

Supplemental Material to:

Gareth A Wilson, Matthias Lechner, Anna Köferle, Helena Caren, Lee M Butcher, Andrew Feber, Tim Fenton, Amrita Jay, Chris Boshoff, and Stephan Beck

Integrated virus-host methylome analysis in head and neck squamous cell carcinoma

Epigenetics 2013; 8(9)

<http://dx.doi.org/10.4161/epi.25614>

**[http://www.landesbioscience.com/journals/epigenetics/
article/25614](http://www.landesbioscience.com/journals/epigenetics/article/25614)**

Sample ID	HPV status	Viral titer (copies/cell by E6 qPCR)	Initial reads (total, pre-alignment)	Uniquely mapped fragments (q \geq 10)
HN_29	neg	0	52,655,678	16,658,305
HN_32	neg	0	58,152,760	18,946,120
HN_39	pos	96.94	40,889,864	11,680,329
HN_96	neg	0	41,538,032	12,317,714
HN_105	pos	110.44	36,259,072	11,135,937
HN_125	pos	0.66	47,522,856	14,868,967

Supplementary Table 1: Summary of read counts and viral counts for each of the MeDIP-seq samples

HPV type	Reference Sequence Accession Code (NCBI)						
		HN29	HN32	HN39	HN96	HN105	HN125
Human_papillomavirus_type_18_uid15506	NC_001357	0	0	0	0	0	0
Human_papillomavirus_type_1_uid15491	NC_001356	0	0	0	0	0	0
Human_papillomavirus_type_2_uid15512	NC_001352	0	0	0	0	0	0
Human_papillomavirus_type_54_uid15466	NC_001676	0	0	0	0	0	0
Human_papillomavirus_type_101_uid17121	NC_008189	0	0	0	0	0	0
Human_papillomavirus_type_103_uid17119	NC_008188	0	0	0	0	0	0
Human_papillomavirus_type_108_uid34847	NC_012213	0	0	0	0	0	0
Human_papillomavirus_type_10_uid15504	NC_001576	0	0	0	0	0	0
Human_papillomavirus_type_16_uid15505	NC_001526	0	0	1449	0	787	17
Human_papillomavirus_type_26_uid15507	NC_001583	0	0	0	0	0	0
Human_papillomavirus_type_32_uid15508	NC_001586	0	0	0	0	0	0
Human_papillomavirus_type_34_uid15509	NC_001587	0	0	0	0	0	0
Human_papillomavirus_type_41_uid15485	NC_001354	0	0	0	0	0	0
Human_papillomavirus_type_48_uid14027	NC_001690	0	0	0	0	0	0
Human_papillomavirus_type_49_uid15455	NC_001591	0	0	0	0	0	0
Human_papillomavirus_type_4_uid15492	NC_001457	0	0	0	0	0	0
Human_papillomavirus_type_50_uid14327	NC_001691	0	0	0	0	0	0
Human_papillomavirus_type_53_uid15510	NC_001593	0	0	0	0	0	0
Human_papillomavirus_type_5_uid15511	NC_001531	0	0	0	0	0	0
Human_papillomavirus_type_60_uid14028	NC_001693	0	0	0	0	0	0

Human_papillomavirus_type_63_uid15486	NC_001458	0	0	0	0	0	0
Human_papillomavirus_type_6b_uid15454	NC_001355	0	0	0	0	0	0
Human_papillomavirus_type_7_uid15450	NC_001595	0	0	0	0	0	0
Human_papillomavirus_type_88_uid28737	NC_010329	0	0	0	0	0	0
Human_papillomavirus_type_90_uid15424	NC_004104	0	0	0	0	0	0
Human_papillomavirus_type_92_uid14406	NC_004500	0	0	0	0	0	0
Human_papillomavirus_type_96_uid15488	NC_005134	0	0	0	0	0	0
Human_papillomavirus_type_9_uid15456	NC_001596	0	0	0	0	0	0

Supplementary Table 2: 29 human papilloma viruses, including Reference Sequence Accession Codes, obtained from NCBI Genomes (<http://www.ncbi.nlm.nih.gov/genomes>) (28/03/2011). Number of reads aligning to each HPV strain for each sample is also shown.

