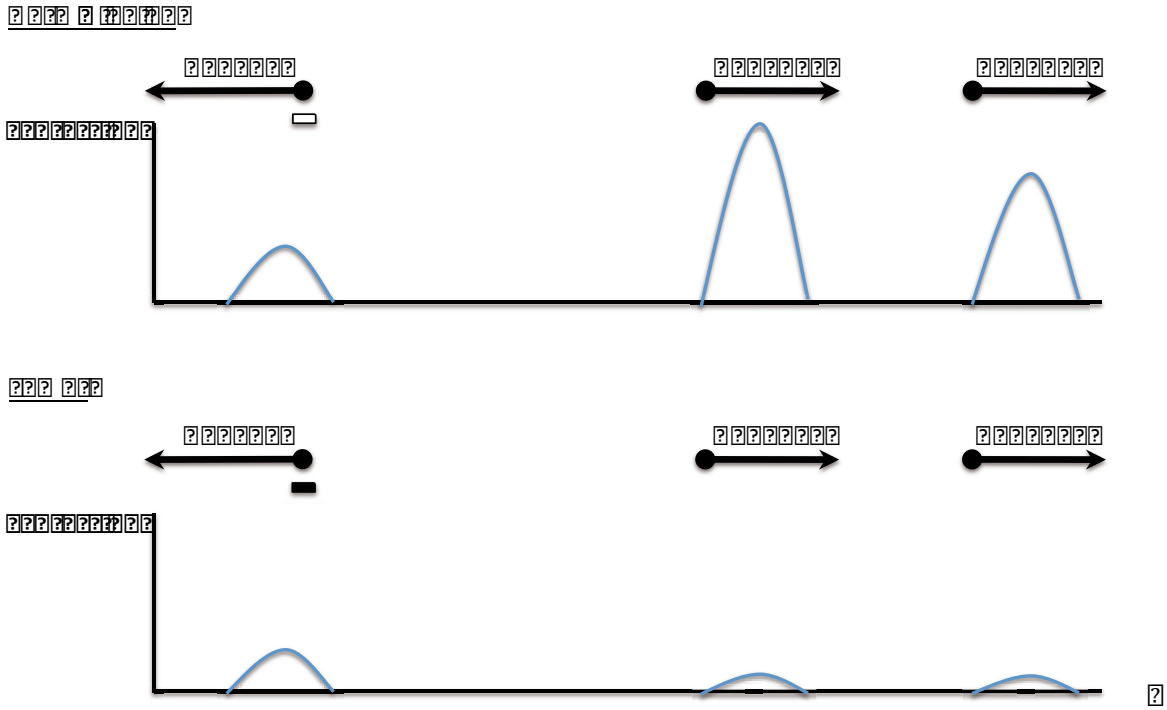
*Many of these gene products have multiple functions. The listed function in this table is the one most commonly cited as the one responsible for CRC formation. ‡MINTs are 'methylated in tumor' loci, and are not specific genes. Abbreviations: ADAM, A Disintegrin and Metalloproteinase; CRC, colorectal cancer; EGFR, epidermal growth factor receptor; MAPK, mitogen-active protein kinase; MMP, matrix metalloproteinase; TNF, tumor necrosis factor.


Luciferase Activity (fold)
0
1
2
3
4
0
0.1
0.2
0.5
Nrf2 (μg)

□

□sei □T□mnsei □ei □□□t 5T□hhs1 vWe3sh□m1Tei □e1□n □ es1 □ei □e□□□□<□<) □□ □□□□□□<□<2□

□T□□□1j□t 5T□hh□□□s □1Tn □o□□ □□eNn 1T□eshhN□vW



□

□□□gE□□v| □A□□□□i□□□p□□□□□□in □□□□□□□B□B□□□□□p□□□□□ □i□c□□i□□□□□□□□□B□□□□□p□□t □E□pp□□□l□□□□ □□E□ □□□□□□□□i□□□□□□□p□□□□□p□□□□ □i□c□□i□□□K□□□E□p□□i□□□n □i□□□□□□ □ic□□□t R□□□□□□□□□□□□E□□i□□□ ig□ □E□p□□ □i□c□□i□□□K□□□□□□□t R□

□e□Tsus Doyb□□i □T□□sh□□□□s u□□□em□□ □ei □□□□□□□h □ei yo□es1 □1□□i □□□□□<□<□D□ □□□ □□ ei □□□□1m T□DNo□es1 □1□□□□□<□<) □□ □□□□□<□<2□D□ □hv□□ □n 1he□□□h□h□ei □□ i y5□Tn □ei yo□es1 □1□□ei □□□□□<□<□□5□□sho□ □□□1TT□o□e□h□msei □hN□he□ es□o□ □1m T□DNo□es1 □1□□ □□□<□<) □□ □□ □□□<□<2□>□sDNT□□ x38v□ □i □T□□1T□□ m□□ i y51ei □hsl□□□ei □□51hhs□o□□T□DNo□e1Ty□T1o□□1□□□□<□<□□5□□sho□ □□s □ei □□ □t 5T□hhs1 □1□□□□<□<) □□ □□□□<□<2v□

ENHANCER ACTIVITY OF AKR1B1 CpG ISLAND

Chromatin conformation in the AKR1B locus

Due to the clear evidences that pointed out a relationship between AKR1B1 CpG island and the regulation of the expression in AKR1B10 and AKR1B15, we decided to investigate the chromatin conformation of AKR1B genes. It is known that enhancers need a close contact with the promoter in order to regulate gene expression (Carter *et al.*,2002; Tolhuis *et al.*, 2014), for this reason we decided to analyzed the structural conformation between AKR1B1 CpG island and AKR1B10, which is more downregulated than AKR1B15.

The Chromosome Conformation Capture (3C) results revealed a close contact between AKR1B1 CpG island surroundings and the AKR1B10 promoter in two cell lines that highly express and do not express AKR1B10 (HepG2 and HCT116, respectively). At that time we wondered if the chromatin interactions of the AKR1B genes promoters would change in carcinogenesis and if this alteration could lead to the observed silencing of AKR1B10 and AKR1B15 genes.

Surprisingly, although the expression of AKR1B10 was very different between HepG2 and HCT116 cell lines, the chromatin conformation of the area was almost identical, what suggested that the silencing of AKR1B10 and AKR1B15 is produced by deregulating the enhancer function of AKR1B1 CpG island without altering the chromatin structure. This result suggests that DNA methylation, and perhaps other epigenetic mechanisms, are the main factors that regulate the expression of AKR1B10 and AKR1B15.

A distant enhancer regulates the expression of AKR1B10 and AKR1B15

To get more insights about the role of AKR1B1 CpG island as distal regulator we decided to use enhancer luciferase approach. The results indicated that whole CpG island already had enhancer activity, however we reduced the size of the fragment to elucidate the TFs responsible for the expression of AKR1B10 and AKR1B15 genes. Finally, we found a single transcription factor binding site (TFBS), used by NF-Y and C/EBP-β, which was associated with an important increase of the luciferase signal. It is worth to mention that within the island,

other fragments also present enhancer activity (F2.1 or F2.3.2), but the mentioned binding site was the one that produced more luciferase signal.

Luciferase experiments provide important clues about the activity of the region, but it is also a limited approach, as the genomic context (nucleosome occupancy, histone marks, DNA methylation, distal interactions and more) is not maintained in this assay. For these reasons we decided to manipulate the original NF-Y and C/EBP-β TFBS, using CRISPR technique (materials and methods) to alter the binding site and study the consequences in the expression of AKR1B10. Finally, we obtained a clone from HepG2 cell line that had a deletion in the TFBS resulting in a downregulation of AKR1B10 gene compared with the control cells. This is a direct evidence that the TFBS identified by enhancer luciferase experiments indeed plays a role in the regulation of AKR1B10 and AKR1B15.

Next step was to find which TF is the responsible for the enhancer activity: NF-y or C/EBP-β. To answer this question we performed CRISPR experiments to truncate both TF separately. Unexpectedly, the truncation of both TF produces the downregulation of AKR1B10. This means that both proteins are required to activate the expression of AKR1B10 and AKR1B15. Intriguingly, AKR1B1 did not display changes in gene expression, although the two TFs bind to their promoter region.

Gathering all the data we could confirm that AKR1B1 CpGi works as an enhancer of AKR1B10 and AKR1B15. Furthermore, the activator role of the enhancer is due to both NF-Y and C/EBP-β binding.

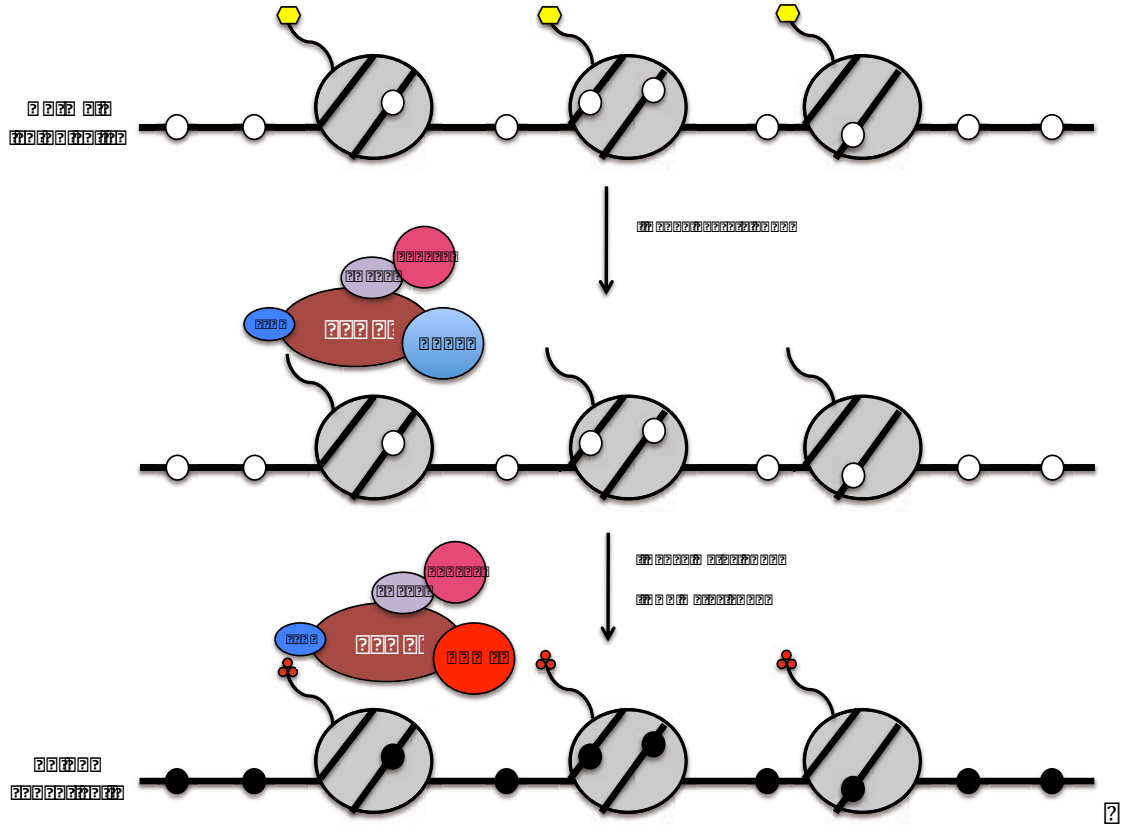
These results are in agreement with a recent study showing that in cancer cell lines, a large number of genes are regulated by DNA methylation changes in enhancer regions rather than in its promoter (Aran *et al.*, 2013).

H3K27ac and PRC2 dynamics in the activation/silencing of AKR1B10 and AKR1B15

Back to 2012, Dr. Batool Akhtar-Zaidi *et al.* analyzed the enhancer histone marks H3K4me1 and H3K27ac to determine gain and loss of enhancer activity at genome-wide level in primary colon cancer cell lines relative to normal colon crypts. They identified thousands of variant enhancer loci (VEL) that comprise a signature that is robustly predictive of the *in vivo* colon cancer transcriptome.

□

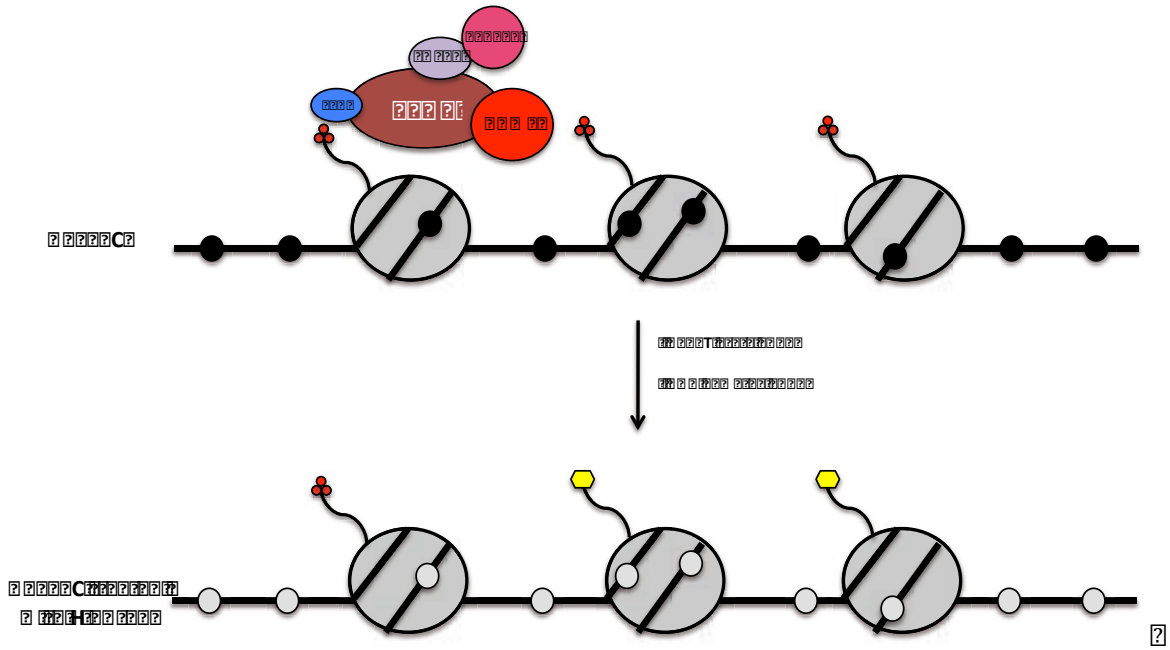
□i □□□i □□jh□9□□ □oyhsh□1□□ei □□□ i □ □□T□□□esA□□h □Tu□□B□0. □□□>□T□yDi e1 □□□□□□□h□
0) <) 8t□1□A□□o□□□ei □e□□□□1Ne□&) 6 □1□□D□s □□□□□□□h□□□9NsT□□□B□0. □□□s □□1o1T□□e□o□
□□ □□Tw□□ □ei □□1ei □T□hse□h□) 6 □1□h□1he□□□□h□□T□□□ Ts□i □□□□1T□□B□0. □□□s □ 1Th □o□
□Ty5eh□□ □□hi 1m□AsTeN□ooy□ 1□□□e□□e□□o□□o□A□dh□1□□□B□0. □□□s □□□□vw□□Nei 1Th□
h5□□s□s□□ooy□□□e□□e□□□03) &□D□s □□□□□□h□□1n n 1 □e1□sA□□1T□n 1T□□□□□□□□oo□s □hl□
□ □□0) &. □o1he□□□□h□□1n n 1 □e1□hst□1T□n 1T□□□□□□s □hw□□


□□□□ □□□
□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□
□□□□□□□□□□□□□□□□□□□

□

□1oy□1n □□ □□5T□hhsA□□ □1n 5o□t□ 0□ >□□□08w□ □□□0□ m1No□□ □eeT□□e□ i she1 □□ □□□□□eyo□h□h□>□□□□8□e1□ei □□T□Ds1 h□mi □e□m1No□□5T1□N□□□ei □□□T□hs D□1□□□B□0 . □□□ □ □□ei □□s □T□□h□□s □□B□0. n □□h □Tu□□y□□□□0□>□□□□□0□hN□N se8w□□i sh□hso□ □s D□ 5T1□□hh□m1No□□s □oN□□□□och1□ei □□□□□□h □ei yoeT□ h□□T□h□h□□ei □e□m1No□□T□hNoe□s □ei □□ o1 D□e□Th □s □□esA□es1 □1□□ei □□□ i □ □□T□>□sDNT□□x. 8w□

□



□□gE□□vvA□□i□e□i□□□□□□□□□□B□B1□□□□□□□□E□□i□E□6a□ □□□□lE□□i □□il□

□ish□y51ei□hsh□sh□N □□T5s □□□□y□1NT□□i □□□T□hNoehw□□ □□□<<3□□1o1T□□e□o□□□ □□T□ □□oo□s □□sh□heT1 Doy□1y5□Tn □ei yo□e□□□s □ei □□□□□<□<□□5□□sho□ □h□□1□h□1e□1□A□□□ □ □B□0. □□□5□□u□□ □□□sh5o□yh□□B□0. n□B□□ □□□□□0□s □ei □□□□□<□<□□5□□sho□ □□ >□sDNT□□3 2 8w□□1m□A□Th□mi □ □m□□eT□□e□ei □□□□ooh□msei □2j□l □□□h□m□□1□h□TA□□□T□j □t 5T□hhs1 □1□□□□□<□<) □□ □□□□□<□<2h□mis□i □sh□□N□□e1□ei □□T□□1 heseNes1 □1□□ □B□0. □□□5□□u□□y□ei □□T□n 1A□o□1□□□□□0□s □ei □□T□Ds1 □□ □□ei □□1□h□TA□□□ □□n □ei yo□es1 □>□sDNT□□x 8w□

□A□ □n 1T□h□mi □ □□□□<<3□□□ooh□m□T□□eT□□e□□□msei □□□□□5□>□□□0□s i s□se1T8h□m□□ 1□h□TA□□□ei □□T□j□t 5T□hhs1 □1□□□□□<□<) h□□Ne□ 1□□i □ D□□s □□□□<□<□□t 5T□hhs1 □ >□sDNT□□ 2 8w□

□□ei □Ts D□□oo□ei □□□5sD□ □es□□□ □□□t 5T□hhs1 □□□e□□1□□□□<□<) □□ □□□□<□<2h□m□□ □1 □sTn □ei □e□ei □□□5sD□ □es□□he□e□h□□h5□□s□oy□ei □□□N□osey□□B□0. □□(□□□0h□1□□ □□□<□<□□5□□sho□ □□sh□u□y□s □□□□s s D□ei □□□□esAsey□he□e□□1□□ei □□□ i □ □□T\w□□ □ei □□

contrary, the epigenetic state of AKR1B10 and AKR1B15 promoters does not seem to play an important role in the regulation of the genes.

Gene regulation by enhancer hypermethylation in cancer

The DNA methylation status of the enhancer has important consequences in gene expression. Dr. Dvir Aran *et al.* analyzed DNA methylation and gene expression data in 58 cell types and developed a machine-learning algorithm for methylation-expression relationships in gene promoters and enhancers. They mapped numerous sites at which DNA methylation was associated with expression of distal genes. Strikingly, some of these identified sites were drastically altered in cancers: hypomethylated enhancer sites associated with upregulation of cancer-related genes and hypermethylated sites with downregulation (Aran *et al.*, 2013). Moreover, changes in the DNA methylation state of the enhancer were better predictors of gene expression than changes in the promoter. Additionally, there was much more hypermethylation in enhancers than in promoters .

Methylation of distal regulatory sites is closely related to gene expression levels across the genome. Single enhancers may modulate ranges of cell-specific transcription levels, from constantly open promoters (Aran & Hellman, 2013).

DNA methylation itself does not produce gene silencing; there are two proposed mechanisms by which DNA methylation hampers gene expression: by avoiding the binding of TFs to its targets or by recruiting DNA methylation binding domain proteins (MBD) that in general do not activate gene expression. Our observations showing an increase of C/EBP-β binding in the AKR1B1 CpG island once demethylated by treatment with 5-AzadC are concordant with Aran *et al.* data showing a depletion of TF in methylated enhancers.

We also tested the importance of DNA methylation in the activity of the AKR1B10 enhancer. Using luciferase enhancer assay, we transfected the CpG island methylated and unmethylated, the result was that when the CpGi was unmethylated has an increased enhancer activity compared with the methylated construct.

Our results are explained by two distinct mechanisms, the first one is based on H3K27ac. This mark is the responsible for the activation of the enhancer and is

fundamental for AKR1B10 and AKR1B15 expression. In those cases where there is no H3K27ac, the enhancer activity is completely lost. Secondly, the enhancer needs to attach some TF to be active. According to our results, NF-Y and C/EBP-β are the most important TFs, but as mentioned before, there are other regions and perhaps other TF that also contribute to express AKR1B10 and AKR1B15. Probably, DNA methylation plays a role in silencing the enhancer hampering the binding of TFs. Although we just describe H3K27ac and DNA methylation mechanisms separately, they are in fact steps of the same silencing process and it is difficult to elucidate the specific contribution of each one, as we always found H3K27ac loss and DNA hypermethylation of the enhancer to occur concomitantly in colorectal cancer.

FUNCTION OF AKR1B10 IN CANCER, LOCATION, LOCATION AND LOCATION

The literature is plenty of studies reporting deregulation of AKR1B genes in cancer. According to previous works, AKR1B10 is most prominently upregulated in liver and lung cancer among others (Fukumoto *et al.*, 2005; Heringlake *et al.*, 2010; Kang *et al.*, 2011; Schmitz *et al.*, 2011; Woenckhaus *et al.*, 2006). AKR1B1 over-expression is more common amongst different tumor types than that of AKR1B10 (Laffin & Petrash, 2012). Alternatively, few studies have reported AKR1B1 and AKR1B10 under-expression in human cancers. Regarding AKR1B15, there is not much literature of this gene and nothing about its alterations in cancer to date.

In many papers, it has been described the advantages of AKR1B10 upregulation (Figure 23). However, AKR1B10 is deeply downregulated in colorectal cancer. Then, what makes colorectal cancer cells downregulate AKR1B10 while other cancers pursuit its upregulation?

One possible explanation would be the one recently proposed by Ohashi *et al.*; they identified AKR1B10 as a direct transcriptional target of the tumor suppressor TF p53. Ohashi *et al* treated HCT116, LoVo and RKO colon cancer cell lines with adriamycin, which activates endogenous p53 protein by inducing DNA damage. As a response to the treatment, they observed an increase in the expression of AKR1B10 protein together with p53 in a dose-dependent manner (Ohashi *et al.*, 2013). Authors claim that p53 activates AKR1B10 by binding to a p53 response element (p53 RE) near the promoter (it is worth to mention that there are also p53 RE in the AKR1B1 CpG island). They proposed that the inactivating mutations of p53 in colorectal cancer could be the responsible for the AKR1B10 downregulation. However, the mechanisms must be more complex, because AKR1B10 is downregulated in most of the colorectal cancers cases, whereas the rate of p53 mutation in colorectal cancer is about 40% (Petitjean *et al.*, 2007). We have also analyzed the relationship between the downregulation of AKR1B10 together with the mutation of p53 using tissue microarrays approach, and we could not observe any significant relationship (Annex XI). Moreover, AKR1B10 was not expressed when they used p53-induced

transactivation directly (Ohashi *et al.*, 2013). Moreover, what it is should be noted is that adriamycin is a substrate of AKR1B10 (Loeffler-Ragg *et al.*, 2009). For this reason, we think that adriamycin triggers AKR1B10 expression as a response of a cellular mechanism that pursuits the inactivation of the harmful agent. Activation of AKR1B10 upon adriamycin treatment needs to be very quick in order to protect the cell against DNA damage. We have seen that adriamycin activates more than 100 fold the expression of AKR1B10 in 24 hours and it is restored to normal levels in about 8 days (data not shown). This prompt activation may differ from the mechanism that regulates the expression of AKR1B10 in normal colon and does not explain the silencing produced in cancer; further experiments are needed to better understand how adriamycin induces the expression of AKR1B10.

