

Online Table 1. Detailed Annotation of Genes with Differentially Methylated CpG Loci in Low and High-risk Anal cancers

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
ADAT3	Adenosine deaminase, tRNA-specific 3	Genomic sequence overlap with SCAMP4; family of enzymes that forms inosine; edits transfer RNA	19		cg20243900	5'UTR	Island
GSG1L	Germ Cell Associated 1 (GSG1)-Like	Glutamate associated receptor subunit; Component of the inner core of AMPA-R complex; member of tetraspanin superfamily	16		cg09528825	Body	Island
KIAA0319	KIAA0319	Large, highly glycosylated, membrane protein with extracellular domain; function in adhesion and attachment	6		cg18428180	TSS1500;TSS200	Island
LOC728392	LOC728392	Uncharacterized; Derived by automated computational analysis using gene prediction method	17		cg18671949 cg06462347 cg27230784 cg22298430 cg01433610	TSS200 1stExon;5'UTR 1stExon;5'UTR 1stExon;5'UTR 1stExon;3'UTR	Island Island Island Island
PARD3	Partitioning defective 3 homolog	Member of the PARD protein; multimers that affect asymmetrical cell division and direct polarized cell growth	10	PAR3	cg26795540	Body	
SALL3	SAL-Like 3	Probable transcription factor; 4 double zinc finger (DZF) domain similar to SAL box motif	18		cg13794993 cg14007067 cg06954658 cg16433156 cg05080154	1stExon TSS200 TSS200 TSS200 TSS1500	Island Island Island Island
SFRP2	secreted frizzled-related protein 2	Tumor Suppressor Gene: cysteine-rich domain homologous to WNT-binding site of Frizzled proteins; SFRPs are soluble inhibitors of WNT signaling	4		cg22178613	TSS1500	Island
ST6GAL2	ST6 beta-galactosamide alpha-2,6-sialyltransferase 2	a type II transmembrane protein; involved in carbohydrate metabolism and cell growth	2		cg13168617	3'UTR	
SCAMP4	Secretory Carrier Membrane Protein 4	Genomic sequence overlaps with ADAT3; Secretory carrier membrane protein	19		cg20243900	5'UTR	Island

Type of file: table

Label: Online Table 2

Filename: Online Table 2.pdf

Online Table 2. Detailed Annotation of Genes with Differentially Methylated CpG Loci by Anal Tumor Size

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
ADAT3	Adenosine deaminase, tRNA-specific 3	Genomic sequence overlap with SCAMP4; family of enzymes that forms inosine; edits transfer RNA	19		cg01397065	5'UTR	Island
AGFG2	ArfGAP with FG repeats 2	Member of the HIV-1 Rev binding protein (HRB) family; contains a zinc finger domain (Arf-GAP) and FG motifs (phe-gly), role in REV nucleocytoplasmic transfer	7	HRBL; RABR	cg03431524	Body	
ALDH1L2	ALDEHYDE DEHYDROGENASE 1 FAMILY, MEMBER L2	Member of both the aldehyde dehydrogenase superfamily and the formyl transferase superfamily; mitochondrial 10-formyltetrahydrofolate dehydrogenase; essential role in the distribution of one-carbon groups between the cytosolic and mitochondrial compartments.	12		cg16527105	Body	
ASPSCR1	alveolar soft part sarcoma (ASPS) chromosome region, candidate 1	Contains a UBX domain and interacts with glucose transporter type 4 (GLUT4); involved in GLUT4 regulation as part of insulin signaling.	17		cg11511084	Body	N_Shore
BDNFOS	brain derived neurotrophic factor (BDNF) antisense RNA	Noncoding natural antisense RNAs that regulated BDNF	11	BDNF-AS	cg23330212	Body	
BLVRA	biliverdin reductase A	Serine/threonine/tyrosine kinase that catalyzes reduction of biliverdin to bilirubin; zinc metalloprotein with antioxidant properties.	7		cg14579118	TSS1500	N_Shore
BRE	brain and reproductive organ-expressed (TNFRSF1A modulator)	Novel stress-responsive gene; death receptor-associated protein in cytoplasm; component of BRCA1/2-containing DNA repair complex in nucleus; anti-apoptotic activity by blocking TNF-a	2	TNFRSF1A modulator and BRCC45	cg13861527	Body	
C2orf74	chromosome 2 open reading frame 74	Uncharacterized protein, predicted single pass membrane protein	2		cg01648237	TSS200	
C6orf136	chromosome 6 open reading frame 136	Uncharacterized protein	6		cg13016528	Body	S_Shore

