Supplementary Methods

Raw DNA methylation data processing

Raw EPIC array intensity values were processed using the Minfi package of the R environment for statistical computing. After background correction we excluded methylation markers with detection p-value >10^{-16} in >20% of samples and samples/patients with a call rate <90% prior to functional normalization.(1) Principal component analysis (PCA) based on gonosomal methylation markers was used to verify correct gender annotation (n=1 sample excluded). Methylation markers on gonosomes, cross-reactive/multi-mapping, non-CpG probes or with an inconsistent base extension were excluded.(2,3) Sensitivity of findings was assessed using the GenomeStudio software (v2011.1) Methylation Module (v1.9.0) for initial quality control: Methylation markers with detection p-value >0.01 in >20% of the samples and samples with a call rate <80% were removed prior to quantile normalization per entity and across entities.

Differential methylation analyses

The 80% most variable probes by median absolute deviation (MAD) were used for nonparametric Jonckheere–Terpstra tests to identify markers with gradient methylation levels (increasing or decreasing) along the sequence GSD→low-grade dysplasia→high-grade dysplasia→GBC. P-values were adjusted for multiple testing using Bonferroni–Holm correction. Adjusted p-values <0.05 were considered to indicate a gradual methylation change. Methylation differences between disease groups were quantified using linear models adjusting for patient's age and gender. Methylation markers were annotated with functional elements and genes using the Illumina Infinium MethylationEPIC BeadChip annotation manifest.(4) PANTHER (v14, www.pantherdb.org) gene ontology (GO) enrichment analyses was done based on genes with differentially methylated markers. Differentially methylated regions (DMRs) are groups of neighboring methylation markers that share a significant methylation difference between two groups of samples (more than one DMR may occur per gene). DMRs in gene/isoform related promoters for GSD plus low-grade dysplasia samples,

compared with high-grade dysplasia plus GBC samples, were identified with the ChAMP R-package. (5)

CNV analysis

Copy number variations (CNVs) were assessed based on methylation array signal intensities using the conumee R package. Copy number segmentation was performed using standard parameters with GSD samples as the reference category and LogR thresholds of ±0.15 to identify altered segments.(6,7) The segmented data from GBC patients were additionally analyzed using GISTIC2.0 with the default parameters to identify segments recurrently altered.(8)

mRNA expression analysis

Total RNA was isolated from OZ and G-415 using standard protocols (Qiagen). cDNA was generated from 1.0 µg DNase I-treated RNA using SuperScript III Reverse Transcriptase (Invitrogen) and random hexamers (QIAGEN) and expression analyzed with a LightCycler 480 real-time PCR system (Roche) using human Universal ProbeLibrary hydrolysis probes (Roche), and LightCycler DNA Probes Master polymerase mix (Roche). Data were normalized to the average of the three housekeeping genes beta actin (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and hypoxanthine phosphoribosyltransferase 1 (*HPRT1*) to calculate relative gene expression. Primer design: Universal ProbeLibrary Assay Design Center (Roche)(**Suppl. Table S1**).

Epigenome-wide DNA methylation and RNA-sequencing data for G-145

For G-415 epigenome-wide DNA methylation data, generated using the same methodology as for tissue samples, and RNA-sequencing data were available. RNA sequencing libraries were prepared from 1.5 µg of total RNA using the NEBNext Ultra II Directional RNA Preparation Kit with the NEBNext Poly A Selection Module, and the NEBNext Multiplex Oligos for Illumina. Libraries were quantified on the Qubit (Thermo Fisher), quality-checked (DNA 1000 Chip; Agilent Bioanalyzer), pooled and then sequenced (Illumina HiSeq 2500, 125bp paired-end mode). Raw reads were quality controlled and filtered with FastQC (v0.11.2) (bioinformatics.bbsrc.ac.uk/projects/fastqc) and PRINSEQ (v0.20.3)

before alignment with STAR (2.5.2b) against GRCh38.90 reference genome and quantification with the Subread R package (v1.6.4).

qPCR expression analysis

cDNA was generated from 1.0 µg DNase I-treated total RNA using SuperScript III Reverse Transcriptase (Invitrogen) and random hexamers (QIAGEN). Expression was analyzed with a LightCycler 480 real-time PCR system (Roche) using human Universal ProbeLibrary hydrolysis probes (Roche), and LightCycler DNA Probes Master polymerase mix (Roche). Data were normalized to the average of the housekeeping genes beta actin (*ACTB*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and hypoxanthine phosphoribosyltransferase 1 (*HPRT1*). **Suppl. Table S1** shows the primers used (Universal ProbeLibrary Assay Design Center, Roche).

Immunohistochemistry assays for *CDO1*

Staining was performed on an automated immunostainer (Ventana BenchMark Ultra, Ventana Medical Systems, Tucson, USA) with the OptiView DAB IHC Detection Kit (Ventana Medical Systems). FFPE tissue sections were deparaffinized and rehydrated prior to heat-induced epitope retrieval. After blocking endogenous peroxidase, slides were incubated with the antibody LS-B12120 (CDO1, Biozol, USA), followed by incubation with OptiView Universal Linker and OptiView HRP Multimer. DAB-Chromogen images were classified by five trained pathologists (IG, FG, MTR, EM and GdT) into the categories negative, weak or strong, and discordant classifications were resolved by majority vote. CDO1 was stained in two separate sets of samples: (1) eleven slides (GSD n=5 and GBC n=6) that were left after conducting the EPIC, MassARRAY, RNAseq and qRT-PCR assays described in the manuscript and (2) twenty slides from an independent group of ten Chilean GBC patients with paired tumor and adjacent non-tumor samples.

Dual-color chromogenic ERBB2 in situ hybridization

Dual-color chromogenic in situ hybridization (dc-CISH) was performed as reported previously using the ZytoDot 2C SPEC ERBB2/CEN 17 Probe Kit

(Zytovision, Bremerhaven, Germany), which contains a digoxigenin-labeled probe specific for the ERBB2 locus at 17q12 and a dinitrophenyl-labeled probe specific for the alpha satellite centromeric region of chromosome 17.(9) In brief, tissue sections were deparaffinized and incubated with H_2O_2 to block endogenous peroxidase followed by incubation in a preheated EDTA pretreatment solution at 98 °C in a water bath. Thereafter, proteins were denatured in a humidity chamber using pepsin solution. After dehydration through an ethanol series, 10 µl of the ZytoDot 2C SPEC *ERBB2*/CEN 17 Probe was applied to each slide, which were then covered with slide glass coverslips and sealed with rubber cement (Fixogum; Marabu, Tamm, Germany). The slides were denatured at 78 °C for 5 min and hybridized at 37 °C in a humidified hybridization chamber (ThermoBrite™, Abbott Molecular, Chicago, IL, USA) overnight. Immunodetection was performed according to the manufacturer's instructions followed by counterstaining with nuclear blue solution.

References

1. Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;15:503.
2. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2016;45:e22.
3. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17:1–17.
4. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8:389–399.
5. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. Bioinformatics. 2014;30:428–430.
6. Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, et al. Practical implementation of DNA methylationand copy-number-based CNS tumor

diagnostics: the Heidelberg experience. Acta Neuropathologica. 2018;136:181–210.

7. Stone TJ, Keeley A, Virasami A, Harkness W, Tisdall M, Delgado EI, et al. Comprehensive molecular characterisation of epilepsy-associated glioneuronal tumours. Acta Neuropathologica. 2017;135:115–129.
8. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011;12:R41.
9. Albrecht T, Rausch M, Roessler S, Geissler V, Albrecht M, Halske C, et al. HER2 gene (ERBB2) amplification is a low-frequency driver with potential predictive value in gallbladder carcinoma. 2019;:1–10.

Suppl. Figure legends

Suppl. Figure S1

A–D) Distribution of normalized methylation values in the four investigated groups of patients for the markers with the smallest p-values from Jonckheere–Terpstra tests.

Suppl. Figure S2

A) Proportion of markers in functional genetic elements such as gene bodies and promoters (TSS200, TSS1500) among those with gradient (red) and nongradient (gray) methylation values as determined using the Jonckheere–Terpstra (JT) test. (Differences in proportions were tested using the proportion z-test per element category.) **B)** Age- and sex-adjusted methylation differences between GBC and gallstone samples for markers in Phantom5 enhancer regions. **C, D)** Enrichment analysis of genes with gradient methylation values along the sequence of GSD → Dysplasia → GBC based on JT test using the PantherDB gene ontology (GO) database and the PantherDB Molecular Pathways.