Regions hypermethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands			Regions hypomethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands		
Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)	Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)
c-Fos	1.65E-03		SETDB1	1.87E-02	
MEF2A	1.82E-03		IRF4_(M-17)	2.02E-02	
E2F4	1.85E-03		MafF_(M8194)	2.41E-02	
IRF3	1.85E-03		SP1	2.75E-02	5.70E-05
BCLAF1_(M33-P5B11)	1.85E-03		BRCA1_(C-1863)	3.63E-02	
PAX5-N19	1.85E-03		MafK_(ab50322)	3.63E-02	
STAT1	2.02E-03		PAX5-N19	3.86E-02	
E2F6	2.38E-03		SIX5	4.23E-02	
HSF1	2.45E-03		SIRT6	4.69E-02	
Pol3	3.35E-03		RPC155	4.94E-02	
FOSL2	5.85E-03		GTF2F1_(RAP-74)	5.93E-02	
TFIIIC-110	5.86E-03		eGFP-JunB	6.25E-02	
KAP1	6.78E-03		c-Jun	6.42E-02	
IRF4_(M-17)	9.98E-03		FOSL1_(SC-183)	6.42E-02	
HNF4A	1.00E-02		MafK_(SC-477)	8.80E-02	
EBF1_(C-8)	1.33E-02	1.50E-04	CEBPB	9.10E-02	
Pol2(b)	1.36E-02		FOXA1_(C-20)	9.10E-02	
SRF	1.37E-02		FOSL2	1.00E-01	
GATA-2	1.56E-02		RFX5_(N-494)	1.06E-01	
MafK_(SC-477)	1.56E-02		eGFP-JunD	1.08E-01	
GR	1.96E-02		HSF1	1.13E-01	
BATF	1.98E-02		HNF4A_(H-171)	1.26E-01	
BRF1	1.98E-02		NF-E2_(H-230)	1.26E-01	
GATA2_(CG2-96)	1.98E-02		p300_(N-15)	1.38E-01	
SUZ12	2.62E-02		eGFP-GATA2	1.57E-01	
TAL1_(SC-12984)	2.82E-02		IRF3	1.57E-01	
BRCA1_(C-1863)	2.83E-02		STAT1	1.57E-01	
ERRA	3.11E-02		Pbx3	1.65E-01	
NF-E2	3.11E-02		BHLHE40	1.69E-01	
SIRT6	3.11E-02		GR	1.76E-01	
TCF4	4.15E-02		MEF2A	2.10E-01	
BRF2	5.14E-02		JunD	2.15E-01	
GRp20	6.07E-02		GATA-2	2.57E-01	