Consequences of AKR1B10 and AKR1B15 loss of function in colorectal cancer

In contrast to the ubiquitously expressed AKR1B1, AKR1B10 is mainly expressed in normal small intestine and colon, which suggests that the physiological function of AKR1B10 is specifically required in intestinal tissues (Cao, 1998). Interestingly, AKR1B10 deregulation in cancer shows an inverted pattern to its condition in normal tissue: silencing in colorectal cancer (where the normal tissue expresses it), and upregulation in but cancers where AKR1B10 is lowly or not expressed in the normal tissue (i.e. pancreas) (Chung *et al.*, 2012; Ma *et al.*, 2012).

Actually, AKR1B10 is more than just another tissue-specific gene. Related to this, Dr. Batlle's group described, in mouse and human, the genes responsible for the differentiation of intestinal stem cells (ISC) to well-differentiated colonocytes (Merlos-Suárez *et al.*, 2011). ISCs are selected by high expression of two genes: the ISC marker EphB2 (Merlos-Suárez *et al.*, 2011) and the classical ISC marker Lgr5 (Barker *et al.*, 2007). Both genes become gradually silenced when ISC differentiate. Using the gradient expression of both genes, they identified the panel of genes activated specifically at each point of differentiation: ISC, transient amplifying cells, late terminal amplifying cells and terminal differentiation. Terminal amplifying cells (TA) result from the expansion of ISC (through several

rounds of mitosis) while they migrate upwards along the crypt axis. Close to the intestinal lumen, TA cells undergo cell-cycle arrest and become terminal differentiated cells. Among all these colon differentiation steps, AKR1B10 appeared in the panel of genes defining the signature for the late-TA step (supplementary table 5 in Merlos-Suárez *et al.*, 2011 paper).

However, due to the impact of AKR1B10 silencing in colorectal cancer, our hypothesis is that this gene is acting as a tumor suppressor gene, hampering the development of the tumor. A recent paper underpinned this hypothesis; Dr. Yao *et al.* described that AKR1B10 is downregulated in gastric cancer and observed that the expression of AKR1B10 was inversely correlated with tumor size, lymph node metastasis, distant metastasis and TNM stage. They concluded that the presence of AKR1B10 in the tumor is a good prognostic indicator in gastric cancer patients (Yao *et al.*, 2013). AKR1B10 is expressed in gastric, colon and rectum normal tissues and in the cancers of these tissues is frequently downregulated. At this point, a question arises: what makes gastric, colon and rectum tissues different from others like pancreas, liver, lung, blood or breast, where this gene is upregulated in cancer?

AKR1B10 is a metabolic enzyme that catalyzes the reduction of a wide range of substrates; the cellular implications of expressing AKR1B10 could vary depending on the available substrate. In other words, each tissue has different accessibility to different compounds, and the reduction of these distinct compounds could be beneficial or detrimental for the cancer cell growth. This may be why AKR1B10 is upregulated in some cancer cells and downregulated in others.

At that moment, we realized that those cancers where AKR1B10 is downregulated (gastric, colon and rectum) are tissues that are in direct contact with the diet. Among all nutrients in the diet, we can find vitamin A ester form, also known as retinal. This compound is efficiently reduced by AKR1B10 to retinol (but not by AKR1B1), which is described as an inhibitor of the cellular growth in colorectal cancer cell lines, through an independent retinoic acid pathway (Dillard & Lane, 2007; Park *et al.*, 2005; Park *et al.*, 2008). Almost all the retinal coming from the diet is converted to retinol in the lumen of the gastrointestinal tissue, specially in the intestines (Fidge *et al.*, 1968). Retinol is

then specifically conjugated to different proteins, distributed across the body and stored in the liver. To summarize, gastrointestinal tissues are distinctively exposed to high amounts of free retinal and retinol.

Our hypothesis is that gastrointestinal cancers specifically silence the activity of AKR1B10 and AKR1B15 to avoid the reduction of the dietary retinal into retinol, which inhibits the growth of the tumor. The inhibition of AKR1B10 is an early alteration already present in adenoma, which would allow uncontrolled growth of tumor cells. In the line of this hypothesis, normal ISC do not express AKR1B10 until the last state of differentiation, when cells do not need to divide anymore. AKR1B10 (and perhaps AKR1B15) is expressed to perform its own functions but also to maintain the non-dividing state of differentiated cells.

Actually, this hypothesis is underpinned by our retinal treatment results. When we performed proliferation assays with colon cancer cell lines transfected with AKR1B10 or AKR1B15 and control, we observed a slight decrease in the proliferation of AKR1B10 and AKR1B15 transfected cells. However, when we do the same experiment adding retinal to the medium, cells that are transfected with AKR1B10 or AKR1B15 are deeply affected and are less proliferative compared with the controls cells. This effect is due to the increased amount of retinol that was produced by retinal treatment in the AKR1B10 and AKR1B15 transfected cells. The increased amount of retinol produced by AKR1B10 and AKR1B15 transfected cells was confirmed by studies performed in Jaume Farrés lab (UAB) (data not shown). These experiments pointed out the crucial antiproliferative role of AKR1B10 and AKR1B15 *per se* but also, and more important, in combination with retinal.

Non-gastrointestinal tissues, which are not exposed to free retinal, tend to overexpress AKR1B10 to acquire the pro-tumor activities described in the literature, without the risk of growth inhibition by retinol production.

DEREGULATION OF THE RETINOIC ACID PATHWAY IN COLORECTAL CANCER

As we have discussed above, gene silencing is a well-known cellular process that commonly occurs in cancer (Khare & Verma, 2012; Nephew & Huang, 2003). The literature is plenty of examples of specific genes silenced in a given cancer type; however, in most of the studies the gene is analyzed overlooking its role in the pathway. Understanding the impact of one single gene alteration in the whole pathway is extremely important, because other genes could easily overcome the function of one silenced gene within the whole pathway.

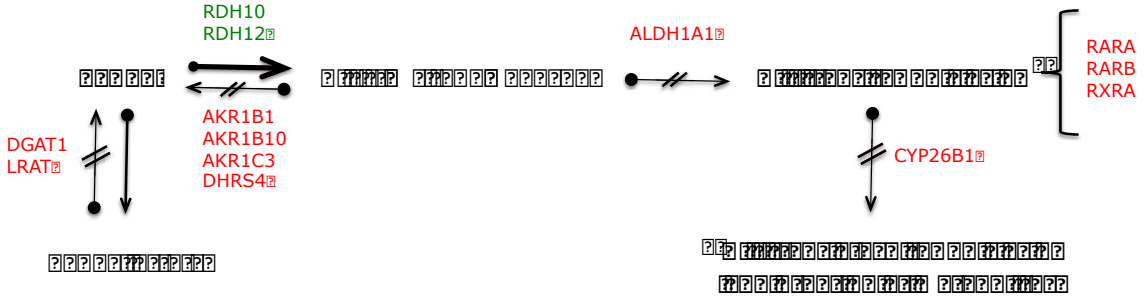
In the present study, we have screened for aberrant gene expression in genes directly involved in the retinoic acid pathway in normal colon, and we have focused our attention on the production of the key metabolites responsible for the adequate function of the pathway (Retinol, All-trans RA and All-trans 4 oxo-RA).

Previous investigations have reported epigenetic silencing of several genes of the retinoic acid pathway. The most well studied is the promoter hypermethylation of RARβ gene, which affects several cancer types (Hayashi *et al.*, 2001). Many other members of the pathway have also been described to undergo silencing through DNA hypermethylation of the promoter (Chim *et al.*, 2005; Chou *et al.*, 2012; Okudela *et al.*, 2013; Tsunoda *et al.*, 2009; Wang *et al.*, 2003). Despite these works, there is not a comprehensive study of the retinoic acid pathway deregulation in cancer. Here we report a deep screening of the genetic and epigenetic alterations of the retinoic acid pathway genes in human colorectal cancer.

We show that 13 genes of the RA pathway were downregulated in colorectal cancer (AKR1B1, AKR1B10, AKR1C3, ALDH1A1, CYP26B1, DGAT1, DHRS4, LRAT, RARA, RARB, RBP1, RETSAT and RXRA) and among them, nine displayed promoter DNA hypermethylation (AKR1B1, AKR1C3, ALDH1A1, CYP26B1, DGAT1, LRAT, RARA, RARB and RBP1) during the cancerous process. We have checked for the importance of promoter DNA methylation and expression on RA pathway genes and, in most cases, the results showed a strict correlation between methylation levels and changes in expression. Taken together, these

□

□□e□□□□n 1 heT□e□□ei □e□□□□□n □ei yo□es1 □shn5o□ys D□□□n □c1T□T1o□□s □ei □□hso□ □s D□1□□ □□□5□ei m□y□D□ □hv□□ 1T□1A□Th□D□ □es□□n Ne□es1 h□□ □□□15y□Nn □□T□□oe□T□es1 h□□1□ 1e□l□A□□h5□□s□o□1□o□A□ □□□1 □ei □□hso□ □□□□D□ □hd□1 □ei □□□1 eT□Tyh□N5T□DNo□es1 □1□□ □□□<) □s □□1o1T□□e□o□□□ □□T□shn□ □s oy□□N□□e1□D□ □□□n 5os□s□□es1 v□□y□1□5T□h□ es D□ □oe□T□□□D□ □hn□s □ei □□□mi 1o□□□□5□ei m□yh□n□□□1No□□D□e□□ □s□□□□□□1Ne□ei □□5□ei h□ei □e□ □T□□□o1□u□□□s □□1o1T□□e□o□□□ □□T□>□sDNT□□xM8v□



□□gE□□vmA□□i□E□i□□p□□E□p□□i□□□E□i□□□□□□□□□□□□i□n □c□□□□□□E□□i□□lg□ □E□□□□□□E□□□□□E□□i□□p□□□□□p□ i□□i□□p□□□cg□E□□g□□i□□□u□□□□□□E□□□□i□□p□□□n □E□□g□□i□□□l□□iEEg□i□□□EE□n p□□□□□□i□E□□□i□□p□ i□□i□□E□□□□□□□□□□□□□□□□□□EE□n p□i□□p□E□□□i□□p□□□□□□□□□□□cn□□□□i□E□□i□□pl□

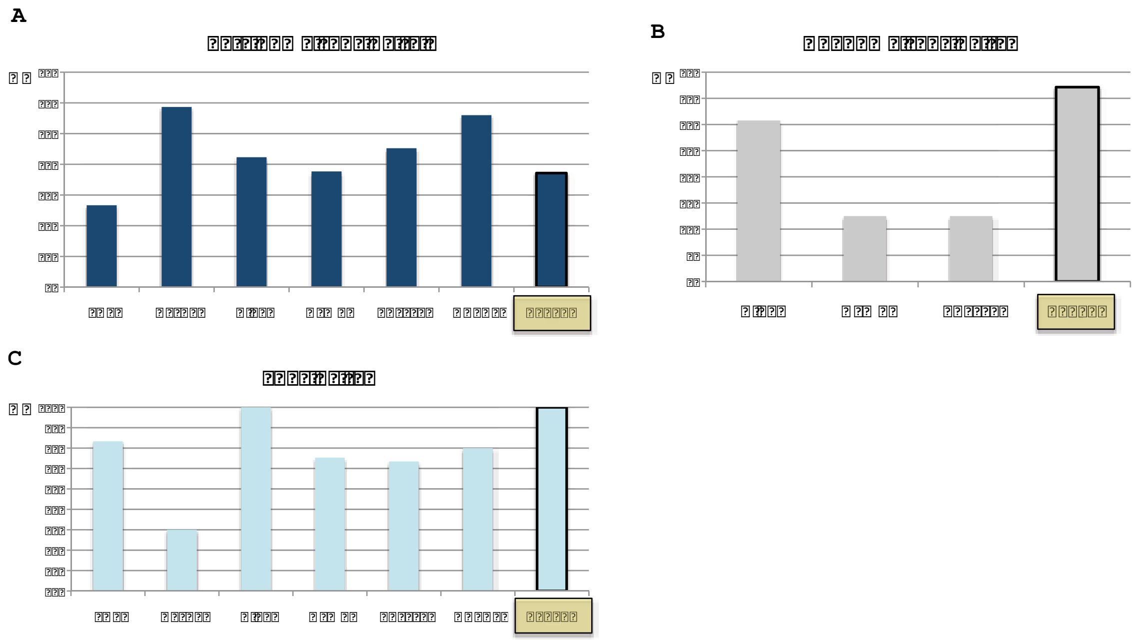
□e□□Sl□he□Do□ □□□b□m□□□1No□□1□h□TA□□ei □e□T□es 1s□□□s□□5□ei m□y□sh□□o□□Toy□N □□o□ □□□□ s □□1o1T□□e□o□□□ □□TiW□i □□□□esA□es1 □1□□ei □□□□ 1 s□□o□T□es 1s□□□s□□5□ei m□y□sh□ □o1□u□□□□e□em1□o□A□dw□□sIhe□□ei □□5T1□N□es1 □1□□ei □□□s1□□esA□□n □e□□1ose□□>□oqjeT□ h□ □□□□ □□ei □sT□51o□T□n □e□□1ose□h8□sh□i □n 5□T□□□□y□ei □□□1m T□DNo□es1 □1□□ei □□ □ lyn □h□□□□<□<□□ □□□□03□<w□□□1 □h□ei □T□□shn□□□o□□T□s i s□ses1 □1□□T□es 1s□□ □□s□□1□□□5e1Th,□□□□□h□□□□□ □□□□□w□

□1n □ei s D□hsn so□T□1□□NTh□e1□ei □□5 1 1Toy□N □□Ihe11□□T□es 1s□□□□s□□T□□□5e1T□ s □□5□ □□ e□n □□i □ shn □□□h□□□1 □T□es 1o□>□soo□T□□r □□□ □h□0)). d□□□Tu□□□□□□h□ 0)) 2h□0)) x8w□□i □□T□□□es1 h□T□h51 hs□o□□□1T□T□es 1o□5T1□N□es1 h□□sei □T□□y□ 1ts□sl s D□T□es yo□□he□Th□1T□T□□N□s D□□ojeT□ h□T□es □o□□i y□□h□□T□□□o1□u□□□s □ □1o1T□□e□o□□□ □□Tw□ 1T□1A□Th□T□es 1o□sh□1□5s□oy□5Nn 5□□□e1□□ojeT□ h□T□es □o□□i y□□□ >□1n n 1 oy□T□□□TT□□□□h□T□es □o8□□y□ei □□1A□T□t 5T□hh□□□□ lyn □h□□□□<) □>□y□D□ □□ □n 5os□s□□es1 8□□ □□□□□<0w□□e□sh□m1Tei □e1□ 1e□□ei □e□T□es 1d□T□es yo□□he□Th□□ □□□ojeT□ h□T□es □o□□i y□□□□ □□□□□sT□□eoy□1□e□s □□□□y□ei □□□s□e□>□i □1□1hs1N□□□□□□h□ 0)<) 8t□mi s□i □h □□ h□ei □e□□1o1T□□e□o□eNn 1Th□□T□□□t 51h□□□e1□ei □n h□□Ne□Nhe□T□es 1o□ sh□□□n 1 heT□e□□□e1□s i s□se□ei □□DT1mei □1□□ei □□eNn 1Tw□1T□ei sh□T□□h1 □ei □5□ei m□y□ sh□N □□o□ □□□□□ □□□1ei b□T□es yo□□he□Th□□ □□□ojeT□ h□T□es □o□□i y□□□□1n 51N □h□ □□ 1e□5T1□N□□□T□es 1d□□□□□Nh□□1□□ei □□s i s□ses1 □1□□ei □□T□h51 hs□o□□□ lyn □h□ >□□□□<□□ □□□□□□□□1T□T□es yo□□he□Th□□ □□□□<□<b□□□□<□<) b□□□□<□B□□ □□□□□□&□

for all-trans retinaldehyde). Even more, if some retinol is produced or absorbed by the tumor through the diet, it would rapidly become All-trans retinaldehyde because of the upregulation of the responsible enzymes (RDH10 and RDH12).

In summary, we demonstrate the inhibition of the biosynthesis of active metabolites of the retinoic acid pathway in colorectal cancer, which results in the impairment of canonical and non-canonical anti-tumorigenic effects of these signals.

Moreover, at the genome scale, other genes are also co-methylated with genes of the retinol acid pathway. A deeper study of these events revealed that processes such as embryo development, neuronal differentiation or cell migration are “co-silenced” with retinoic acid pathway genes. Interestingly these processes are known to be driven by retinoic acid. These results indicate that the production of bioactive metabolites from the retinoic acid pathway is prevented in colorectal cancer cells.


A
B
C

AKR1B10 protein downregulation in colorectal cancer

Tissue microarray experiments revealed loss of AKR1B10 protein (downregulation or complete absence) in the cancer tissue of 96.4% of the tumors analyzed (162 out of 168). Due to its almost universal occurrence, it is unlikely that the identification of this alteration could have prognostic utility. Nevertheless, it remains to be investigated whether this analysis could be applied to distinguish between a benign lesion and cancer in cases where there is a lesion in the epithelium but it is difficult to determine its malignancy by other methods. Further experiments are required until this application could be used, and different types of lesions including adenomas and bowel disease samples are required to have a complete profile of AKR1B10 in the different possible scenarios involving the intestinal tissue.

RETINOIC ACID PATHWAY GENES AS PROGNOSTIC MARKERS

The discovery of new biomarkers is crucial for the future of medicine. Biomarkers could be related with the tumor biology and be essential for the election of the treatment. Biomarkers could also predict the outcome of the disease and help to improve the management of the disease. In this work we described some gene expression alterations of retinoic acid pathway genes that could also provide several confident prognosis markers. The alteration (downregulation or upregulation) of the normal gene expression of ALDH1A2, CYP26B1, DGAT1, DHRS4, LRAT, RARA, RDH10 and RXRA, in colorectal cancer is significantly related with worse overall survival. These results suggest that alterations in RA pathway genes expression could be used for prognostic prediction in colorectal cancers.

Regarding AKR1B1 protein expression, we obtain some interesting results with the tissue microarrays analysis, the few cases that have no expression of AKR1B1 display a worse overall and disease free survival than the cases where it is not altered. This result is very motivating but we need to validate it with an independent cohort. In summary, we have shown that transcriptional disruption of the retinoic acid pathway (partly due to epigenetic alterations) is an important feature of most colorectal cancers, and its analysis might have important applications in the clinical management of patients.

CONCLUSIONS

1. AKR1B1 CpG island is deeply and frequently hypermethylated in colorectal cancer, but with minimal or no effect on the expression of the associated gene.
2. Hypermethylation of AKR1B1 CpG island is strongly associated with downregulation of AKR1B10 and AKR1B15, which promoter region is 60Kb upstream of the CpG island.
3. AKR1B1 CpG island presents enhancer properties in regard to AKR1B10 and AKR1B15 genes, which explains the functional association.
4. AKR1B10 and AKR1B15 are tumor suppressor genes in colorectal cancer, especially in the presence of retinal, a form of vitamin A and precursor of retinoic acid.
5. *Akr1b* enzymes are also dowregulated in a murine colorectal cancer, suggesting that the impairment of its function in colorectal cancer is not an epiphenomenon.
6. The retinoic acid pathway is deregulated in most colorectal cancers by epigenetic mechanisms resulting in an overall decrease of the antitumor compounds of the retinoic acid pathway.
7. Detection of AKR1B1 CpG island hypermethylation (resulting in downregulation of AKR1B10) in stools DNA is a non-invasive, sensitive and specific marker of colorectal adenomas and carcinomas, rendering it as a proper candidate for colorectal cancer screenings.
8. Low expression levels of AKR1B1, AKR1B10, ALDH1A2, CYP26B1, DGAT1, DHRS4, LRAT, RARA and RXRA; and high expression levels of RDH10 are associated with tumor aggressiveness, which might be useful as a prognostic factor.

BIBLIOGRAPHY

Ahlquist, D. A., Skoletsky, J. E., Boynton, K. A., Harrington, J. J., Mahoney, D. W., Pierceall, W. E., … Shuber, A. P. (2000). Colorectal cancer screening by detection of altered human DNA in stool: Feasibility of a multitarget assay panel. *Gastroenterology,(119) 1219-1227*.

Akhtar-Zaidi, B., Cowper-Sal-lari, R., Corradin, O., Saiakhova, A., Bartels, C. F., Balasubramanian, D., … Scacheri, P. C. (2012). Epigenomic enhancer profiling defines a signature of colon cancer. *Science*, (*336*) 736–739.

Akulenko, R., & Helms, V. (2013). DNA co-methylation analysis suggests novel functional associations between gene pairs in breast cancer samples. *Human Molecular Genetics*, (*22)* 3016–3022.

Albanes, D., Heinonen, O. P., Taylor, P. R., Virtamo, J., Edwards, B. K., Rautalahti, M., … Huttunen, J. K. (1996). Alpha-Tocopherol and beta-carotene supplements and lung cancer incidence in the alpha-tocopherol, beta-carotene cancer prevention study: effects of base-line characteristics and study compliance. *Journal of the National Cancer Institute*, (88) 1560–1570.

Aldini, G., Dalle-Donne, I., Facino, R. M., Milzani, A., & Carini, M. (2007). Intervention strategies to inhibit protein carbonylation by lipoxidation-derived reactive carbonyls. *Medicinal Research Reviews*, (*27*) 817–868.

Ando, Y., Saka, H., Ando, M., Sawa, T., Muro, K., Ueoka, H., … Hasegawa, Y. (2000). Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. *Cancer Research*, (*60*) 6921–6926.

Antequera, F., & Bird, a. (1993). Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America*, (90), 11995–11999.

Aran, D., & Hellman, A. (2013). DNA methylation of transcriptional enhancers and cancer predisposition. *Cell*, (154) 11-13.

Aran, D., Sabato, S., & Hellman, A. (2013). DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biology*, (*14*), R21.

Azuara, D., Rodriguez-Moranta, F., de Oca, J., Soriano-Izquierdo, A., Mora, J., Guardiola, J., … Capellá, G. (2010). Novel methylation panel for the early detection of colorectal tumors in stool DNA. *Clinical Colorectal Cancer*, (9), 168–76.

Baba, S. P., Barski, O. A., Ahmed, Y., O'Toole, T. E., Conklin, D. J., Bhatnagar, A., & Srivastava, S. (2009). Reductive Metabolism of AGE Precursors: A Metabolic Route for Preventing AGE Accumulation in Cardiovascular Tissue. *Diabetes*, (58).

Bailey, C. E., Eschrich, S., Kis, C., Levy, S., & Kay, M. (2012). Predicts Recurrence and Death in Patients With Colon Cancer, (138), 958–968.

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, (21) 381–95.

Barker, N., van Es, J. H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., … Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, (449) 1003–1007.

Baron, J. M., Heise, R., Blaner, W. S., Neis, M., Joussen, S., Dreuw, A., … Jugert, F. K. (2005). Retinoic acid and its 4-oxo metabolites are functionally active in

human skin cells in vitro. *The Journal of Investigative Dermatology*, (125), 143–53.

Barski, O. A., Tipparaju, S. M., & Bhatnagar, A. (2008). The Aldo-Keto Reductase Superfamily and its Role in Drug Metabolism and Detoxification. *Drug Metabolism Reviews*, (40) 553–624.

Bastide, N. M., Pierre, F. H. F., & Corpet, D. E. (2011). Heme iron from meat and risk of colorectal cancer: a meta-analysis and a review of the mechanisms involved. *Cancer Prevention Research.* (4) 177–184.

Bavinck, J. N., Tieben, L. M., Van der Woude, F. J., Tegzess, A. M., Hermans, J., ter Schegget, J., & Vermeer, B. J. (1995). Prevention of skin cancer and reduction of keratotic skin lesions during acitretin therapy in renal transplant recipients: a double-blind, placebo-controlled study. *Journal of Clinical Oncology*, (13) 1933–1938.

Baxter, N., & Goldwasser, M. (2009). Association of Colonoscopy and Death From Colorectal Cancer. *Annals of Internal medicine,(150)* 1–9.

Bell, O., Tiwari, V. K., Thomä, N. H., & Schübeler, D. (2011). Determinants and dynamics of genome accessibility. *Nature Reviews. Genetics*, (12) 554–564.

