

**Epigenetic deregulation of GATA3 in neuroblastoma is associated with increased
GATA3 protein expression and with poor outcomes**

Bader Almutairi, Jessica Charlet, Anthony R. Dallosso, Marianna Szemes,

Heather C. Etchevers, Karim T. A. Malik, Keith W. Brown.

Supplementary data:
Supplementary figures S1 to S16
Supplementary tables S1 to S6

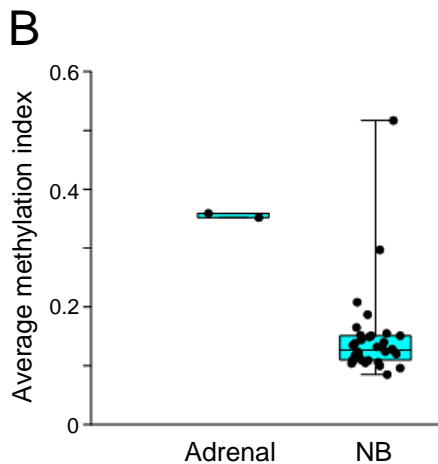
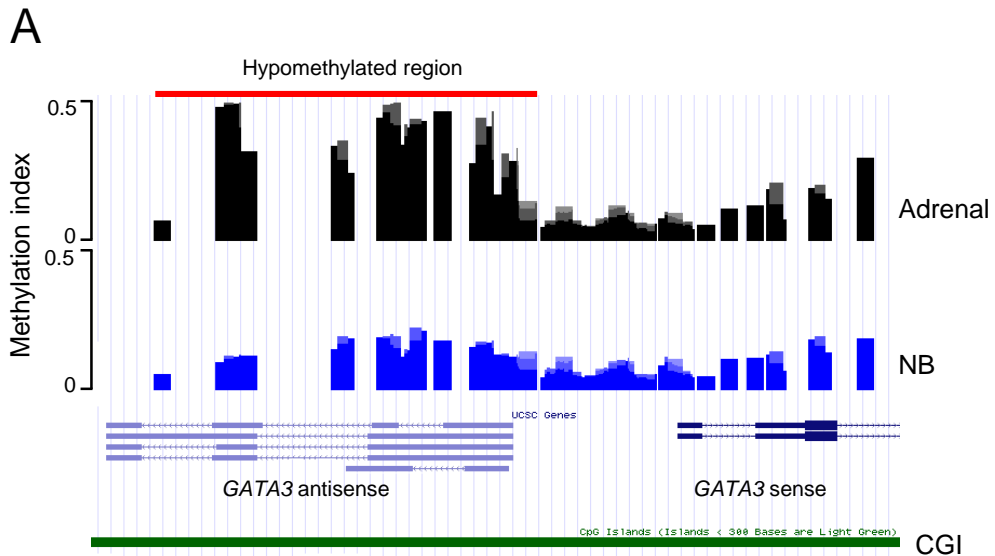


Figure S1: GATA3 DNA methylation from public dataset GSE54719

A: DNA methylation index from Illumina Human Methylation 450K Beadchip analysis, showing the mean results from two adrenal gland samples (Adrenal) and the mean results from 35 neuroblastoma tumours (NB), across the *GATA3* locus, taken from GSE54719. Results were uploaded as a Bedgraph file to the UCSC genome browser (<http://genome.ucsc.edu/>), using GRCh37/hg19.

B: Boxdotplot of the methylation index of each sample, averaged across the hypomethylated region (shown in red at the top of part A).

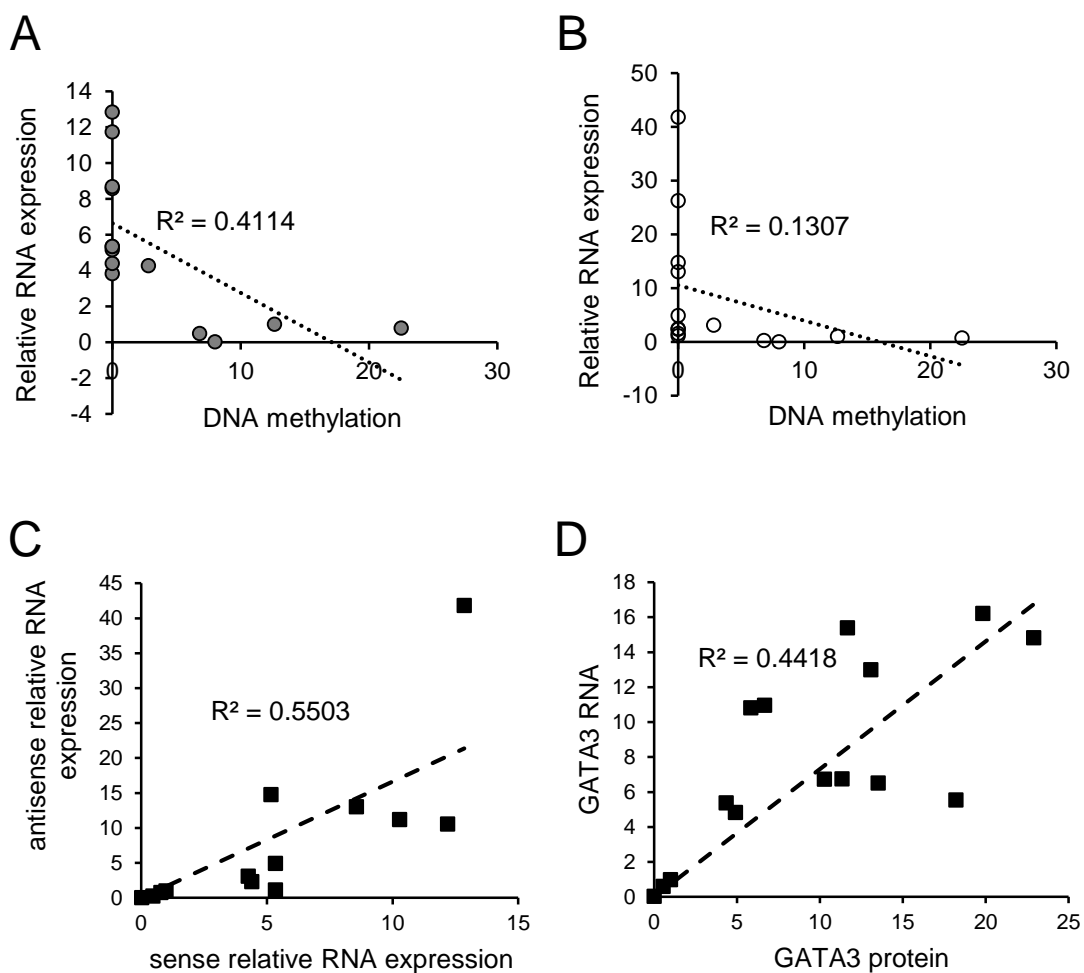


Figure S2: Correlations between GATA3 DNA methylation, RNA and protein expression

GATA3 sense and antisense RNA expression, protein expression and DNA methylation in normal tissues (FA, hNCC, DRG/SG) and neuroblastoma cell lines. DNA methylation is the average of pyrosequencing assays 01 and 02. RNA levels were assayed by real-time PCR and normalised to the endogenous levels of *TBP*, and expressed relative to hNCC RNA. Protein levels were assayed by Western blotting relative to ACTIN, and expressed as a ratio of the level in FA.

A: GATA3 sense RNA expression and DNA methylation.

B: GATA3 antisense RNA expression and DNA methylation.

C: GATA3 antisense and sense RNA expression.

D: GATA3 sense RNA and protein expression.

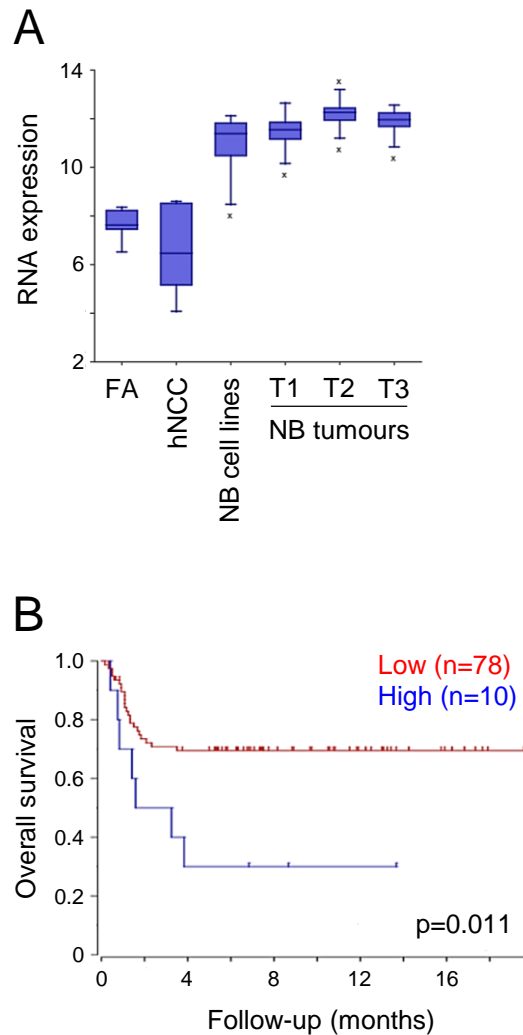


Figure S3: GATA3 expression and survival data from R2

A: Boxplot of GATA3 sense RNA expression determined by Affymetrix microarray, taken from R2 (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi>), using 13 samples of FA collected from different datasets (GSE3525, GSE7307, GSE8514), 5 samples of hNCC (GSE1430), 24 samples of NB cell lines (GSE28019) and three sets of NB tumours, containing 51 samples in T1 (GSE16237), 30 samples in T2 (GSE13136) and 88 samples in T3 (GSE16476). GATA3 RNA expression was significantly upregulated in NB cell lines and three sets of NB tumours compared to normal tissues (FA and hNCC) (One Way ANOVA; $p = 6.4 \times 10^{-50}$).

