UPTEC X 07 015
FEB 2007

ISSN 1401-2138

LINDA SOOMAN

Epigenetic studies of SHP1, SHP2, SOCS1, SOCS3 and STAT1 in esophageal and lung cancer

Master's degree project

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 07 015	**Date of issue 2007-02**
Author **Linda Sooman**	
Title (English) **Epigenetic studies of SHP1, SHP2, SOCS1, SOCS3 and STAT1 in esophageal and lung cancer**	
Title (Swedish)	
Abstract Aberrant DNA methylation is a hallmark of cancer. The tumor suppressor genes SHP1, SHP2, SOCS1 and SOCS3 and the transcription factor STAT1 have shown aberrant methylation in several tumors. In this report we have investigated their methylation and protein expression status in human cancer cell lines from esophageal (EC) and lung cancer (LC). We show that SHP1 SOCS1 and SOCS3 were highly methylated in several EC and LC cell lines, whereas SHP2 and STAT1 were not methylated in any EC or LC cell line. Furthermore, the methylation of the promoter region 1 of SHP1 was associated with protein expression reduction in both the EC and LC cell lines.	
Keywords EC, LC, CpG methylation, protein expression, SHP1, SHP2, SOCS1, SOCS3, STAT1.	
Supervisors **Monica Pettersson (Biotage AB, Uppsala), Simon Ekman and Michael Bergqvist (Dep. of Oncology, Uppsala University Hospital)**	
Scientific reviewer **Joachim Gullbo (Clinical Pharmacology, Uppsala University Hospital)**	
Project name	Sponsors
Language **English**	Security **1 year**
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **42**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Introduction

Esophageal cancer

Esophageal cancer (EC) is ranked seventh of the cancers causing most deaths in the world [1], responsible for over 400 000 new cases per year [2]. The incidence in most countries is however fairly low, about 5 cases per 100 000 persons per year, but in some particular areas, such as northern Iran, southern and eastern Africa and south central Asia the incidence can be as high as 30-800 cases per 100 000 persons per year [1,2]. EC is an aggressive cancer with early metastasis and rapid growth. It is rarely curable and the overall 5-year survival rate for patients receiving treatment ranges from 5% to 30% [3]. It can be categorized into two main histological types; adenocarcinoma (EAC) and squamous cell carcinoma (ESCC). The most common type is ESCC, which constitutes 75-90% of all ECs [4]. Hence, this study is focused on ESCC.

Lung cancer

Lung cancer (LC) is the cancer causing most deaths in the world [5], responsible for 1.2 million new cases per year [6]. The incidence is highest in developed countries. Like EC, LC is also an aggressive, fast growing cancer which commonly metastasizes. LC is a curable disease when discovered at an early stage, but due to commonly late diagnoses the 5-year survival rate is among the lowest of all cancers at 10-15% [7]. It can be categorized into two main histological types; non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). The most common type is NSCLC, which constitutes 80% of all LCs [7]. This study includes both types of LCs.

DNA methylation

DNA methylation is a naturally occurring event that, in higher eukaryotes, acts as a regulation of gene expression. It involves addition of a methyl group to the fifth carbon of the cytosine pyrimidine ring. In mammals this occurs most commonly on 5′-CpG-3′ dinucleotides (CpG sites). DNA methylation can impact the transcription of genes in two ways. The methyl group may physically prevent the binding of transcription factors to the gene and thus block transcription. Also, methylated DNA attracts binding of histone acetylases and other chromatin remodeling proteins that forms compact and transcriptionally inactive chromatin [8]. About 80% of all CpG sites in mammalian genomes are methylated. The majority of the unmethylated 20% are grouped in clusters called CpG islands that are situated in promoters or in the first exons of genes [9].

Aberrant DNA methylation, including hypermethylation of tumor suppressor genes and genome wide hypomethylation, is a hallmark of cancer and can be found in almost all cancer types [10]. Information about which methylation events are disease specific and also have effect on gene expression has a great potential in diagnostics and drug development. Genetic abnormalities of proto-oncogenes and tumor suppressor genes have been demonstrated to be changes that are frequently involved in esophageal cancer pathogenesis. However, hypermethylation of CpG islands is coming more and more into focus in carcinogenesis of the esophagus [11].