Benelli, R., Monteghirfo, S., Venè, R., Tosetti, F., & Ferrari, N. (2010). The chemopreventive retinoid 4HPR impairs prostate cancer cell migration and invasion by interfering with FAK/AKT/GSK3beta pathway and beta-catenin stability. *Molecular Cancer*, (*9)* 142.

Berg, V., & Kinzler, K. W. (1997). Gatekeepers and caretakers. *Nature, (386) 761-763*.

Bergman, Y., & Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nature Structural & Molecular Biology*, (*20*) 274–281.

Bettington, M., Walker, N., Clouston, A., Brown, I., Leggett, B., & Whitehall, V. (2013). The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology*, (*62*), 367–386.

Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E. W., Wu, B., … Fan, J.-B. (2006). High-throughput DNA methylation profiling using universal bead arrays. *Genome Research*, (*16*), 383–393.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, (16) 6–21.

Bollag, W. (1972). Prophylaxis of chemically induced benign and malignant epithelial tumors by vitamin A acid (retinoic acid). *European Journal of Cancer*, (8) 689–693.

Bosch, L. J. W., Mongera, S., Terhaar Sive Droste, J. S., Oort, F. a, van Turenhout, S. T., Penning, M. T., … Meijer, G. a. (2012). Analytical sensitivity and stability of DNA methylation testing in stool samples for colorectal cancer detection. *Cellular Oncology (Dordrecht)*, (35) 309–315.

Boutanaev, A. M., Kalmykova, A. I., Shevelyov, Y. Y., & Nurminsky, D. I. (2002). Large clusters of co-expressed genes in the Drosophila genome. *Nature*, (420) 666–669.

Boyle, P., & Langman, J. (2000). ABC of colorectal cancer Epidemiology. *BMJ: British Medical Journal*, (321) 805–808.

Boyle, P., & Leon, M. E. (2002). Epidemiology of colorectal cancer. *British Medical Bulletin*, (64) 1–25.

Bozic, I., Antal, T., Ohtsuki, H., Carter, H., Kim, D., Chen, S., … Nowak, M. a. (2010). Accumulation of driver and passenger mutations during tumor progression.

Proceedings of the National Academy of Sciences of the United States of America, (107), 18545–50.

Bray, G. A. (2002). The Underlying Basis for Obesity: Relationship to Cancer. *The Journal of Nutrition*, (132) 3451–3455.

Brock, H. W., & van Lohuizen, M. (2001). The Polycomb group — no longer an exclusive club? *Current Opinion in Genetics & Development*, *(11*) 175–181.

Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafrenière, R. G., Xing, Y., Lawrence, J., & Willard, H. F. (1992). The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell (71) 527-542*.

Brown, J. C., Winters-Stone, K., Lee, A., & Schmitz, K. H. (2012). Cancer, Physical Activity, and Exercise. *Comprehensive Physiology, (4) 2775-27809.*

Bushue, N., & Wan, Y.-J. Y. (2010). Retinoid pathway and cancer therapeutics. *Advanced Drug Delivery Reviews*, (62) 1285–1298.

Campanero, M. R., Armstrong, M. I., & Flemington, E. K. (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proceedings of the National Academy of Sciences* , (97) 6481–6486.

Campos, E. I., & Reinberg, D. (2009). Histones: annotating chromatin. *Annual Review of Genetics*, *(43*) 559–99.

Cao, D. (1998). Identification and Characterization of a Novel Human Aldose Reductase-like Gene. *Journal of Biological Chemistry*, (273) 11429–11435.

Carmona, F. J., Azuara, D., Berenguer-Llergo, A., Fernández, A. F., Biondo, S., de Oca, J., … Moreno, V. (2013). DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer. *Cancer Prevention Research*, (6) 656–665.

Carter, D., Chakalova, L., Osborne, C. S., Dai, Y., & Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nature Genetics*, (32) 623–626.

Casey, P. J. (1994). Lipid modifications of G proteins. *Current Opinion in Cell Biology*, (6) 219–225.

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews. Genetics*, (10) 295–304.

Chan, A. T., & Giovannucci, E. L. (2010). Primary prevention of colorectal cancer. *Gastroenterology*, (136) 2029–2043

Chen, J., Guo, L., Zhang, L., Wu, H., Yang, J., Liu, H., … Pei, D. (2013). Vitamin C modulates TET1 function during somatic cell reprogramming. *Nature Genetics*, (45) 1504–1509.

Chia, W. K., Ali, R., & Toh, H. C. (2012). Aspirin as adjuvant therapy for colorectal cancer--reinterpreting paradigms. *Nature Reviews. Clinical Oncology*, (9) 561–570.

Chim, C.-S., Wong, S.-Y., Pang, A., Chu, P., Lau, J. S., Wong, K.-F., & Kwong, Y.-L. (2005). Aberrant promoter methylation of the retinoic acid receptor alpha gene in acute promyelocytic leukemia. *Leukemia*, (19) 2241–2246.

Chomczynski, P., & Sacchi, N. (2006). The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature Protocols*, (1) 581–585.

Chou, A. P., Chowdhury, R., Li, S., Chen, W., Kim, A. J., Piccioni, D. E., … Lai, A. (2012). Identification of retinol binding protein 1 promoter hypermethylation in isocitrate dehydrogenase 1 and 2 mutant gliomas. *Journal of the National Cancer Institute*, (104) 1458–1469.

Chuang, J. C., Warner, S. L., Vollmer, D., Vankayalapati, H., Redkar, S., Bearss, D. J., … Jones, P. a. (2010). S110, a 5-Aza-2'-deoxycytidine-containing dinucleotide, is an effective DNA methylation inhibitor in vivo and can reduce tumor growth. *Molecular Cancer Therapeutics*, (9) 1443–1450.

Chung, Y. T., Matkowskyj, K. a, Li, H., Bai, H., Zhang, W., Tsao, M.-S., … Yang, G.-Y. (2012). Overexpression and oncogenic function of aldo-keto reductase family 1B10 (AKR1B10) in pancreatic carcinoma. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology*, (25) 758–766.

Comb, M., & Goodman, H. M. (1990). CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Research* , (13), 3975–3982.

Cremer, T., & Cremer, M. (2010). Chromosome territories. *Cold Spring Harbor Perspectives in Biology*, (3).

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., … Jaenisch, R. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America*, (107) 21931–21936.

Cunningham, D., Atkin, W., Lenz, H.-J., Lynch, H. T., Minsky, B., Nordlinger, B., & Starling, N. (2010). Colorectal cancer. *The Lancet*, (375) 1030–1047.

Dawson, M. a, & Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell*, (150) 12–27.

De Carvalho, D. D., Sharma, S., You, J. S., Su, S.-F., Taberlay, P. C., Kelly, T. K., … Jones, P. A. (2012). DNA Methylation Screening Identifies Driver Epigenetic Events of Cancer Cell Survival. *Cancer Cell*, (21) 655–667.

Dekker, J. (2006). The three " C " s of chromosome conformation capture : controls , controls , controls, (3) 17–21.

Dekker, J., Marti-Renom, M. a, & Mirny, L. a. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews. Genetics*, (14) 390–403.

Deng, G., Chen, A., Pong, E., & Kim, Y. S. (2001). Methylation in hMLH1 promoter interferes with its binding to transcription factor CBF and inhibits gene expression. *Oncogene*, (48) 7120–7127.

Di Ruscio, A., Ebralidze, A. K., Benoukraf, T., Amabile, G., Goff, L. a, Terragni, J., … Tenen, D. G. (2013). DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature*, (503) 371-376.

Dillard, A., & Lane, M. (2007). Retinol Decreases β-Catenin Protein Levels in Retinoic Acid-Resistant Colon Cancer Cell Lines. *Molecular Carcinogenesis*, (329), 315–329.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., … Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, (485) 1–5.

Djebali, S., Davis, C. a, Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., … Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, (489) 101–108.

Down, T. a, Rakyan, V. K., Turner, D. J., Flicek, P., Li, H., Kulesha, E., … Beck, S. (2008). A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nature Biotechnology*, (26) 779–785.

Eden, A., Gaudet, F., Waghmare, A., & Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*, (300) 405.

Ehrlich, M., Woods, C. B., Yu, M. C., Dubeau, L., Yang, F., Campan, M., … Laird, P. W. (2006). Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. *Oncogene*, (18) 2636–2645.

Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., … Bernstein, B. E. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, (473) 43–49.

Esteller, M. (2011). Non-coding RNAs in human disease. *Nature Reviews. Genetics*, (12), 861–74.

Fairley, T. L., Cardinez, C. J., Martin, J., Alley, L., Friedman, C., Edwards, B., & Jamison, P. (2006). Colorectal cancer in U.S. adults younger than 50 years of age, 1998–2001. *Cancer*, (107) 1153–1161.

Faivre, J., Dancourt, V., Lejeune, C., Tazi, M. a., Lamour, J., Gerard, D., … Bonithon-Kopp, C. (2004). Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study1. *Gastroenterology*, (126) 1674–1680.

Fearon, E. R., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, (61) 759–767.

Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F., … Jenab, M. (2011). Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, *(22*) 1958–1972.

Feinberg, A. P., & Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, *(301*) 89–92.

Ferrari, K. J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., … Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Molecular Cell* (53) 49–62.

Fidge, N. H., Shiratori, T., Ganguly, J., & Goodman, D. S. (1968). Pathways of absorption of retinal and retinoic acid in the rat. *Journal of Lipid Research*, *(9)* 103–109.

Fijten, G. H., Starmans, R., Muris, J. W. M., Schouten, H. J. A., Blijham, G. H., & Knottnerus, J. A. (1995). Predictive value of signs and symptoms for colorectal cancer in patients with rectal bleeding in general practice. *Family Practice*, (12) 279–286.

Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., … Esteller, M. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics*, (37) 391–400.

Fraser, P., & Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature*, (447) 413–417.

Friedl, W., Caspari, R., Sengteller, M., Uhlhaas, S., Lamberti, C., Jungck, M., … Propping, P. (2001). Can APC mutation analysis contribute to therapeutic decisions in familial adenomatous polyposis. Experience from 680 FAP families. *Gut* , (48) 515–521.

Frigola, J., Song, J., Stirzaker, C., Hinshelwood, R. a, Peinado, M. a, & Clark, S. J. (2006). Epigenetic remodeling in colorectal cancer results in coordinate

gene suppression across an entire chromosome band. *Nature Genetics*, (38) 540–539.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., … Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, (89) 1827–1831.

Fu, Z., Shrubsole, M. J., Li, G., Smalley, W. E., Hein, D. W., Cai, Q., … Zheng, W. (2013). Interaction of cigarette smoking and carcinogen-metabolizing polymorphisms in the risk of colorectal polyps. *Carcinogenesis*, (34) 779–786.

Fujii, Y., Watanabe, K., Hayashi, H., Urade, Y., Kuramitsu, S., Kagamiyama, H., & Hayaishi, O. (1990). Purification and characterization of rho-crystallin from Japanese common bullfrog lens. *Journal of Biological Chemistry*, (265) 9914–9923.

Fukumoto, S., Yamauchi, N., Moriguchi, H., Hironaka, M., Ishikawa, Y., Kodama, T., & Nishimura, M. (2005). Overexpression of the Aldo-Keto Reductase Family Protein AKR1B10 Is Highly Correlated with Smokers ' Non − Small Cell Lung Carcinomas Overexpression of the Aldo-Keto Reductase Family Protein AKR1B10 Is Highly Correlated with Smokers ' Non – Small Cell Lung. *Clinical Cancer Research,* (11) 1776–1785.

Gabbay, K. (2004). Aldose reductase inhibition in the treatment of diabetic neuropathy: Where are we in 2004? *Current Diabetes Reports*, (6) 405–408.

Galiatsatos, P., & Foulkes, W. (2006). Familial adenomatous polyposis. *The American Journal of Gastroenterology*, *(18)* 7–30.

Gama-Sosa, M., & Slagel, V. (1983). The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Research*, (19) 6883–6894.

Gao, J., Aksoy, B. a., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., … Schultz, N. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Science Signaling*, (269) pl1.

Garinis, G. a, Patrinos, G. P., Spanakis, N. E., & Menounos, P. G. (2002). DNA hypermethylation: when tumour suppressor genes go silent. *Human Genetics*, (111) 115–127.

Gaspar, C., Cardoso, J., Franken, P., Molenaar, L., Morreau, H., Möslein, G., … Fodde, R. (2008). Cross-species comparison of human and mouse intestinal polyps reveals conserved mechanisms in adenomatous polyposis coli (APC)-driven tumorigenesis. *The American Journal of Pathology*, (172) 1363–1380.

Gaudet, F., Hodgson, J. G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J. W., … Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. *Science*, (300) 489–492.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., … Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, (5) R80.

Gervin, K., Hammero, M., Akselsen, H. E., Moe, R., Nygard, H., Brandt, I., … Lyle, R. (2011). Extensive variation and low heritability of DNA methylation identified in a twin study. *Genome Research*, (11) 1813–1821.

Gibbs, J., Kohl, N., Koblan, K., Omer, C., Sepp-Lorenzino, L., Rosen, N., … Oliff, A. (1996). Farnesyltransferase inhibitors and anti-Ras therapy. *Breast Cancer Research and Treatment*, (38) 75–83.

Giovannucci, E. (2001). An Updated Review of the Epidemiological Evidence that Cigarette Smoking Increases Risk of Colorectal Cancer. *Cancer Epidemiology, Biomarkers & Prevention*, (7) 725–731.

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., … Carell, T. (2010). Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. *PLoS ONE*, (12) e15367.

Grady, W. M., Parkin, R. K., Mitchell, P. S., Lee, J. H., Kim, Y.-H., Tsuchiya, K. D., … Tewari, M. (2008). Epigenetic silencing of the intronic microRNA hsa-miR-342 and its host gene EVL in colorectal cancer. *Oncogene*, (27) 3880–3888.

Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics*, (13) 343-357.

Grölund, M.-M., Lehtonen, O.-P., Eerola, E., & Kero, P. (1999). Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *Journal of Pediatric Gastroenterology and Nutrition*, (28) 19-25.

Grunau, C., Clark, S. J., & Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Research*, (29) E65–5.

Guarner, F., & Malagelada, J.-R. (2003). Gut flora in health and disease. *Lancet*, (361) 512–519.

Guo, J. U., Su, Y., Zhong, C., Ming, G., & Song, H. (2011, April 29). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, (145) 423-434.

Haggar, F., & Boushey, R. (2009). Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clinics in Colon and Rectal Surgery*, (212) 191–197.

Halpern, A. C., Schuchter, L. M., Elder, D. E., Guerry, D., Elenitsas, R., Trock, B., & Matozzo, I. (1994). Effects of topical tretinoin on dysplastic nevi. *Journal of Clinical Oncology* , (5) 1028–1035.

Hanahan, D., & Weinberg, R. a. (2011). Hallmarks of cancer: the next generation. *Cell*, (144) 646–674.

Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell*, (*100*) 57–70.

Hardin, J., & Mydlarski, P. R. (2010). Systemic retinoids: chemoprevention of skin cancer in transplant recipients. *Skin Therapy Letter*, (15) 1–4.

Harisiadis, L., Miller, R. C., Hall, E. J., & Borek, C. (1978). A vitamin A analogue inhibits radiation-induced oncogenic transformation. *Nature*, (274) 486–487.

Harrow, J., Nagy, A., Reymond, A., Alioto, T., Patthy, L., Antonarakis, S. E., & Guigó, R. (2009). Identifying protein-coding genes in genomic sequences. *Genome Biology*, (10) 201.

Hayashi, K., Yokozaki, H., Goodison, S., Oue, N., Suzuki, T., Lotan, R., … Tahara, E. (2001). Inactivation of retinoic acid receptor beta by promoter CpG hypermethylation in gastric cancer. *Differentiation; Research in Biological Diversity*, (68) 13–21.

He, J., & Efron, J. E. (2011). Screening for Colorectal Cancer. *Advances in Surgery*. Mosby Year Book.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., … Xu, G.-L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* , (333) 1303–1307.

Heibein, A. D., Guo, B., Sprowl, J. a, Maclean, D. a, & Parissenti, A. M. (2012). Role of aldo-keto reductases and other doxorubicin pharmacokinetic genes in doxorubicin resistance, DNA binding, and subcellular localization. *BMC Cancer*, (12) 381.

Hellebrekers, D. M. E. I., Lentjes, M. H. F. M., van den Bosch, S. M., Melotte, V., Wouters, K. a D., Daenen, K. L. J., … van Engeland, M. (2009). GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, (15) 3990–3997.

Henikoff, S., & Shilatifard, A. (2011). Histone modification: cause or cog? *Trends in Genetics*, (27) 389–396.

Herbst, A., & Kolligs, F. T. (2012). Detection of DNA hypermethylation in remote media of patients with colorectal cancer: new biomarkers for colorectal carcinoma. *Tumour Biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine*, (33) 297–305.

Heringlake, S., Hofdmann, M., Fiebeler, A., Manns, M. P., Schmiegel, W., & Tannapfel, A. (2010). Identification and expression analysis of the aldo–ketoreductase1-B10 gene in primary malignant liver tumours. *Journal of Hepatology,* (2) 220-227.

Hiraoka, S., Kato, J., Horii, J., Saito, S., Harada, K., Fujita, H., … Yamamoto, K. (2010, January 1). Methylation status of normal background mucosa is correlated with occurrence and development of neoplasia in the distal colon. *Human Pathology*, (1) 38-47.

Hol, L., van Leerdam, M. E., van Ballegooijen, M., van Vuuren, A. J., van Dekken, H., Reijerink, J. C. I. Y., … Kuipers, E. J. (2010). Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. *Gut*, (59) 62–68.

Holliday, R., & Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. *Science,* (187) 226–232.

Hong, L., & Ahuja, N. (2013). DNA methylation biomarkers of stool and blood for early detection of colon cancer. *Genetic Testing and Molecular Biomarkers*, (17) 401–406.

Hong, W. K., Endicott, J., Itri, L. M., Doos, W., Batsakis, J. G., Bell, R., … Strong, S. (1986). 13-cis-Retinoic Acid in the Treatment of Oral Leukoplakia. *New England Journal of Medicine*, (24) 1501–1505.

Hong, W. K. I., Lippman, S. M., Itri, L. M., Karp, D. D., Lee, J. S., Byers, R. M., … Goepfert, H. (1990). Prevention of Second Primary Tumors with Isotretinoin in Squamous-Cell Carcinoma of the Head and Neck. *New England Journal of Medicine*, (323) 795–801.

Hotchkiss, R. (1948). The quantitative separation of purines, pyrimidiines, and nucleosides by paper chromatography. *Journal of Biological Chemistry,* (175) 315-332.

Hu, H., Krasinskas, A., & Willis, J. (2011). Perspectives on current tumor-node-metastasis (TNM) staging of cancers of the colon and rectum. *Seminars in Oncology*, (38) 500–510.

Huang, D. W., Sherman, B. T., & Lempicki, R. a. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, (37) 1–13.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, (4) 44–57.

Huxley, R. R., Woodward, M., & Clifton, P. (2012). The Epidemiologic Evidence and Potential Biological Mechanisms for a Protective Effect of Dietary Fiber on the Risk of Colorectal Cancer. *Current Nutrition Reports*, (2) 63–70.

Hyndman, D. J., & Flynn, T. G. (1998). Sequence and expression levels in human tissues of a new member of the aldo-keto reductase family. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* (2–3), 198–202.

Iida, N., Dzutsev, A., Stewart, C. a., Smith, L., Bouladoux, N., Weingarten, R. a., … Goldszmid, R. S. (2013). Commensal Bacteria Control Cancer Response to Therapy by Modulating the Tumor Microenvironment. *Science*, (342) 967–970.

Imperiale, T. F., Ransohoff, D. F., Itzkowitz, S. H., Turnbull, B. A., & Ross, M. E. (2004). Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *New England Journal of Medicine*, (26) 2704–2714.

Innocenti, F., Undevia, S. D., Ramirez, J., Mani, S., Schilsky, R. L., Vogelzang, N. J., … Ratain, M. J. (2004). A phase I trial of pharmacologic modulation of irinotecan with cyclosporine and phenobarbital. *Clin Pharmacol Ther*, (76) 490–502.

Irizarry, R. A., Ladd-acosta, C., Carvalho, B., Wu, H., Brandenburg, S. A., Jeddeloh, J. A., … Feinberg, A. P. (2008). Comprehensive high-throughput arrays for relative methylation (CHAR). *Genome research,* (18) 780–790.

Issa, J.-P. J., & Ahuja, N. (2000). Aging , methylation and cancer. *Histology and Histopathology*, (15) 835–842.

Issa, J.-P. J., Vertino, P. M., Wu, J., Sazawal, S., Celano, P., Nelkin, B. D., … Baylin, S. B. (1993). Increased Cytosine DNA-Methyltransferase Activity During Colon Cancer Progression. *Journal of the National Cancer Institute*, (15) 1235–1240.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., … Zhang, Y. (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* , (333) 1300–1303.

Iyer, L. M., Tahiliani, M., Rao, A., & Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle*, (11) 1698–1710.

Jin, S.-G., Cai Guo,& Pfeifer G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS Genetics*, (3) p1.

Jin, Y., & Penning, T. M. (2007). Aldo-keto reductases and bioactivation/detoxication. *Annual Review of Pharmacology and Toxicology*, (47) 263–292.

Johnson, I. T., & Lund, E. K. (2007). Review article: nutrition, obesity and colorectal cancer. *Alimentary Pharmacology & Therapeutics,* (26) 161–181.

Jones, P. A. (1999). The DNA methylation paradox. *Trends in Genetics*, (15) 34–37.

Jorissen, R. N., Gibbs, P., Christie, M., Prakash, S., Lipton, L., Desai, J., … Sieber, O. M. (2009). Metastasis-Associated Gene Expression Changes Predict Poor

Outcomes in Patients with Dukes Stage B and C Colorectal Cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, (24) 7642–7651.

Kaaij, L. T., van de Wetering, M., Fang, F., Decato, B., Molaro, A., van de Werken, H. J., … Ketting, R. F. (2013). DNA methylation dynamics during intestinal stem cell differentiation reveals enhancers driving gene expression in the villus. *Genome Biology*, (14) R50.

Kamakaka, R. T., & Biggins, S. (2005). Histone variants: deviants? *Genes & Development*, (19) 295–310.

Kang, M.-W., Lee, E.-S., Yoon, S. Y., Jo, J., Lee, J., Kim, H. K., … Kim, H. (2011). AKR1B10 is associated with smoking and smoking-related non-small-cell lung cancer. *The Journal of International Medical Research*, (39) 78–85.

Kelly, V., Sherratt, P., Crouch, D., & Hayes, J. (2002). Novel homodimeric and heterodimeric rat γ-hydroxybutyrate synthases that associate with the Golgi apparatus define a distinct subclass of aldo-keto reductase 7. *Biochem. J.*, (366) 847-861.

Kennison, J. a. (1995). The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. *Annual Review of Genetics*, (*29)* 289–303.

Kershaw, E. E., & Flier, J. S. (2004). Adipose Tissue as an Endocrine Organ. *The Journal of Clinical Endocrinology & Metabolism*, (89) 2548–2556.

Khare, S., & Verma, M. (2012). Cancer Epigenetics, (*863)* 177–185.

Kim, H., Lapointe, J., Kaygusuz, G., Ong, D. E., Li, C., van de Rijn, M., … Pollack, J. R. (2005). The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer. *Cancer Research*, (65) 8118–8124.

Kim, M. S., Louwagie, J., Carvalho, B., Terhaar Sive Droste, J. S., Park, H. L., Chae, Y. K., … Sidransky, D. (2009). Promoter DNA methylation of oncostatin m receptor-beta as a novel diagnostic and therapeutic marker in colon cancer. *PloS One*, (8) e6555.

Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., … Greenberg, M. E. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature*, (465) 182–187.

Kim, Y.-H., Petko, Z., Dzieciatkowski, S., Lin, L., Ghiassi, M., Stain, S., … Grady, W. M. (2006). CpG island methylation of genes accumulates during the adenoma progression step of the multistep pathogenesis of colorectal cancer. *Genes, Chromosomes and Cancer*, (45) 781–789.