B: Kaplan-Meier analysis demonstrates overall survival probability in NB tumour data set GSE16476. Tumours with a high level of GATA3 RNA (“high”) (n=10) showed poorer outcomes compared to tumours expressing a low level of GATA3 RNA (“low”) (n=78).

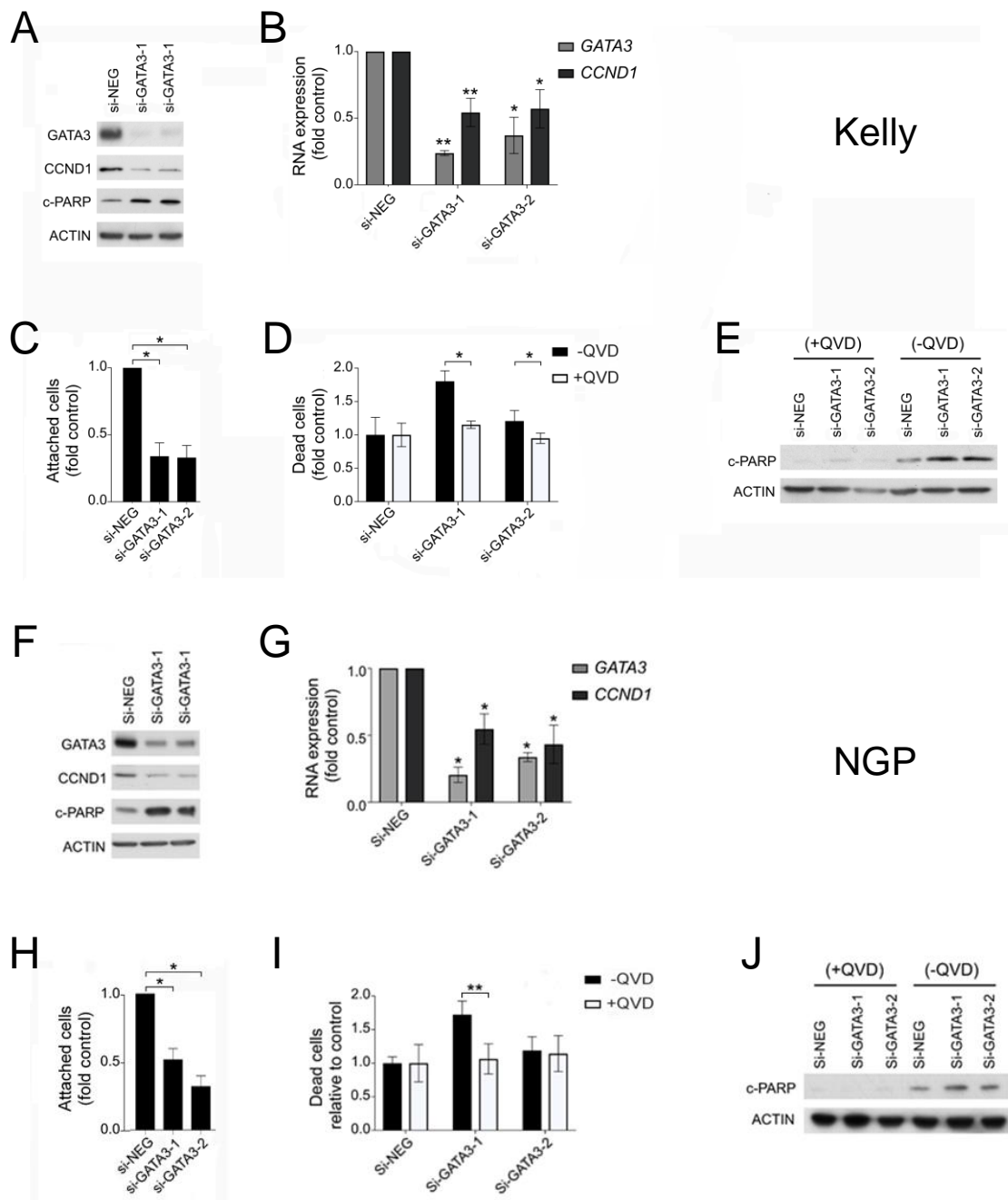


Figure S4: GATA3 biological functions

Growth of Kelly (A-E) or NGP (F-J) neuroblastoma cell lines after transfection with non-targeting siRNA (siNEG), or GATA3 siRNAs (si-GATA3-1 and si-GATA3-2).

A, F: Western Blot of GATA3, cleaved PARP (c-PARP) and CCND1 protein expression, with ACTIN as a loading control.

B, G: GATA3 and CCND1 RNA expression assayed by QPCR expressed relative to siNEG.

C, H: Attached cell counts relative to siNEG controls.

D, I: Dead cell counts (floating cells that were trypan blue permeable) expressed relative to siNEG control in cells treated caspase 3 inhibitor QVD (+QVD, unfilled bars) or with drug solvent (DMSO; -QVD, filled bars).

E, J: Western Blot of c-PARP protein expression in transfected cells with (+QVD) or without (-QVD) QVD- treatment, with ACTIN as a loading control.

A, B, C, D, F, G, H, I; mean \pm S.E.M, n=3; * p<0.05; ** p<0.01, t test. E, J; representative of n=3.

Uncropped blots are shown in supplementary figures S13 and S14.