c-Myc	3.86E-06	5.90E-10	SMC3_(ab9263)	5.36E-06	
EBF	4.31E-06		POU2F2	6.41E-06	
ZBTB7A_(SC-34508)	4.43E-06		Oct 2, 2013	7.12E-06	
p300	8.47E-06		USF-1	8.48E-06	
TAF7_(SQ-8)	9.23E-06		HMG3	1.05E-05	
RFX5_(N-494)	1.03E-05		NRSF	1.19E-05	
HDAC2_(SC-6296)	1.50E-05		ETS1	1.25E-05	
IRF1	1.64E-05		GABP	3.44E-05	1.20E-05
NF-YA	1.74E-05		USF2	4.88E-05	
NRSF	2.31E-05		CtBP2	6.62E-05	
RXRA	3.37E-05		EBF	7.23E-05	
E2F1	4.33E-05		NFKB	1.13E-04	4.16E-06
BDP1	4.42E-05		ELK4	1.58E-04	
HNF4G_(SC-6558)	6.09E-05		Ini1	1.71E-04	
Max	6.22E-05		STAT2	2.92E-04	
ATF3	6.22E-05		USF1_(SC-8983)	7.80E-04	
ETS1	7.29E-05		Pol2(b)	1.21E-03	
GABP	8.44E-05	8.99E-03	Pol2(phosphoS2)	1.49E-03	
eGFP-JunD	1.00E-04		Mxi1_(bHLH)	1.50E-03	
CHD2_(N-1250)	1.19E-04		CHD2_(N-1250)	2.47E-03	
JunD	1.19E-04		TAL1_(SC-12984)	2.64E-03	
PU.1	1.42E-04		E2F4	3.20E-03	
USF2	1.77E-04		c-Fos	5.79E-03	
ERalpha_a	2.20E-04		SRF	7.55E-03	
p300_(N-15)	2.60E-04		NELFe	8.34E-03	
POU2F2	3.94E-04		ATF3	1.02E-02	
Oct2	4.40E-04		eGFP-FOS	1.26E-02	
THAP1_(SC-98174)	5.38E-04		IRF1	1.54E-02	
SP2_(SC-643)	6.12E-04		SP2_(SC-643)	1.55E-02	
HNF4A_(H-171)	9.07E-04		GTF2B	1.55E-02	
GATA3_(SC-268)	9.07E-04		BCLAF1_(M33-P5B11)	1.76E-02	
RPC155	9.07E-04		PGC1A	1.76E-02	
ZEB1_(SC-25388)	9.44E-04		TCF4	1.76E-02	
SIX5	1.09E-03		ZBTB33	1.77E-02	

Regions hypermethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands			Regions hypomethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands		
Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)	Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)
SREBP1	6.07E-02		STAT3	2.61E-01	
STAT3	6.07E-02		Brg1	2.64E-01	
CEBPB	8.29E-02		p300	2.76E-01	
NF-E2_(H-230)	8.41E-02		p300_(F-4)	2.76E-01	
Pol2(phosphoS2)	9.83E-02		NANOG_(SC-33759)	2.92E-01	
MafK_(ab50322)	1.05E-01		KAP1	3.11E-01	
BAF170	1.43E-01		NF-YB	3.15E-01	
BCL11A	1.43E-01		BCL3	3.18E-01	
CtBP2	1.43E-01		E2F1	3.26E-01	
TCF12	1.55E-01		GRp20	3.26E-01	
AP-2gamma	1.86E-01		GATA2_(CG2-96)	3.33E-01	
c-Jun	2.02E-01		WHIP	3.33E-01	
SETDB1	2.50E-01		ZZZ3	3.33E-01	
FOSL1_(SC-183)	2.77E-01		RXRA	3.38E-01	
TR4	2.78E-01		NF-YA	3.53E-01	
eGFP-JunB	3.13E-01		HDAC2_(SC-6296)	3.81E-01	
FOXA1_(SC-101058)	3.33E-01		ERalpha_a	4.34E-01	
Mxi1_(bHLH)	3.33E-01		TR4	4.34E-01	
POU5F1_(SC-9081)	3.36E-01		GATA-1	4.60E-01	
ELK4	3.48E-01		AP-2gamma	4.80E-01	
FOXA2_(SC-6554)	3.55E-01		BCL11A	4.84E-01	
GTF2B	3.68E-01		MEF2C_(SC-13268)	4.84E-01	
ZBTB33	3.72E-01		BATF	5.23E-01	
MafF_(M8194)	4.23E-01		TFIIIC-110	5.33E-01	
AP-2alpha	4.43E-01		BAF155	6.11E-01	
BCL3	4.43E-01		HNF4A	6.18E-01	
BAF155	5.33E-01		GATA3_(SC-268)	6.66E-01	
BHLHE40	5.75E-01		THAP1_(SC-98174)	6.71E-01	
GATA-1	5.77E-01		FOXA1_(SC-101058)	8.97E-01	
Ini1	7.40E-01		FOXA2_(SC-6554)	9.78E-01	
FOXA1_(C-20)	8.91E-01		AP-2alpha	1	
Brg1	1		BAF170	1	
eGFP-FOS	1		BDP1	1	