Suppl. Figure S3

A,B) Pooled methylation values for markers in the promoter-associated CpG islands of the indicated genes for GBC patients grouped by T stage **(A)** and M stage **(B)**. **C)** Distribution of methylation values across markers of the CpG island in the *TWIST1* promoter. **D)** Pooled methylation values of TWIST1 CpG markers in C) for GBC patients grouped by tumor grade (left) and T stage (right). **E)** Methylation values across markers of the CpG island in the *HBE1* promoter (left) and pooled beta values for these markers in GBC patients by T stage (right). **F)** Methylation values across markers of the CpG island in the *RPL22* promoter. (All p-values from ANOVA tests adjusting for baseline CpG differences; green = GSD, light green = LG-dysplasia, light red = HG-dysplasia, red = GBC)

Suppl. Figure S4

A) Comparison of age- and sex-adjusted methylation difference between GSD and GBC samples after Minfi- or GS-based selection for markers significant in Minfi-based JT tests. **B)** Comparison of JT test p-values after Minfi- or GS-based selection for all shared markers. **C)** Results of differentially methylated region analysis in candidate genes based on GS-based marker selection/pre-processing. In CDO1 and ZSCAN18 two DMRs were detected.

Suppl. Figure S5

Results from GISTIC2.0 analyses displaying regions with copy number gains **(A)** and losses **(B)** in GBC patients. False discovery rate (FDR)-corrected p-values are indicated on the bottom axis; the names of potentially relevant candidate genes are given (green line: FDR threshold at 0.1).

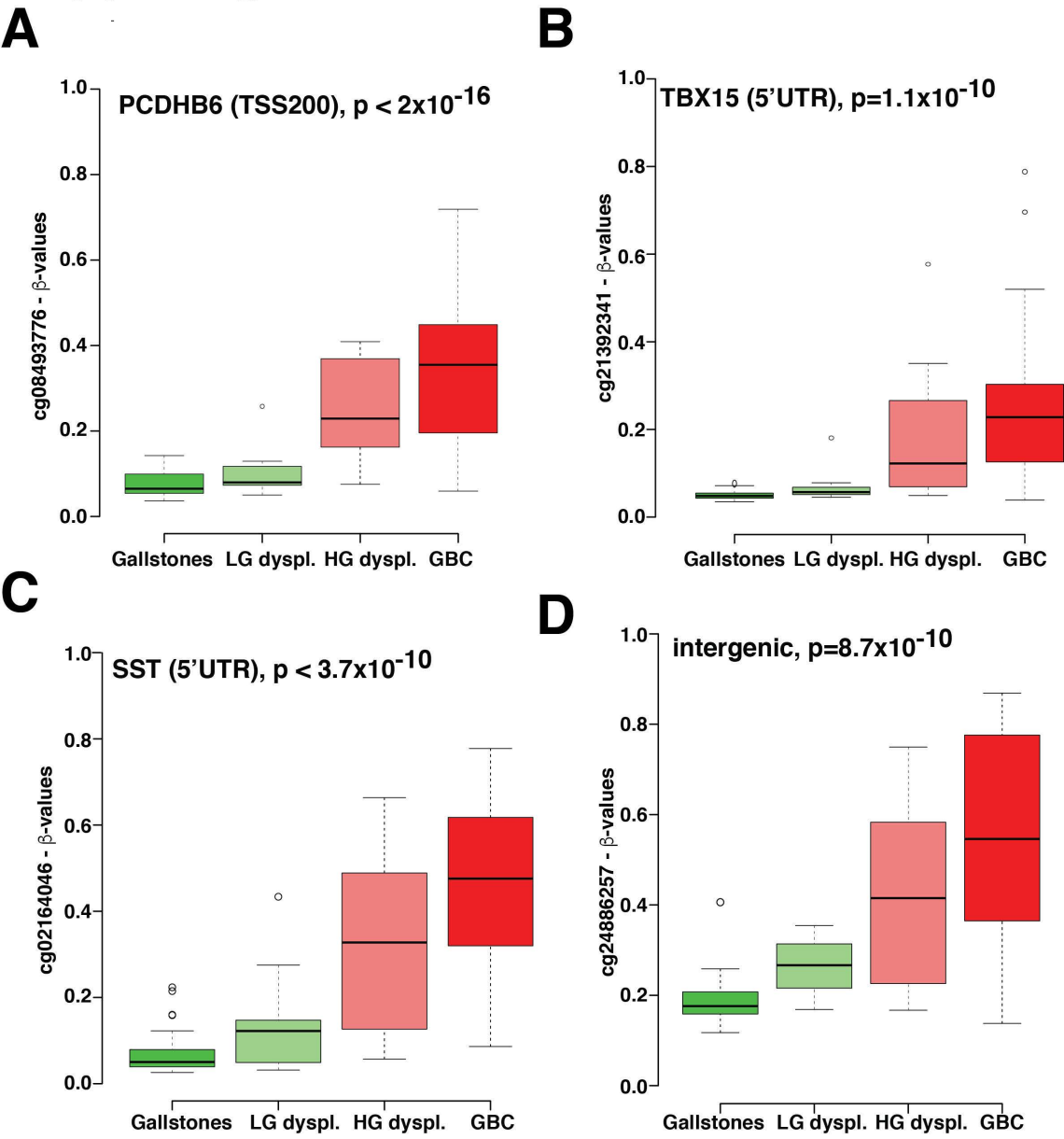
Suppl. Figure S6

A) Number of genes per sample with RNAseq mRNA expression data that fulfilled the quality control filters across disease groups. The p-value is based on a Kruskal-Wallis test. **B)** Correlation between β values of CpGs shown in **Figure 4B** and RNAseq mRNA expression in samples with high quality data (n=51). Pearson correlation coefficients and the corresponding p-values are shown **C)** RNAseq mRNA expression for three genes previously reported as hypermethylated and underexpressed in Indian GBC patients (10). P-values are based on Jonckheere-Terpstra tests across disease groups.

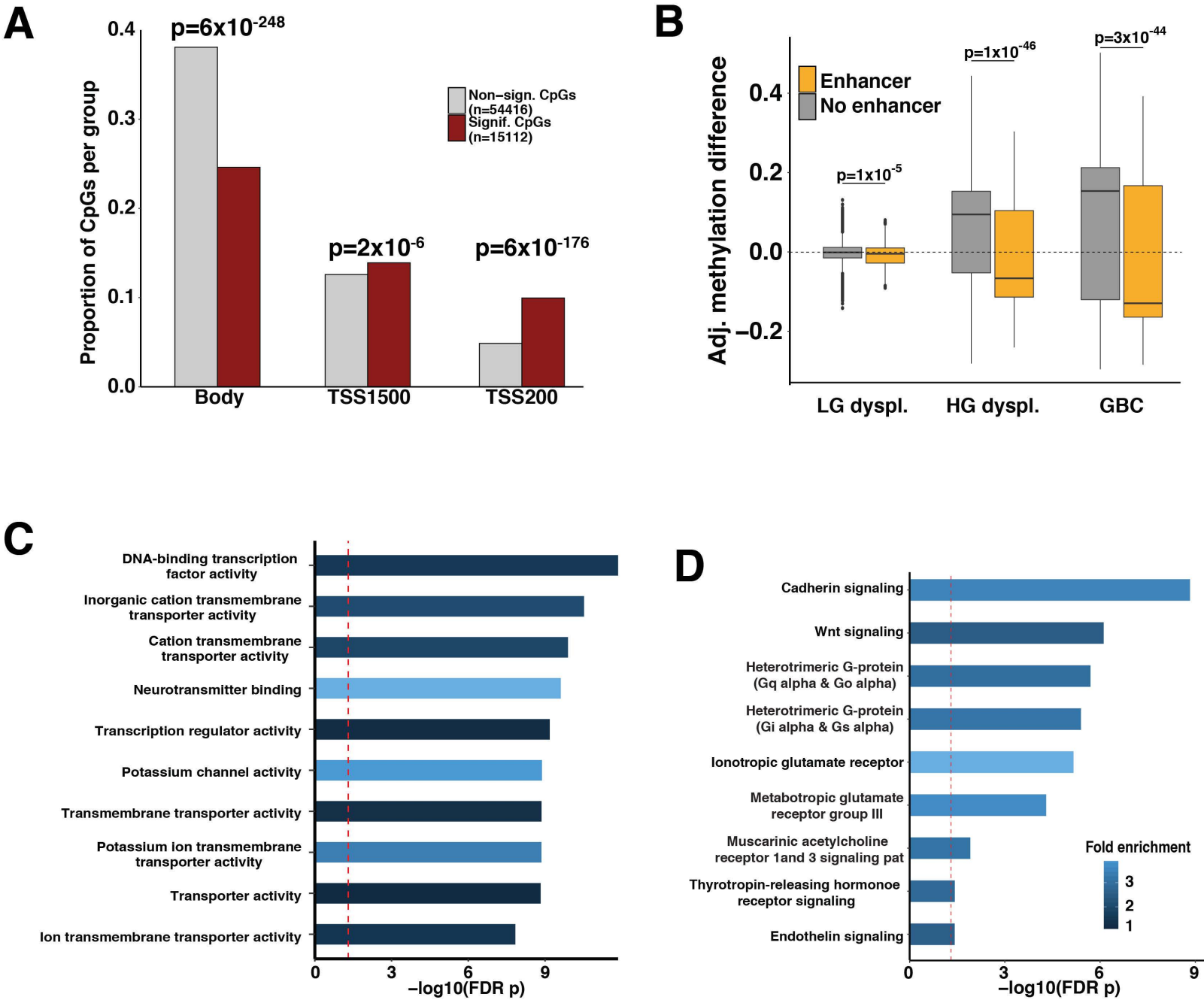
Suppl. Figure S7

A) Representative images showing CDO1 IHC staining in GSD (top) and GBC (bottom) samples. **B)** Association between CDO1 IHC staining categories and CDO1 methylation for CpG cg16707405 for samples with available EPIC methylation data (n=10), which were also investigated in the validation analyses, please see **Figure 4A**, **Figure 4C** and **Suppl. Figure S6B**. P-values based on Jonckheere-Terpstra tests across IHC categories. **C)** Representative images showing CDO1 IHC staining in an independent group of Chilean GBC patients with paired adjacent non-tumor (top) and tumor (bottom) samples.