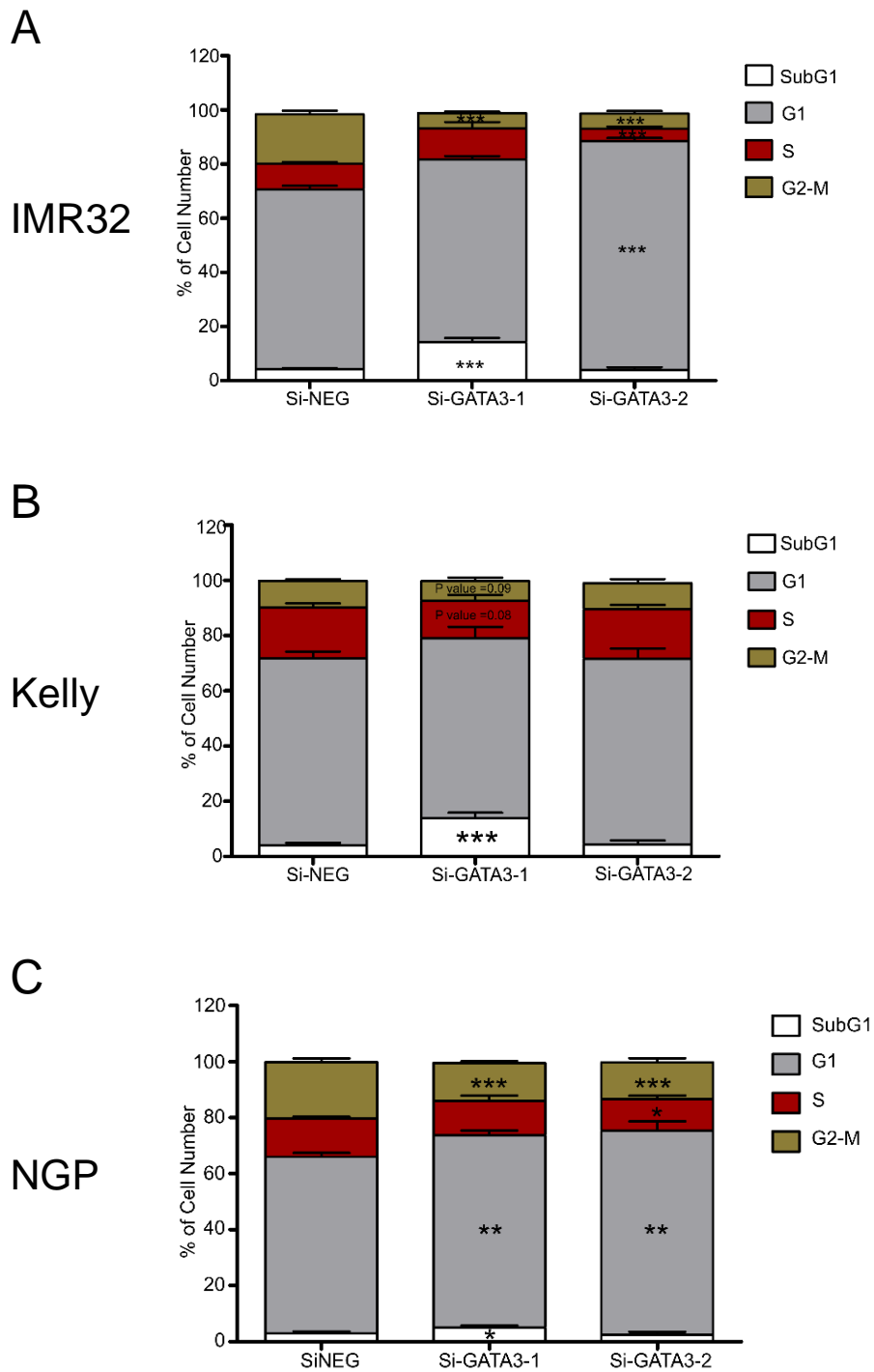


Figure S5: Effect of GATA3 depletion on cell cycle distribution

Cell cycle analysis by flow cytometry of cell lines 72hr after transfection with non-targeting siRNA (si-NEG) or with siRNAs targeting GATA3 (si-GATA3-1 and -2). Results shown are the mean \pm standard deviation of three experiments.

A: IMR32, B: Kelly, C: NGP.

***p<0.001; unpaired t test; in comparison to si-NEG.

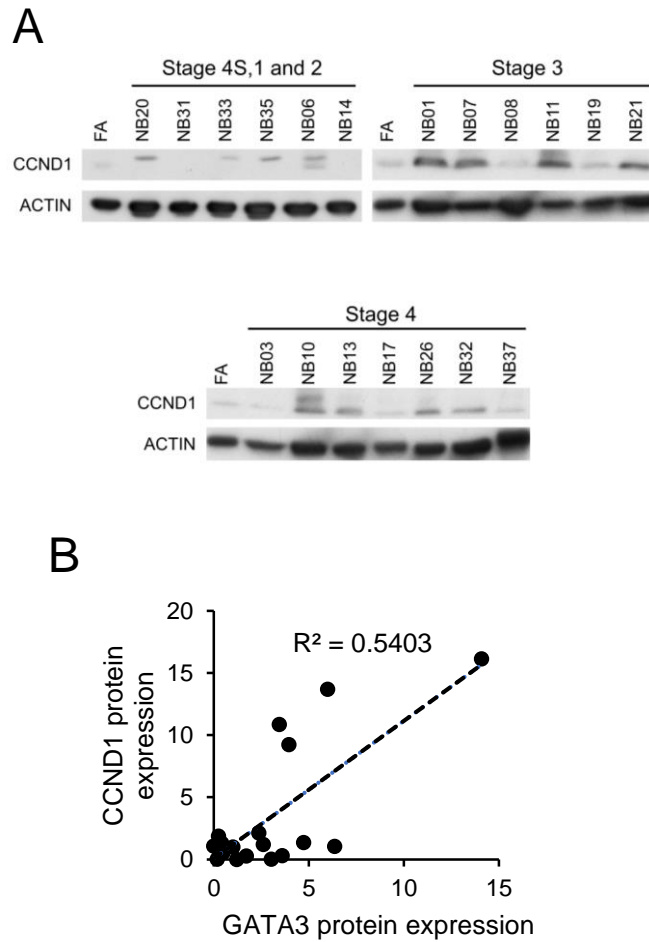


Figure S6: CCND1 expression in neuroblastoma

A: CCND1 protein levels assayed by Western blot in fetal adrenal (FA) and 19 neuroblastomas (NB) with ACTIN as loading control. Uncropped blots are shown in supplementary figure S15.

B: Correlation between GATA3 and CCND1 protein expression in neuroblastoma tumours (n=19).

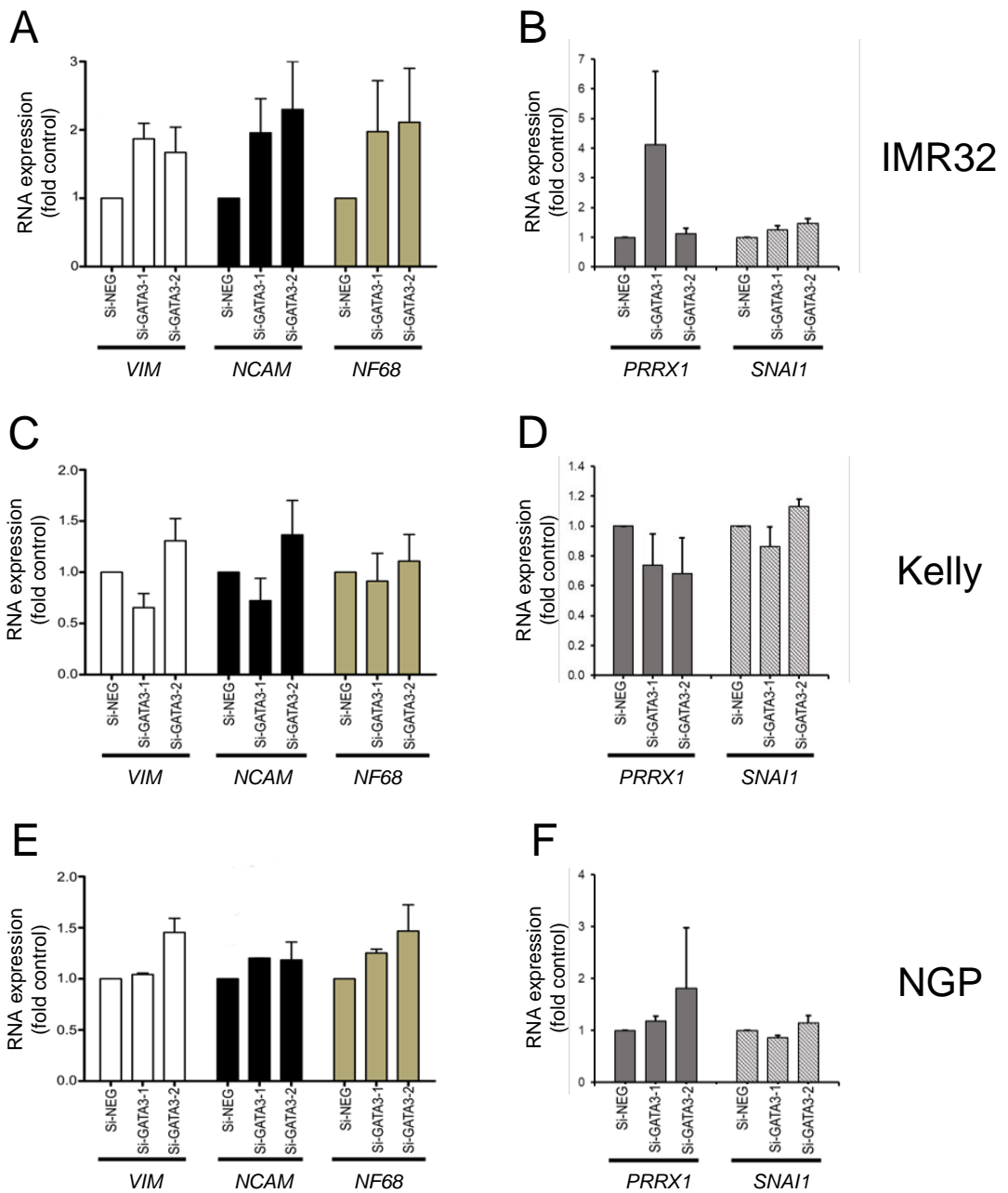
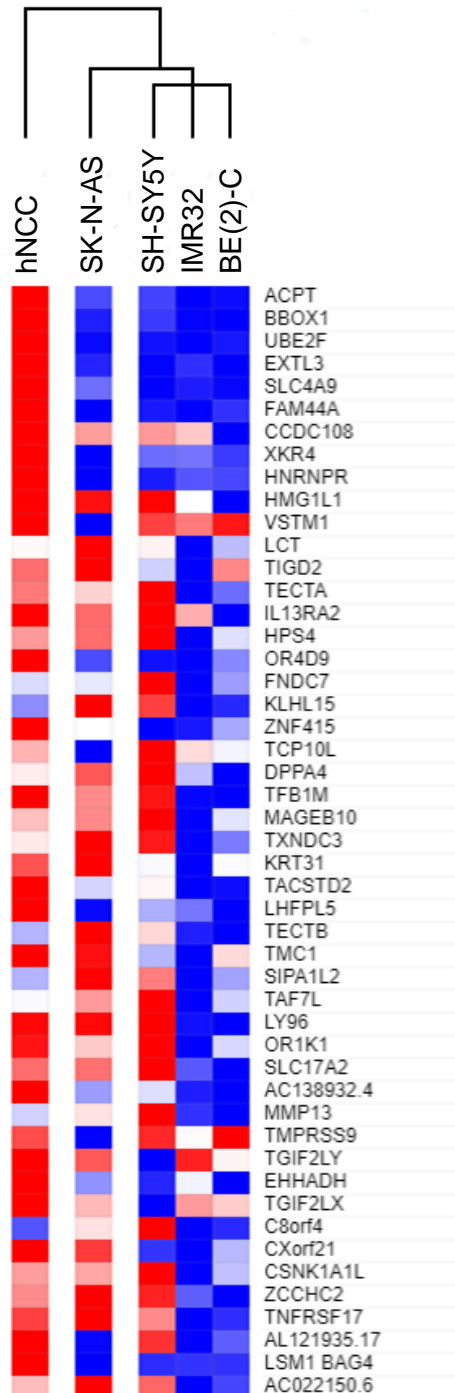