Regions hypermethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands			Regions hypomethylated in HPV+ rel. to HPV- HNSCC falling into CpG Islands		
Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)	Encode Factor	Fisher's Exact Test (1-tailed, FDR corrected)	TRAP (FDR corrected)
eGFP-GATA2	1		BRF1	1	
eGFP-HDAC8	1		BRF2	1	
eGFP-NR4A1	1		eGFP-HDAC8	1	
GCN5	1		eGFP-NR4A1	1	
MEF2C_(SC-13268)	1		ERRA	1	
NELFe	1		GCN5	1	
p300_(F-4)	1		HNF4G_(SC-6558)	1	
PGC1A	1		NF-E2	1	
PRDM1_(Val90)	1		Pol3	1	
SPT20	1		POU5F1_(SC-9081)	1	
SREBP2	1		PRDM1_(Val90)	1	
STAT2	1		SPT20	1	
WHIP	1		SREBP1	1	
XRCC4	1		SREBP2	1	
ZNF274	1		XRCC4	1	
ZZZ3	1		ZNF274	1	

Supplementary Table 3: Putative associations between DMRs and transcription factor binding sites. All transcription factors considered in the analysis are shown here.

Genes with E2F1 binding sites targeted by hypermethylated DMRs (HPV+ rel. to HPV- HNSCC)		Genes with E2F1 binding sites targeted by hypomethylated DMRs (HPV+ rel. to HPV- HNSCC)	
Ensembl Gene ID	Gene Name	PID	Gene Name
ENSG00000010270	STARD3NL	ENSG00000006025	OSBPL7
ENSG000000021645	NRXN3	ENSG000000058335	RASGRF1
ENSG000000072135	PTPN18	ENSG000000087510	TFAP2C
ENSG000000104714	ERICH1	ENSG000000111490	TBC1D30
ENSG000000106006	HOXA6	ENSG000000115935	WIPF1
ENSG000000107562	CXCL12	ENSG000000118257	NRP2
ENSG000000108175	ZMIZ1	ENSG000000119919	NKX2-3
ENSG000000111052	LIN7A	ENSG000000123388	HOXC11
ENSG000000119139	TJP2	ENSG000000129255	MPDU1
ENSG000000136866	ZFP37	ENSG000000132182	NUP210
ENSG000000138293	NCOA4	ENSG000000136630	HLX
ENSG000000140941	MAP1LC3B	ENSG000000139508	SLC46A3
ENSG000000141526	SLC16A3	ENSG000000143727	ACP1
ENSG000000144681	STAC	ENSG000000152128	TMEM163
ENSG000000150201	FXYP4	ENSG000000152804	HHEX
ENSG000000151025	GPR158	ENSG000000154358	OBSCN
ENSG000000164093	PITX2	ENSG000000162086	ZNF75A
ENSG000000164904	ALDH7A1	ENSG000000162545	CAMK2N1
ENSG000000169629	RGPD8	ENSG000000162639	HENMT1
ENSG000000169862	CTNND2	ENSG000000162728	KCNJ9
ENSG000000170190	SLC16A5	ENSG000000163701	IL17RE
ENSG000000172264	MACROD2	ENSG000000166669	ATF7IP2
ENSG000000173404	INSM1	ENSG000000167363	FN3K
ENSG000000174469	CNTNAP2	ENSG000000171119	NRTN
ENSG000000177519	RPRM	ENSG000000173530	TNFRSF10D
ENSG000000177853	ZNF518A	ENSG000000174197	MGA
ENSG000000182963	GJC1	ENSG000000183963	SMTN
ENSG000000183783	KCTD8	ENSG000000184860	SDR42E1
ENSG000000186479	RGS7BP	ENSG000000185332	TMEM105
ENSG000000188002		ENSG000000185442	FAM174B
ENSG000000198879	SFMBT2	ENSG000000199287	
ENSG000000204186	ZDBF2	ENSG000000203995	ZYG11A
ENSG000000223491		ENSG000000204950	LRRC10B
ENSG000000235978		ENSG000000213903	LTB4R
ENSG000000237575	PYY2	ENSG000000213994	
ENSG000000240280	TCAM1P	ENSG000000225611	
ENSG000000248597		ENSG000000228340	
ENSG000000249328		ENSG000000228672	C5orf65