Kinzler, K. W., & Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell*, (87) 159–170.

Ko, C.-Y., Hsu, H.-C., Shen, M.-R., Chang, W.-C., & Wang, J.-M. (2008). Epigenetic silencing of CCAAT/enhancer-binding protein delta activity by YY1/polycomb group/DNA methyltransferase complex. *The Journal of Biological Chemistry*, (283) 30919–30932.

Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., … Rao, A. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, (468) 839–843.

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, (502) 472–479.

Kota, S. K., & Feil, R. (2010). Epigenetic transitions in germ cell development and meiosis. *Developmental Cell*, (19) 675–86.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* (128) 693–705.

Kozma, E., Brown, E., Ellis, E. M., & Lapthorn, A. J. (2002). The Crystal Structure of Rat Liver AKR7A1: a dimeric member of the aldo-keto reductase superfamily. *Journal of Biological Chemistry* , (18) 16285–16293.

Kriaucionis, S., & Heintz, N. (2009). The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*, (324) 929–930.

Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews. Genetics*, (11) 597–610.

Kucherlapati, R., & Wheeler, D. A. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* (487) 330–337.

Kurie, J. M., Lee, J. S., Khuri, F. R., Mao, L., Morice, R. C., Lee, J. J., … Hong, W. K. (2000). N-(4-Hydroxyphenyl) Retinamide in the Chemoprevention of Squamous Metaplasia and Dysplasia of the Bronchial Epithelium, (8) 2973–2979.

Laffin, B., & Petrash, J. M. (2012). Expression of the Aldo-Ketoreductases AKR1B1 and AKR1B10 in Human Cancers. *Frontiers in Pharmacology*, (3) 104.

Laird, P. W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nature Reviews. Genetics*, (11) 191–203.

Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S. L., Edward Jung, W., Li, E., … Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell*, (81) 197-205.

Lao, V. V., & Grady, W. M. (2011). Epigenetics and colorectal cancer. *Nature Reviews. Gastroenterology & Hepatology*, (12) 686–700.

Lasnitzki, I. (1955). The influence of A hypervitaminosis on the effect of 20-methylcholanthrene on mouse prostate glands grown in vitro. *British Journal of Cancer*, (3) 434–41.

Law, J. a, & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews. Genetics*, (3) 204–220.

Lee, J. T. (2012). Epigenetic Regulation by Long Noncoding RNAs. *Science*, (338) 1435–1439.

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* (69) 915-926

Lim, D. H. K., & Maher, E. R. (2010). Genomic Imprinting Syndromes and Cancer. *Advances in Genetics*, (70) 145–175.

Limketkai, B. N., Lam-Himlin, D., Arnold, M. a, & Arnold, C. a. (2013). The cutting edge of serrated polyps: a practical guide to approaching and managing serrated colon polyps. *Gastrointestinal Endoscopy*, (77) 360–375.

Lister, R., & Ecker, J. R. (2009). Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Research*, (19) 959–66.

Lister, R., Mukamel, E. a, Nery, J. R., Urich, M., Puddifoot, C. a, Johnson, N. D., … Ecker, J. R. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science,* (341).

Loeffler-Ragg, J., Mueller, D., Gamerith, G., Auer, T., Skvortsov, S., Sarg, B., … Zwierzina, H. (2009). Proteomic identification of aldo-keto reductase AKR1B10 induction after treatment of colorectal cancer cells with the

proteasome inhibitor bortezomib. *Molecular Cancer Therapeutics*, (7) 1995–2006.

López-Abente, G., Ardanaz, E., Torrella-Ramos, a, Mateos, a, Delgado-Sanz, C., & Chirlaque, M. D. (2010). Changes in colorectal cancer incidence and mortality trends in Spain. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, (21) iii76–82.

Luger, K. (2006). Dynamic nucleosomes. *Chromosome Research*, (14) 5–16.

Lynch, H. T., Lynch, P. M., Lanspa, S. J., Snyder, C. L., Lynch, J. F., & Boland, C. R. (2009). Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clinical Genetics*, (76) 1–18.

Ma, J., & Cao, D. (2011). Human aldo-keto reductases: structure, substrate specificity and roles in tumorigenesis. *BioMolecular Concepts*, (2) 115–126.

Ma, J., Luo, D.-X., Huang, C., Shen, Y., Bu, Y., Markwell, S., … Cao, D. (2012). AKR1B10 overexpression in breast cancer: association with tumor size, lymph node metastasis and patient survival and its potential as a novel serum marker. *International Journal of Cancer. Journal International Du Cancer*, (6).

Majumdar, S. R., Fletcher, R. H., & Evans, A. T. (1999). How does colorectal cancer present[quest] symptoms, duration, and clues to location. *Am J Gastroenterol*, (10) 3039–3045.

Mali, P., Yang, L., Esvelt, K., Aach, J., & Guell, M. (2013). RNA-guided human genome engineering via Cas9. *Science*, (339) 823–826.

Mann, I. K., Chatterjee, R., Zhao, J., He, X., Weirauch, M. T., Hughes, T. R., & Vinson, C. (2013). CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo. *Genome,* (23) 988-997.

Marcuello, E., Altes, A., Menoyo, A., del Rio, E., Gomez-Pardo, M., & Baiget, M. (2004). UGT1A1 gene variations and irinotecan treatment in patients with metastatic colorectal cancer. *Br J Cancer*, (91) 678–682.

Marx, V. (2012). Epigenetics: Reading the second genomic code. *Nature*, (491) 143–147.

Mayor, R., Casadomé, L., Azuara, D., Moreno, V., Clark, S. J., Capellà, G. & Peinado, M. A. (2009). Long-range epigenetic silencing at 2q14.2 affects most human colorectal cancers and may have application as a non-invasive biomarker of disease. *British Journal of Cancer*, (100) 1534–1539.

Maze, I., Noh, K.-M., Soshnev, A. a., & Allis, C. D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nature Reviews Genetics*, (15) 259-271.

McVicker, G., Geijn, B. van de, & Degner, J. (2013). Identification of genetic variants that affect histone modifications in human cells. *Science*, (342) 747–749.

Meinsma, R., Fernandez-Salguero, P., Van Kuilenburg, a B., Van Gennip, a H., & Gonzalez, F. J. (1995). Human polymorphism in drug metabolism: mutation in the dihydropyrimidine dehydrogenase gene results in exon skipping and thymine uracilurea. *DNA and Cell Biology*, (14) 1–6.

Ménard, S., Camerini, T., Mariani, L., Tomasic, G., Pilotti, S., Costa, A., … Veronesi, U. (2001). Re: Randomized Trial of Fenretinide to Prevent Second Breast

Malignancy in Women With Early Breast Cancer. *Journal of the National Cancer Institute* , (93) 240–241.

Merlos-Suárez, A., Barriga, F. M., Jung, P., Iglesias, M., Céspedes, M. V., Rossell, D., … Batlle, E. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell*, (8) 511–24.

Milicic, A., Harrison, L.-A., Goodlad, R. A., Hardy, R. G., Nicholson, A. M., Presz, M., … Jankowski, J. A. Z. (2008). Ectopic Expression of P-Cadherin Correlates with Promoter Hypomethylation Early in Colorectal Carcinogenesis and Enhanced Intestinal Crypt Fission In vivo. *Cancer Research*, (68) 7760–7768.

Mills, A. a. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. *Nature Reviews. Cancer*, (10), 669–82.

Minucci, S., & Pelicci, P. G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer*, (6) 38–51.

Mizuno, H., Kitada, K., Nakai, K., & Sarai, A. (2009). PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Medical Genomics*, *(2)* 18.

Moise, A. R. (2010). Pharmacology of Retinoid Receptors.

Mongan, N. P., & Gudas, L. J. (2007). Diverse actions of retinoid receptors in cancer prevention and treatment. *Differentiation; Research in Biological Diversity*, (75) 853–70.

Muto, Y., Moriwaki, H., Ninomiya, M., Adachi, S., Saito, A., Takasaki, K. T., … Kojima, T. (1996). Prevention of Second Primary Tumors by an Acyclic Retinoid, Polyprenoic Acid, in Patients with Hepatocellular Carcinoma. *New England Journal of Medicine*, (334) 1561–1568.

Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., … Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, (502) 59–64.

Nagaraj, N. S., Beckers, S., Mensah, J. K., Waigel, S., Vigneswaran, N., & Zacharias, W. (2006). Cigarette smoke condensate induces cytochromes P450 and aldo-keto reductases in oral cancer cells. *Toxicology Letters*, (165) 182–194.

Nagasaka, T., Tanaka, N., Cullings, H. M., Sun, D.-S., Sasamoto, H., Uchida, T., … Goel, A. (2009). Analysis of fecal DNA methylation to detect gastrointestinal neoplasia. *Journal of the National Cancer Institute*, (101) 1244–1258.

Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., … Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, (272) 263–267.

Narayan, G., Arias-Pulido, H., & Koul, S. (2003). Frequent promoter methylation of CDH1, DAPK, RARB, and HIC1 genes in carcinoma of cervix uteri: its relationship to clinical outcome. *Molecular Cancer*, (12) 1–12.

Nephew, K. P., & Huang, T. H.-M. (2003). Epigenetic gene silencing in cancer initiation and progression. *Cancer Letters*, (190) 125–133.

Newton, K. F., Newman, W., & Hill, J. (2012). Review of biomarkers in colorectal cancer. *Colorectal Disease : The Official Journal of the Association of Coloproctology of Great Britain and Ireland*, (14) 3–17.

Niles, R. M. (2004). Signaling pathways in retinoid chemoprevention and treatment of cancer. *Mutation Research*, (555) 81–96.

Nishinaka, T., & Yabe-Nishimura, C. (2005). Transcription factor Nrf2 regulates promoter activity of mouse aldose reductase (AKR1B3) gene. *Journal of Pharmacological Sciences*, (51) 43–51.

Nosho, K., Shima, K., Irahara, N., Kure, S., Baba, Y., Kirkner, G. J., … Ogino, S. (2009). DNMT3B Expression Might Contribute to CpG Island Methylator Phenotype in Colorectal Cancer. *Clinical Cancer Research* , (15) 3663–3671.

Novak, P., Jensen, T. J., Garbe, J. C., Stampfer, M. R., & Futscher, B. W. (2009). Stepwise DNA methylation changes are linked to escape from defined proliferation barriers and mammary epithelial cell immortalization. *Cancer Research*, (69) 5251–5258.

O'Keefe, S. J. D., Kidd, M., Espitalier-Noel, G., & Owira, P. (1999). Rarity of colon cancer in Africans is associated with low animal product consumption, not fiber. *Am J Gastroenterol*, (94) 1373–1380.

Ohashi, T., Idogawa, M., Sasaki, Y., Suzuki, H., & Tokino, T. (2013). AKR1B10, a transcriptional target of p53, is Downregulated in Colorectal Cancers associated with poor prognosis. *Molecular Cancer Research,* (11) 1554-1563.

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell*, (99) 247-257.

Okudela, K., Woo, T., Mitsui, H., Suzuki, T., Tajiri, M., Sakuma, Y., … Ohashi, K. (2013). Downregulation of ALDH1A1 expression in non-small cell lung carcinomas--its clinicopathologic and biological significance. *International Journal of Clinical and Experimental Pathology*, (6) 1–12.

Olek, A., & Walter, J. (1997). The pre-implantation ontogeny of the H19 methylation imprint. *Nature Genetics,* (17) 275-276.

Ong, C.-T., & Corces, V. G. (2014). CTCF: an architectural protein bridging genome topology and function. *Nature Reviews Genetics*, (15) 234–246.

Osanai, M., Sawada, N., & Lee, G.-H. (2010). Oncogenic and cell survival properties of the retinoic acid metabolizing enzyme, CYP26A1. *Oncogene*, (29) 1135–1144.

Oshima, M., Dinchuk, J. E., Kargman, S. L., Oshima, H., Hancock, B., Kwong, E., … Taketo, M. M. (1996). Suppression of intestinal polyposis in Apc delta716 knockout mice by inhibition of cyclooxygenase 2 (COX-2). *Cell*, (87) 803–809.

Ozdağ, H., Teschendorff, A. E., Ahmed, A. A., Hyland, S. J., Blenkiron, C., Bobrow, L., … Caldas, C. (2006). Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics*, (7).

Park, E. Y., Dillard, A., Williams, E. a, Wilder, E. T., Pepper, M. R., & Lane, M. a. (2005). Retinol inhibits the growth of all-trans-retinoic acid-sensitive and all-trans-retinoic acid-resistant colon cancer cells through a retinoic acid receptor-independent mechanism. *Cancer Research*, (65) 9923–9933.

Park, E. Y., Wilder, E. T., Chipuk, J. E., & Lane, M. a. (2008). Retinol decreases phosphatidylinositol 3-kinase activity in colon cancer cells. *Molecular Carcinogenesis*, (47) 264–274.

Pastel, E., Pointud, J.-C., Volat, F., Martinez, A., & Lefrançois-Martinez, A.-M. (2012). Aldo-Keto Reductases 1B in Endocrinology and Metabolism. *Frontiers in Pharmacology*, (3) 148.

Pastor, W. a, Aravind, L., & Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature Reviews. Molecular Cell Biology*, (14) 341–356.

Pastorino, U., Infante, M., Maioli, M., Chiesa, G., Buyse, M., Firket, P., … Valente, M. (1993). Adjuvant treatment of stage I lung cancer with high-dose vitamin A. *Journal of Clinical Oncology* , (11) 1216–1222.

Pennacchio, L. a., Bickmore, W., Dean, A., Nobrega, M. a., & Bejerano, G. (2013). Enhancers: five essential questions. *Nature Reviews Genetics*, (14) 288–295.

Penning, T. M., & Drury, J. E. (2007). Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms. *Archives of Biochemistry and Biophysics*, (464) 241–250.

Pennisi, E. (2011). Does a Gene's Location in the Nucleus Matter? *Science*, (334) 1050–1051.

Pennisi, E. (2013). The CRISPR craze. *Science*, (341) 833–836.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., Brugman, W., … van Steensel, B. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular Cell*, (38) 603–613.

Petitjean, A., Mathe, E., Kato, S., Ishioka, C., Tavtigian, S. V, Hainaut, P., & Olivier, M. (2007). Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Human Mutation*, (28) 622–629.

Plebuch, M., Soldan, M., Hungerer, C., Koch, L., & Maser, E. (2007). Increased resistance of tumor cells to daunorubicin after transfection of cDNAs coding for anthracycline inactivating enzymes. *Cancer Letters*, (255) 49–56.

Pox, C. (2011). Colon cancer screening: which non-invasive filter tests? *Digestive Diseases (29)*, 56–59.

Quinn, A., Harvey, R., & Penning, T. (2008). Oxidation of PAH trans-dihydrodiols by human aldo-keto reductase AKR1B10. *Chemical Research in Toxicology,* (21) 2207–2215.

Ramana, K. V, Bhatnagar, A., Srivastava, S., Yadav, U. C., Awasthi, S., Awasthi, Y. C., & Srivastava, S. K. (2006). Mitogenic Responses of Vascular Smooth Muscle Cells to Lipid Peroxidation-derived Aldehyde 4-Hydroxy-trans-2-nonenal (HNE): Role of aldose reductase-catalized reduction of the HNE-glutathione conjugates in regulation cell growth. *Journal of Biological Chemistry* , (281) 17652–17660.

Ramana, K. V, Fadl, A. a, Tammali, R., Reddy, A. B. M., Chopra, A. K., & Srivastava, S. K. (2006). Aldose reductase mediates the lipopolysaccharide-induced release of inflammatory mediators in RAW264.7 murine macrophages. *The Journal of Biological Chemistry*, (281) 33019–33029.

Ramírez, N., Bandrés, E., Navarro, A., Pons, A., Jansa, S., Moreno, I., … García-Foncillas, J. (2008). Epigenetic events in normal colonic mucosa surrounding colorectal cancer lesions. *European Journal of Cancer,* (26) 2689-2695.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. a, & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, (8) 2281–2308.

Rashid, A., Shen, L., Morris, J. S., Issa, J.-P. J., & Hamilton, S. R. (2001). CpG Island Methylation in Colorectal Adenomas. *The American Journal of Pathology*. American Society for Investigative Pathology.

Ravindranath, T. M., Mong, P. Y., Ananthakrishnan, R., Li, Q., Quadri, N., Schmidt, A. M., … Wang, Q. (2009). Novel role for aldose reductase in mediating acute

inflammatory responses in the lung. *Journal of Immunology*, (183) 8128–8137.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, (447) 425–432.

Renehan, A. G., O'Connell, J., O'Halloran, D., Shanahan, F., Potten, C. S., O'Dwyer, S. T., & Shalet, S. M. (2003). Acromegaly and Colorectal Cancer: A Comprehensive Review of Epidemiology, Biological Mechanisms, and Clinical Implications. *Horm Metab Res*, (35) 712–725.

Rhee, I., Bachman, K. E., Park, B. H., Jair, K.-W., Yen, R.-W. C., Schuebel, K. E., … Vogelstein, B. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, (416) 552–556.

Rideout, W. M., Coetzee, G. A., Olumi, A. F., & Jones, P. A. (1990). 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science*, (249) 1288–1290.

Ries, L. A. G., Melbert, D., Krapcho, M., Stinchcomb, D. G., Howlader, N., Horner, M. J., … Edwards, B. K. (2008). SEER cancer statistics review, 1975-2005 . Bethesda, MD : U.S. National Institutes of Health, National Cancer Institute .

Riggs, A. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research*, (14) 9-25.

Rishi, V., Bhattacharya, P., Chatterjee, R., Rozenberg, J., Zhao, J., Glass, K., … Vinson, C. (2010). CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes. *Proceedings of the National Academy of Sciences of the United States of America*, (107) 20311–20316.

Roberts, D. L., Dive, C., & Renehan, A. G. (2010). Biological mechanisms linking obesity and cancer risk: new perspectives. *Annual Review of Medicine*, (61) 301–316.

Rodriguez, J., Muñoz, M., Vives, L., Frangou, C. G., Groudine, M., & Peinado, M. a. (2008). Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*, (105) 19809–19814.

Rodriguez, J., Vives, L., Jordà, M., Morales, C., Muñoz, M., Vendrell, E., & Peinado, M. a. (2008). Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells. *Nucleic Acids Research*, (36) 770–784.

Rodríguez-Paredes, M., & Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nature Medicine*, (17) 330–339.

Rose, D. P., Komninou, D., & Stephenson, G. D. (2004). Obesity , adipocytokines , and insulin resistance in breast cancer. *Obesity Reviews*, (7) 153–165.

Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, a., Plachot, M., Paldi, a., & Viegas-Pequignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes & Development,* (12) 2108–2113.

Rouits, E., Boisdron-Celle, M., Dumont, A., Guérin, O., Morel, A., & Gamelin, E. (2004). Relevance of different UGT1A1 polymorphisms in Irinotecan-induced toxicity a molecular and clinical study of 75 patients. *Clinical Cancer Research*, (10) 5151–5159.

Ruiz, F. X., Gallego, O., Ardèvol, A., Moro, A., Domínguez, M., Alvarez, S., … Farrés, J. (2009). Aldo-keto reductases from the AKR1B subfamily: retinoid specificity

and control of cellular retinoic acid levels. *Chemico-Biological Interactions*, (178) 171–177.

Ruiz, F. X., Porté, S., Parés, X., & Farrés, J. (2012). Biological role of aldo-keto reductases in retinoic Acid biosynthesis and signaling. *Frontiers in Pharmacology*, (3) 58.

Russo, V. E. A., Martienssen, R. A., & Riggs, A. D. (1997). Epigenetic Mechanisms of Gene Regulation. Edited. *Genetics Research*, (69) 159–162.

Ruthenburg, A. J., Li, H., Patel, D. J., & Allis, C. D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature Reviews. Molecular Cell Biology*, (8) 983–994.

Salabei, J. K., Li, X.-P., Petrash, J. M., Bhatnagar, A., & Barski, O. a. (2011). Functional expression of novel human and murine AKR1B genes. *Chemico-Biological Interactions*, (191) 177–184.

Salabei, J., Li, X., & Petrash, J. (2011). Functional expression of novel human and murine AKR1B genes. *Chemico Biological Interactions*, (*191*) 177–184.

Sales, K. J., Maldonado-Pérez, D., Grant, V., Catalano, R. D., Wilson, M. R., Brown, P., … Jabbour, H. N. (2009). Prostaglandin F2α-F-prostanoid receptor regulates CXCL8 expression in endometrial adenocarcinoma cells via the calcium–calcineurin–NFAT pathway. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, (12) 1917–1928.

Sandoval, J., & Esteller, M. (2012). Cancer epigenomics: beyond genomics. *Current Opinion in Genetics & Development*, (22) 50–55.

Sanyal, A., Lajoie, B. R., Jain, G., & Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature*, (489) 109–113.

Schefe, J. H., Lehmann, K. E., Buschmann, I. R., Unger, T., & Funke-Kaiser, H. (2006). Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. *Journal of Molecular Medicine*, (84) 901–910.

Schmidl, C., Klug, M., Boeld, T. J., Andreesen, R., Hoffmann, P., Edinger, M., & Rehli, M. (2009). Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Research*, (19) 1165–1174.

Schmidt, W. M., Sedivy, R., Forstner, B., Steger, G. G., Zöchbauer-Müller, S., & Mader, R. M. (2007). Progressive up-regulation of genes encoding DNA methyltransferases in the colorectal adenoma-carcinoma sequence. *Molecular Carcinogenesis*, (46) 766–772.

Schmitz, K. J., Sotiropoulos, G. C., Baba, H. a, Schmid, K. W., Müller, D., Paul, A., … Loeffler-Ragg, J. (2011). AKR1B10 expression is associated with less aggressive hepatocellular carcinoma: a clinicopathological study of 168 cases. *Liver International : Official Journal of the International Association for the Study of the Liver*, (31) 810–816.

Schübeler, D. (2012). Molecular biology. Epigenetic islands in a genetic ocean. *Science*, (338) 756–757.

Schuettengruber, B., Martinez, A.-M., Iovino, N., & Cavalli, G. (2011). Trithorax group proteins: switching genes on and keeping them active. *Nature Reviews. Molecular Cell Biology*, (12) 799–814.

Scott, A. M., Wolchok, J. D., & Old, L. J. (2012). Antibody therapy of cancer. *Nature Reviews. Cancer*, (12) 278–87.

Seligson, D. B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., & Kurdistani, S. K. (2005). Global histone modification patterns predict risk of prostate cancer recurrence. *Nature*, (435) 1262–1266.

Shaib, W., Mahajan, R., & El-Rayes, B. (2013). Markers of resistance to anti-EGFR therapy in colorectal cancer. *Journal of Gastrointestinal Oncology,* (4) 308-318.

Sheikhnejad, G., Brank, a, Christman, J. K., Goddard, a, Alvarez, E., Ford, H., … Cheng, X. (1999). Mechanism of inhibition of DNA (cytosine C5)-methyltransferases by oligodeoxyribonucleotides containing 5,6-dihydro-5-azacytosine. *Journal of Molecular Biology*, (285) 2021–2034.

Shen, L., Kondo, Y., Hamilton, S. R., Rashid, A., & Issa, J. J. (2003, March 1). p14 methylation in human colon cancer is associated with microsatellite instability and wild-type p53. *Gastroenterology,* (124) 626-633.

Shen, L., Kondo, Y., Rosner, G. L., Xiao, L., Hernandez, N. S., Vilaythong, J., … Issa, J.-P. J. (2005). MGMT promoter methylation and field defect in sporadic colorectal cancer. *Journal of the National Cancer Institute*, (97) 1330–1338.

Shlyueva, D., Stampfel, G., & Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nature Reviews Genetics* (15), 272–286.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., … Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* (479) 74-79.

Siegel, R. (2013). Cancer statistics, 2013. *… Cancer Journal for Clinicians*, (63) 11–30.

Simon, J. (1995). Locking in stable states of gene expression: transcriptional control during Drosophila development. *Current Opinion in Cell Biology*, (7) 376–385.

Simon, J. a, & Kingston, R. E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Reviews. Molecular Cell Biology*, (10) 697–708.