Analyzed genes

In this study we have chosen to investigate the effect of aberrant CpG methylation patterns on gene expression, of the tumor suppressor genes SHP1, SHP2, SOCS1 and SOCS3 and the transcription factor STAT1. These genes, with exception of SHP2, have previously shown hypermethylation-associated tumor occurrence in several types of cancers, such as chronic myeloid leukemia [12], human hepatocellular carcinoma [13, 14], T cell lymphoma [15], non-muscle invasive bladder carcinoma [16], myelodysplastic syndromes and acute myeloid leukemia [17]. SHP1 has also shown hypomethylation,

compared to normal epithelial cells, associated with occurrence of psoriasis [18]. There are no previously reported methylation studies on SHP2, but mutations in this gene have been implicated with several human diseases. One example is the Noonan syndrome, which is a human developmental disorder characterized by proportionate short stature, facial dysmorphia, increased risk of leukemia, and congenital heart defects in 50% of cases [19].

Normal activity of these genes is required for a functional regulation of cell growth. SHP1 is a negative regulator of cell signaling expressed in haematopoietic and epithelial cells [20]. SHP2 is a positive regulator of cell proliferation expressed in most cell types [20], but has also been shown to function as a negative effector in interferon-induced growth-inhibitory and apoptotic pathways [21]. SOCS1 and SOCS3 are negative regulators of cytokine signaling expressed in various epithelial tissues [22]. STAT1 negatively regulates cell proliferation and angiogenesis and thereby inhibits tumor formation. Consistent with its tumor suppressive properties, STAT1 and its downstream targets have been shown to be reduced in a variety of human tumors and STAT1 deficient mice are highly susceptible to tumor formation [23].

Methods

To study methylation patterns, bisulfite treatment of total DNA followed by PCR amplification and Pyrosequencing® analysis was employed. The gene regions to be analyzed were determined according to the regions that previously had shown aberrant methylation patterns in other types of cancers [12-17] and psoriasis [18]. In the bisulfite treatment non-methylated cytosines are deaminated to uracil, which are amplified as thymidine in the PCR reaction. Methylated cytosines remain unconverted. In the Pyrosequencing reaction the proportion of C's and T's in each CpG site is determined, and thus the percent of methylated DNA in each CpG site. Pyrosequencing uses a sequencing-by-synthesis principle where an enzymatic cascade releases light proportional to the amount of incorporated deoxynucleotides (dNTP's) in each position in the sequence. The measured light signals are displayed as peaks in a Pyrogram®, with the peak height proportional to the light intensity. It gives a display of the sequence and the percentage of methylation in each CpG site. This is a fast and simple method for analysis of CpG methylation in multiple sites in a single assay.

Cell lines used in this study were ten ESCC cell lines, four NSCLC cell lines, five SCLC cell lines and one uncategorized LC cell line. The NSCLC cell lines represent one squamous cell carcinoma (U-1752), two adenocarcinomas (NCI H-23 and NCI H-611) and two large cell carcinomas (U-1810 and NCI H-157). In the protein expression studies only three SCLC cell lines were used due to poor growth of the cell line 2050. Primary human foreskin fibroblasts were used as a control cell line.

Results and conclusions

We showed that SHP1, SOCS1 and SOCS3 were highly methylated in several EC and LC cell lines, whereas SHP2 and STAT1 were not methylated in any EC or LC cell line. Moreover, the methylation of the promoter region 1 of SHP1 was associated with protein expression reduction in both the EC and LC cell lines.