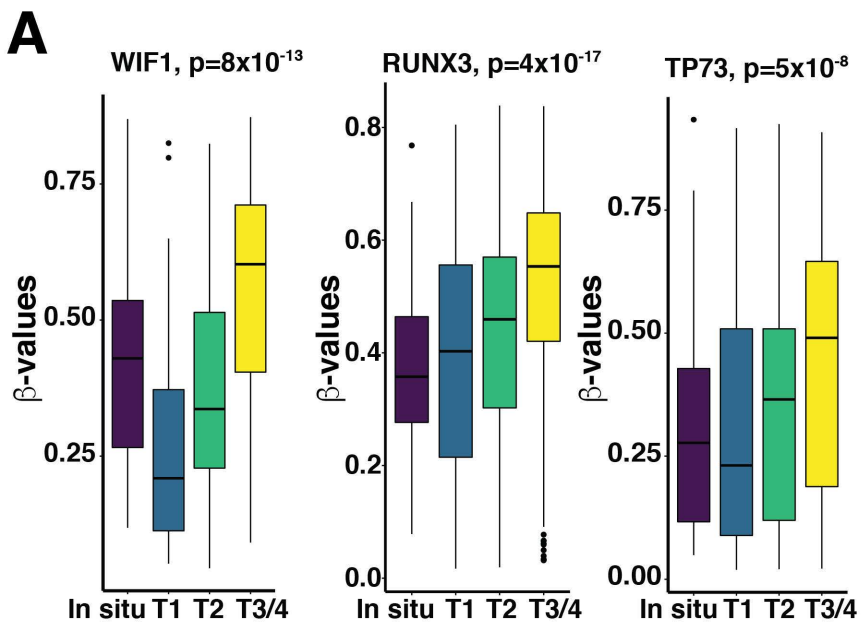
Suppl. Figure S1

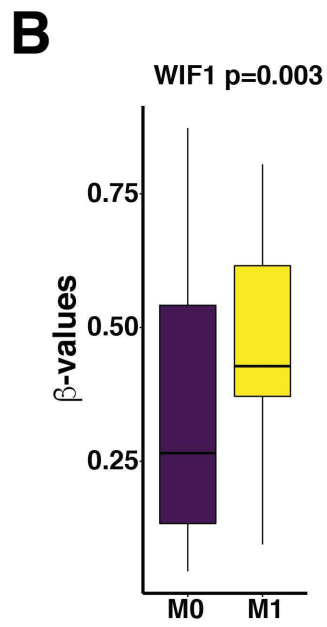


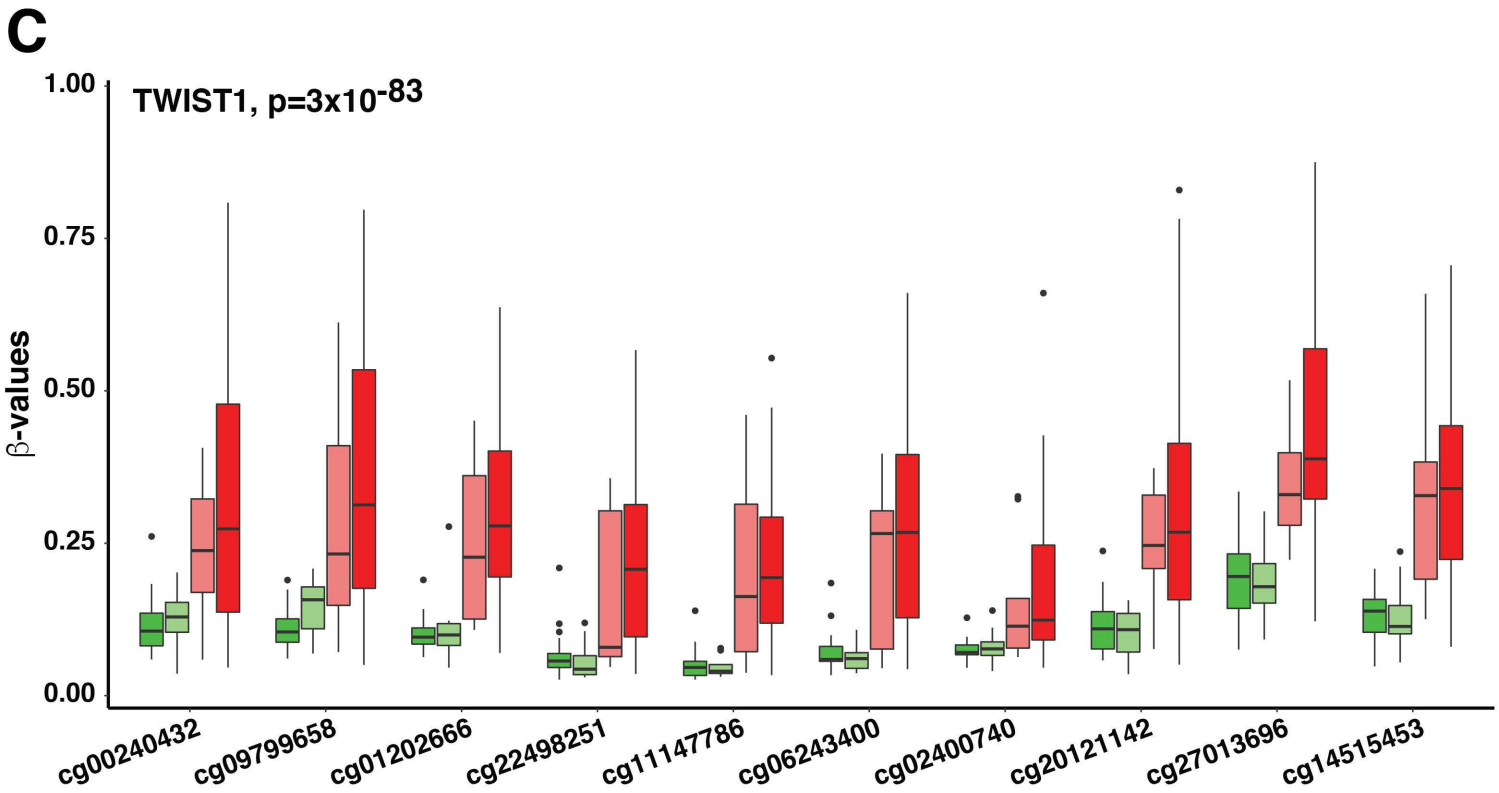
Suppl. Figure S2

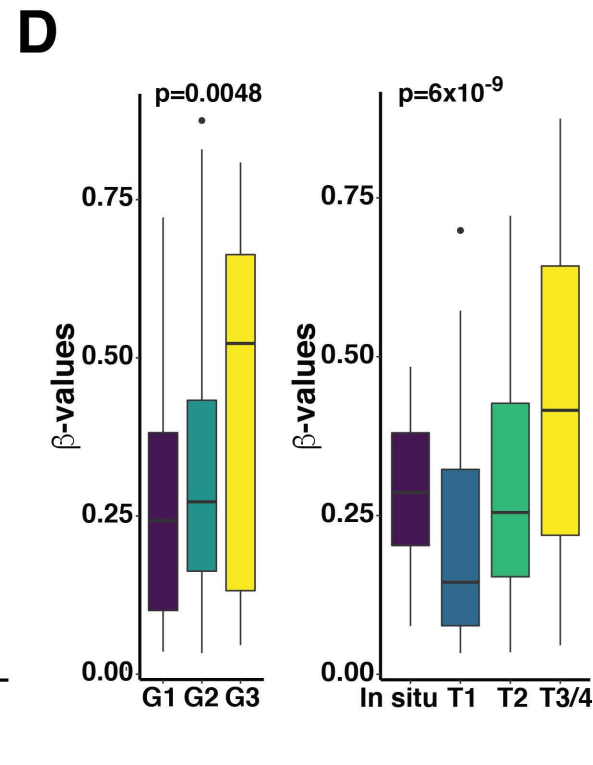


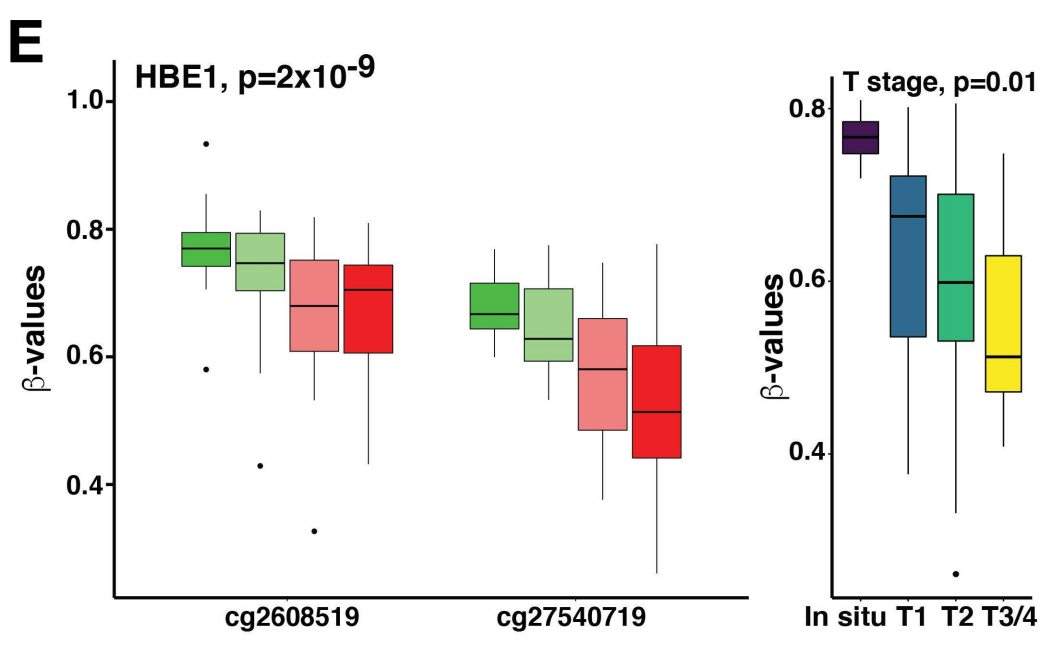