Genes with E2F1 binding sites targeted by hypermethylated DMRs (HPV+ rel. to HPV- HNSCC)		Genes with E2F1 binding sites targeted by hypomethylated DMRs (HPV+ rel. to HPV- HNSCC)	
Ensembl Gene ID	Gene Name	PID	Gene Name
ENSG00000253982		ENSG00000230489	VAV3-AS1
		ENSG00000231651	DLG3-AS1
		ENSG00000233384	
		ENSG00000254732	
		ENSG00000255542	
		ENSG00000261251	
		ENSG00000265683	

Supplementary Table 4: Genes with E2F1 binding sites targeted by DMRs.

Genes with E2F6 binding sites targeted by hypermethylated DMRs (HPV+ rel. to HPV- HNSCC)		Genes with E2F6 binding sites targeted by hypomethylated DMRs (HPV+ rel. to HPV- HNSCC)	
Ensembl Gene ID	Gene Name	PID	Gene Name
		ENSG00000182013	PNMAL
		ENSG00000183248	
		ENSG00000183963	SMTN
		ENSG00000185332	TMEM105
		ENSG00000185338	SOCS1
		ENSG00000185442	FAM174B
		ENSG00000185507	IRF7
		ENSG00000196660	SLC30A10
		ENSG00000199287	
		ENSG00000203808	BVES-AS1
		ENSG00000203995	ZYG11A
		ENSG00000204389	HSPA1A
		ENSG00000204950	LRRC10B
		ENSG00000213903	LTB4R
		ENSG00000225446	
		ENSG00000225611	
		ENSG00000241159	
		ENSG00000241238	
		ENSG00000254732	
		ENSG00000256806	
		ENSG00000257702	
		ENSG00000259285	
		ENSG00000259753	
		ENSG00000260669	
		ENSG00000261251	

Supplementary Table 5: Genes with E2F6 binding sites targeted by DMRs.

	HPV+ (n=3)	HPV- (n=3)
Age	58.3 years (53-62)	73.6 years (66-80)
Gender	M: 2 F: 1	M: 3 F: 0
Tumour site	Oropharynx: 3	Oropharynx: 3
Tumour grade	Well diff: 0 Mod diff: 0 Poorly diff: 3	Well diff: 0 Mod diff: 1 Poorly diff: 2
Tumour stage (T)	T1: 0 T2: 3 T3: 0 T4: 0 N/a: 0	T1: 0 T2: 0 T3: 0 T4: 2 N/a: 1
Cervical lymph node involvement (N)	Yes: 3 No: 0 N/a: 0	Yes: 0 No: 2 N/a: 1
Smoking	Ever: 3 Never: 0 N/a: 0	Ever: 2 Never: 0 N/a: 1
Alcohol	Heavy drinker: 1 Occ. alcohol: 2 Never: 0 N/a: 0	Heavy drinker: 0 Occ. alcohol: 2 Never: 1 N/a: 0

Supplemental Table 6: Patient characteristics of 3 HPV+ and 3 HPV- fresh frozen HNSCC samples used for MeDIP-Seq and BS-Seq analysis.

Cell line	Origin (site)	HPV status
UPCI:SCC90	oropharynx	HPV+ (integrated)
UM:SCC47	oropharynx	HPV+ (integrated)
93-VU-147T	oral cavity	HPV+ (integrated)
UPCI:SCC003	oropharynx	HPV-
UPCI:SCC036	oropharynx	HPV-
PCI-30	oral cavity	HPV-

Supplemental Table 7: HPV+ and HPV- HNSCC cell lines used for experiments.

Name of primer	Primer Sequence (5'-3')
forward L1/2 PCR primer vol. 2 (5541-5569)	AATATATAATTATTGTTGATGTAGGTGAT
reverse L1/2 PCR primer vol. 2 (5984-5962)	ACCCACACCTAATAACTAACCCAC

Supplemental Table 8: L1 and L2 PCR primers used for validation of MeDIP-Seq results by BS-Seq.