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
GET4	golgi to ER traffic protein 4 homolog (S. cerevisiae)	Member of a multiprotein complex involved in the post-translational delivery of tail-anchored membrane proteins from ribosomes to endoplasmic reticulum (ER) membrane; Involved in ubiquitination.	7	CEE; TRC35; CGI-20; C7orf20	cg24580076	TSS1500	N_Shore
CCDC63	coiled-coil domain containing 63	Uncharacterized; identified by genome-wide evaluation of expression transcripts	12	ODA5	cg02855409	TSS200	
					cg10995082	TSS200	
					cg19006003	1stExon;5'UTR	
CDK6	cyclin-dependent kinase 6	Member of the cyclin-dependent protein kinase (CDK) family; regulates cell cycle and activity of pRb; regulated by miRNAa.	7		cg23628117	3'UTR	
COX8A	cytochrome c oxidase subunit VIIIA (ubiquitous)	Nuclear-encoded subunit of the cytochrome c oxidase complex; joined by >3 mitochondrial subunits; terminal enzyme of the respiratory chain; oxidative stress destabilizes the complex and can lead to disease, including cancer.	11	COX, COX8, COX8-2, COX8L, VIII, VIII-L	cg17292384	3'UTR	S_Shore
CYP27C1	cytochrome P450, family 27, subfamily C, polypeptide 1	A member of the Cytochrome P450 family with ~15 families and several subfamilies. Cytochrome P450 proteins catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids.	2		cg08022717	Body	
DMRT3	Doublesex and MAB-3 related transcription factor 3	Predicted to regulate transcription during sexual development.	9	DMRTA3	cg14176274	Body	Island
EBF3	early B-cell factor 3	Tumor suppressor gene; member of the early B-cell factor (EBF) family of DNA binding transcription factors	10	OE-2; EBF-3;	cg04804618	Body	Island
					cg14737286	Body	Island
					cg07076175	TSS200	Island
FAM150A	Family With Sequence Similarity 150, Member A	Increased methylation by Infinium in more aggressive renal cell cancers	8		cg10094616	TSS200	Island
					cg22862746	Body	Island
					cg09442654	5'UTR;1stExon	Island

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
FOS	FBJ murine osteosarcoma viral oncogene homolog	Oncogene, Member of AP-1 transcription factor complex; regulates cell proliferation, differentiation, and transformation	14		cg23404711	Body	S_Shore
FZD10	Frizzled family receptor 10	Member of the frizzled gene family; 7-transmembrane domain proteins that are receptors for the WNT family of signaling proteins; coupled to the beta-catenin canonical signaling pathway	12		cg13859208	3'UTR;1stExon	S_Shore
GSG1L	Germ Cell Associated 1 (GSG1)-Like	Glutamate associated receptor subunit; Component of the inner core of AMPA-R complex; member of tetraspanin superfamily	16		cg03921753	1stExon	Island
					cg09528825	Body	Island
					cg03394150	Body	Island
GULP1	engulfment adaptor PTB domain containing 1	Adapter protein for engulfment of apoptotic cells by phagocytes	2	CED6; GULP	cg19202813	5'UTR	
HDAC2	histone deacetylase 2	Forms transcriptional repressor complexes responsible for the deacetylation of lysine residues at the N-terminal regions of core histones (H2A, H2B, H3 and H4). Epigenetic regulation of transcription, cell cycle progression, and development.	6		cg15069235	Body	N_Shore
HORMAD2	HORMA domain containing 2	Proposed involvement in meiotic progression; protein that localizes to unsynapsed meiotic chromosomes.	22		cg24211826	3'UTR	
HOXA6	homeobox A6	Encodes a DNA-binding transcription factor from HOXA cluster on chromosome 7: regulates gene expression, morphogenesis, and differentiation.	7	HOX1; HOX1B; HOX1.2	cg23129930	1stExon	Island
HS6ST1	heparan sulfate 6-O-sulfotransferase 1	Member of the heparan sulfate biosynthetic enzyme family; type II membrane protein for 6-O-sulfation of heparin sulfate	2	HH15, HS6ST	cg08472795	Body	