Suppl. Figure S3

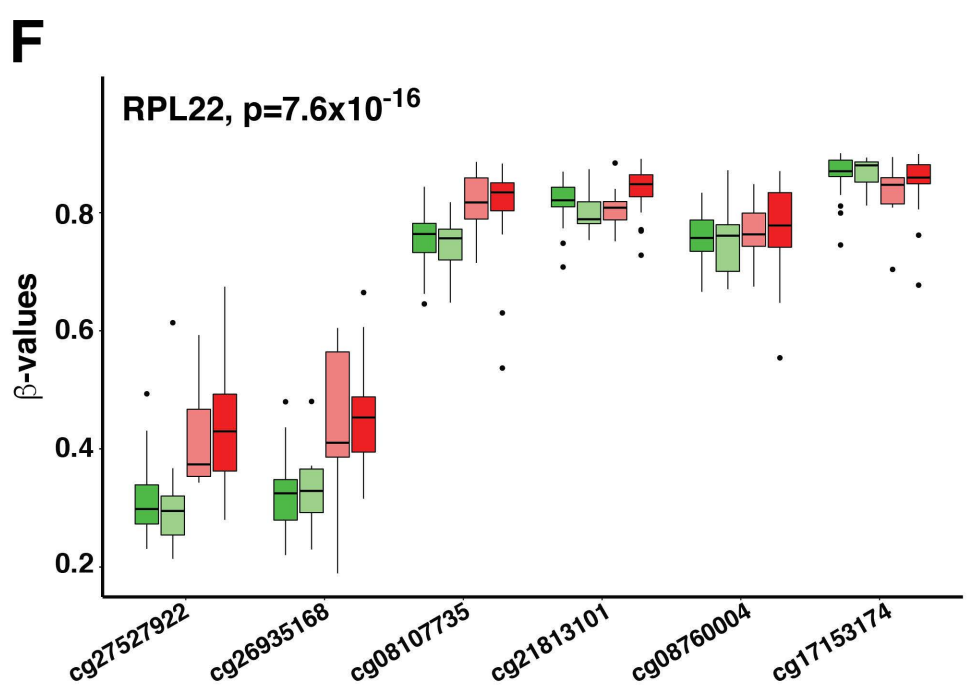




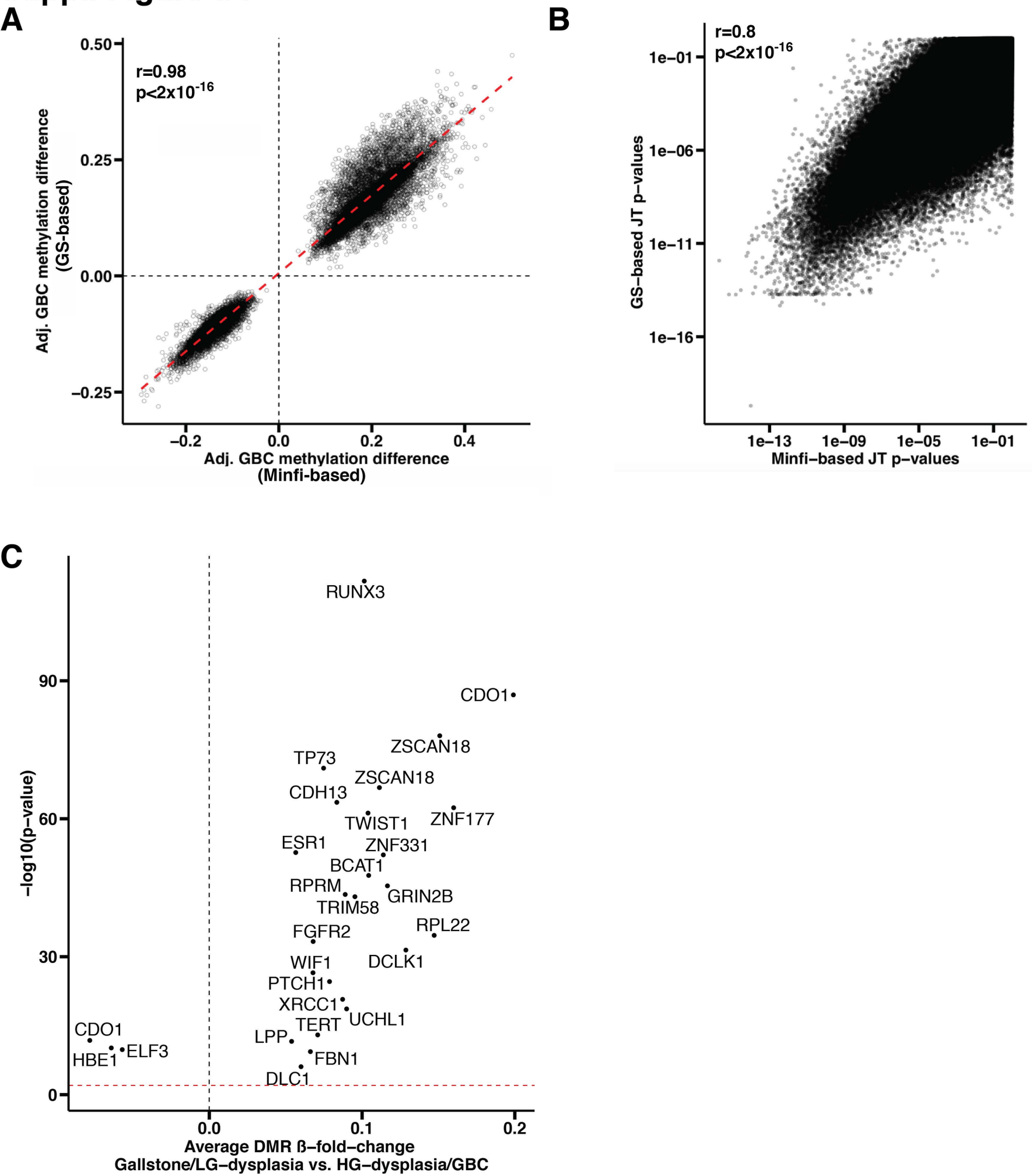




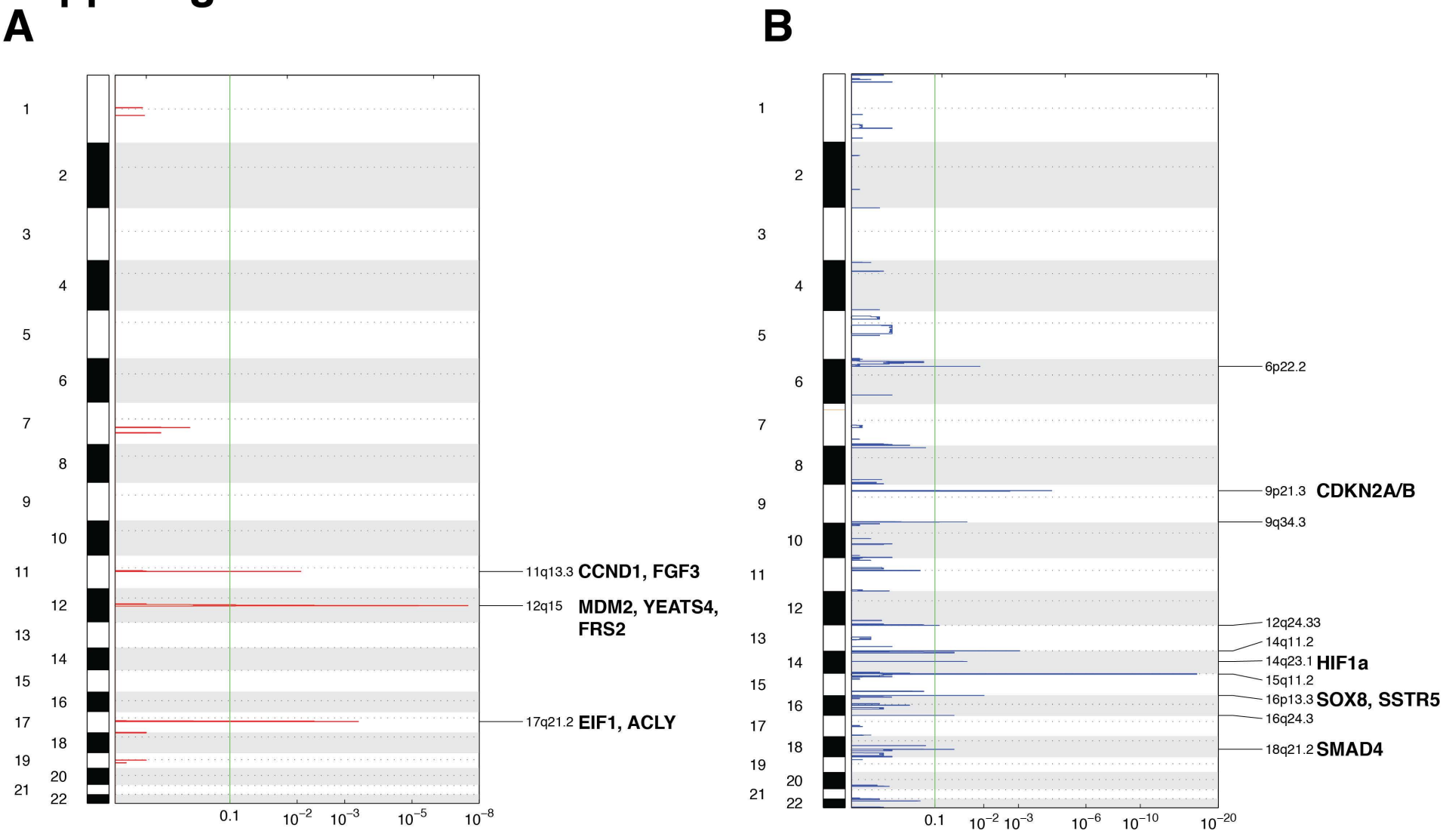




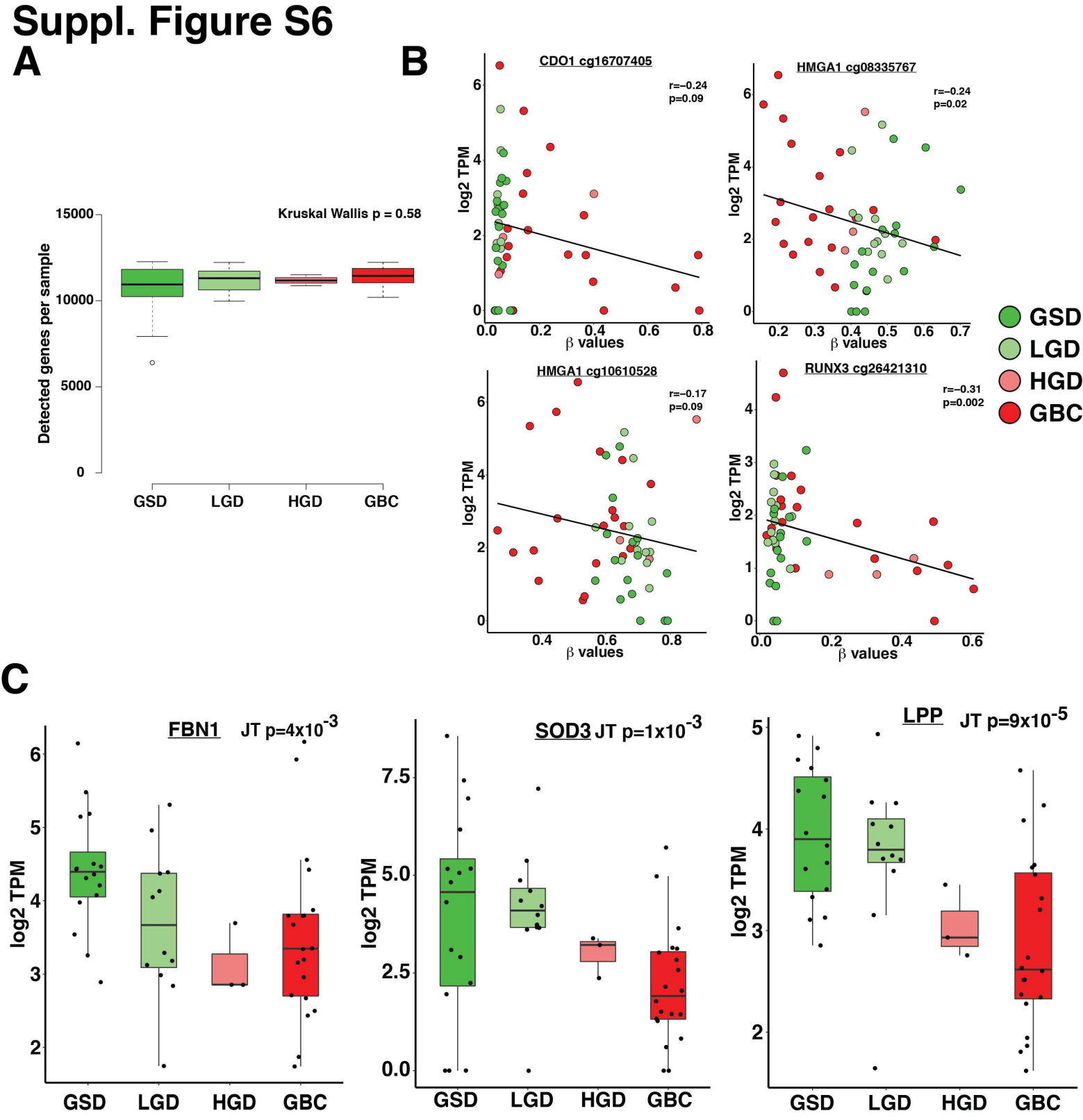
Suppl. Figure S4