Singh, H., Nugent, Z., Demers, A. a, Kliewer, E. V, Mahmud, S. M., & Bernstein, C. N. (2010). The reduction in colorectal cancer mortality after colonoscopy varies by site of the cancer. *Gastroenterology*, (139) 1128–1137.

Skrzypczak, M., Goryca, K., Rubel, T., Paziewska, A., Mikula, M., Jarosz, D., … Ostrowsk, J. (2010). Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability. *PloS One*, (5).

Smith, H. W. (1965). Observations on the flora of the alimentary tract of animals and factors affecting its composition. *The Journal of Pathology and Bacteriology*, (89) 95–122.

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature Reviews. Genetics*, (14) 204–220.

Smyth, G. K., Michaud, J., & Scott, H. S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics*, (21) 2067–2075.

Srivastava, S., Spite, M., Trent, J. O., West, M. B., Ahmed, Y., & Bhatnagar, A. (2004). Aldose Reductase-catalyzed Reduction of Aldehyde Phospholipids. *Journal of Biological Chemistry* , (279) 53395–53406.

Stadler, M. B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., … Schübeler, D. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* (480) 490-495.

Staub, E., Groene, J., Heinze, M., Mennerich, D., Roepcke, S., Klaman, I., … Rosenthal, A. (2009). An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types. *Journal of Molecular Medicine (Berlin, Germany)*, (87) 633–644.

Sun, S.-Y., & Lotan, R. (2002). Retinoids and their receptors in cancer development and chemoprevention. *Critical Reviews in Oncology/hematology*, (41) 41–55.

Suzuki, K., Suzuki, I., Leodolter, A., Alonso, S., Horiuchi, S., Yamashita, K., & Perucho, M. (2006). Global DNA demethylation in gastrointestinal cancer is age dependent and precedes genomic damage. *Cancer Cell*, (9) 199–207.

Syngal, S., Stoffel, E., Chung, D., Willett, C., Schoetz, D., Schroy, P., … Ross, M. (2006). Detection of stool DNA mutations before and after treatment of colorectal neoplasia. *Cancer*, (106) 277–283.

Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F., & Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Research*, (38) e181.

Tagore, K. S., Lawson, M. J., Yucaitis, J. A., Gage, R., Orr, T., Shuber, A. P., & Ross, M. E. (2003). Sensitivity and Specificity of a Stool DNA Multitarget Assay Panel for the Detection of Advanced Colorectal Neoplasia. *Clinical Colorectal Cancer*, (3) 47–53.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., … Rao, A. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science*, (324) 930–935.

Talbert, P. B., & Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. *Nature Reviews. Molecular Cell Biology*, (11) 264–275.

Tammali, R., Ramana, K. V, Singhal, S. S., Awasthi, S., & Srivastava, S. K. (2006). Aldose Reductase Regulates Growth Factor-Induced Cyclooxygenase-2 Expression and Prostaglandin E2 Production in Human Colon Cancer Cells. *Cancer Research* , (66) 9705–9713.

Tammali, R., Srivastava, S. K., & Ramana, K. V. (2011). Targeting aldose reductase for the treatment of cancer. *Current Cancer Drug Targets*, (11) 560–571.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., … Zhao, Y. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, (146) 1016–1028.

Tang, X.-H., & Gudas, L. J. (2011). Retinoids, retinoic acid receptors, and cancer. *Annual Review of Pathology*, (6) 345–364.

Theodosiou, M., Laudet, V., & Schubert, M. (2010). From carrot to clinic: an overview of the retinoic acid signaling pathway. *Cellular and Molecular Life Sciences*, (67) 1423–1445.

Thun, M. J., Jacobs, E. J., & Patrono, C. (2012). The role of aspirin in cancer prevention. *Nature Reviews. Clinical Oncology*, (9) 259–267.

Tiscornia, G., Singer, O., & Verma, I. M. (2006). Production and purification of lentiviral vectors. *Nature Protocols*, (1) 241–245.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., & de Laat, W. (2014). Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. *Molecular Cell*, (10) 1453–1465.

Tonus, C., Sellinger, M., Koss, K., & Neupert, G. (2012). Faecal pyruvate kinase isoenzyme type M2 for colorectal cancer screening: a meta-analysis. *World Journal of Gastroenterology,* (18) 4004–4011.

Triantafillidis, J. K., Nasioulas, G., & Kosmidis, P. a. (2009). Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. *Anticancer Research*, (29) 2727–2737.

Tropberger, P., & Schneider, R. (2013). Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nature Structural & Molecular Biology*, (20) 657–661.

Tsai, H.-C., & Baylin, S. B. (2011). Cancer epigenetics: linking basic biology to clinical medicine. *Cell Research*, (21) 502–517.

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., … Chang, H. Y. (2010). Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science*, (329) 689–693.

Tsunoda, S., Smith, E., Young, N. J. D. E., Wang, X., Tian, Z., Liu, J., … Drew, P. A. (2009). TMEFF2 in esophageal squamous cell carcinoma, (21) 1067–1073.

Van Engeland, M., Derks, S., Smits, K. M., Meijer, G. A., & Herman, J. G. (2011). Colorectal Cancer Epigenetics: Complex Simplicity. *Journal of Clinical Oncology* , (29) 1382–1391.

Van Steensel, B. (2011). Chromatin: constructing the big picture. *The EMBO Journal*, (30) 1885–1895.

Vanneman, M., & Dranoff, G. (2012). Combining immunotherapy and targeted therapies in cancer treatment. *Nature Reviews. Cancer*, (12) 237–251.

Vermeulen, L., Morrissey, E., van der Heijden, M., Nicholson, a. M., Sottoriva, a., Buczacki, S., … Winton, D. J. (2013). Defining Stem Cell Dynamics in Models of Intestinal Tumor Initiation. *Science*, (342) 995–998.

Viaud, S., Saccheri, F., Mignot, G., Yamazaki, T., Daillere, R., Hannani, D., … Zitvogel, L. (2013). The Intestinal Microbiota Modulates the Anticancer Immune Effects of Cyclophosphamide. *Science*, (342) 971–976.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., … Fuks, F. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, (439) 871–874.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. a, & Kinzler, K. W. (2013). Cancer genome landscapes. *Science*, (339) 1546–1558.

Vona-Davis, L., & Rose, D. P. (2007). Adipokines as endocrine, paracrine, and autocrine factors in breast cancer risk and progression. *Endocrine-Related Cancer*, (14) 189–206.

Wald, N., Idle, M., Boreham, J., & Bailey, A. (1980). Low serum-vitamin-A and subsequent risk of cancer: Preliminary results of a prospective study. *The Lancet*, (316) 813–815.

Walker, J. M., & Trygve O.F. (2004). DNA Methylation. *Protocols Epigenetics.* Human Press.

Wallis, Y. L., Morton, D. G., McKeown, C. M., & Macdonald, F. (1999). Molecular analysis of the APC gene in 205 families: extended genotype-phenotype correlations in FAP and evidence for the role of APC amino acid changes in colorectal cancer predisposition. *Journal of Medical Genetics* , (36) 14–20.

Wang, Y., Fang, M. Z., & Liao, J. (2003). Hypermethylation-Associated Inactivation of Retinoic Acid Receptor β in Human Esophageal Squamous Cell Carcinoma, *Clinical Cancer Research,* (9) 5257–5263.

Wei, P.-L., Chang, Y.-J., Ho, Y.-S., Lee, C.-H., Yang, Y.-Y., An, J., & Lin, S.-Y. (2009). Tobacco-specific carcinogen enhances colon cancer cell migration through alpha7-nicotinic acetylcholine receptor. *Annals of Surgery*, (249) 978–985.

Weissman, S., Bellcross, C., Bittner, C., Freivogel, M., Haidle, J., Kaurah, P., … Daniels, M. (2011). Genetic Counseling Considerations in the Evaluation of Families for Lynch Syndrome—A Review. *Journal of Genetic Counseling*, (20) 5–19.

Weng, J., Cao, Y., Moss, N., & Zhou, M. (2006). Modulation of Voltage-dependent Shaker Family Potassium Channels by an Aldo-Keto Reductase. *Journal of Biological Chemistry*, (281) 15194–15200.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., … Laird, P. W. (2007). Epigenetic stem cell signature in cancer. *Nature Genetics*, (39) 157–158.

Winawer, S. J., Zauber, A. G., Gerdes, H., O'Brien, M. J., Gottlieb, L. S., Sternberg, S. S., … Bishop, D. T. (1996). Risk of Colorectal Cancer in the Families of Patients with Adenomatous Polyps. *New England Journal of Medicine*, (334) 82–87.

Woenckhaus, M., Klein-Hitpass, L., Grepmeier, U., Merk, J., Pfeifer, M., Wild, P. J., … Dietmaier, W. (2006). Smoking and cancer-related gene expression in bronchial epithelium and non-small-cell lung cancers. *The Journal of Pathology*, (210) 192–204.

Wolbach, S., & Howe, P. (1925). Tissue changes following deprivation of fat-soluble A vitamin. *The Journal of Experimental Medicine* (42) 753-777.

Wollowski, I., Rechkemmer, G., & Pool-zobel, B. L. (2001). Protective role of probiotics and prebiotics in colon cancer. *The American Journal of Clinical Nutrition*, (73) 451S-455S.

Wu, S. C., & Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. *Nature Reviews. Molecular Cell Biology*, (11) 607–20.

Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet*, (12) 542–553.

Yadav, U. C. S., Ramana, K. V, Aguilera-Aguirre, L., Boldogh, I., Boulares, H. a, & Srivastava, S. K. (2009). Inhibition of aldose reductase prevents experimental allergic airway inflammation in mice. *PloS One*, (4) e6535.

Yao, H.-B., Xu, Y., Chen, L.-G., Guan, T.-P., Ma, Y.-Y., He, X.-J., … Shao, Q.-S. (2013). AKR1B10, a good prognostic indicator in gastric cancer. *European Journal of Surgical Oncology,* (40) 318-324.

Young, L., Sung, J., Stacey, G., & Masters, J. R. (2010). Detection of Mycoplasma in cell cultures. *Nature Protocols*, (5) 929–934.

Yu, M., Hon, G. C., Szulwach, K. E., Song, C.-X., Zhang, L., Kim, A., … He, C. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell*, (149) 1368–1380.

Zentner, G. E., & Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. *Nature Structural & Molecular Biology*, (20) 259–66.

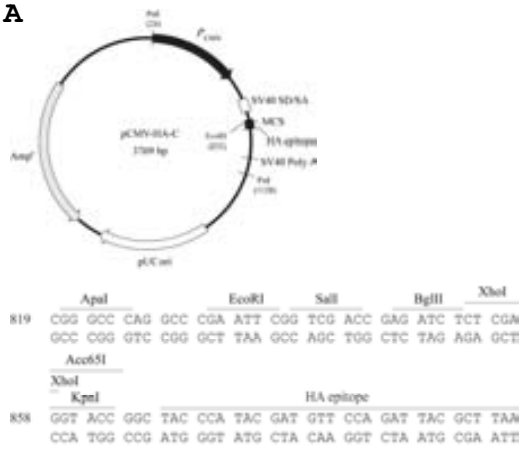
Zhang, F. L., & Casey, P. J. (1996). Protein Prenylation: Molecular Mechanisms and Functional Consequences. *Annual Review of Biochemistry*, (65) 241–269.

?

? ? ? ?? ??????? ? ??? ? ? ??? ??????? ? ?? ????? ???

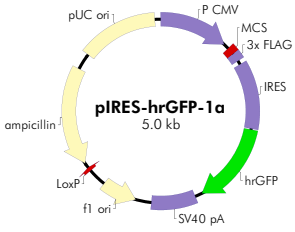
?

A

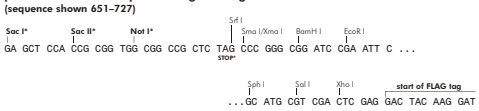


B

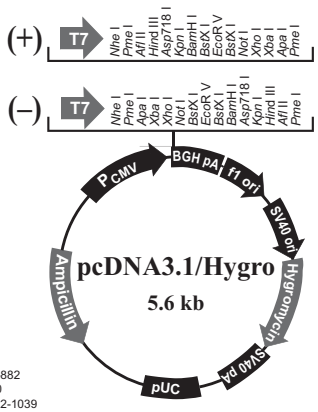
CMV promoter 1–602
multiple cloning site 651–715
3× FLAG tag 716–787
internal ribosome entry site 823–1397
hrGFP ORF 1407–2123
SV40 polyA 2188–2571
f1 origin 2709–3015
LoxP sequence 3178–3211
ampicillin resistance (*bla*) ORF 3256–4113
pUC origin 4260–4927



pIRES-hrGFP-1a Multiple Cloning Site Region
(sequence shown 651–727)



C



Comments for
pcDNA™3.1/Hygro (+):
5597 nucleotides

CMV promoter: bases 209-863
T7 promoter/priming site: bases 863-882
Multiple cloning site: bases 895-1010
BGH reverse priming site: bases 1022-1039
BGH polyadenylation signal: bases 1021-1235
f1 origin: bases 1298-1711
SV40 promoter and origin: bases 1776-2100
Hygromycin resistance gene: bases 2118-3141
SV40 early polyadenylation signal: bases 3154-3526
pUC origin: bases 3786-4456 (complementary strand)
Ampicillin resistance gene: bases 4601-5461 (complementary strand)

?

?AA?b??u? ????nA??? ??c??BnI | ?a | ???AB?? ?| ?Ba?ov?i e??? ?r? ?N?i e????r?I ???ry ??A??i ? e????) ty s?o?I n?- i t?

?
?
?
?
?
?
?
?
?
?
?
?

Primer name Sequence

AKR1B1 CpGi C F1	GAGTATTTATTTTTTAGGTATT
AKR1B1 CpGi C R1	TTTCCCACCAAATACAACA
AKR1B1 CpGi C F2	AATTTTAGGATGGGTATTTTG
AKR1B1 CpGi C R2	AACTAAAAAACTCCTTTCTAC
AKR1B1 CpG A F1	ATTACCTCACTTATAAATAATTAA
AKR1B1 CpG A R unique	TTTTTTGAATTATTAATTAGTTA
AKR1B1 CpG A F2	GGGTGTTTAGATTTTTTTTTT
AKR1B1 CpGi B F1	AGGGAAAGGAGGTTGGGTTT
AKR1B1 CpGi B R1	CCATCCTAAAATTAAATACCTAA
AKR1B1 CpGi B F2	GGAAAGGAGGTTGGGTTTG
AKR1B1 CpGi B R2	ATTAAATACCTAAAAAATAAATACT
AKR1B10 F1	TTATGTTGGTTAGGTTGGTTTT
AKR1B10 R1	CTTACTTTCCCTATAATAATTTCCT
AKR1B10 F2	ATTATAGGTGTGAGATATTATGTT
AKR1B10 R2	CTTACTTTCCCTATAATAATTTCCT
AKR1B15 F1	GGTTTATATTATTTTTAAGAGTTAT
AKR1B15 R unique	ACCTTATACCCAATTAAATATCCT
AKR1B15 F2	TTGTAATATATATTGTGAAGGTTTATG
Mouse AKR1B3 CpGi F1	AGAGGATAGAGAAATAAGTTGG
Mouse AKR1B3 CpGi R1	AAACCAAAACTTCTCCCCAATCTC
Mouse AKR1B3 CpGi F2	TGGGTGTGTTTTTATTTGGGG
Mouse AKR1B3 CpGi R2	TATCTCTAATCTACTAATTATTTTC
EN1 CpGi C F1	AGAATAATAAAGATAAGAGAT
EN1 CpGi C R1	ACTATCCTACTTATAAACTC
EN1 CpGi C F2	GTTTTAGGGATTTAGAGTTT
EN1 CpGi C R2	CTACTTATAAACTCAACCAA
INHBB F1	AAGTTTTGAGGTTTAGTGGTTTT
INHBB R1	AAACCCTACCTCTATCCCAA
INHBB F2	GGTTAGGTTTTTAGTGGTTATT
INHBB R2	CCTCCCCAACCAACAAAAA

□

□□□ □□□□□□□□□□□□□□□□□□ □ □□□□ □□□□

□

Primer name	Sequence	Genome
AKR1B1 Fw AKR1B1 Rv	CCGTCTCCTGCTCAACAAC TACACATGGGCACAGTCGAT	Human
AKR1B10 Fw AKR1B10 Rv	GCTTCTCGATCTGGAAGTGG GTAATGCCATCGGTGGAAAA	Human
AKR1B15 Fw AKR1B15 Rv	CCCTTTGACTGGCCTAAAGAG AATGTGGCGATATTCTGCATCA	Human
EN1 Fw EN1 Rv	TGGGTGTACTGCACACGTTATTC CTTGTCCTCCTTCTCGTTCTTCTT	Human
INHBB Fw INHBB Rv	CGCGTTTCCGAAATCATCA GGACCACAAACAGGTTCTGGTT	Human
CPLX2 Fw CPLX2 Rv	GAGGCGGAGCGGGAGAAGGTC GCCCGGGCAGGTATTTGAGCA	Human
18S Fw 18S Rv	GCGAAAGCATTTGCCAAGAA CATCACAGACCTGTTATTGC	Human
PP1A Fw PP1A Rv	CTCCTTTGAGCTGTTTGCAG CACCACATGCTTGCCATCC	Human
Beta-2-Micro Fw Beta-2-Micro Rv	CCAGCAGAGAATGGAAAGTC GATGCTGCTTACATGTCTCG	Human
PSMC4 Fw PSMC4 Rv	TGTTGGCAAAGGCGGTGGCA TCTCTTGGTGGCGATGGCAT	Human
PUM1 Fw PUM1 Rv	CGGTCGTCCTGAGGATAAAA CGTACGTGAGGCGTGAGTAA	Human
MRPL19 Fw MRPL19 Rv	CAGTTTCTGGGGATTTGCAT TATTCAGGAAGGGCATCTCG	Human
Akr1b3 mice Fw Akr1b3 mice Rv	AGGCCGTGAAAGTTGCTATTG ATGCTCTTGTCATGGAACGTG	Mouse
Akr1b8 mice Fw Akr1b8 mice Rv	TCTGATTCGGTTTCACATCCAG CCAGTTTCTGTTGAAGCTAAGGA	Mouse
Akr1b10 mice Fw Akr1b10 mice Rv	CTAGTGCCAAACCAGAGGACC TCCTGTATTCGAGAAGGTGTCA	Mouse
Akr1b7 mice Fw Akr1b7 mice Rv	AGGCTGGGAATGCGTTATTAC GGGTGAGATAAGGGTGGCTCT	Mouse
mGAPDH Fw mGAPDH Rv	TGCACCACCAACTGCTTAG GATGCAGGGATGATGTTC	Mouse
mB2M Fw mB2M Rv	GCTATCCAGAAAACCCCTCAA CATGTCTCGATCCCAGTAGACGGT	Mouse
m18S Fw m18S Rv	TTGACGGAAGGGCACCACCAG GCACCACCACCCACGGAATCG	Mouse

□AA□b □□ ı□beI□|□nA□eI□ □I|ı

□

??????

?

? ? ? ??? ??? ????? ????? ??? ? ????? ????

?

Primer name	Sequence	Genome coordinates	Amplicon size
GCGTTGCATCATGCTTAGAA	Fw. Fragment 1 Luciferase	chr7:134142082+134143103	1022 bp
TCTGGAACTCGGGAGAAAAA	Rv. Fragment 1 Luciferase		
GCGTTGCATCATGCTTAGAA	Fw. Fragment 2 Luciferase	chr7:134142082+134143103	1022 bp
TCTGGAACTCGGGAGAAAAA	Rv. Fragment 2 Luciferase		
CAACCCCAGTTTGATGCTCT	Fw. Fragment 3 Luciferase	chr7:134144202+134145228	1027 bp
GCAACCTCCAAGTCAAGAGG	Rv. Fragment 3 Luciferase		
ATTTTTCTCCCGAGTTCCAGA	Fw. Fragment 2.1 Luciferase	chr7:134143083+134143531	449 bp
AATCCTGCGAGTGGGGTTTG	Rv. Fragment 2.1 Luciferase		
TTCGCTTTCCCACCAGATAC	Fw. Fragment 2.2 Luciferase	chr7:134143736-134144109	374 bp
CACGAGCACCTACCTTCCAG	Rv. Fragment 2.2 Luciferase		
CCAAGCTTTTCGCTTTCCCACCAGATAC	Fw. Fragment 2.3.1 Luciferase	chr7:134143927-134144117	191 bp
CCAAGCTTGGAGCCTTCTGATTGGTTGC	Rv. Fragment 2.3.1 Luciferase		
CCAAGCTTTTCGCTTTCCCACCAGATAC	Fw. Fragment 2.3.2 Luciferase	chr7:134143985-134144117	133 bp
CCAAGCTTGGGATGCTTCTTCCGCCT	Rv. Fragment 2.3.2 Luciferase		
CCAAGCTTTTCGCTTTCCCACCAGATAC	Fw. Fragment 2.3.1 Mut Luciferase	chr7:134143927-134144117	191 bp
CCAAGCTTGGAGCCTTCTGAATCGTTGC	Rv. Fragment 2.3.1 Mut Luciferase		

?

?AA?b??a????I ?| ??eI ? ?I | ?N???Bc????An ????nnI ??A?B??A?? e???nA?| ?D?h??????l ?? ?AB?a | ??? ?A?a????I ?| ???be?I ? ?AB¦t?