Figure S7: Differentiation markers in NB cell lines

RNA expression of *VIM*, *NCAM* and *NF68* (A, C, E), *PRRX1* and *SNAI1* (B, D, F) in IMR32 (A, B), Kelly (C, D) and NGP (E, F) cell lines, 72 hours after siGATA3 transfection. RNA expression was assayed by QPCR, normalised to the housekeeping gene *TBP* and expressed relative to control (siNEG). Results shown are mean \pm S.E.M, n=3.

A



B

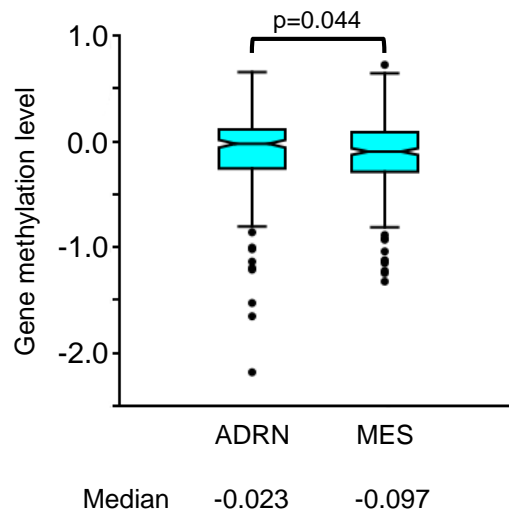


Figure S8: DNA methylation clustering of neural crest and neuroblastoma cell lines, and DNA methylation of ADRN and MES signature genes in neural crest cells

A: Hierarchical clustering of neural crest cells (hNCC) and four neuroblastoma cells lines by DNA methylation, using one minus Pearson clustering (<https://software.broadinstitute.org/morpheus/>). Only a portion of the heat map is shown below the hierarchical cluster.

B: Boxplot of DNA methylation levels in hNCC of adrenergic (ADRN) and mesenchymal (MES) signature genes (gene lists from van Groningen et al (2017) Nat Genet 49:1261-1266). p value from Mann-Whitney test.

Gene methylation levels were derived from the mean probe ratios of the differentially methylated probes within 700bp of the transcriptional start site (\log_2), from our previously published MCIP study (Charlet et al (2017) Mol Carcinogenesis 56: 1290-1301).

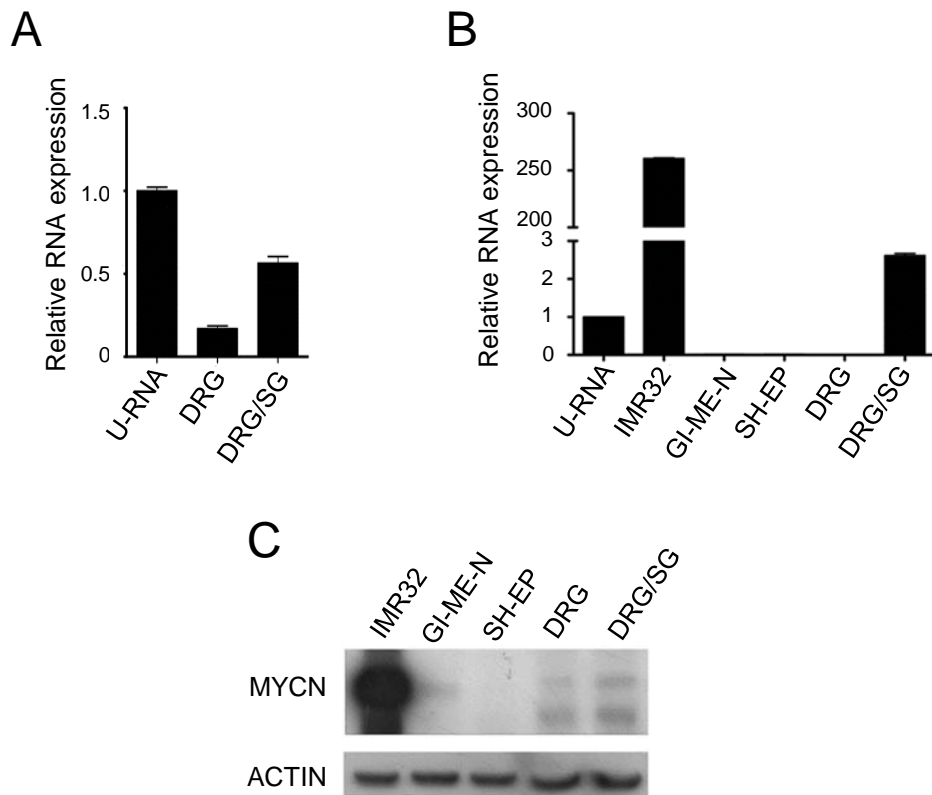


Figure S9: DRG/SG express *DBH* and *MYCN*

A: *DBH* RNA expression in universal RNA (U-RNA), dorsal root ganglia cultured cells (DRG) and dorsal root ganglia/sympathetic ganglia cultured cells (DRG/SG).

B: *MYCN* RNA expression in U-RNA, IMR32 (amplified *MYCN*), GI-ME-N (unamplified *MYCN*), SH-EP (unamplified *MYCN*), DRG and DRG/SG.

In A and B, RNA levels were assayed by real-time PCR and normalised to the endogenous levels of *TBP*, and expressed relative to U-RNA.

C: Western blot of *MYCN* protein expression in IMR32 (amplified *MYCN*), GI-ME-N (unamplified *MYCN*), SH-EP (unamplified *MYCN*), DRG and DRG/SG, with ACTIN as the loading control. Uncropped blots are shown in supplementary figure S16.