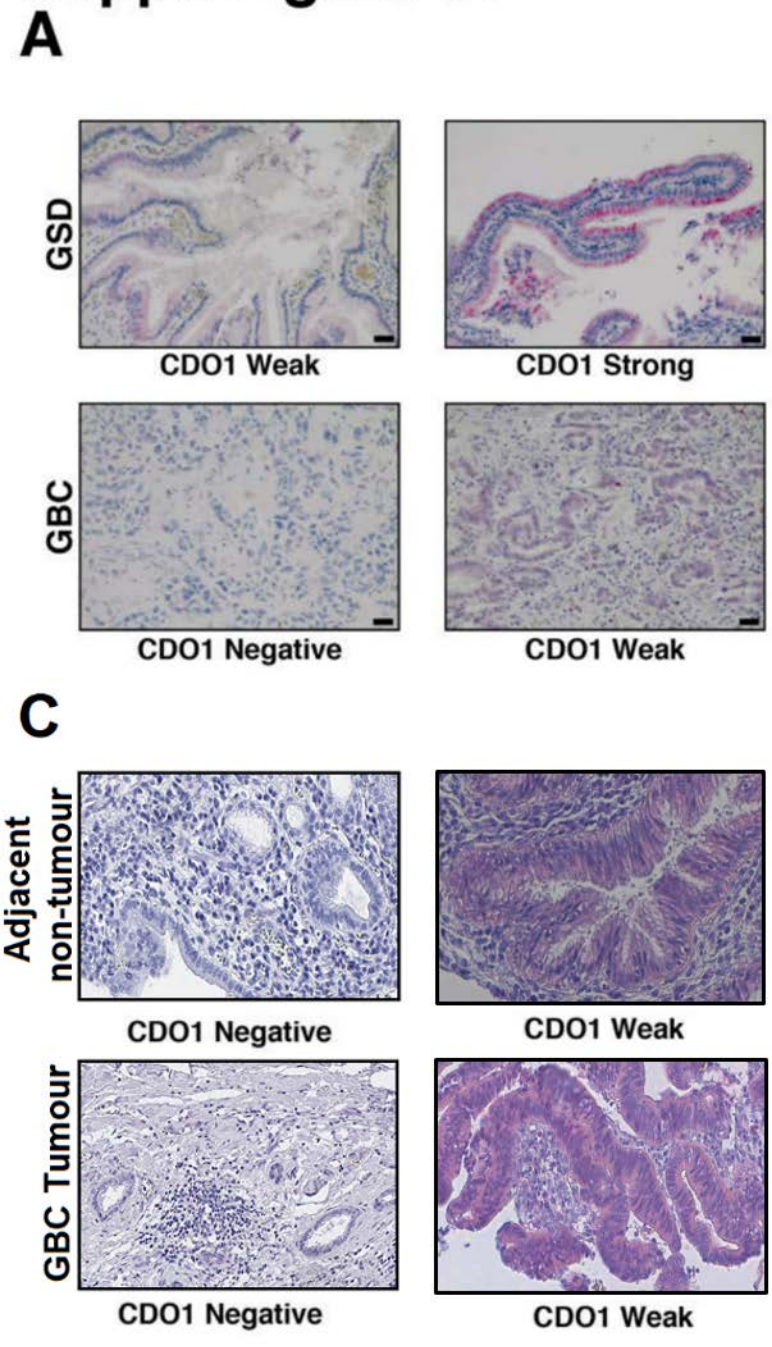
Suppl. Figure S5



Suppl. Figure S6



Suppl. Figure S7



Suppl. Table S1. DNA oligonucleotides used for DNA methylation analysis and qRT-PCR. UPL: LightCycler 480 Universal ProbeLibrary, Roche