Supplementary Figures

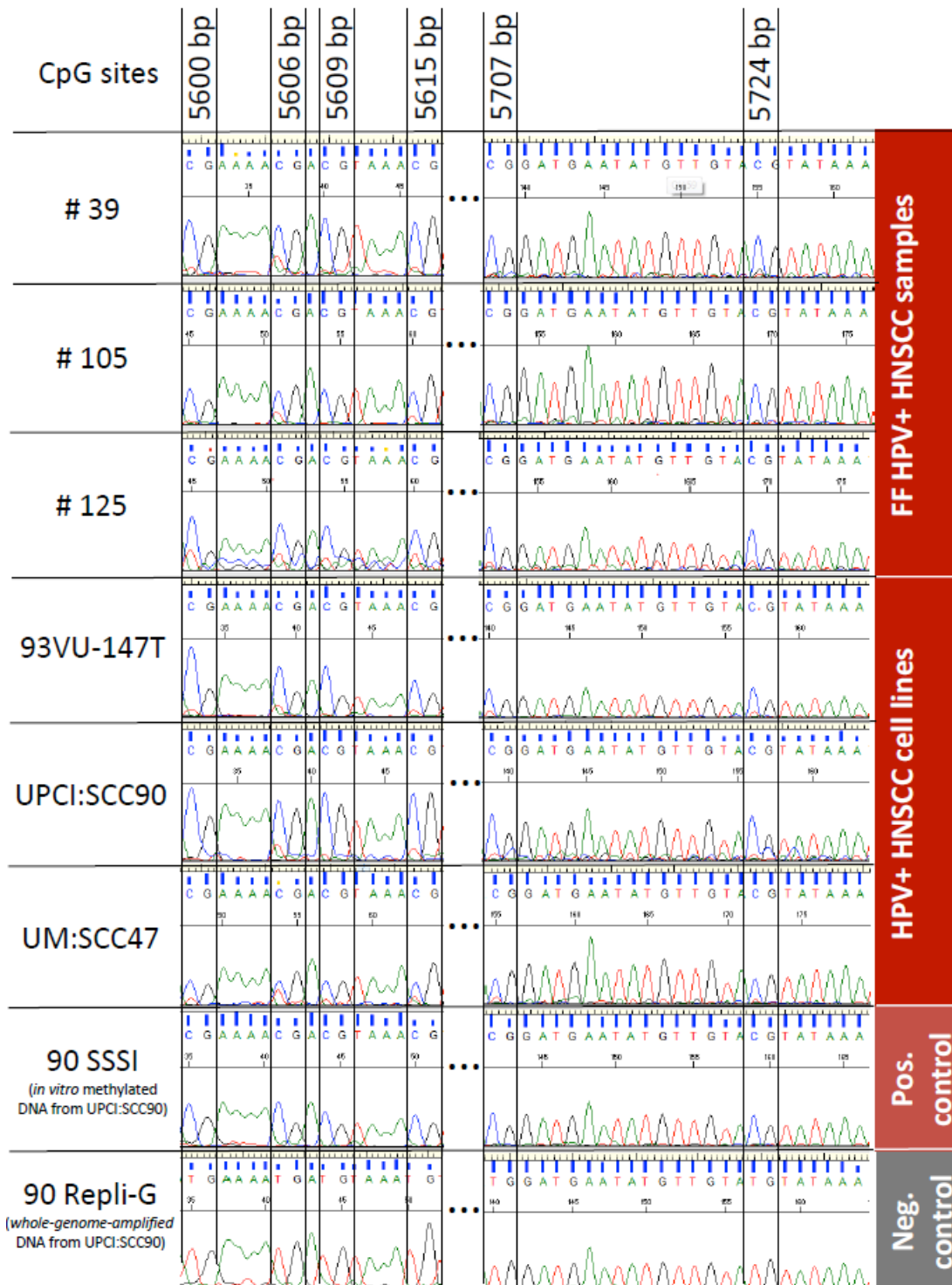
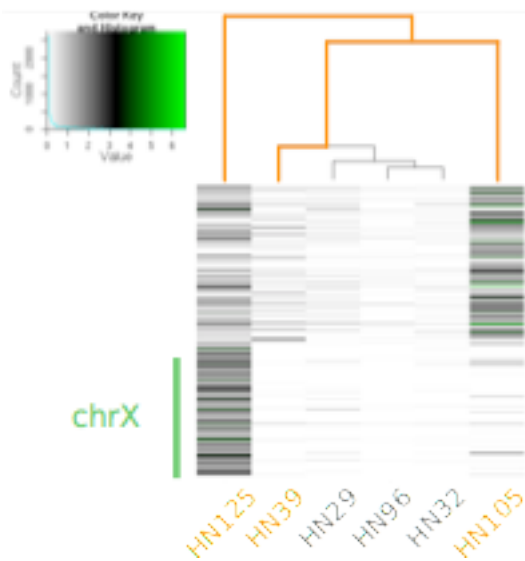