Figure 3B

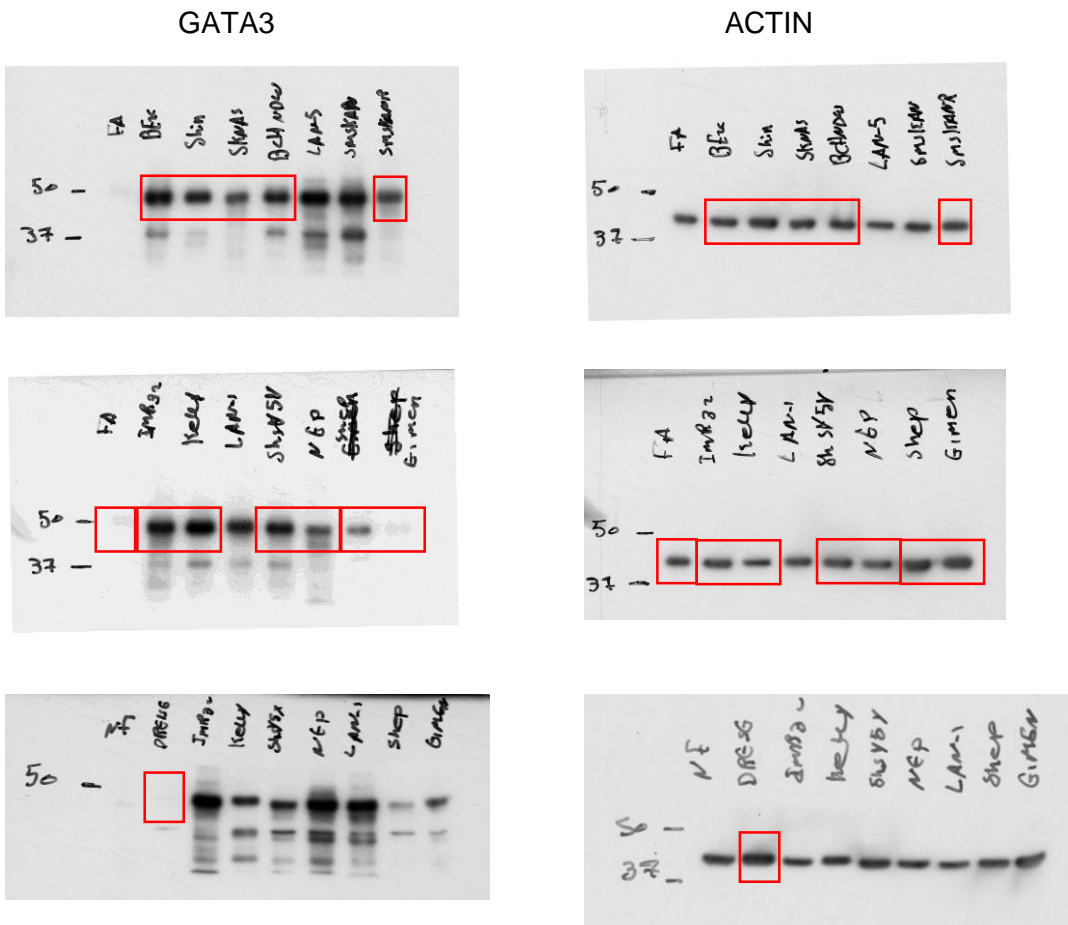


Figure 3C

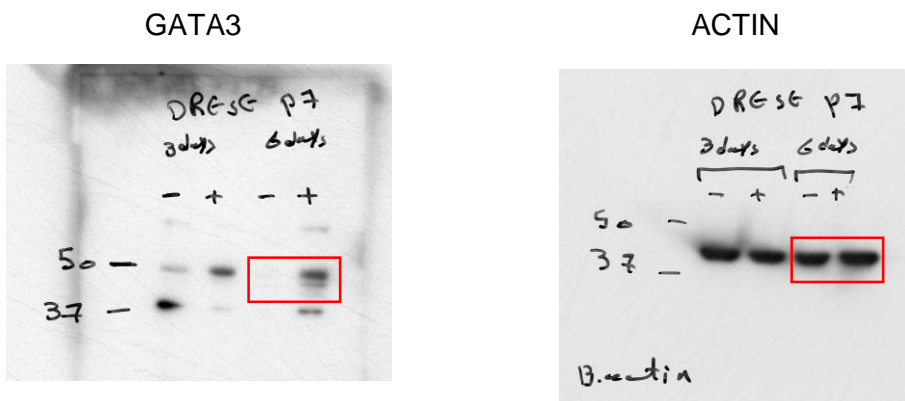


Figure S10: Uncropped blots for figure 3
Areas used in the figure are outlined by red boxes

Figure 4C

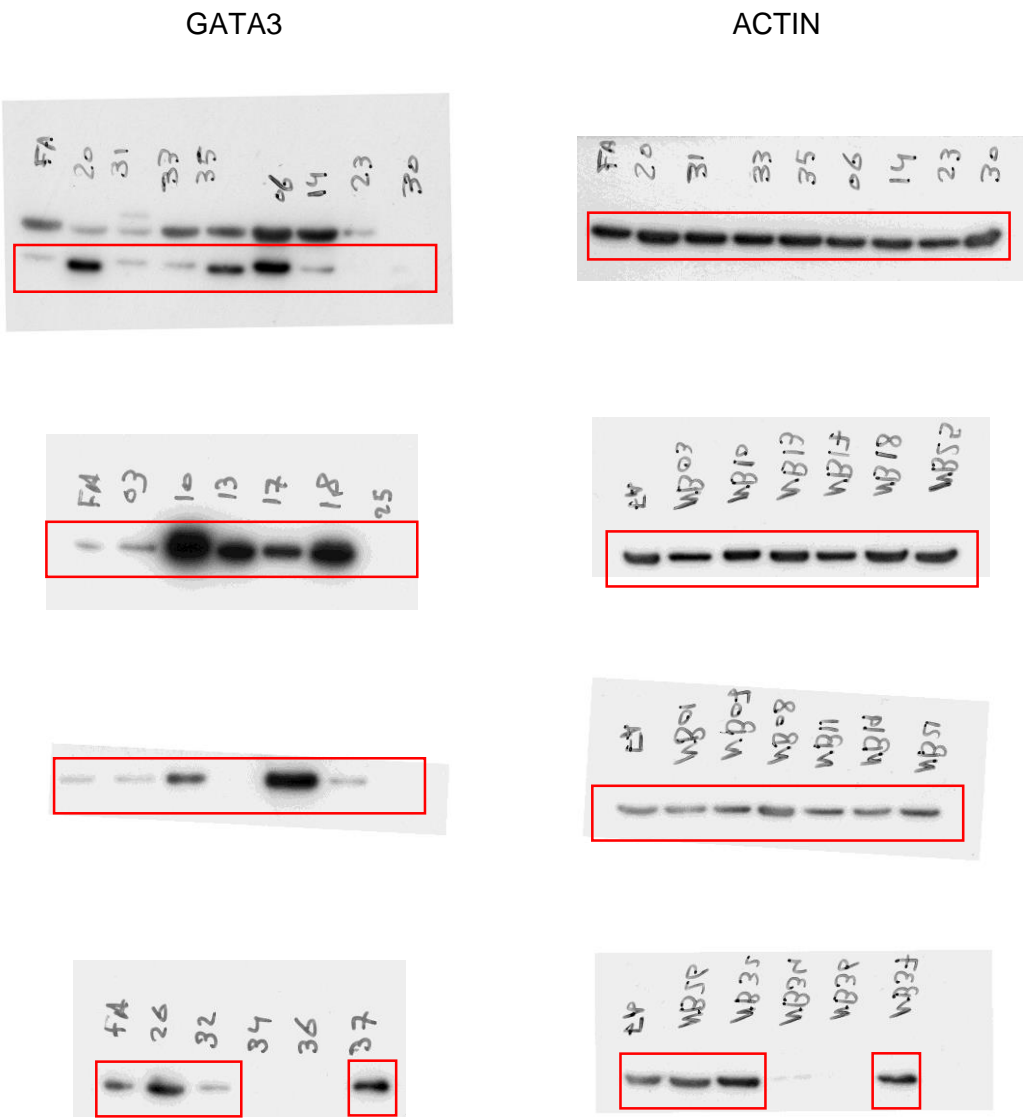


Figure S11: Uncropped blots for figure 4
Areas used in the figure are outlined by red boxes