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
KIR3DX1	killer cell immunoglobulin-like receptor, three domains, X1	Member of the KIR family (17 members total) of receptors on natural killer cells; involved in innate response to infection and cancer; ancestral KIR3DX1 gene was identified by homology to rhesus monkey; expression in humans unknown; may be pseudogene.	19	LENG12; KIR3DL0	cg10731960	TSS1500	
KLHL29	kelch-like family member 29	Complete sequence is not known with certainty; Predicted domains (Kelch repeats; BTB/POZ) infer homo- or heterodimer interactions. Kelch repeats associated with actin tails.	2	KBTBD9; KIAA1921	cg00537210	Body	S_Shore
MECOM	MDS1 and EVI1 complex locus	Transcriptional regulator and oncoprotein	3	EVI1, MDS1, MDS1-EVI1, PRDM3	cg20528780	Body	
MF12	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5; aka Melanotransferrin	Cell-surface glycoprotein; similarities to members of the transferrin superfamily; involved in iron transport from cell surface	3	MTF1; CD228; MAP97	cg00477017	Body	S_Shore
MIR200B;MIR200A	microRNA 200b / MicroRNA 200a	Members of the MIR200 family; cluster of MIR200B/200A/429 jointly regulated; predicted microRNA stem-loop; master mediators of the epithelial phenotype	1		cg14161399	TSS200;TS S1500	S_Shore
					cg02825344	TSS200;TS S1500	S_Shore
MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	Member of the matrix metalloproteinase (MMP) family; degrades type IV and V collagens in extracellular matrix; involved in tissue remodeling and development	20	GELB; CLG4B; MMP-9; MANDP2	cg17664577	Body	Island
MRPS22	Mitochondrial ribosomal protein S22	28S subunit protein of mitochondrial ribosomal (mitoribosome); no related prokaryotic proteins; Located in telemetric region	3		cg11277156	Body	S_Shelf
MSI2	musashi RNA-binding protein 2	RNA binding protein and progenitor stem-cell marker	17	MSI2H	cg04486382 cg19810715	Body Body	

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
NPTX1	neuronal pentraxin I	Encodes a binding protein for the snake venom toxin, taipoxin.	17	NP1	cg20853771	3'UTR	N_Shore
OR2C3	olfactory receptor, family 2, subfamily C, member 3	Receptor in G-protein mediated transduction of odorant signals.	1	OR2C4; OST742	cg20320823	TSS1500	S_Shelf
PAK1	p21 protein (Cdc42/Rac)-activated kinase 1	Target for small GTP binding proteins (Cdc42/Rac); regulated cell motility and morphology; oncogenic	11	PAKalpha	cg24536703	5'UTR	N_Shore
PDE3B	phosphodiesterase 3B, cGMP-inhibited	Genomic sequence overlaps with PSMA1. Member of the cyclic nucleotide phosphodiesterase family; dual-specificity for cAMP/cGMP; involved in adipose tissue metabolism and angiogenesis.	11	HcGIP1; cGIPDE1	cg21901307	TSS1500;5'UTR;1stExon	N_Shore
PSMA1	proteasome (prosome, macropain) subunit, alpha type, 1	Genomic sequence overlaps with PDE3B. member of the peptidase T1A family, a 20S core alpha subunit of the multicatalytic proteinase complex; Proteasome cleaves peptides in an ATP/ubiquitin-dependent process and processes class I MHC peptides.	11	NU; HC2; PROS30	cg21901307	TSS1500;5'UTR;1stExon	N_Shore
PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)	Involved in regulation of the phospholipid metabolism in biomembranes.	1	MOM1; PLA2; PLA2B; PLA2L	cg13211559	TSS1500;TSS1500;TSS1500	
PLEKHG1	pleckstrin homology domain containing, family G (with RhoGef domain) member 1	Member of the Dbl family of Rho guanine nucleotide exchange factors (RhoGEFs) (70 members). RhoGEFs activate Rho GTPases. Targets for PLEKHG1 uncharacterized.	6	ARHGEF41	cg26852242	5'UTR	
PLEKHG3	pleckstrin homology domain containing, family G (with RhoGef domain) member 3	Member of the Dbl family of Rho guanine nucleotide exchange factors (RhoGEFs) (70 members). RhoGEFs activate Rho GTPases. Targets for PLEKHG3 uncharacterized.	14	ARHGEF43; KIAA0599	cg11802553	Body	
PON3	paraoxonase 3	Rapidly hydrolyzes lactones; inhibits oxidation of low-density lipoprotein (LDL).	7		cg04080282 cg08898155 cg11435506	TSS1500 TSS1500 TSS1500	Island Island Island
PPP1R9A	protein phosphatase 1, regulatory subunit 9	Regulatory subunit of protein phosphatase I; controls actin cytoskeleton reorganization.	7	NRB1; NRBI; Neurabin-I	cg09724492	TSS1500;5'UTR	S_Shore