Primer name	Sequence
Rv.CRISPR_AKR1B1.CpGi.GUIDE A	CACCGAGCCTTCTGATTGGTTGCGCTGG
Rv.CRISPR_AKR1B1.CpGi.GUIDE A	AAACCCAGCGCAACCAATCAGAAGGCTC
Fw.CRISPR_AKR1B1.CpGi.GUIDE B	CACCGCCGCTGCGCGAAGGAGCCTTCTG
Rv.CRISPR_AKR1B1.CpGi.GUIDE B	AAACCAGAAGGCTCCTTCGCGCAGCGGC
Fw.CRISPR_NF-Y.Gene.GUIDE A	CACCGCCACTGACCTGCACCATTAATGG
Rv.CRISPR_NF-Y.Gene.GUIDE B	AACCCATTAATGGTGCAGGTCAGTGGC
Fw.CRISPR_NF-Y.Gene.GUIDE B	CACCGTGGCCAACCCATCATGGTCCAGG
Rv.CRISPR_NF-Y.Gene.GUIDE B	AACCCTGGACCATGATGGGTTGGCCAC
Fw.CRISPR_CEBPB.Gene.GUIDE A	CACCGTCGGCACCGCCTGGTAGCCGAGG
Rv.CRISPR_CEBPB.Gene.GUIDE A	AACCCTCGGCTACCAGGCGGTGCCGAC
Fw.CRISPR_CEBPB.Gene.GUIDE B	CACCGTCCTCCTCGTCCAGCCCGCCCGG
Rv.CRISPR_CEBPB.Gene.GUIDE B	AACCCGGGCGGGCTGGACGAGGAGGAC

Gene	450k Methylation Probe
ABCG5&8	cg26706238
ADH1	cg23949936
ADH4	cg12011299
AKR1B1	cg02215070
AKR1B10	cg11693019
AKR1B15	cg27018005
AKR1C1	cg10568634
AKR1C2	cg18841653
AKR1C3	cg12530994
AKR1C4	cg09272256
ALDH1A1	cg01812894
ALDH1A2	cg02888131
ALDH1A3	cg21359747
AWAT1	cg20220865
AWAT2	cg23904115
BCDO2	cg10503334
BCMO1	cg01678878
CD36C	cg06719671
CRABP1	cg22496102
CRABP2	cg03734660
CYP26A1	cg22536554
CYP26B1	cg20097440
CYP26E1	cg14250048
DHRS3	cg01083689
LPLC	cg07964216
LRAT	cg26084529
NPC1L1	cg06907626
RARA	cg24957657
RARB	cg12479047
RARG	cg08104462
RBP1	cg11027570
RBP4	cg18585903
RBP7	cg18086187
RDH10	cg20179892
RDH11	cg02218134
RDH12	cg21238221
RDH5	cg02192520
RDH8	cg25661884
RETSAT	cg02301651
RXRA	cg14654324
RXRB	cg08949668
RXRG	cg22675486
SCARB1	cg00894440
SULT1A1	cg15812873

Primer name	Sequence
Fw.Fragment1.AKR1B1.ChIP	GCTCTTTCTGCTCCATCACC
Rv.Fragment1.AKR1B1.ChIP	TCCAGGAGGGAGCACTTTTA
Fw.Fragment2.AKR1B1.ChIP	CGGCGCGTACCTTTAAATAG
Rv.Fragment2.AKR1B1.ChIP	GAAAGAATCCGCTGCCACTA
Fw.Fragment3.AKR1B1.ChIP	TCTACCGGCTTGAGATGCTT
Rv.Fragment3.AKR1B1.ChIP	GGAAAGTGGGCTTCACAGAC
Fw.Fragment4.AKR1B1.ChIP	TTAAGGGCCTGGTGTTTCAC
Rv.Fragment4.AKR1B1.ChIP	AATGGAAGAATTGCCCACAC
Fw.Fragment5.AKR1B1.ChIP	TTTTCTGAGCCCCAACAATC
Rv.Fragment5.AKR1B1.ChIP	CCTTCACAAGAGGGAAGCAC
Fw.Fragment6.AKR1B1.ChIP	CCTCCCTCCTCCAACTCTTC
Rv.Fragment6.AKR1B1.ChIP	ATCCATTCCAGGCTTTTGTG
Fw.Fragment7.AKR1B1.ChIP	TTGCTCGGTTACGACATTTG
Rv.Fragment7.AKR1B1.ChIP	TTCCCCTCCAGAAGACAGTG
Fw.Fragment8.AKR1B1.ChIP	CATGCTGCCATCTGTCAACT
Rv.Fragment8.AKR1B1.ChIP	CCAACCCCTCCTGTGAGATT
Fw.Fragment9.AKR1B1.ChIP	CCTGCCAGAATGAACCCTAA
Rv.Fragment9.AKR1B1.ChIP	TGGAATTAAGCCTGCCAATC
Fw.Fragment10.AKR1B1.ChIP	CATTCCCATTTCCAGGTTTG
Rv.Fragment10.AKR1B1.ChIP	TGGGCTAGAAATCGTTGACC
THOC3 Fw ChIP	TTCTTCTGGGGTTGTTCCGTAATC
THOC3 Rv ChIP	CACACCCGCTAGCCCTTTTCAT
CPLX2 Fw ChIP	CCTCCAGACCCCCACCCCATCC
CPLX2 Rv ChIP	AGGCTTCCCCGCGGCTTCTCAG
GADPH Fw ChIP	TTGCAACCGGGAAGGAAA
GAPDH Rv ChIP	TAGCCTCCGTCCAGCTGACTT
16CEN Fw ChIP	CAGTTTGCCAGATAGTCTCTTT
16CEN Rv ChIP	GAGACATTTGGGAAGGTCACTGAAT

□

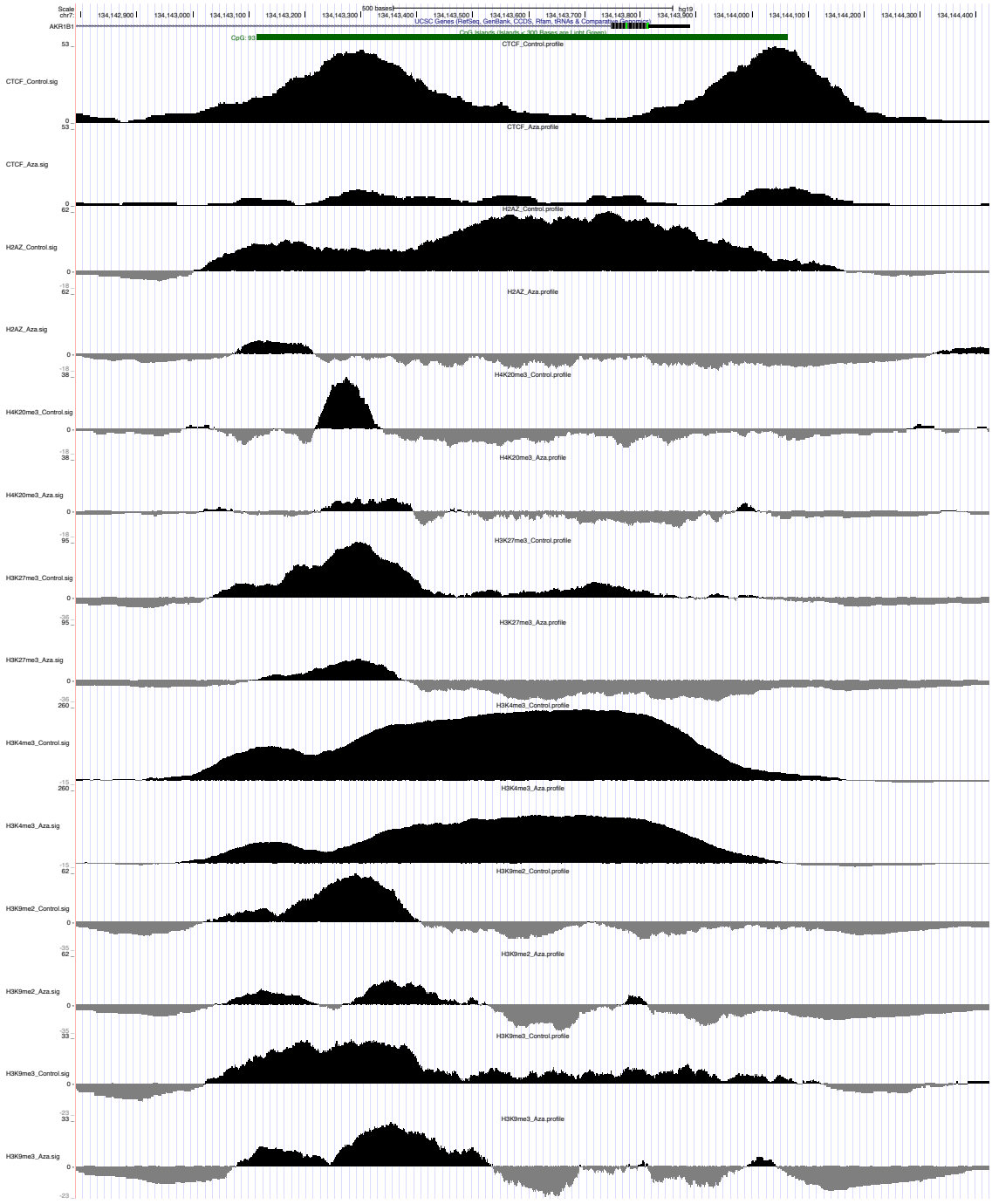
□□□ □□□□□□□□□□□□□□□□□□□□ □ □□□□ □□□□

□

Gene	Expression probe	
AKR1B1	A_23_P258190	□
AKR1B10	A_24_P129341	□
AKR1C3	A_23_P138541	□
ALDH1A1	A_23_P83098	□
ALDH1A2	A_24_P73577	□
CRABP1	A_33_P3326483	□
CYP26A1	A_23_P138655	□
CYP26B1	A_23_P210101	□
DGAT1	A_23_P112162	□
DHRS4	A_23_P14376	□
LRAT	A_32_P113066	□
RARA	A_32_P5251	□
RARB	A_24_P3243	□
RBP1	A_23_P257649	□
RETSAT	A_23_P209944	□
RXRA	A_23_P423197	□
RXRG	A_23_P23292	□
SULT1A1	A_24_P262201	□

□AA□b□□□□□□□□□□□□AB□□pCh; □□ □□In□II□o□ □beI□||□nA□eIn□□|t□□ □ □ □

ANNEX X. CHIP-SEQ DATA



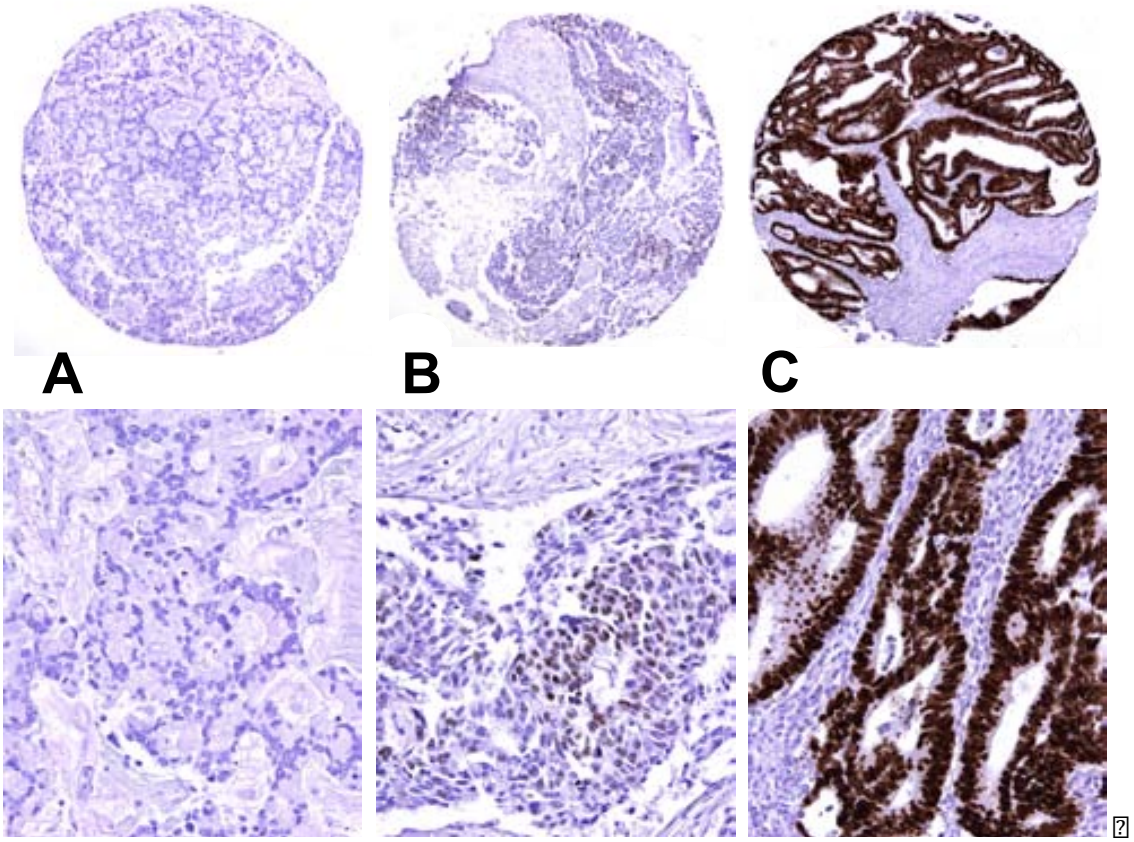
Annex X | ChIP-seq data on HCT116 control and treated with 5-AzadC.

□

□□□ □□□□ □□□□□□□□□□ □□□

□

□



□AA□b□□□□□C) □B□□A□A□□□AB□□□□□ □| □A□□oD□□□□hI □□□□y□yh□□A□□□□y□y□eInB□□A□□beI□| | □nAt□

□

□

□

□

□

ANNEX XII. ARTICLE

Regina Mayor, Mar Muñoz, Marcel W. Coolen, **Joaquin Custodio**, Manel Esteller, Susan J.Clark and Miguel A. Peinado. *Dynamics of bivalent chromatin domains upon drug induced reactivation and resilencing in cancer cells.* Epigenetics 2011, 9: 1138-1148.

Epigenetics 6:9, 1138-1148; September 2011; © 2011 Landes Bioscience

Dynamics of bivalent chromatin domains upon drug induced reactivation and resilencing in cancer cells

Regina Mayor,[1] Mar Muñoz,[1] Marcel W. Coolen,[2,3] Joaquin Custodio,[1] Manel Esteller,[4] Susan J. Clark[2,5] and Miguel A. Peinado[1,*]

[1]Institut de Medicina Predictiva i Personalitzada del Càncer (IMPPC); Badalona; Barcelona, Spain; [2]Epigenetics Group, Cancer Program; Garvan Institute of Medical Research; Sydney, Australia; [3]Department of Human Genetics; Nijmegen Centre for Molecular Life Sciences (NCMLS); Radboud University Nijmegen Medical Centre; Nijmegen, The Netherlands; [4]Cancer Epigenetics and Biology Program (PEBC); Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet; Barcelona, Spain; [5]St. Vincent's Clinical School; University of NSW; Sydney, New South Wales, Australia

Key words: chromatin remodeling, DNA methylation, epigenetic drugs, colorectal cancer, epigenetic silencing

Epigenetic deregulation revealed by altered profiles of DNA methylation and histone modifications is a frequent event in cancer cells and results in abnormal patterns of gene expression. Cancer silenced genes constitute prime therapeutic targets and considerable progress has been made in the epigenetic characterization of the chromatin scenarios associated with their inactivation and drug induced reactivation. Despite these advances, the mechanisms involved in the maintenance or resetting of epigenetic states in both physiological and pharmacological situations are poorly known. To get insights into the dynamics of chromatin regulation upon drug-induced reactivation, we have investigated the epigenetic profiles of two chromosomal regions undergoing long range epigenetic silencing in colon cancer cells in time-course settings after exposure of cells to chromatin reactivating agents. The DNA methylation states and the balance between histone H3K4 methylation and H3K27 methylation marks clearly define groups of genes with alternative responses to therapy. We show that the expected epigenetic remodeling induced by the reactivating drugs, just achieves a transient disruption of the bivalent states, which overcome the treatment and restore the transcriptional silencing approximately four weeks after drug exposure. The interplay between DNA methylation and bivalent histone marks appears to configure a plastic but stable chromatin scenario that is fully restored in silenced genes after drug withdrawal. These data suggest that improvement of epigenetic therapies may be achieved by designing strategies with long lasting effects.

Introduction

The combination of genetic and epigenetic lesions in cancer cells results in altered gene expression profiles. DNA methylation and different modifications on histone tails are the two principal forms of epigenetic regulation and both are largely disturbed in cancer cells.[1,2] Many unmethylated CpG islands often become hypermethylated during cancer progression resulting in epigenetic inactivation of the associated gene. This appears to be a principal mechanism of tumor suppressor inactivation, but not all silenced genes are considered to have antitumor activities and heterogeneous profiles have been identified for different tumor types.[1,2] Most CpG island hypermethylations appear as isolated and independent events, but the concurrent hypermethylation of neighboring CpG islands accompanied by global remodeling of the chromatin in large chromosomal regions may also occur.[3,4] This is a phenomenon known as Long Range Epigenetic Silencing (LRES) and has been reported to affect multiple chromosomal regions in different tumor types, including colorectal, prostate and breast cancer.[3-10]

Histone modifications can lead to either activation or repression depending upon which residues are modified and the type of modification.[11] For instance, in histone 3, dimethylaton and trimethylation of lysine 4 (H3K4me3) are associated with transcriptional activity, while trimethylation of lysine 27 (H3K27me3) is characteristic of silenced promoters. This later mark is driven by the presence of the polycomb repressor complex and is believed to direct de novo methylation in cancer-related silenced genes.[12,13] Different studies have reported that in genes that become silenced in cancer cells, the repressive mark H3K27me3 co-exists with active marks (H3K4me3 and H3K4me2).[6,14,15] These two opposite modifications participate in the mitotic inheritance of lineage-specific gene expression programs and have key developmental functions. Its co-localization is considered characteristic of stem cells and is believed to keep developmental regulator genes poised for induction.[16-21] More recently, it has been shown that bivalent chromatin domains are also prone to DNA hypermethylation

*Correspondence to: Miguel A. Peinado; Email: map@imppc.org
Submitted: 04/18/11; Accepted: 07/15/11
DOI: 10.4161/epi.6.9.16066

in aging, providing an epigenetic link between the processes of aging and cancer.[22,23]

Stem-cell like signatures are mimicked by cancer cells and contribute to define their properties.[24-26] Resetting of epigenetic profiles appears as a novel and promising therapeutic strategy in cancer,[1,27] but a better understanding of the mechanisms setting epigenetic memory and how the chromatin landscapes are either maintained or modified is required to make it a broad reality. This is especially important, since most of the critical genes in cell programming and targets of epigenetic therapies show bivalent chromatin signatures in both stem and cancer cells.

Genome-scale approaches just illustrate the ubiquitous and complex reorganization of epigenomic profiles upon treatment with gene-reactivating drugs, as shown in a recent study.[28] Due to the heterogeneity of chromatin landscapes before and after the treatment, insights into the regulatory mechanisms can be only obtained from a detailed analysis of specific loci. An interesting example is a recent investigation in which a clear picture of chromatin and gene expression dynamics has been achieved by analysis of green fluorescent protein expression under the control of a methylated cytomegalovirus promoter after treatment with 5-AzadC.[29] This artificial setting illustrates the different epigenetic events associated with changes in gene activity, but it remains unsolved which mechanism or driving signal retains the "switch-off" memento and is able to reset the original state of the chromatin. Si et al. suggest that residual DNA methylation near the CMV-GFP locus could play a role. It is also unknown if the dynamics of chromatin remodeling and DNA methylation observed in this system apply to endogenous genes.[29]

Although important insights into the mechanisms of epigenetic control have been made, the interaction between different coexisting epigenetic marks and the dynamics of the events responsible for the repression/activation of chromatin are still poorly known.[11] Here we have characterized the epigenetic profiles of two chromosomal regions undergoing long range epigenetic silencing in most colorectal cancers, how these profiles are affected by epigenetic drugs and how they are reset upon drug withdrawal and maintenance in in vitro culture. Our results illustrate the dominant nature of DNA methylation and bivalent histone marks, which endure drug-induced changes in transcriptional activity.

Results

DNA methylation, expression and histone patterns define three types of genes in long range epigenetic silencing (LRES) regions. In previous studies we have shown that LRES affects chromosomal regions 2q14.2 and 5q35.2 in most colorectal cancers.[3,6,30] Although genetic activity is downregulated all along the LRES region, uneven epigenetic profiles are likely to define different chromatin domains. To better characterize the epigenetic regulation of the diverse chromatin domains in LRES we have investigated gene expression and epigenetic profiles in nine and six genes embedded in chromosomal regions 2q14.2 and 5q35.2, respectively, in HCT116 colon cancer cells. An overview of the chromosome maps and the molecular profiles of both regions are depicted in **Figure 1 and Supplemental Figure 1**.

According to the expression levels and the epigenetic state of their promoter, the genes could be classified into three groups (**Fig. 2**). First group consists of silenced genes exhibiting methylated CpG island-promoter; this group includes EN1, SCTR and INHBB (2q14.2) and HRH2, CPLX2 and SNCB (5q35.2) that express at low levels in normal colon cells.[3,6] A second group consisted of genes with unmethylated CpG island-promoter, and includes DDX18, INSIG2, PTPN4, RALB, TSN (2q14.2), SFXN1 and THOC3 (5q35.2). These genes tend to be downregulated in colorectal tumors (as compared with the normal colon cells) but still retain high expression levels.[3,6] Finally, a third group includes MARCO (2q14.2) and PCLKC (5q35.2) genes that are low expressed and do not contain a CpG island in the promoter region (**Fig. 1 and Sup. Fig. 1**).

Histone modification profiles were also analyzed by chromatin immunoprecipitation (ChIP) and each one of the groups exhibited characteristic profiles (**Fig. 2**). A good agreement between H3K4me2 and H3K4me3 profiles was observed (**Sup. Fig. 2**) and for simplification, H3K4me2 data have been represented in figures and Tables throughout the manuscript.

The concurrent presence of repressive H3K27me3 and active H3K4me2/me3 chromatin marks was observed in genes of the first group containing a DNA hypermethylated promoter (**Fig. 2**). This is consistent with a bivalent state, as it has been reported previously for some of these genes[6] and in other genes that become hypermethylated in cancer.[14,15,31] Genes with these features (N1, SCTR and INHBB, HRH2, CPLX2 and SNCB) are referred as MBV (methylated and bivalent). Noteworthy, the same genes that showed higher levels of H3K4me2/me3 within this group, presented the highest levels of H3K27me3, as it is the case of CPLX2 and INHBB (**Fig. 2**). H3K9Ac was absent from MBV genes and some of them (i.e., EN1, SCTR) exhibited low levels of H3K9me2 (**Sup. Fig. 3**). The H3K27me3/H3K4me2 ratio was above 0.5 in all MBV genes.

The activated state of genes of the second group was associated with high levels of H3K4me2/me3 (**Fig. 2**) and H3K9Ac marks (**Sup. Fig. 3**). However, there was no correlation between the gene expression levels and the amount of active marks when comparing genes among them. As expected, H3K27me3 mark was absent on the promoters of these genes. The H3K27me3/H3K4me2 ratio was at least ten fold lower than the minimum value observed in MBV genes. Genes of this group are referred as ACTIVE, because they were expressed in all experimental situations, although the levels were not maintained.

MARCO and PcLKC (also known as PCDH24) genes displayed histone modification profiles compatible with bivalent chromatin and accordingly their gene expression levels were very low. MARCO exhibited the highest levels of repressive marks H3K27me3 and H3K9me2 among all analyzed genes (**Fig. 2 and Sup. Fig. 3**). Genes of this group do not contain a CpG island in the promoter and are referred as NoCpGi.

Dynamics of epigenetic profiles in LRES regions upon 5-AzadC/TSA treatment and drug withdrawal. It has been repeatedly shown that epigenetically silenced genes can be

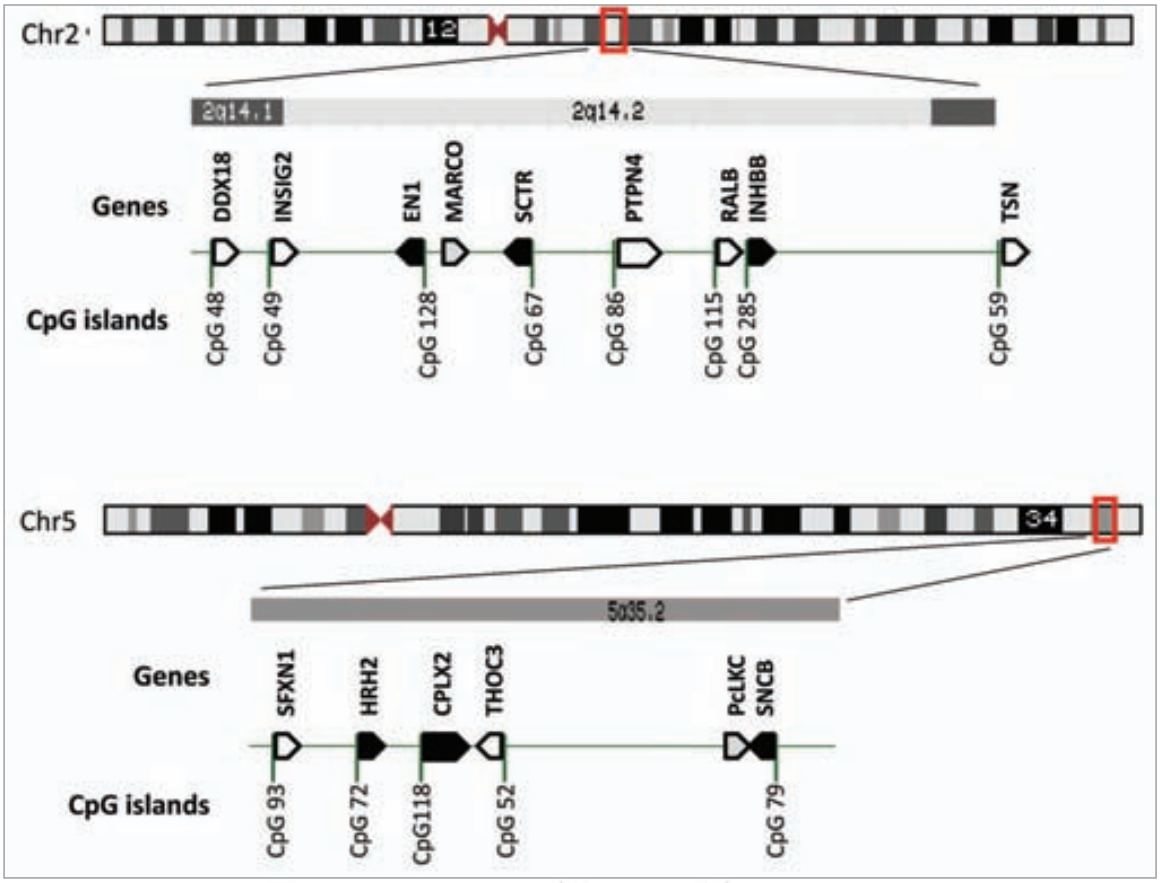


Figure 1. Scheme of the chromosomal regions analyzed at epigenetic level. Genes are depicted as arrowheads (indicating the transcription direction) and coded according to its epigenetic signature in HCT116 cells. Closed black arrowhead: DNA methylated and bivalent chromatin marks (MBV), Open arrowhead: Active genes, Grey arrowhead: No CpG island.

reactivated by treating the cells with the demethylating agent 5-Aza-2'-deoxycytidine (5AzadC) and the inhibitor of histone deacetylases Trichostatin A (TSA).[27,32] More recently it has also been shown that bivalent domains in silenced genes are not resolved in spite of the reactivation upon drug treatment.[6] To get a better picture of the events associated with gene activation and silencing in these chromosomal regions we performed a time course experiment after treating the cells separately with 5AzadC or TSA, and with a combination of both. Cells were analyzed at days 1 (the same day the drugs were removed), 16, 22 and 29 after treatment. As expected, 5AzadC alone or in combination with TSA resulted in global demethylation of the seven hypermethylated genes as determined by direct bisulfite sequencing and melting curve analysis (data not shown). For a subset of genes (EN1, SCTR and INHBB), a quantitative determination was performed using the MassCLEAVE™ DNA methylation assay system. A 60% of loss of methylation was observed just after the treatment (day 1) in 5AzadC treatment alone (data not shown) and after the co-treatment that affected most CpG sites (**Fig. 3**). Some CpG sites (SCTR 6_7 and SCTR 40) exhibiting lower methylation levels (about 50%) in untreated cells did not display changes upon treatment.