Figure 5A

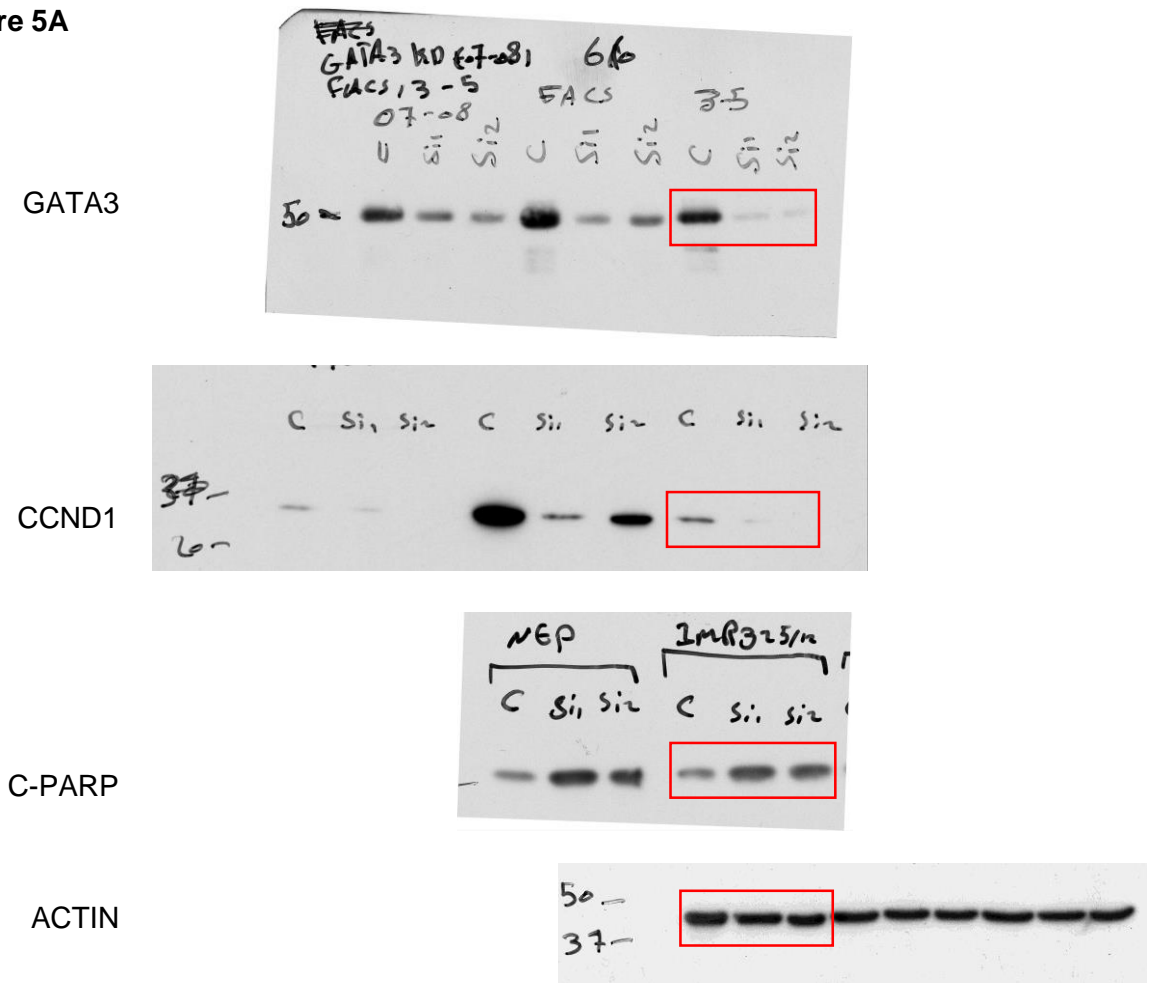


Figure 5E

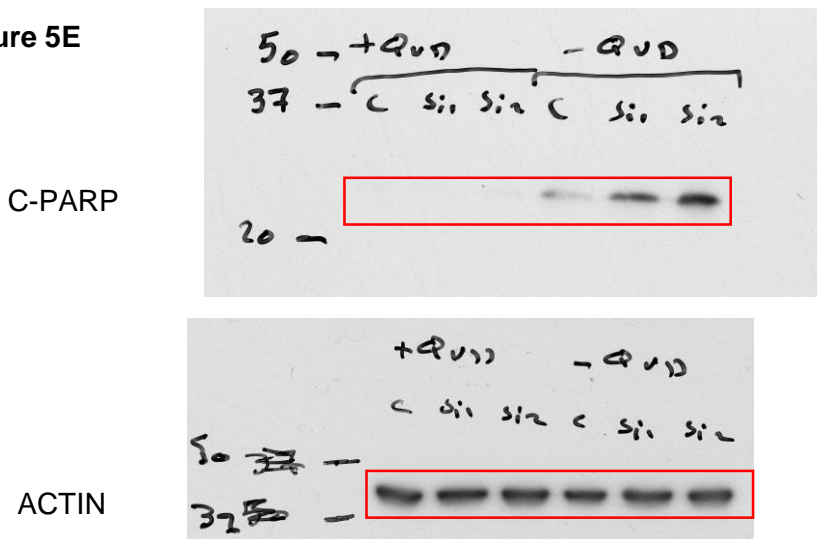


Figure S12: Uncropped blots for figure 5
Areas used in the figure are outlined by red boxes

Figure S4A

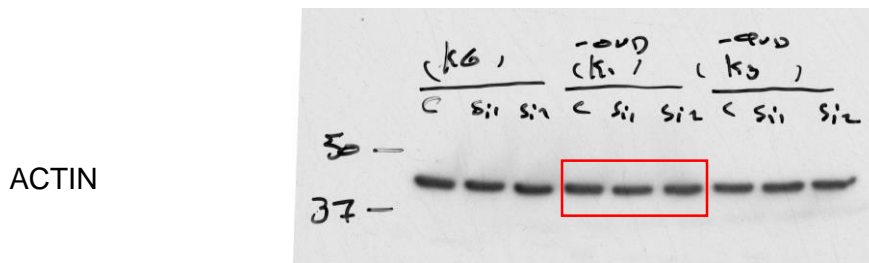
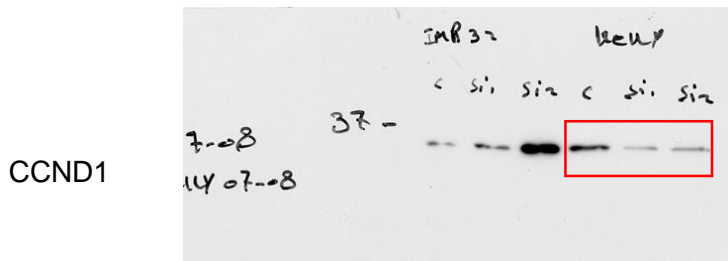
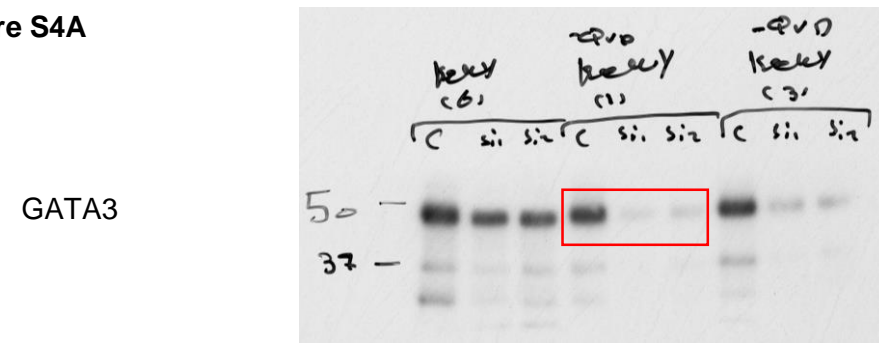


Figure S4E

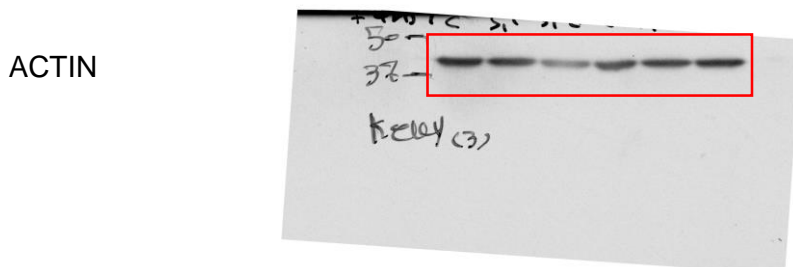
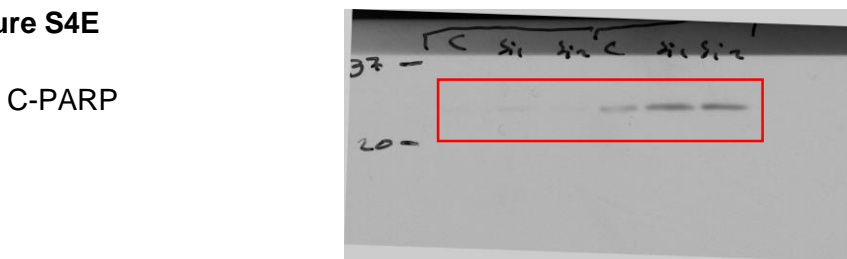


Figure S13: Uncropped blots for supplementary figure S4A, E
Areas used in the figure are outlined by red boxes