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
RERE	arginine-glutamic acid dipeptide (RE) repeats	Encoded protein co-localizes with a transcription factor in the nucleus; pro-apoptotic; caspase-dependent	1	ARG; ARP; DNB1; ATN1L	cg19679865	5'UTR;Body	
RNASEN	ribonuclease type III, nuclear	miRNA processing enzyme	5	DROSHA	cg04590036	Body	
RREB1	ras responsive element binding protein 1	Zinc finger transcription factor that binds to RAS-responsive elements (RREs) of gene promoters.	6	HNT; FINB; LZ321; Zep-1	cg03137792	Body	
SCAMP4	Secretory Carrier Membrane Protein 4	Genomic sequence overlaps with ADAT3; Secretory carrier membrane protein; Not involved in endocytosis	19		cg01397065	5'UTR	Island
SEPP1	selenoprotein P, plasma, 1	Extracellular glycoprotein; binds heparin; antioxidant activities in extracellular space	5	SeP; SELP	cg08626131	TSS1500	
SLC9A3	SLC9A3 solute carrier family 9, subfamily A, member 3	Sodium/hydrogen exchanger 3; Involved in pH regulation and sodium balance; essential to maintain the epithelial barrier function.	5	NHE3	cg06058576	Body	Island
SNTB1	syntrophin, beta 1 (dystrophin-associated protein A1, 59kDa, basic component 1	AKA dystrophin-associated protein A1, 59kDa, basic component 1. membrane protein associated with dystrophin and dystrophin-related proteins. May link receptors to the actin cytoskeleton and the dystrophin glycoprotein complex	8	A1B; SNT2; BSYN2; 59-DAP; DAPA1B	cg04318006	1stExon	Island
STMN2	stathmin-like 2	Stathmin proteins function in microtubule dynamics and signal transduction; STMN2 reported in neuronal growth and osteogenesis.	8	SCG10; SCGN10	cg00398130	Body	Island
SURF6	Surfeit 6	Located in the surfeit gene cluster in telomere region (60Kb region spanning 6 Surfeit genes, all with 5' CpG island); supports nucleolar matrix structure and function(s) via its association with nucleic acids.	9	RRF14	cg01832218	Body	N_Shore
TMEM196	transmembrane protein 196	No publications; function unknown	7		cg18505401	TSS200	Island
TRA2B	transformer 2 beta homolog	Nuclear sequence-specific splicing factor; role in mRNA processing, splicing patterns, and gene expression.	3	SFRS10; SRFS10; TRAN2B; TRA2-BETA; Htra2-beta	cg12825509	Body	

Symbol	Product	Annotation	Chromosome	Alternative Symbol	CpG Number*	Genomic Region	UCSC CpG Island Annotation
TTC9	Tetratricopeptide Repeat Domain-Containing 9	Associated with the endoplasmic reticulum; contains 3 tetratricopeptide repeats (degenerate 34 amino acid sequence motif)	14	KIAA0227	cg01634544	Body	
VWDE	von Willebrand factor D and EGF domains	Contains 7 epidermal growth factor (EGF) like domains and 1 vWF D' domain; both domains found in secreted proteins and promote protein-protein interactions; Identified in genome-wide sequencing.	7		cg02935154	5'UTR;1stExon	Island
					cg16278512	5'UTR;1stExon	Island
WNT9A	wingless-type MMTV integration site family, member 9A	Member of the WNT gene family of 19 cysteine-rich secreted glycoproteins; extracellular signaling molecules when bound to Frizzles receptors; clustered by WNT3A. WNT9a involved in canonical signaling pathway leading to stabilization of b-catenin.	1		cg26452081	Body	N_Shore
ZHX2	ZHX2 zinc fingers and homeoboxes 2	Member of the zinc fingers and homeoboxes gene family; nuclear homo- and heterodimeric transcriptional repressor	8	RAF; AFR1	cg01801603	5'UTR	S_Shore