Name	Forward primer sequence	Reverse primer sequence	Chromosomal location (hg19)
EpiTYPER primers for methylation analysis[1] covered cg probes			
RUNX3 cg26421310	GAAGTAGTAAGAGTTGGGGAAGTT	CAATACCACAACCCAAAACCC	chr1:25,256,961-25,257,156 chr1:25,257,058
CDO1 cg16707405	TTTAGATTTGTGGGGTTTATTTTTT	CCACCTTTCTTAAATAACTCTCC	chr5:115,152,346-115,152,459 chr5:115,152,414
TP73 promotor cg04493946 cg27340829	GTTTTGGTGGGTTTAATTATGGAGT	AAACAAACCTCAACTTACCC	chr1:3,607,072-3,607,179 chr1:3,607,098 chr1:3,607,143
TP73 distant promotor cg24678611 cg10038618 cg07382920	GGGTTTAGTTTTGGTAATAGAAAAGG	CCTAACCCTAAATTCTAAAAACT	chr1:3,567,593-3,567,766 chr1:3,567,732 chr1:3,567,719 chr1:3,567,647
HMGA1 cg08335767 cg10610528	GGTTTTTGATGTTTAGAAATAGGTGG	TACTAAAAACCCTTTAATCCCCTCA	chr6:34,206,459-3,420,6578 chr6:34,206,495 chr6:34,206540
LINE-1	TTTATATTTTGGTATGATTTTGTAG	TTTATCACCACCAAACCTACCCT	chrX:146,737,074-146,737,176
RT-qPCR primers (UPL hydrolysis probe)			
RUNX3 (#71)	TCAGCACCACAAGCCACTT	AATGGGTTCAGTTCCGAGGT	
CDO1 (#74)	TCTCTGTTGGGGTGAAGGAC	AAGCAGTGGGAGTTGGTATGA	
TP73 (#19)	GGAGGGACTTCAACGAAGG	TCATCCACATACTGCGAGAGA	
HMGA1 (#75)	GGAAAAGGACGGCACTGA	ACTTCGCTGGGCTCCTTC	
ACTB (#11)	ATTGGCAATGAGCGGTTC	GGATGCCACAGGACTCCAT	
GAPDH (#60)	GCCCAATACGACCAAATCC	AGCCACATCGCTCAGACAC	
HPRT1 (#73)	TGACCTTGATTTATTTTGCATACC	CGAGCAAGACGTTCAGTCCT	

[1]Primers were 5'-tagged with tags for EpiTYPER analysis
(forward primer: AGGAAGAGAG, reverse primer: CAGTAATACGACTCACTATAGGGAGAAGGCT).

***Supplementary Table S2*:** Clinical characteristics of the investigated Chilean cohort (total n=81 patients)

Gallstone disease (n=32 patients)		
	Median age	48.5 years
	Gender (f/m)	30/2
Low-grade dysplasia (n=13 patients)		
	Mean age	53 years
	Gender (f/m)	11/2
	With gallstones	13
High-grade dysplasia (n=9 patients)		
	Mean age	67.5 years (missing information for 1 patient)
	Gender (f/m)	7/2
	With gallstones	9
Gallbladder cancer (n=27 patients)		
	Mean age	62 years (missing information for 2 patients)
	Gender (f/m)	21/5 (missing information for 1 patient)
	With gallstones	25 (missing information for 2 patients)
T stage	In situ	2
	T1	6
	T2	10
	T3	4
	T4	2
	Not available	3
N stage	NX	6
	N0	15
	N1	2
	N2	1
	Not available	3
M stage	MX	9
	M0	11
	M1	4
	Not available	3
Grading	G1	7
	G2	11
	G3	2
	Not available	7

***Supplementary Table S3**:* Top significant CpG markers with differential methylation in Jonckheere-Terpstra (JT) test. p-values were Bonferroni-Holm adjusted, b-differences were calculated from linear models adjusting for patient age and gender.

No.	Methylation marker	Gene	adj. JT p-value	Mean β-value in gallstone samples	Mean β-difference LG-dyspl. vs gallstone	Mean β-difference HG-dyspl. vs gallstone	Mean β-difference GBC vs gallstone
1	cg08493776	PCDHB6	0	0.04 (0;0.14)	0.02 (-0.05;0.1)	0.19 (0.09;0.28)	0.26 (0.19;0.33)
2	cg21392341	TBX15	1.1e-10	0.03 (0;0.14)	0.01 (-0.07;0.1)	0.15 (0.04;0.26)	0.21 (0.13;0.28)
3	cg02164046	SST	3.7e-10	0 (0;0.09)	0.06 (-0.04;0.16)	0.26 (0.13;0.38)	0.36 (0.27;0.44)
4	cg24886257		8.7e-10	0.12 (0;0.26)	0.07 (-0.04;0.18)	0.23 (0.09;0.36)	0.35 (0.26;0.45)
5	cg11823511	BARHL2	1.6e-09	0 (0;0.11)	0.02 (-0.07;0.12)	0.24 (0.12;0.36)	0.29 (0.21;0.38)
6	cg00656990	WWOX	3.2e-09	0.89 (0.8;0.98)	-0.02 (-0.08;0.05)	-0.17 (-0.26;-0.08)	-0.21 (-0.27;-0.15)
7	cg24503966	NOL4	3.8e-09	0.04 (0;0.15)	0.03 (-0.05;0.1)	0.13 (0.03;0.23)	0.23 (0.17;0.3)
8	cg02950416	BCAN	4.3e-09	0 (0;0.1)	0.03 (-0.05;0.1)	0.14 (0.04;0.24)	0.28 (0.22;0.35)
9	cg26958783	SALL3	4.3e-09	0.1 (0.02;0.17)	0.03 (-0.03;0.08)	0.16 (0.09;0.24)	0.16 (0.11;0.21)
10	cg18359578	KCNMA1	5.5e-09	0.4 (0.35;0.45)	-0.04 (-0.08;0)	-0.11 (-0.17;-0.06)	-0.14 (-0.18;-0.1)
11	cg26296488	DRD5	6.2e-09	0.02 (0;0.18)	-0.01 (-0.12;0.11)	0.2 (0.05;0.34)	0.36 (0.26;0.46)
12	cg12665460	ZNF578	6.2e-09	0.13 (0.05;0.22)	0 (-0.06;0.07)	0.11 (0.03;0.2)	0.21 (0.16;0.27)
13	cg19274890	DPP6	6.2e-09	0 (0;0.09)	-0.01 (-0.11;0.08)	0.26 (0.13;0.38)	0.37 (0.29;0.46)
14	cg05928342	ZNF177	7.2e-09	0.04 (0;0.14)	0.01 (-0.07;0.08)	0.07 (-0.03;0.17)	0.23 (0.16;0.3)
15	cg02519751	ZIC1	7.7e-09	0.08 (0;0.22)	0.05 (-0.05;0.15)	0.27 (0.14;0.4)	0.35 (0.26;0.44)
16	cg15885148	CFAP61	9.4e-09	0.83 (0.73;0.92)	-0.05 (-0.12;0.02)	-0.13 (-0.22;-0.04)	-0.2 (-0.26;-0.14)
17	cg03254451	EN1	9.6e-09	0.04 (0;0.14)	0.02 (-0.05;0.09)	0.17 (0.08;0.27)	0.23 (0.16;0.29)
18	cg03653841	SFTA3	9.7e-09	0.03 (0;0.1)	0.01 (-0.05;0.07)	0.12 (0.04;0.19)	0.17 (0.12;0.22)
19	cg17857974	PCDHGA4;PCDHGA9;PCDHGA1;...	1.1e-08	0.14 (0.06;0.23)	0.05 (-0.02;0.11)	0.15 (0.06;0.23)	0.18 (0.13;0.24)
20	cg14457782	WNK4	1.1e-08	0.08 (0;0.17)	0 (-0.07;0.07)	0.05 (-0.04;0.14)	0.22 (0.16;0.28)
21	cg15582891	TAAR1	1.1e-08	0.87 (0.79;0.94)	-0.06 (-0.12;-0.01)	-0.06 (-0.13;0.01)	-0.15 (-0.2;-0.1)
22	cg20859731		1.2e-08	0.05 (0;0.16)	0.01 (-0.06;0.09)	0.14 (0.04;0.24)	0.26 (0.19;0.33)
23	cg12612118	IRF4	1.4e-08	0.08 (0;0.21)	0.02 (-0.08;0.12)	0.29 (0.16;0.42)	0.34 (0.25;0.43)
24	cg21183256		2.1e-08	0 (0;0.07)	0.04 (-0.05;0.13)	0.17 (0.06;0.29)	0.3 (0.22;0.38)