Figure S4F

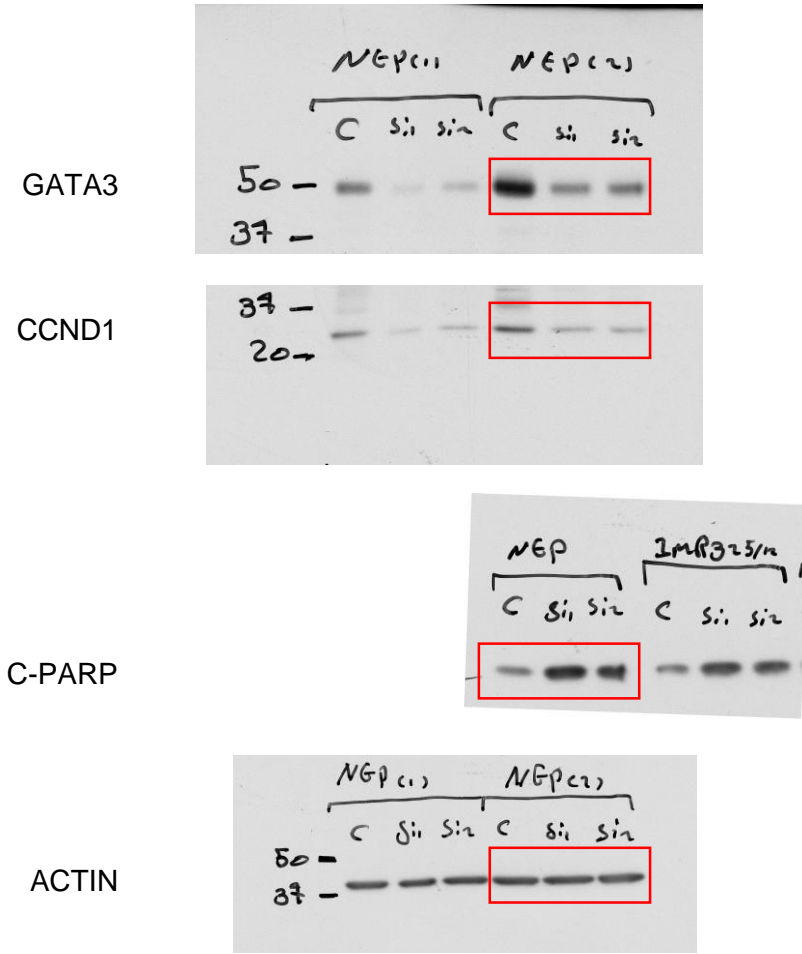


Figure S4J

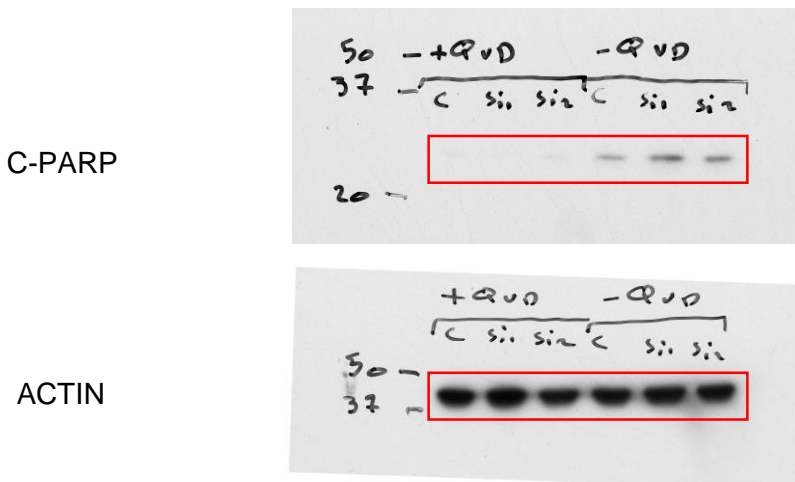


Figure S14: Uncropped blots for figure S4F, J
Areas used in the figure are outlined by red boxes

Figure S6A

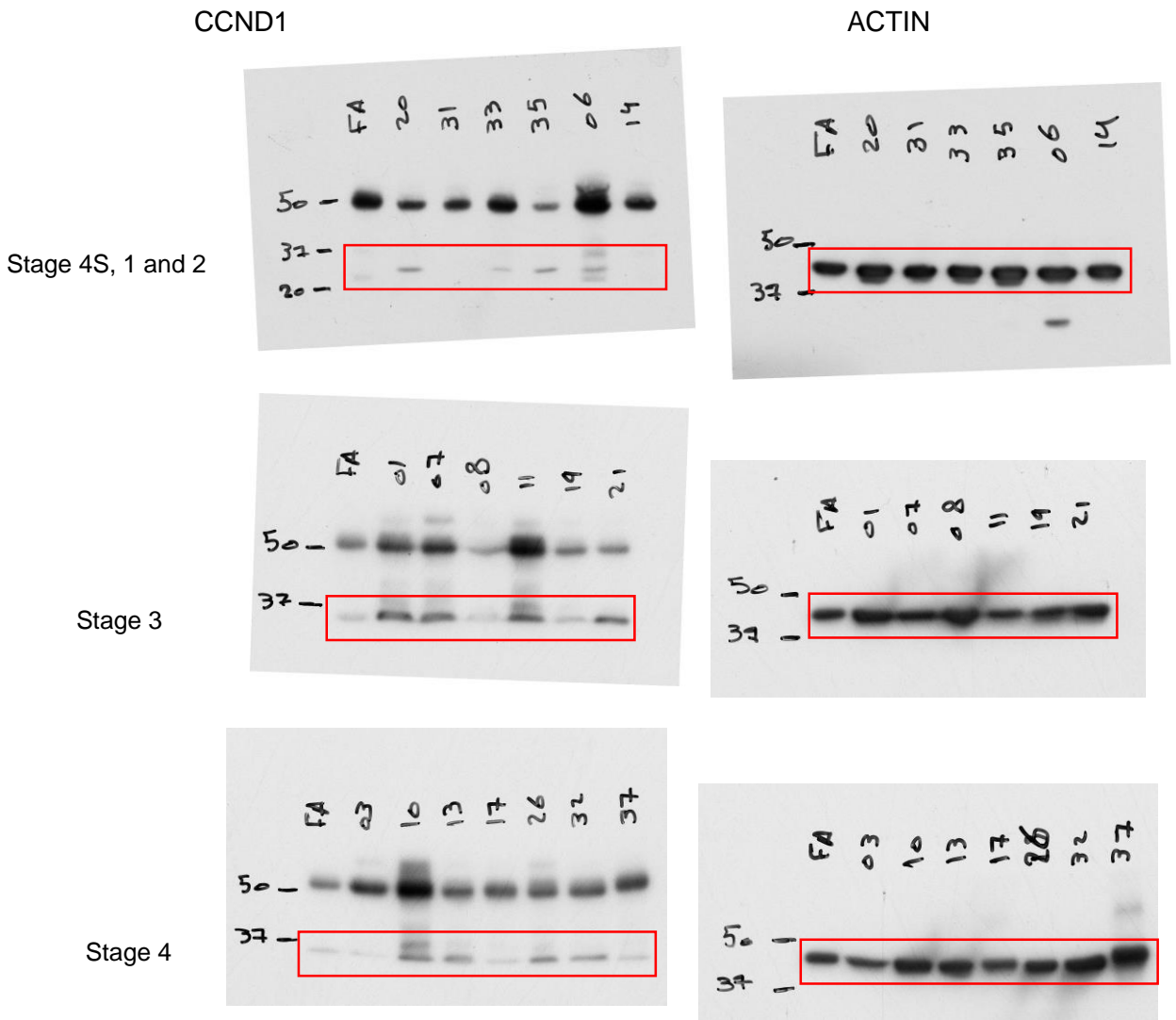


Figure S15: Uncropped blots for figure S6A
Areas used in the figure are outlined by red boxes

Figure S9C

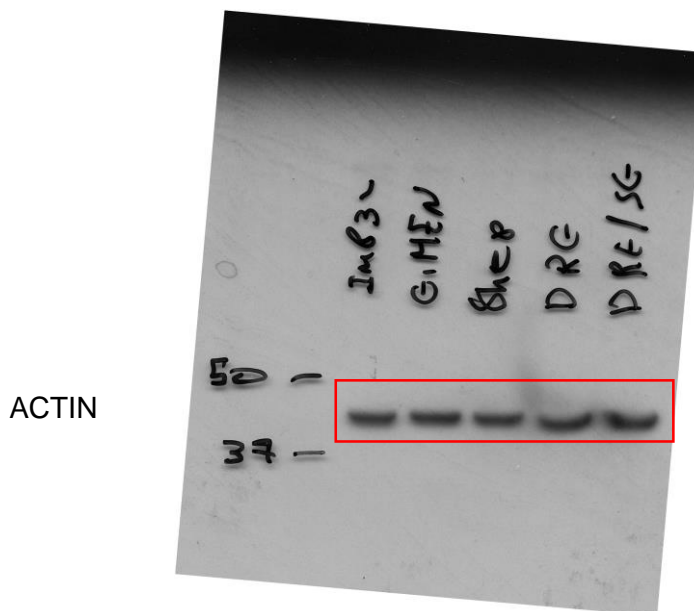
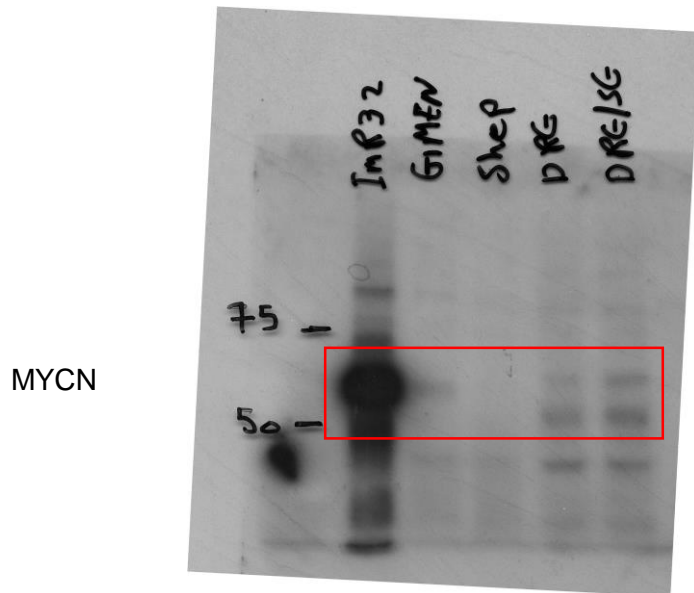


Figure S16: Uncropped blots for supplementary figure S9
Areas used in the figure are outlined by red boxes

Table S2: Gene ontology of hypomethylated genes

Gene ontology of the genes identified by Chipmonk software as being hypomethylated between neural crest and neuroblastoma cell lines - see Table S5 for full list. Gene ontology was assessed using the statistical overrepresentation test in PANTHER (<http://www.pantherdb.org/>). Only results with P<0.05 are shown.