Figure S1: Methylation of the L1/2 boundary region is observed in all tested FF HPV type 16+ HNSCC tissue samples and 3 HPV type 16+ HNSCC cell lines (UPCI:SCC90, UM:SCC47, 93VU-147T). Selected BS-Seq traces, covering the 6 CpG sites of interest, are illustrated along with the positive (*in vitro* methylated DNA from UPCI:SCC90) and negative (whole-genome amplified DNA from UPCI:SCC90) control at six CpG sites at genomic position 5600, 5606, 5609, 5615, 5707, and 5724 of the HPV type 16 genome.

A Hypermethylated DMRs
n=782



B Hypomethylated DMRs
n=386

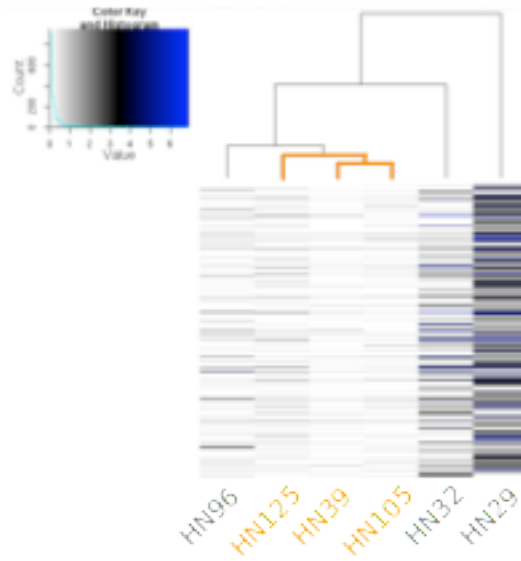


Figure S2: Hypermethylation of female chromosome X drives many of the hypermethylated differentially methylated regions (DMRs).

A. Heatmap of MeDIP-signal in reads per million (rpm) for each sample in the two cohorts in the hypermethylated (HPV+ rel. to HPV-) set of DMRs. For the purpose of visualisation, outliers with signal greater than 7 rpm were removed. Samples in the HPV- cohort are shown in black, samples in the HPV+ cohort are highlighted in orange. The heatmap is clustered according to column. DMRs labelled by the green line map to chrX. **B.** Same as above for hypomethylated DMRs (HPV+ rel. to HPV-).