Upon drug withdrawal, a remethylation to >90% was reached in 5AzadC treated cells after 16 days (data not shown), while the co-treatment appeared to delay the full remethylation, that was achieved two weeks later (day 29) (**Fig. 3**). TSA treatment alone did not affect DNA methylation (data not shown).

5AzadC treatment restored the expression of genes with a hypermethylated promoter (MBV group) (data not shown), although the co-treatment with TSA resulted in a more substantial reactivation, that was maintained for about two weeks but returned to original levels 4 weeks later (**Fig. 4**), consistent with the DNA methylation profiles. H3K4me2/3 and H3K27me3 marks were also increased in response to the drugs, which is suggestive of balanced dynamics. After drug withdrawal, the active chromatin mark H3K4me2/me3 exhibited a time dependent decrease that, after four weeks, reached levels similar to the untreated cells for genes of the chromosomal band 5q35.2 and even lower for those of the 2q14.2 region. On the other hand, the repressed chromatin mark H3K27me3 displayed a heterogeneous profile and in some genes retained higher levels all along the time course study (**Fig. 4**). This result could be interpreted as a partial disruption of the balance between active and inactive marks during the resilencing process. Nevertheless, the alternation of different cell populations with distinct chromatin signatures might also contribute

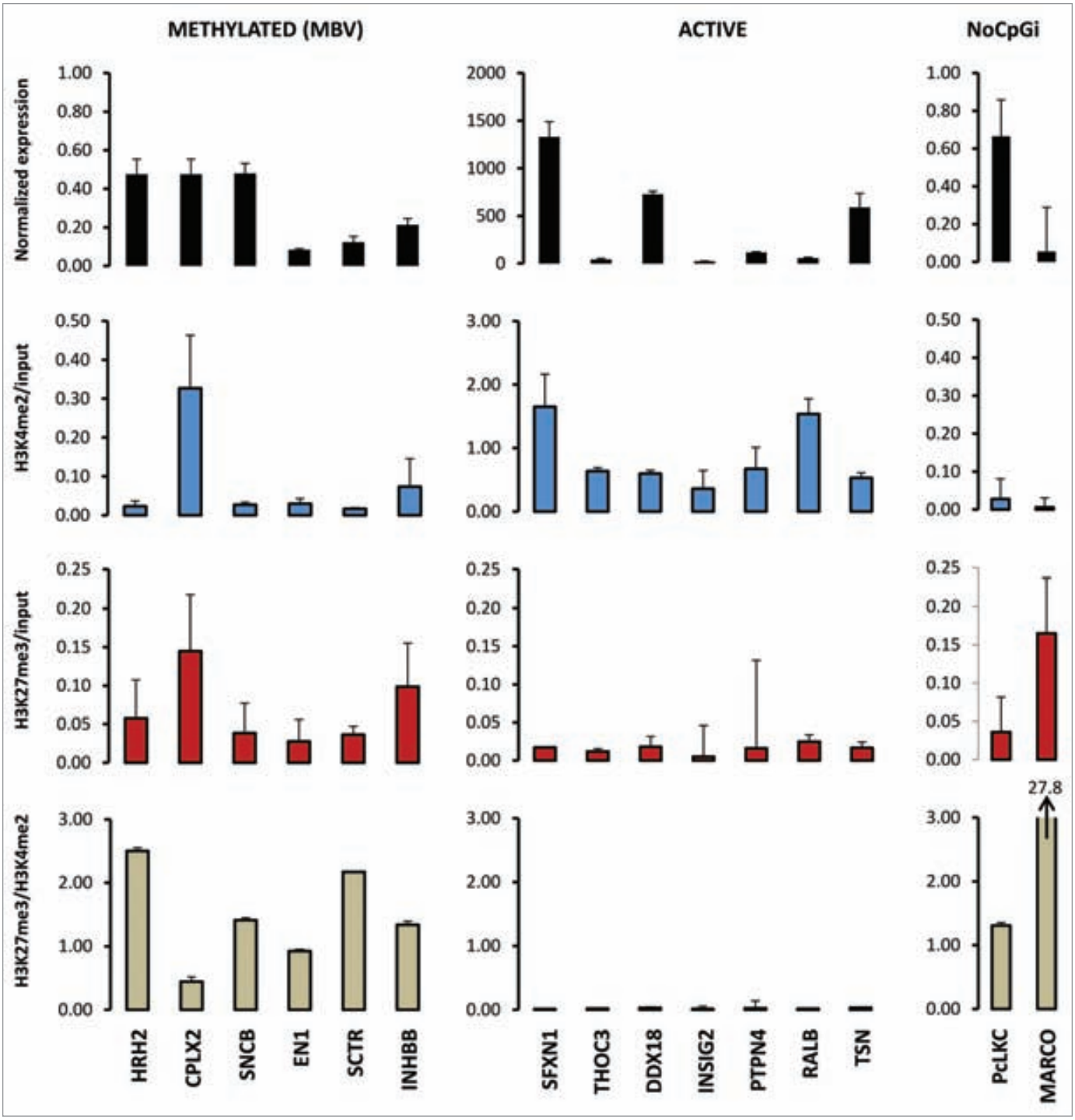


Figure 2. Gene expression and chromatin modification patterns of genes classified into three groups according to their epigenetic signatures in HCT116 cells. MBV corresponds to silenced genes exhibiting a DNA methylated promoter embedded in bivalent chromatin (presence of methylation in H3K27 and H3K4); ACTIVE corresponds to unmethylated and expressed genes; NoCpGi corresponds to genes without CpG island. Gene expression levels were normalized to 18S. H3K4me2 and H3K27me3 ChIP levels were normalized to input and the ratios between both marks are shown. All quantifications were performed by real-time PCR in triplicate. Error bars indicate standard deviation. Note that gene expression and H3K4me2 values are represented in a different scale for ACTIVE genes due to their high values as compared with the other two groups. The regional profiles of genes arranged by chromosomal position in shown in Supplemental Figure 1.

to the observed changes. H3K4me2 and H3K4me3 active marks showed parallel changes in most of the experiments in which both were analyzed, although H3K4me2 changes appeared to occur at a slower pace than H3K4me3 (**Sup. Fig. 2**).

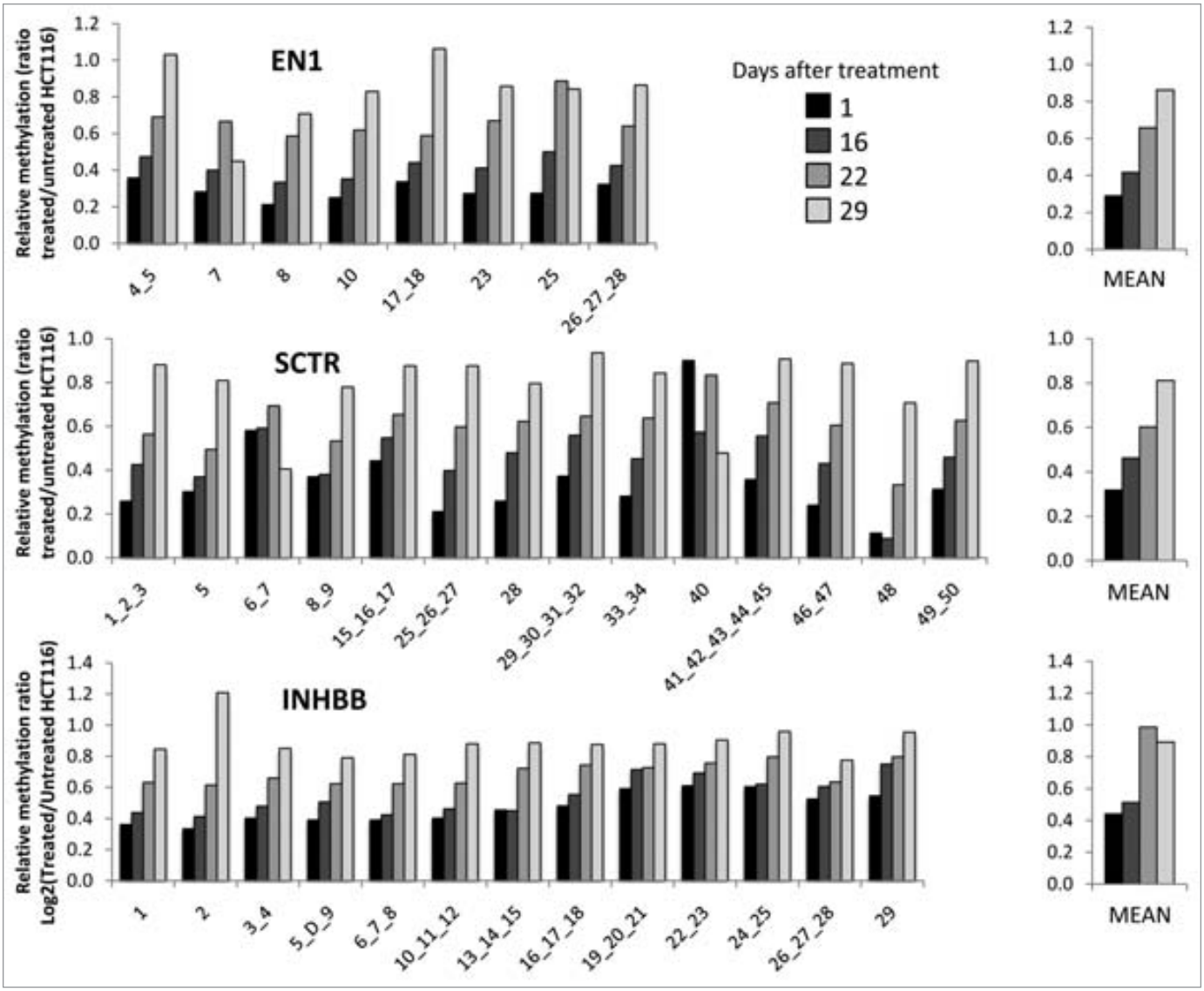


Figure 3. Relative methylation levels of the named CpG sites contained in the EN1, SCTR and INHBB amplicons as analyzed by MassCLEAVE™ after 5AzadC/TSA co-treatment. A pool of three direct bisulfite PCR products was used for analysis. Represented values correspond to the ratio of treated versus untreated HCT116 cells (Y axis). The gradient scale represents the different time points corresponding to days 1, 16, 22 and 29 after treatment (from dark to light). Mean methylation values of all the CpG sites are represented in the right part.

On the other hand, ACTIVE genes exhibited a completely different behavior in response to the various treatments. Most of them suffered a clear decrease in the transcription rates by day 1, either with 5AzadC or TSA alone (data not shown), or with the combination of both (**Fig.** 4). Reduction of expression was paralleled by an initial decrease in H3K4me2/me3 mark. The initial expression levels were recovered 16 days after the drug withdrawal as the H3K4me2/me3 mark did. H3K27me3 mark remained low or undetectable in these genes along the time course, although its sporadic presence could be detected in some genes (**Fig.** 4).

Finally, MARCO and PcLKC gene expression was slightly induced by the treatment. This reactivated status was maintained over 16 days. Afterwards, transcription rates started to decrease to the low or undetectable levels of untreated cells (**Fig.** 4). Changes in chromatin modifications were also minor, with slight increases of H3K4me2/3 and H3K27me3 in PcLKC along the time course (**Fig.** 4).

Notably, for all three groups of genes, an enrichment of H3K4me2/me3 levels was found after treating the cells with TSA alone (**Sup. Fig. 2B** and data not shown), but this effect had disappeared by day 16 for most of the genes, and was not reflected on the expression patterns. Regarding other histone modifications analyzed, H3K9Ac was increased in most genes after TSA or the combined treatment but not when treated with 5dAzaC alone as expected (data not shown). H3K9me2 and total H3 levels displayed small variations without a consistent pattern.

Maintenance of chromatin states in DNA methylation deficient cells. Since DNA demethylation appeared to reactivate silenced genes but was not able to resolve bivalent chromatin,

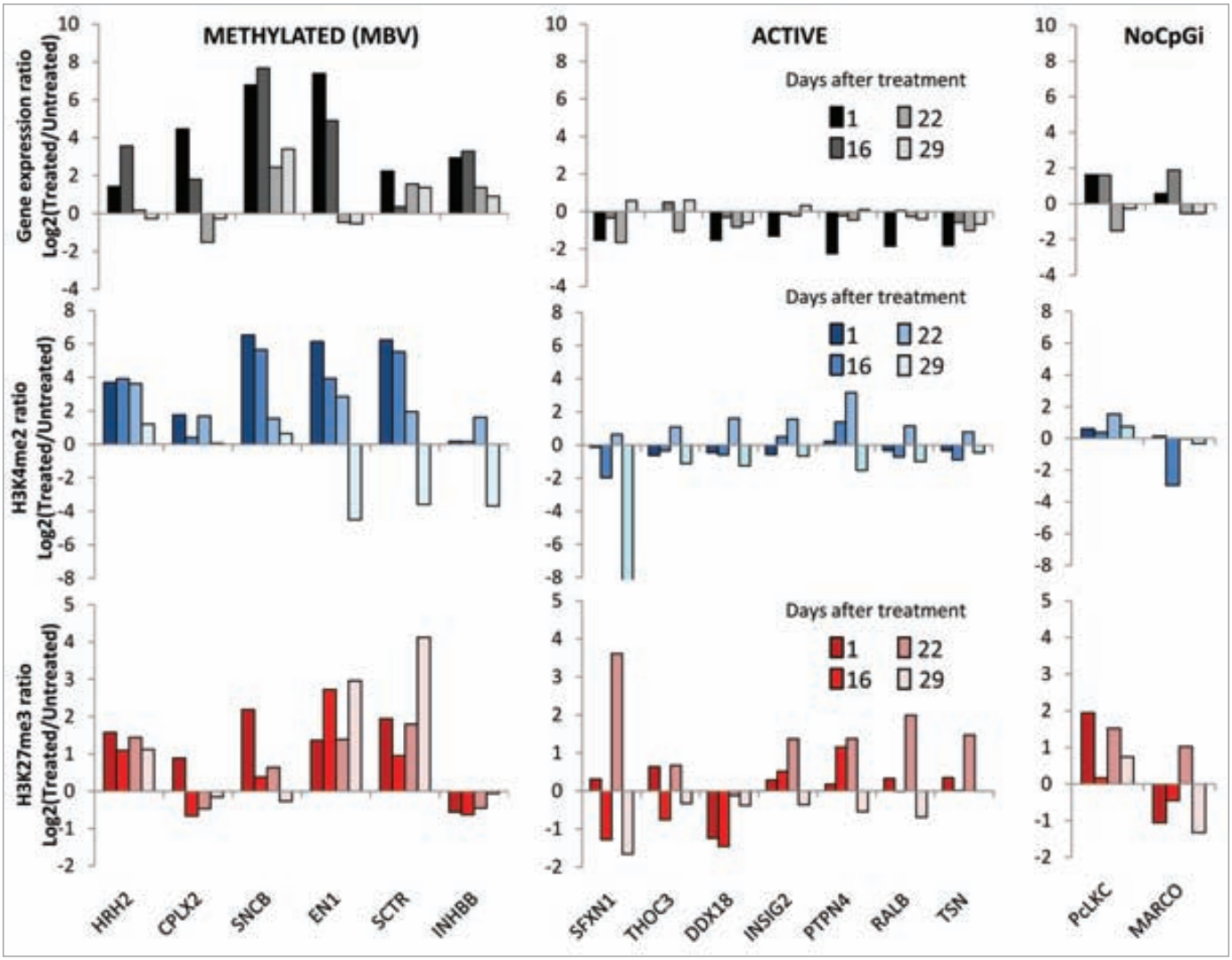


Figure 4. Relative expression and histone mark (H3K4me2 and H3K27me3) levels in the three classes of genes after 5-AzadC and TSA co-treatment. Relative values were calculated as the log2 of the treated/untreated ratio. The gradient scale indicates the different time points corresponding to days 1, 16, 22 and 29 after treatment (from dark to light).

we wondered if stable demethylation will be sufficient to override the epigenetic memory. We analyzed the DKO cells, a variant of HCT116 colon cancer cells which is deficient for DNA methyltransferases,[33] although retaining some residual DNMT1 activity.[34] As expected, hypermethylated genes (MBV) in the wild-type HCT116 were fully demethylated in DKO cells (data not shown), and their expression was restored to higher levels than in the parental cells, although two of the genes, EN1 and CPLX2, exhibited a limited reactivation (**Fig. 5A**). As in drug induced gene reactivation, DKO cells exhibited an overall increase of the H3K4me2/me3 and H3K27me3 marks in the demethylated MBV genes (**Fig. 5A**), which suggests a dynamic rebalancing of these two histone modifications enduring the retention of the bivalent states. These results are in agreement with previous observations in other genes.[15] Genes with unmethylated CpG islands (ACTIVE group) in the HCT116 cells and those without CpG island (noCpGi group) exhibited null or minimal changes in gene expression and histone modification marks (**Fig. 5A** and data not shown).

The treatment of DKO cells with TSA resulted in an increased expression of all the MBV genes (**Fig. 5B**) that was paralleled by a partial disruption of the balance between H3K4me2

Figure 5 (See opposite page). (A) Relative gene expression and histone modification profiles in DKO cells in regard to parental HCT116. Relative values were calculated as log2 of DKO/HCT116 ratio. Gene expression was normalized to 18S. ChIP results were normalized to the respective input fractions. Major differences are seen in the methylated bivalent genes (MBV) that exhibit DNA demethylation, higher expression and increased levels of histone marks H3K4me2 and H3K27me3. (B) Relative gene expression and modification profiles in DKO cells treated with TSA in regard to untreated DKO cells. Relative values were calculated as log2 of treated/untreated cells ratio. Gene expression was normalized to 18S. ChIP results were normalized to the respective input fractions.

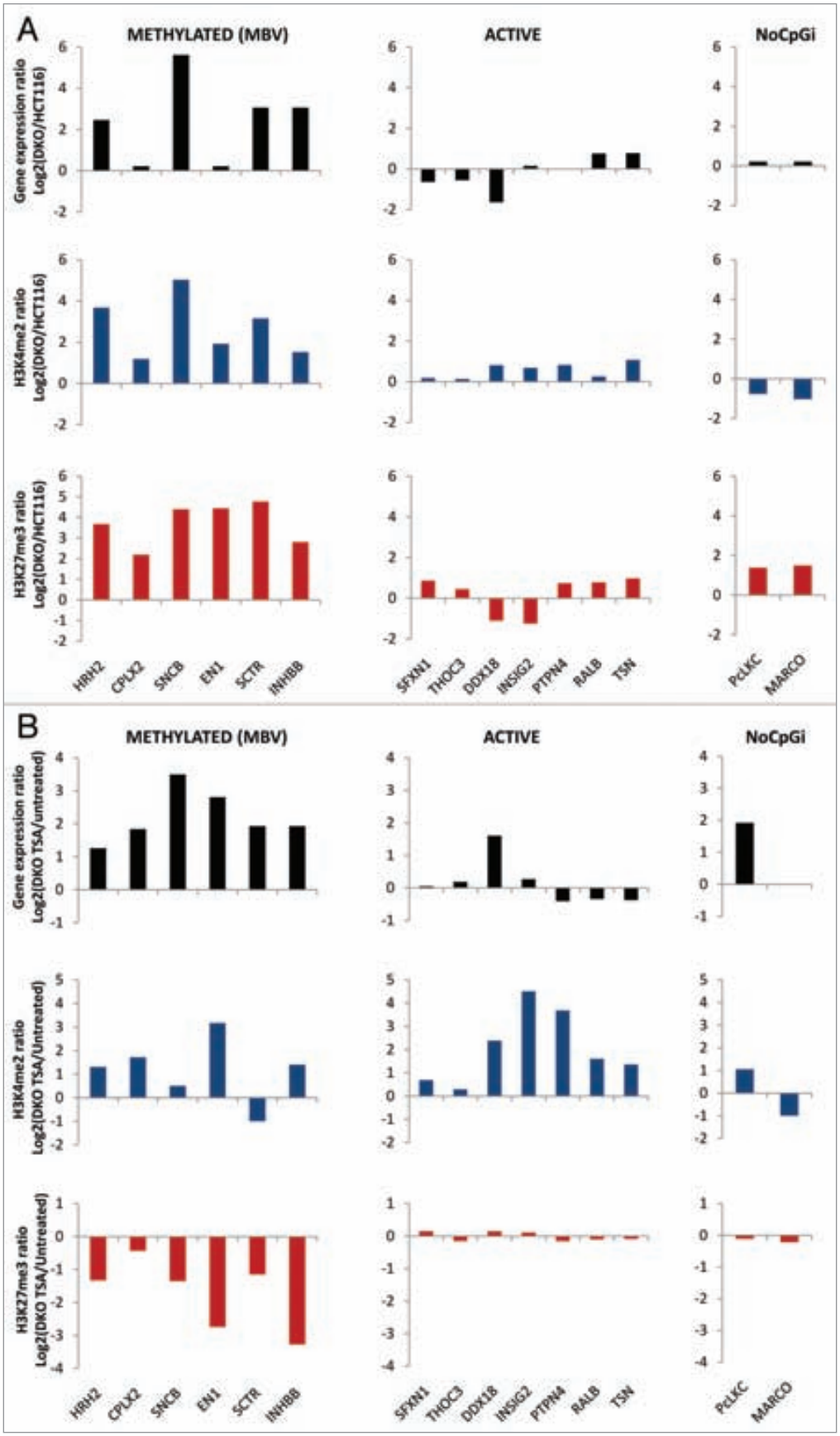


Figure 5. For figure legend, see page 6.