	Reference list	Client Text Box Input
Mapped IDs:	16906 out of 16844	65 out of 70
Unmapped IDs:	7125	2
Multiple mapping information:	1509	5

	2006-11-02_HG18_CpG_Promo gene list.txt (REF)	Client Text Box Input					
PANTHER GO-Slim Biological Process	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
regulation of biological process	2143	23	8.9	2.6	+	0.0000112	0.000303
biological regulation	2637	26	11.0	2.4	+	0.0000168	0.00041
response to stimulus	2337	31	9.7	3.2	+	7.54E-10	6.13E-08
G-protein coupled receptor signaling pathway	404	11	1.7	6.6	+	0.000001	0.0000489
cell surface receptor signaling pathway	1033	16	4.3	3.7	+	0.00000442	0.000154
sensory perception of smell	210	12	0.9	13.8	+	1.06E-10	2.58E-08
sensory perception of chemical stimulus	248	12	1.0	11.6	+	6.52E-10	7.95E-08
sensory perception	387	13	1.6	8.1	+	8.34E-09	0.00000509
neurological system process	856	15	3.6	4.2	+	2.13E-06	0.0000865
system process	945	15	3.9	3.8	+	6.98E-06	0.000213
single-multicellular organism process	1474	18	6.1	2.9	+	2.56E-05	0.000568
multicellular organismal process	1488	18	6.2	2.9	+	2.90E-05	0.00059

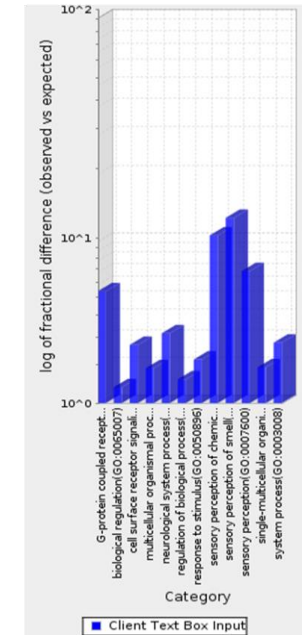
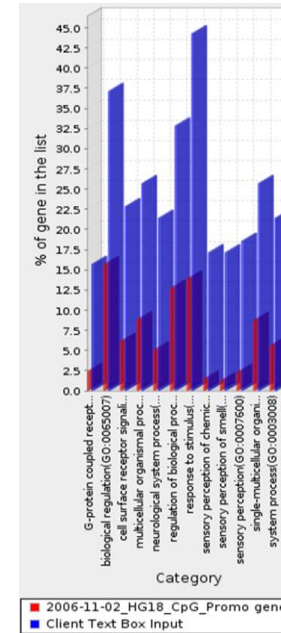


Table S3: Neuroblastoma cell lines used in the study

N = neuronal subtype, S = substrate adherent subtype, I = intermediate subtype
M = male, F = female
LN = lymph node, BM = bone marrow
Blank = not known.

Data taken from: Thiele, C. J. (1998) Neuroblastoma cell lines, *Journal of Human Cell Culture* 1: 21-53
Reynolds CP, et al (1986) Human Neuroblastoma Cell-Lines Established before and after Therapy. *Journal of the National Cancer Institute*, 76(3):375-87
Cornaglia-Ferraris P, et al (1990) A new human highly tumorigenic neuroblastoma cell line with undetectable expression of N-myc. *Pediatr Res*, 27(1):1-6
BCH-N-DW: personal communication, C. McConville

Cell line	Cell type	Parental cell line	Patient age (months)	Sex	Primary site	Metastatic site	Cell line origin	Treated	MYCN amplification	Stage
BE(2)-C	I	SK-N-BE(2)	26	M		BM	BM	Yes	Yes	4
BE(2)-M17	N									
BCH-N-DW	N & S	-	17	M	Abdomen	BM	BM	No	Yes	4
GI-ME-N	S	-	24	F	Adrenal	LN, BM	BM	Yes	No	4
IMR32	N	-	13	M	Abdomen		Abdomen	No	Yes	
Kelly	N	-	12	F					Yes	
LAN-1	N	-	24	M		BM	BM		Yes	4
LAN-5	N	-	5	M		BM	BM		Yes	4
NGP	N	-	30	M		Lung	Lung		Yes	4
SH-EP	S	SK-N-SH	48	F	Thorax	BM	BM	Yes	No	4
SH-IN	I									
SH-SY5Y	N									
SK-N-AS	I	-	96	F	Adrenal	BM	BM	Yes	No	4
SMSKAN	N	-	36	F	Pelvic	BM	Pelvis	No	Yes	4
SMSKANR	N	-	44	F	Pelvic	BM	BM	Yes	Yes	4

Table S4: Neuroblastoma tumours used in the study

Blank = not known, ND = not done.

Lab code	Diagnosis	Stage	Relapse	Death	MYCN amplification
NB01	NB	3	No	No	No
NB03	NB	4	Yes	Yes	ND
NB06	NB	2	Yes	No	No
NB07	NB	3	Yes	Yes	Yes
NB08	Ganglio NB	3	No	No	No
NB10	NB	4	Yes	Yes	Yes
NB11	NB	3	Yes	Yes	No
NB13	NB	4	No	No	Yes
NB14	Ganglio NB	2	Yes	No	No
NB16	NB in situ	-----	-----	Yes ¹	No
NB17	NB	4	Yes	Yes	No
NB18	NB	4	Yes	Yes	Yes
NB19	NB	3	No	No	No
NB20	NB	4S	Yes	No	No
NB21	NB	3	No	No	No
NB23	Ganglio NB	2	No	No	No
NB25	NB	4	Yes	Yes	No
NB26	NB	4	Yes	Yes	No
NB27	Ganglio NB		No	No	No
NB30	NB	2	No	No	No
NB31	NB	4S	No	No	No
NB32	NB	4	Yes	Yes	No
NB33	Ganglio NB	1	Yes	No	No
NB35	NB	1	No	No	ND
NB37	NB	4	No	No	ND

¹Sudden infant death syndrome

Table S5: siRNAs used in the study

siRNA sequences taken from:

Shan L, et al (2014) GATA3 cooperates with PARP1 to regulate CCND1 transcription through modulating histone H1 incorporation. *Oncogene*, 33(24):3205-16

siRNA	Sequence 5'-3'
si-GATA3-NEG	UUCUCCGAACGUGUCACGU
si-GATA3-1	CUCUGGAGGAGGAAUGCCA
si-GATA3-2	CUACAAGCUUCACAAUAAU

Table S6: Oligonucleotide primers used in the study

Application	Gene		5' - 3' sequence
QPCR	<i>CCND1</i>	F	ACTACCGCCTCACACGCTTC
		R	TTCGATCTGCTCCTGGCAG
	<i>DBH</i>	F	CCAATATCCCCGAACCGGAG
		R	GAAGCCCTTTGGAAGCTCCT
	<i>GATA3 sense</i>	F	TACGTGCCCGAGTACAGCTC
		R	ACAGTTCACACACTCCCTGC
	<i>GATA3 antisense</i>	F	TAGAAGAACCGTCCCTGAA
		R	CTTAGCAAAAATGCGTGTGC
	<i>MYCN</i>	F	TTTTAGATCTATTAACGAACGGGGCGG
		R	AAAGATATCAAAGTGCTATAAGATGCAC
	<i>NCAM</i>	F	TCACCCTGGTGTGCGATG
		R	ATGTA CT TCTCATCGTCTTCCTCTTG
	<i>NF68</i>	F	GTGACCAAGCCCGACCTTT
		R	ATTCCTCAGCGTTCTGCATGT
	<i>PRRX1</i>	F	CATCGTACCTCGTCCTGCTC
		R	GAATCCGTTATGAAGCCCCT
	<i>SNAI1</i>	F	TTCTCACTGCCATGGAATTCC
		R	GCAGAGGACACAGAACCAGAAA
	<i>TBP</i>	F	GCCCGAAACGCCGAATAT
		R	CCGTGGTTCGTGGCTCTCT
<i>VIM</i>	F	CTGCCAACCGGAACAATGA	
	R	GTA CT CAGTGGACTCCTGCTTT	

F = forward, R = Reverse