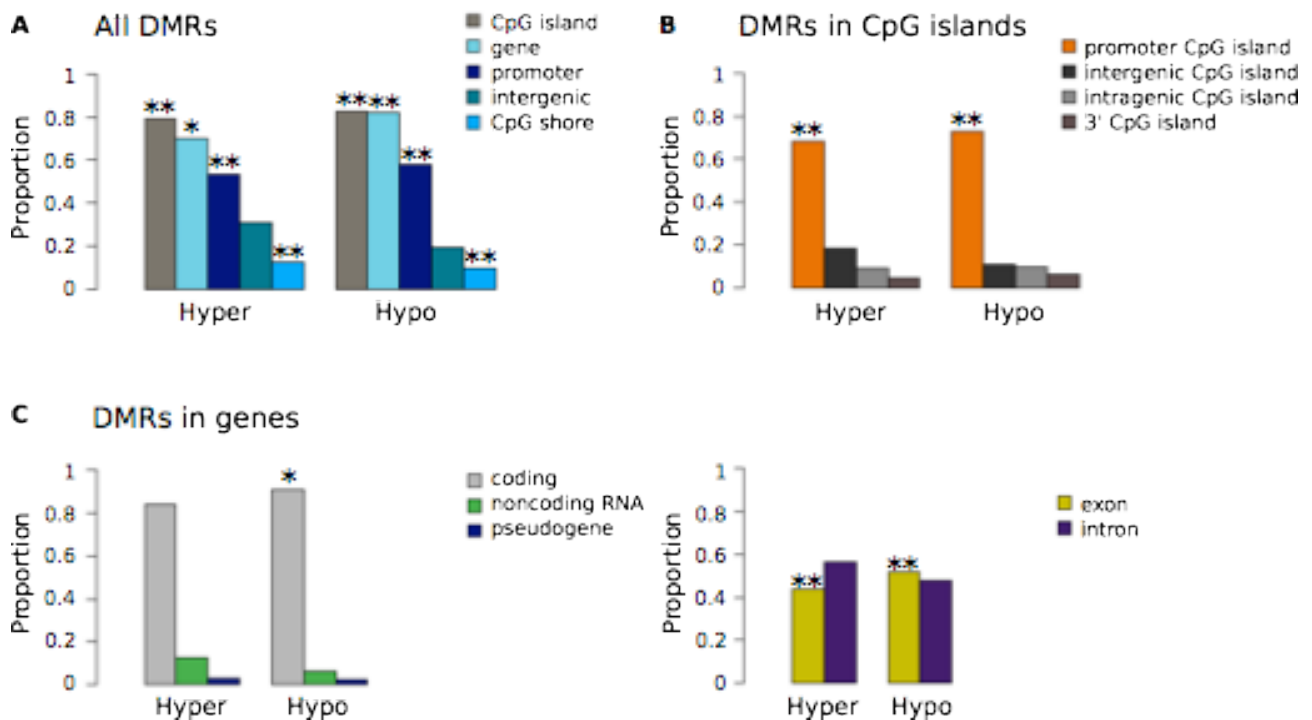


Figure S3: DMRs were annotated with a genomic feature if minimum 51% of the DMR overlapped the feature or 51% of the feature overlapped the DMR. Fisher's Exact test was used to determine the level of significance of the association as described in Materials & Methods. (*) indicates $P \leq 0.05$; (**) indicates $P \leq 0.001$. **A.** Overlap for the final set of DMRs (Hypermethylated $n=209$, Hypomethylated $n=244$) with genomic features. **B.** DMRs that fall into CpG Islands are significantly associated with promoter CpG islands but not intergenic, intragenic or 3' CpG islands. **C.** (left panel) Hypomethylated DMRs that fall into genes show a weak significant association with coding regions; no preference for a particular subclass of genic regions could be identified among hypermethylated DMRs. Both hyper- and hypomethylated DMRs that fall into genes show significant association with exons (right panel).

```
5521          aatatacaat tattgctgat gcaggtgact tttatttaca
5581 tcctagttat tacatgtta c gaaaa cga cg taaa cg tta ccatatttt tttcagatgt
5641 ctctttggct gcctagtgag gccactgtct acttgcctcc tgtoccagta tctaaggttg
5701 taagca cg ga tgaatatgtt gca cg cacia acatatatta tcatgcagga acatccagac
5761 tacttgcagt tggacatccc tattttccta ttaaaaaacc taacaataac aaaatattag
5821 ttctaaaagt atcaggatta caatacaggg tatttagaat acatttacct gacccaata
5881 agtttggttt tcttgacacc tcattttata atccagatac acag cg gctg gtttgggcct
5941 gtgtaggtgt tgaggtaggc cg tggtcagc cattaggtgt gggc
```

Figure S4: 421 bp viral DNA segment of interest.