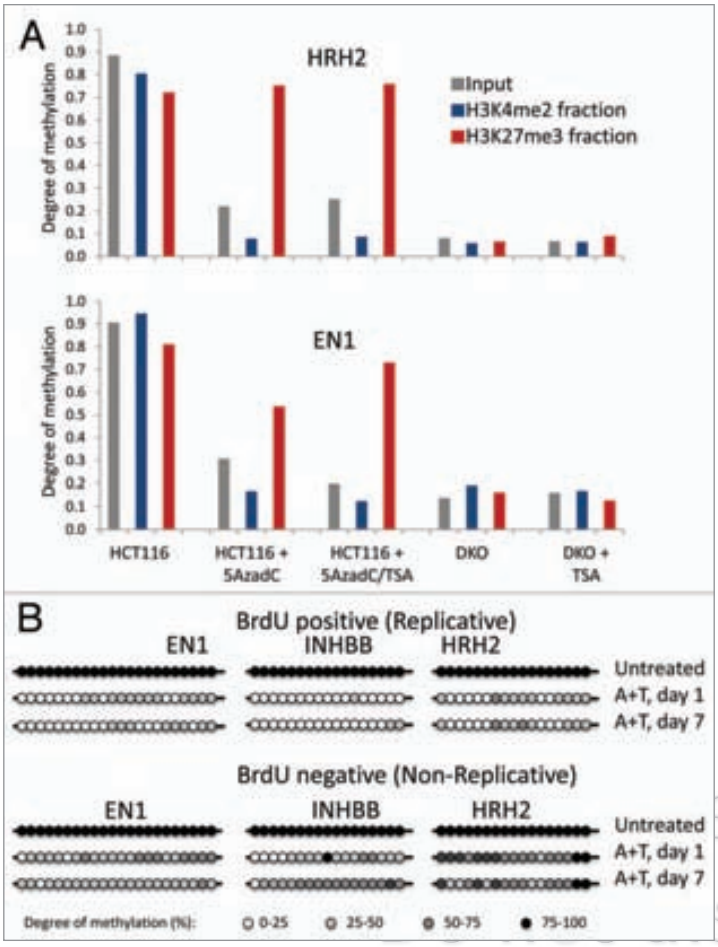


Figure 6. DNA methylation analysis of cell fractions after 5AzadC/TSA treatment. (A) MassCLEAVE™ DNA methylation analysis of input, H3K4me2 and H3K27me3 immunoprecipitated DNA fractions in EN1 and HRH2 gene CpG islands. Each bar represents the mean methylation (1.0 is full methylation) of all CpG sites contained in the amplicon analyzed in untreated HCT116, 5AzadC and 5AzadC/TSA treated cells and DKO and TSA DKO treated cells. (B) Direct bisulfite sequencing of CpG island regions corresponding to three MBV genes in HCT116 cells after BrdU labeling and FACS.

and H3K27me3 marks (higher H3K4me2 levels but lower H3K27me3) (**Fig. 5B**). The rest of genes were not affected by TSA, with the exception of DDX18 and PcLKC, which were also upregulated. A global increase of H3K4me2 levels was also seen in most genes (**Fig. 5B**).

These results suggest that retention of the bivalent nature of chromatin in the MBV genes is preserved even in the absence of an efficient DNA methylation system, although the balance between H3K4me2/me3 and H3K27me3 marks is partially disrupted after inhibition of histone deacetylases with TSA.

Epigenetic profiles of different chromatin and cell fractions. In vitro treatments with epigenetic drugs clearly show that silenced genes can be reactivated but do not illustrate the direct associations among the different components of the epigenetic code and gene activity in response to the drug. This is because changes observed in response to drugs may be driven by only a fraction of the cells and the fraction of cells exhibiting changes may be different for every assessed parameter. To determine whether drug induced DNA demethylation is directly associated with the changes detected in active and repressive histone modifications, we analyzed DNA methylation in H3K4me2 and H3K27me3 chromatin fractions in untreated HCT116 and DKO cells, 5-AzadC and 5-AzadC/TSA treated HCT116 cells and TSA treated DKO cells. DNA methylation profiles of five MBV genes (EN1, SCTR, INHBB, CPLX2 and HRH2) were analyzed in the Input, H3K4me2 and H3K27me3 ChIP fractions by bisulfite sequencing and MassCLEAVE™.

High levels of DNA methylation were observed in all chromatin fractions analyzed in HCT116 cells, suggesting that DNA methylation was homogeneous and ubiquitous in the five genes analyzed. Illustrative results of HRH2 and EN1 genes have been represented (**Fig. 6A**). Upon treatment (5AzadC alone or in combination with TSA), marked differences between H3K4me2/me3 and H3K27me3 fractions were appreciated. A dramatic decrease in DNA methylation was observed in the input and the H3K4me2/me3 fraction of treated cells (20–30% and 10–30% of that found in untreated cells), but the H3K27me3 fraction retained a higher level of DNA methylation (40 to 100% of that found in untreated cells). This discordance suggests that drug treatment ensues in at least two different cell populations. The responsive cell population would be characterized by the loss of DNA methylation, gene re-expression, increased levels of H3K4me2, and probably by loss of the H3K27me3 mark. A second population of resistant cells would retain DNA methylation and probably increased levels of H3K27me3.

Next we wondered if cells retaining DNA methylation underwent replication. We analyzed the DNA methylation and histone marks in replicating and non-replicating cells after fluorescence assisted cell sorting (FACS) of cells cultured in medium containing bromodeoxyuridine (BrdU). All fractions exhibited DNA demethylation, although it was deeper in cells that have incorporated high levels of BrdU. After 1 week, the demethylation was maintained and the differences between BrdU positive and negative cells were still apparent (**Fig. 6B**). Combined analysis of histone modifications and DNA methylation was performed in cells labeled with propidium iodide. Cells were fractioned by FACS into two groups (G_0/G_1 and S/G_2). Both groups retained balanced levels of H3K27me3 and H3K4me3, although DNA demethylation tended to be deeper in H3K4me3 fractions and cells in S/G_2 phases. Equivalent results were obtained 1 day and 7 days after treatment (data not shown).

We do not know if gene reactivation of silenced genes takes place in all cells, but a deeper DNA demethylation is consistently associated with enrichment of the H3K4me2/me3 marks and, as expected, DNA replication. H3K27me3 mark was present in all cell fractions. DKO cells exhibited low methylation levels similar to drug treated cells and no differences among treatments and chromatin fractions.

Discussion

Reactivation of tumor suppressor genes that have undergone epigenetic silencing appears as a promising therapeutic strategy.[32,35] It has been shown that reactivated promoters are often slowly remethylated and the gene is resilenced after withdrawal of the DNA methylation inhibitor.[36-40] Recent discoveries have revealed chromatin scenarios linking stem cell and cancer biology [6,12,14,15,25] and have underscored the impact of neighboring epigenetic states on the regulation of chromatin stability.[3,6,9] To get insights into the dynamics of chromatin reactivation and resilencing in cancer cells submitted to epigenetic treatments, we have investigated the epigenetic profiles of two of these chromosome regions, 2q14.2 and 5q35.2, containing genes silenced in colorectal cancer. We have defined three different genetic compartments based on the epigenetic signatures of the genes.

The first group is represented by six genes undergoing epigenetic silencing in cancer cells (as denoted by DNA hypermethylation) and the retention of bivalent chromatin epigenetic marks (MBV genes). It is of note the bivalent nature of these genes in murine colon cells, that has been demonstrated by co-immunoprecipitation of H3K4me3 and H3K27me3 (unpublished data). When reactivated, the expression of these genes is restored but at low levels, similar to those detected in non-tumorigenic cells,[3,6] and histone modifications profiles are also changed to a more pronounced bivalent state (elevated H3K4me2/3 and H3K27me3 signals) as it has been previously reported in reference 6 and 40. Upon drug withdrawal, the bivalent marks are maintained, but the balance is partially disrupted by a predominance of the repressive mark H3K27me3, suggesting its contribution to the recovery of the silenced state. The overall data are consistent with a coordinated dynamics of bivalent chromatin signatures among the MBV genes Nevertheless, it should be noted that our analysis is limited to specific regions inside the CpG island. Regional profiles along and outside the promoter region may be also affected by the treatments and may exhibit different dynamics during the resilencing process. Future studies a genome scale should address this issue. The effects of the treatment are reversed at mid-term and a complete resilencing of the reactivated genes is accomplished three weeks after drug withdrawal, while the DNA remethylation is fully restored one week later for most of the CpG islands analyzed. The combined treatment resulted in an extended recovery period as compared with 5-AzadC alone. These results are consistent with a secondary nature of DNA methylation in regard to gene activity.[39,41] The percentages of demethylation and the rate of remethylation were of the same order as previously described for other genes.[36,38,39] While treatment with the histone deacetylases inhibitor agent TSA alone did neither induce DNA demethylation nor gene reactivation, it had a synergistic effect with 5-AzadC, in agreement with other studies.[42,43]

The use of cells deficient in DNA methyltransferase activity (DKO) has allowed us to contrast the results presented above in a more stable system. As it has been described before, most hypermethylated tumor suppressor genes in HCT116 cells, such as p16INK4a or TIMP3, are found demethylated and re-expressed in DKO cells as compared with the parental HCT116 cells.[33,44] Most HCT116 hypermethylated genes are reexpressed and exhibit higher levels of H3K4me2 and H3K27me3 marks in the DKO cells (**Fig. 5A**). These results are similar to what we obtain when HCT116 cells are treated with 5-AzadC alone (our data and ref. 40). As expected, DKO cells treated with TSA exhibited increased expression of genes hypermethylated in HCT116. Interestingly, TSA had deep effects on DKO cells by disrupting the balance between bivalent marks (**Fig. 5B**). While H3K4me2 is increased in most genes, the repressive H3K27me3 mark is lost from most of the genes in which promoter hypermethylation was present (MBV). Our results suggest that DNA methylation also contributes to maintain the repressive histone code in genes poised for silencing and confirm the role of DNA methylation as the foremost player in preserving the inactivation of silenced genes.[38] Besides that, retention of bivalent signatures probably acts as a memento and is likely to participate in the restoration of the silencing.[6]

Another interesting finding arises from the analysis of DNA methylation in different chromatin fractions and in various experimental conditions. At global level, all the MBV genes analyzed in HCT116 untreated cells show coexistence of histone marks H3K4me2 and H3K27me3 with fully methylated DNA (**Fig. 6A**). Nevertheless, upon DNA demethylation by treatment with 5AzaC, H3K27me3 chromatin fraction retains higher levels of DNA methylation, while the H3K4me2/me3 fractions are extensively DNA demethylated. This result could be interpreted as a partial disruption of bivalent domains that become active in a fraction of the cells (unmethylated DNA and H3K4me2), while reaching a new balance in "drug-resistant cells" (methylated DNA and bivalent marks).

It remains to be elucidated if the recovery of the silenced state after drug removal is due to chromatin re-remodeling or to a cell population renewal in which a reservoir of cells resistant to the drug (denoted by the presence of high levels of methylation associated with the H3K27me3 chromatin fraction) become prevalent in the cell population. While it seems clear the homogeneity of the parental cells, it could be hypothesized that the treatment induces the presence of mixed cell populations with different chromatin landscapes (active, silenced, bivalent). We tried to resolve this conundrum by coimmunoprecipitation of both histone marks H3K4me2 and H3K27me3, but it did not work in our hands, although we have been able to do it in murine cells (data not shown). Our data indicate that the retention of repressive marks is unlikely to be driven by a resting cell population, since epigenetic profiles of bivalent chromatin domains are similarly affected in replicating and non-replicating cells. In agreement with previous studies,[6,15] H3K27me3 is retained and even increased in the reactivated loci. Hence, this repressing histone modification could trigger the resilencing and the DNA remethylation in the affected regions. Additional studies using inhibitors of H3K27 methyltransferase, as the DZNep,[45] may contribute to better understand the role of H3K27me3 mark in this process. Residual DNA methylation in the locus or near the locus is also likely to play a role.[29]

As we have shown, treated cells recover the original chromatin states after a few weeks of drug withdrawal. Moreover, they show

similar sensitivity and response to the drugs upon retreatment as never-treated cells (data not shown), suggesting that selection of drug-resistant cells is not sufficient to explain our results. Epigenetic drugs used here appear to induce plastic effects on chromatin. The transient disruption of the silenced states achieved by these treatments is overcome by a dominant repressive chromatin landscape determined by DNA methylation and bivalent chromatin marks. Design of new drugs with long lasting effects should improve the efficiency of epigenetic therapies.

In summary, our work illustrates the complexity of the epigenetic changes occurring upon drug-induced activation of silenced genes. While DNA methylation appears to play a dominant role in long lasting silencing, other players, characterized here as bivalent histone marks, contribute to maintain the epigenetic memory of silenced genes. The histone modification H3K27me3 is likely to play an important role in maintaining the memory of silenced genes that are in a bivalent context. Novel epigenetic strategies in cancer should target this repressive mark in order to get a more efficient effect of current DNA demethylating therapies.

Materials and Methods

Cell culture and 5-Aza-2'-deoxycytidine and trichostatin A treatments. The HCT116 colorectal carcinoma cell line was obtained from the American Type Culture Collection (ATCC). Cell culture and treatment with 5-Aza-2'-deoxycytidine (5AzadC) and Trichostatin A (TSA) were performed as described in reference 3, with minor modifications. Briefly, 0.65 x 10^6 cells were seeded in 10-cm cell culture dishes and 24 h later treated with 0.5 μM 5-Aza-2'-deoxycytidine (5AzadC; Sigma) for 48 h. Cells were treated with Trichostatin A (TSA) (Sigma) at 0.3 μM for 16 h. For co-treatment of cells with 5AzadC and TSA, 0.8 x 10^6 cells were seeded and treated initially with 5AzadC for 48 h, afterwards the medium was removed and cells were treated with TSA for an additional 16 h; then, the medium was changed and cells were harvested at the indicated time points. For control samples, half a million cells were seeded and cultured at the same time in which the drugs were omitted. DKO cells (HCT116 deficient in both DNA methyltransferases),[46] were cultured under the same conditions as wild type HCT116 cells. A complete new treatment of HCT116 cells with 5AzadC and TSA was done in parallel as control of drug efficiency.

Fluorescence-activated cell sorting (FACS). Cell sorting of replicating and non-replicating cells was performed using a modification of the method described previously in reference 47. Briefly, untreated and treated (5AzadC + TSA) cells were incubated for 24 h with 30 μM BrdU (Sigma) immediately and 7 days after the treatment. Cells were fixed with 3.7% formaldehyde at room temperature for 10 min, washed 3 times with cold PBS. The blocking solution (15% goat serum in 0.1% TritonX100) was added to the cells and incubated for 30 minutes at room temperature. After removal of blocking solution, fixed cells were incubated for 1 h at room temperature with anti-BrdU mouse monoclonal antibody 1:250 (Becton Dickinson, Ref. 555627), washed 3 times with PBS, and incubated again with anti-mouse FITC-conjugated antibody 1:250 (Invitrogen, Ref. A10543) for 45 minutes at room temperature in a dark chamber. Cells were washed twice with PBS and Tween (0.02%) and 2 more times with PBS. Flow cytometry and sorting was performed on an Influx flow cytometer (BD FACSAria™ II). Standard negative controls were used to set up the threshold and to fractionate cells positive and negative for BrdU.

Cell cycle analysis. Cell cycle analysis using PI was performed in 5dAzadC+TSA treated cells 1 and 7 days after the treatment. Cells were resuspended in 0.9 ml PBS, permeabilized with 2.1 ml 100% ethanol and kept at -20°C for 30 min. Then, cells were washed twice to eliminate the ethanol and 1 ml of the analysis solution compound (Propidium iodide 500 μg/ml (Sigma), sodium citrate 38 mM and Ribonuclease A (Sigma) was added. Flow cytometry and sorting was performed as described above. Fractions for G_0/G_1 and S/G_2 were collected.

Bisulfite sequencing and massCLEAVE™ analysis. Genomic DNA was obtained using standard protocols. Bisulfite treatment was performed as previously described in reference 41, or using the EZ DNA methylation kit™ (Zymo Research). Three independent PCR reactions were carried out and products were pooled to ensure a representative methylation profile. The primers used for the bisulfite PCR amplifications are listed in **Supplemental Table 1**. MassCLEAVE™ methylation analyses were performed as previously described in reference 48. Data was analyzed using the MassCLEAVE™ technology as previously reported in reference 48.

RNA extraction and quantitative real-time RT-PCR. RNA and the corresponding cDNA were obtained using standard protocols. Expression was quantified using the ABI PRISM 7900 HT sequence detection system (Applied Biosystems) or the Light Cycler 2.0 real time PCR system (Roche Diagnostics). The primers used for RT-PCR amplification are listed in **Supplemental Table 2**. The reactions were performed in triplicate. Gene expression levels were normalized using 18S determinations.

Chromatin immunoprecipitation (ChIP) assays. ChIP assays were carried out using the Chromatin Immunoprecipitation Assay Kit (Upstate Biotechnology) according to the manufacturer's instructions. The complexes were immunoprecipitated with antibodies specific for total histone H3 from Abcam (ab1791) and acetylation of histone H3 (Lys 9) (no. 07-352), dimethyl-histone H3 (Lys 4) (no. 07-030), dimethyl-histone H3 (Lys 9) (no. 07-441), trimethyl-histone H3 (Lys 27) (no. 07-449) from Upstate Biotechnology. As negative control we used rabbit IgG serum (Jackson Immunoresearch). The amount of immunoprecipitated target was measured by real-time PCR as described above. Positive and negative controls for each histone modification were used to set the lower limits.[3,6,49] Amplification primers for gene promoters are listed in **Supplemental Table 3**. PCRs were performed in triplicate.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

R.M. was supported by a FPI fellowship from the Spanish Ministry of Science and Innovation. This work was supported by grants from the Spanish Ministry of Science and Innovation

(SAF2008/1409, CSD2006/49), National Health and Medical Research Council (NHMRC) and Cancer Institute NSW, Australia. M.W.C. is supported by the Netherlands Organization for Scientific Research NWO/ZonMW (916-10-108) and a Marie Curie grant (IRG: 248397).

Note

Supplemental materials can be found at: www.landesbioscience.com/journals/epigenetics/article/16066

References

1. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet 2007; 8:286-98.
2. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128:683-92.
3. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark S. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. Nat Genet 2006; 38:540-9.
4. Clark SJ. Action at a distance: epigenetic silencing of large chromosomal regions in carcinogenesis. Hum Mol Genet 2007; 16:88-95.
5. Hitchins MP, Lin VA, Buckle A, Cheong K, Halani N, Ku S, et al. Epigenetic inactivation of a cluster of genes flanking MLH1 in microsatellite-unstable colorectal cancer. Cancer Res 2007; 67:9107-16.
6. Rodriguez J, Muñoz M, Vives L, Frangou CG, Groudine M, Peinado MA. Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells. Proc Natl Acad Sci USA 2008; 105:19809-14.
7. Novak P, Jensen T, Oshiro MM, Watts GS, Kim CJ, Futscher BW. Agglomerative epigenetic aberrations are a common event in human breast cancer. Cancer Res 2008; 68:8616-25.
8. Dallosso AR, Hancock AL, Szemes M, Moorwood K, Chilukamarri L, Tsai HH, et al. Frequent long-range epigenetic silencing of protocadherin gene clusters on chromosome 5q31 in Wilms' tumor. PLoS Genet 2009; 5:1000745.
9. Coolen MW, Stirzaker C, Song JZ, Statham AL, Kassir Z, Moreno CS, et al. Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. Nat Cell Biol 2010; 12:235-46.
10. Devaney J, Stirzaker C, Qu W, Song JZ, Statham AL, Patterson KI, et al. Epigenetic deregulation across 2q14.2 differentiates normal from prostate cancer and provides a regional part of novel DNA methylation cancer biomarkers. Cancer Epidemiol Biomarkers Prev 2010; 10.1158/055-9965.EPI-10-0719.
11. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128:693-705.
12. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 2007; 39:232-6.
13. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006; 439:871-4.
14. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007; 39:237-42.
15. McGarvey KM, Van Neste L, Cope L, Ohm JE, Herman JG, Van Criekinge W, et al. Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. Cancer Res 2008; 68:5753-9.
16. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006; 125:315-26.
17. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007; 448:553-60.
18. Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, et al. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell 2007; 1:299-312.
19. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 2007; 1:286-98.
20. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007; 129:823-37.
21. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 2008; 4:1000242.
22. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res 2010; 20:434-9.
23. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res 2010; 20:440-6.
24. Clarke MF, Fuller M. Stem cells and cancer: two faces of eve. Cell 2006; 124:1111-5.
25. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet 2008; 40:499-507.
26. Spivakov M, Fisher AG. Epigenetic signatures of stem-cell identity. Nat Rev Genet 2007; 8:263-71.
27. Sharma S, Kelly TK, Jones PA. Epigenetics in Cancer. Carcinogenesis 2010; 31:27-36.
28. Komashko VM, Farnham PJ. 5-azacytidine treatment reorganizes genomic histone modification patterns. Epigenetics 2010; 5:229-40.
29. Si J, Boumber YA, Shu J, Qin T, Ahmed S, He R, et al. Chromatin remodeling is required for gene reactivation after decitabine-mediated DNA hypomethylation. Cancer Res 2010; 70:6968-77.
30. Mayor R, Casadome L, Azuara D, Moreno V, Clark SJ, Capella G, et al. Long-range epigenetic silencing at 2q14.2 affects most human colorectal cancers and may have application as a non-invasive biomarker of disease. Br J Cancer 2009; 100:1534-9.
31. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, et al. Epigenetic stem cell signature in cancer. Nat Genet 2007; 39:157-8.
32. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature 2004; 429:457-63.
33. Rhee I, Bachman KE, Park BH, Jair KW, Yen RW, Schuebel KE, et al. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature 2002; 416:552-6.
34. Egger G, Jeong S, Escobar SG, Cortez CC, Li TW, Saito Y, et al. Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. Proc Natl Acad Sci USA 2006; 103:14080-5.
35. Das PM, Singal R. DNA methylation and cancer. J Clin Oncol 2004; 22:4632-42.
36. Bender CM, Gonzalgo ML, Gonzales FA, Nguyen CT, Robertson KD, Jones PA. Roles of cell division and gene transcription in the methylation of CpG islands. Mol Cell Biol 1999; 19:6690-8.
37. Liang G, Gonzalgo ML, Salem C, Jones PA. Identification of DNA methylation differences during tumorigenesis by methylation-sensitive arbitrarily primed polymerase chain reaction. Methods 2002; 27:150-5.
38. Egger G, Aparicio AM, Escobar SG, Jones PA. Inhibition of histone deacetylation does not block resilencing of p16 after 5-aza-2'-deoxycytidine treatment. Cancer Res 2007; 67:346-53.
39. Kagey JD, Kapoor-Vazirani P, McCabe MT, Powell DR, Vertino PM. Long-term stability of demethylation after transient exposure to 5-aza-2'-deoxycytidine correlates with sustained RNA polymerase II occupancy. Mol Cancer Res 2010; 8:1048-59.
40. McGarvey KM, Fahrner JA, Greene E, Martens J, Jenuwein T, Baylin SB. Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. Cancer Res 2006; 66:3541-9.
41. Stirzaker C, Song JZ, Davidson B, Clark SJ. Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. Cancer Res 2004; 64:3871-7.
42. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat Genet 1999; 21:103-7.
43. Zhu B, Benjamin D, Zheng Y, Angliker H, Thiry S, Siegmann M, et al. Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc Natl Acad Sci USA 2001; 98:5031-6.
44. Paz MF, Wei S, Cigudosa JC, Rodriguez-Perales S, Peinado MA, Huang TH, et al. Genetic unmasking of epigenetically silenced tumor suppressor genes in colon cancer cells deficient in DNA methyltransferases. Hum Mol Genet 2003; 12:2209-19.
45. Miranda TB, Cortez CC, Yoo CB, Liang G, Abe M, Kelly TK, et al. DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. Mol Cancer Ther 2009; 8:1579-88.
46. Rhee I, Jair KW, Yen RW, Lengauer C, Herman JG, Kinzler KW, et al. CpG methylation is maintained in human cancer cells lacking DNMT1. Nature 2000; 404:1003-7.
47. Ormerod MG. Analysis of cell proliferation using the bromodeoxyuridine/Hoechst-ethidium bromide method. Methods Mol Biol 1997; 75:357-65.
48. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res 2007; 35:119.
49. Rodriguez J, Vives L, Jorda M, Morales C, Munoz M, Vendrell E, et al. Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells. Nucleic Acids Res 2008; 36:770-84.