25	cg21277995	IRF4	2.3e-08	0.03 (0;0.16)	0.01 (-0.09;0.11)	0.16 (0.03;0.28)	0.28 (0.19;0.36)
26	cg11941811	TMEM108	2.4e-08	0.84 (0.74;0.94)	-0.05 (-0.12;0.02)	-0.07 (-0.17;0.02)	-0.23 (-0.29;-0.16)
27	cg20040743	NECAB1	2.7e-08	0.09 (0.02;0.15)	0 (-0.04;0.05)	0.08 (0.02;0.14)	0.14 (0.1;0.18)
28	cg01337836		2.9e-08	0.81 (0.74;0.89)	-0.03 (-0.09;0.02)	-0.06 (-0.13;0.01)	-0.17 (-0.21;-0.12)
29	cg03692651		3,00E-08	0 (0;0.15)	0.04 (-0.06;0.15)	0.26 (0.12;0.4)	0.4 (0.3;0.49)
30	cg11563680		3.6e-08	0.04 (0;0.14)	0.02 (-0.06;0.1)	0.17 (0.07;0.28)	0.31 (0.24;0.38)
31	cg13724788	NCAN	3.8e-08	0 (0;0.12)	0 (-0.09;0.09)	0.13 (0.02;0.25)	0.23 (0.15;0.31)
32	cg02623400	ELAVL4	3.9e-08	0 (0;0.05)	0.03 (-0.05;0.11)	0.24 (0.13;0.34)	0.28 (0.21;0.35)
33	cg24909548		3.9e-08	0.03 (0;0.18)	0 (-0.11;0.11)	0.27 (0.13;0.41)	0.37 (0.28;0.47)
34	cg19893751		4.6e-08	0.04 (0;0.16)	-0.01 (-0.09;0.08)	0.19 (0.08;0.3)	0.31 (0.24;0.38)
35	cg05380019		4.6e-08	0.08 (0;0.17)	0.02 (-0.05;0.09)	0.12 (0.03;0.21)	0.23 (0.17;0.3)
36	cg03181248	SP9	4.6e-08	0 (0;0.11)	0.04 (-0.04;0.12)	0.23 (0.13;0.34)	0.25 (0.17;0.32)
37	cg13677149	EVX1	4.6e-08	0 (0;0.13)	0.02 (-0.09;0.12)	0.19 (0.06;0.32)	0.37 (0.28;0.46)
38	cg22166999		4.8e-08	0.9 (0.83;0.98)	-0.02 (-0.08;0.03)	-0.1 (-0.18;-0.03)	-0.16 (-0.21;-0.11)
39	cg21433231		5,00E-08	0 (0;0.08)	-0.01 (-0.08;0.06)	0.11 (0.02;0.2)	0.22 (0.16;0.28)
40	cg05638493	SLC25A2	5.5e-08	0.29 (0.21;0.37)	0.03 (-0.03;0.08)	0.18 (0.1;0.25)	0.21 (0.16;0.26)
41	cg14559259	CCDC140	5.5e-08	0.09 (0;0.21)	0.02 (-0.07;0.11)	0.09 (-0.03;0.21)	0.3 (0.22;0.38)
42	cg16508480	ADAMTS16	6,00E-08	0 (0;0.1)	0 (-0.08;0.07)	0.2 (0.1;0.29)	0.26 (0.2;0.33)
43	cg18780257	FGF12	6.6e-08	0 (0;0.11)	0.01 (-0.08;0.09)	0.17 (0.06;0.28)	0.33 (0.26;0.41)
44	cg16073378		6.6e-08	0.01 (0;0.09)	0.01 (-0.05;0.07)	0.11 (0.02;0.19)	0.24 (0.18;0.29)
45	cg12788878	SLC12A8	7,00E-08	0.01 (0;0.12)	0.06 (-0.02;0.14)	0.24 (0.14;0.34)	0.27 (0.2;0.34)
46	cg12459904	PCDHGA4;PCDHGA6;PCDHGA1;...	7.1e-08	0.18 (0.09;0.28)	0.02 (-0.05;0.09)	0.19 (0.1;0.28)	0.25 (0.19;0.32)
47	cg11283429		7.1e-08	0.06 (0;0.17)	0.05 (-0.03;0.13)	0.18 (0.07;0.28)	0.27 (0.2;0.35)
48	cg07873041		7.1e-08	0.03 (0;0.16)	0 (-0.1;0.1)	0.24 (0.11;0.37)	0.4 (0.31;0.48)
49	cg03498096	INSM1	7.1e-08	0.08 (0;0.17)	0.02 (-0.05;0.09)	0.14 (0.04;0.23)	0.22 (0.16;0.29)
50	cg27252696	SIM1	7.5e-08	0 (0;0.08)	0.01 (-0.07;0.09)	0.11 (0;0.22)	0.27 (0.2;0.35)
51	cg26811372	PCDHGA4;PCDHGA6;PCDHGA1;...	7.7e-08	0.18 (0.08;0.28)	0.05 (-0.03;0.12)	0.26 (0.16;0.35)	0.26 (0.2;0.33)

52	cg15260349	MTNR1B	7.7e-08	0.15 (0.06;0.24)	0.05 (-0.01;0.12)	0.12 (0.03;0.2)	0.22 (0.16;0.28)
53	cg00824767		8.4e-08	0.06 (0;0.17)	0.02 (-0.06;0.1)	0.19 (0.08;0.29)	0.28 (0.2;0.35)
54	cg15672437	F7	8.5e-08	0 (0;0.09)	0 (-0.09;0.08)	0.14 (0.03;0.25)	0.29 (0.21;0.37)
55	cg11494930		8.6e-08	0.85 (0.79;0.91)	-0.03 (-0.07;0.01)	-0.07 (-0.13;-0.01)	-0.13 (-0.17;-0.09)
56	cg14189141		9.1e-08	0 (0;0.08)	0 (-0.08;0.08)	0.11 (0.01;0.21)	0.28 (0.21;0.35)
57	cg15138339	COASY	9.3e-08	1 (0.9;1)	-0.06 (-0.13;0.02)	-0.12 (-0.22;-0.02)	-0.24 (-0.31;-0.18)
58	cg05845376	SLC25A2	1,00E-07	0.17 (0.03;0.3)	0.03 (-0.07;0.14)	0.2 (0.07;0.33)	0.3 (0.21;0.4)
59	cg18490522	PRAC2;PRAC1	1.1e-07	0.1 (0;0.2)	0.02 (-0.05;0.1)	0.14 (0.04;0.23)	0.31 (0.24;0.38)
60	cg17759274		1.1e-07	0 (0;0.1)	0.12 (0.02;0.22)	0.26 (0.13;0.39)	0.34 (0.24;0.43)
61	cg18253016	NR2E1	1.1e-07	0.08 (0;0.17)	0 (-0.07;0.07)	0.17 (0.08;0.26)	0.24 (0.18;0.3)
62	cg16557178	HOXB13	1.1e-07	0.06 (0;0.16)	-0.01 (-0.09;0.07)	0.2 (0.1;0.3)	0.28 (0.21;0.35)
63	cg05785155		1.1e-07	0.04 (0;0.17)	0.03 (-0.06;0.13)	0.25 (0.13;0.38)	0.3 (0.21;0.39)
64	cg19509330	TFAP2D	1.1e-07	0.08 (0.01;0.16)	0.01 (-0.05;0.07)	0.12 (0.05;0.19)	0.18 (0.13;0.23)
65	cg13525197	ZSCAN23	1.1e-07	0.07 (0;0.15)	0.02 (-0.04;0.08)	0.13 (0.05;0.2)	0.2 (0.14;0.25)
66	cg23141355		1.1e-07	0.06 (0;0.2)	0.01 (-0.1;0.11)	0.15 (0.02;0.29)	0.28 (0.19;0.38)
67	cg26980244	NEFM	1.2e-07	0.05 (0;0.18)	0.03 (-0.06;0.13)	0.28 (0.16;0.4)	0.39 (0.3;0.47)
68	cg19671120	CNGA3	1.2e-07	0.13 (0.06;0.2)	0.02 (-0.03;0.07)	0.11 (0.05;0.18)	0.17 (0.13;0.22)
69	cg07920503	FAM123A	1.2e-07	0.07 (0;0.22)	0.04 (-0.07;0.15)	0.28 (0.14;0.42)	0.41 (0.31;0.51)
70	cg21684012	SIM1	1.2e-07	0 (0;0.07)	0.01 (-0.05;0.07)	0.13 (0.04;0.21)	0.25 (0.19;0.31)
71	cg00851770	CIDEA	1.2e-07	0.04 (0;0.16)	0.01 (-0.08;0.1)	0.18 (0.07;0.3)	0.29 (0.21;0.37)
72	cg21733531		1.3e-07	0.11 (0;0.23)	0.03 (-0.06;0.12)	0.26 (0.15;0.38)	0.29 (0.21;0.37)
73	cg04637478	PCDHGA4;PCDHGA1;PCDHGA6;...	1.3e-07	0.08 (0;0.18)	0.03 (-0.04;0.11)	0.18 (0.08;0.28)	0.25 (0.18;0.32)
74	cg14563954		1.3e-07	0.09 (0.01;0.17)	0 (-0.06;0.06)	0.07 (0;0.15)	0.2 (0.14;0.25)
75	cg02387803	ZIC4	1.3e-07	0.05 (0;0.18)	0.01 (-0.08;0.11)	0.11 (-0.02;0.23)	0.3 (0.21;0.39)
76	cg21790369	CACNA1E	1.3e-07	0.01 (0;0.14)	0.06 (-0.04;0.16)	0.19 (0.06;0.32)	0.23 (0.14;0.32)
77	cg00232292	UNC5D	1.3e-07	0.82 (0.74;0.89)	-0.01 (-0.06;0.05)	-0.1 (-0.17;-0.02)	-0.18 (-0.23;-0.13)
78	cg02328239	GDNF	1.3e-07	0.1 (0.02;0.17)	0.01 (-0.04;0.07)	0.16 (0.09;0.23)	0.18 (0.13;0.22)

79	cg25209842	FGF8	1.4e-07	0.01 (0;0.12)	0.03 (-0.05;0.12)	0.16 (0.04;0.27)	0.25 (0.18;0.33)
80	cg17018096	ADAMTS10	1.4e-07	0.1 (0.04;0.16)	0.01 (-0.04;0.05)	0.1 (0.04;0.16)	0.15 (0.11;0.19)
81	cg22601348		1.4e-07	0.04 (0;0.16)	0 (-0.08;0.09)	0.24 (0.13;0.35)	0.27 (0.2;0.35)
82	cg15347189	SST	1.4e-07	0.03 (0;0.13)	-0.01 (-0.09;0.07)	0.1 (0;0.2)	0.23 (0.16;0.3)
83	cg11885396	SLC6A17	1.4e-07	0 (0;0.07)	0.02 (-0.04;0.08)	0.15 (0.07;0.22)	0.2 (0.14;0.25)
84	cg12878812	SRRM4	1.5e-07	0.03 (0;0.12)	0.04 (-0.03;0.1)	0.17 (0.09;0.25)	0.26 (0.2;0.31)
85	cg18688293	NMBR	1.5e-07	0.08 (0;0.16)	0.01 (-0.05;0.07)	0.09 (0.01;0.17)	0.2 (0.15;0.25)
86	cg15191830	CCAT2;CASC8	1.5e-07	0.91 (0.8;1)	-0.04 (-0.12;0.04)	-0.12 (-0.23;-0.01)	-0.23 (-0.3;-0.16)
87	cg00495860		1.6e-07	0.04 (0;0.16)	0.01 (-0.08;0.09)	0.11 (-0.01;0.22)	0.28 (0.2;0.36)
88	cg18235734		1.6e-07	0.02 (0;0.13)	0 (-0.08;0.08)	0.22 (0.11;0.33)	0.36 (0.28;0.44)
89	cg21215550		1.6e-07	0.03 (0;0.18)	0.02 (-0.09;0.13)	0.08 (-0.07;0.22)	0.29 (0.19;0.39)
90	cg13160071		1.7e-07	0.9 (0.82;0.99)	-0.02 (-0.08;0.05)	-0.08 (-0.16;0)	-0.22 (-0.28;-0.16)
91	cg08942049	FGFR1	1.7e-07	0.85 (0.79;0.9)	-0.01 (-0.05;0.02)	-0.1 (-0.15;-0.05)	-0.13 (-0.16;-0.09)
92	cg01524853	HOXC4	1.8e-07	0.07 (0;0.17)	0.01 (-0.06;0.09)	0.17 (0.07;0.26)	0.23 (0.16;0.29)
93	cg10184983	SRCIN1	1.8e-07	0 (0;0.09)	0.03 (-0.06;0.12)	0.11 (-0.01;0.23)	0.26 (0.18;0.34)
94	cg07962143		1.8e-07	0.11 (0;0.24)	0.11 (0;0.21)	0.23 (0.1;0.36)	0.33 (0.24;0.42)
95	cg23421023	ZMAT4	1.9e-07	0.03 (0;0.15)	0.03 (-0.06;0.12)	0.16 (0.05;0.28)	0.23 (0.15;0.31)
96	cg13481132	KCNT1	1.9e-07	0.04 (0;0.13)	0.01 (-0.06;0.08)	0.03 (-0.06;0.12)	0.15 (0.09;0.21)
97	cg23612932		2,00E-07	0.41 (0.33;0.48)	0.07 (0.02;0.13)	0.13 (0.06;0.2)	0.17 (0.12;0.22)
98	cg24454829	SLC17A6	2,00E-07	0.09 (0.02;0.16)	0.05 (-0.01;0.1)	0.25 (0.18;0.32)	0.21 (0.16;0.26)
99	cg27331241	PRKAR1B	2,00E-07	0.02 (0;0.14)	0 (-0.08;0.09)	0.06 (-0.05;0.18)	0.23 (0.15;0.31)
100	cg27029821	PRDM14	2.1e-07	0.11 (0.01;0.2)	0.01 (-0.06;0.08)	0.16 (0.07;0.25)	0.25 (0.18;0.31)

***Supplementary Table S4**:* Validation of methylation measurements in patient samples and the assessment of the effect of methylation on gene expression by 5-aza-dC treatment of the GBC cell lines G-415 and OZ.

		Validation of differential methylation				**Assessment of the effect of methylation on gene expression by treatment of GBC cell lines with 5-aza-dC or DMSO as a negative control**							
		EPIC Array		**MassARRAY**		**GBC cell line G-415**				**GBC cell line OZ**			
Gene	**CpG**	**β diff. GBC-GSD**	**p-value[1]**	**β diff. GBC-GSD**	**p-value[1]**	**β diff. Aza-DMSO**	**p-value[2]**	**expr. fold change**	**p-value[2]**	**β diff. Aza-DMSO**	**p-value[2]**	**expr. fold change**	**p-value[2]**
CDO1	cg16707405	0.27	0.004	0.14	0.002	-0.17	0.04	32.3	4e-3	-0.20	0.04	29.5	0.002
RUNX3	cg26421310	0.11	0.40	0.07	0.20	-0.31	0.01	33.7	7e-7	-0.17	0.01	14.7	0.003
TP73	cg04493946	-0.08	0.02	-0.08	0.60	-0.10	0.30			-0.13	0.02		
TP73	cg07382920	0.37	4e-04	0.20	3e-04	-0.22	0.02			-0.08	0.48		
TP73	cg10038618	0.24	0.002	0.26	0.007	-0.22	0.01	4.6	3e-3	-0.10	0.19	1.0	0.78
TP73	cg24678611	0.22	0.003	0.14	0.04	-0.13	0.37			-0.23	0.11		
TP73	cg27340829	-0.08	0.20	-0.10	0.30	-0.07	0.04			-0.11	0.03		
HMGA1	cg08335767	-0.23	3e-04	-0.15	0.006	0.00	0.65	1.1	0.80	0.00	1.00	1.6	0.14
HMGA1	cg10610528	-0.09	0.07	-0.14	0.02	0.00	0.74			0.00	1.00		

β diff. GBC-GSD: Mean methylation difference between gallbladder cancer and gallstone disease samples; expr. fold change: fold change in expression

[1] Two-sample U-test p-value

[2] Two-sample t-test (single CpG) or ANOVA (multiple CpGs) p-value

Suppl. Table S5: Immunohistochemistry analysis of CDO1 expression in GSD and GBC patients.

	Negative	Weak	Strong
GSD	**0**	**1**	**4**
GBC	**4**	**2**	**0**
Fisher test p=0.013			

Suppl. Table S6: Jonckheere-Terpstra test p-values of RNA-seq expression values across disease stages for genes detected as being significantly methylated along the progression GSD->LGD->HGD->GBC if sRNA-seq data for these genes passed quality control thresholds (see **Table 1**).

Gene	JT p-value
WWOX	2e-4
KCNMA1	3e-7
ZNF578	0.55
DPP6	0,067
ZNF177	0,46
EN1	0,007
PCDHGA4	4e-6
RPL22	0.7