Materials and methods

Cell culturing and counting

The ESCC cell lines KYSE30, KYSE70, KYSE140, KYSE150, KYSE180, KYSE270, KYSE410, KYSE450, KYSE510 and KYSE520, the NSCLC cell lines U-1752, U-1810, NCI H-23, NCI H-157, NCI H-611, the SCLC cell lines 1906-E, 1906-L, U-2020 and U-2050, the uncategorized LC cell line YA (from Caucasian male, 60 years) and the primary human foreskin fibroblast cell line AG60 were cultivated in RPMI-1640 Medium (Sigma), supplemented with 10% FBS (Sigma), L-Glutamine (2 mM, Sigma) and Penicillin-Streptomycin (0.05 U Penicillin and 50 µg Streptomycin/ml, Sigma), in 37°C, 5% pCO_2 and 95% pH_2O. Adhesion growing cells were split when 100% confluent and suspension growing cells when the medium turned yellow. To collect the adherent cell lines cells were incubated in 6 µl/cm^2 5xTrypsin-EDTA Solution (25 mg porcine trypsin and 10 mg EDTA/ml, Sigma) in 37°C until detachment from the surface. Suspension growing cells were collected by centrifugation for 10 min at 1200 rpm in room temperature (RT).

Cells were counted using a Bürker chamber, to obtain approx. 5 million cells of each cell line for methylation studies and 10-15 million cells/lysate of each suspension growing cell line for the protein expression studies. Cell suspension and 0.4% Trypan Blue Solution (Sigma Chemical Co.) were mixed in a ratio of 1:1 in a Bürker chamber with the dimensions 0.0025 mm^2 x 0.100 mm, cells were counted and cell concentrations were calculated.

DNA purification and quantification

Most cell lines used for methylation studies were obtained as cell pellets from Dr Simon Ekman (Department of Oncology, Uppsala University Hospital, 751 85 Uppsala, Sweden). Total DNA samples were purified line using the DNeasy Blood and Tissue kit from Qiagen (07/2006). DNA concentrations were measured using the Quant-iT™ PicoGreen® dsDNA Kit from Molecular Probes™, Invitrogen. The interference of RNA on the fluorescence signal was considered negligible.

In vitro methylation

DNA samples from normal blood donors (average conc. 40 ng/µl), obtained from Monica Pettersson (Biotage AB, Kungsgatan 76, 753 18 Uppsala, Sweden), were used for *in vitro* methylation. For one 20 µl *in vitro* methylation reaction 1 µg DNA was mixed with 1X NEBuffer 2 (New England Biolabs® Inc.), S-adenodylmethionine (SAM, 160 µM, New England Biolabs® Inc.) and SssI CpG Methylase (100 U/ml, New England Biolabs® Inc.) and incubated for 1 h in 37°C. Another addition of SAM (160µM) and SssI Methylase (100 U/ml) was done to the reaction mixture followed by an additional 1 h incubation in 37°C.

Bisulfite treatment and PCR

The DNA samples were precipitated before bisulfite treatment to obtain a concentration of 25 ng/µl. Sodium acetate (0.3 M) and ethanol (3 volumes) was added to the samples, which were incubated in -80°C for 20-30 min or -20°C overnight and thereafter centrifuged for 20 min at 14 000 rpm. The supernatants were removed and the DNA was dried and redissolved in dH_2O. The DNA samples were bisulfite treated using the EZ DNA Methylation-Gold Kit™ from Zymo Research. The DNA was eluted in 10 µl elution buffer and thereafter diluted 1:2 in dH_2O.

The bisulfite treated DNA was PCR amplified using a PCR Kit from Qiagen. For each 25 µl reaction a PCR mix containing the following components was used: 1 x PCR buffer (containing 1.5 mM $MgCl_2$), dNTP's (200 µl of each), forward primer (0.2 µM), 5'-biotinylated reverse primer (0.2 µM), HotStarTaq DNA polymerase (0.8 U), template DNA

(1 µl) and H_2O. Additional $MgCl_2$ was added to the PCR mix for some primer pairs (table 3), to give a final concentration of 3mM.

DNA gel electrophoresis

5 µl of a 100 bp DNA ladder and 5 µl of each PCR-product mixed with 2 µl 6XLoading Dye Solution (#R0611, MBI Fermentas) were run on a 2% agarose gel in 1XTBE buffer at 140 V for 20 min. The gel was incubated in an ethidium bromide bath (0.35 mg/ml) for 10 min followed by rinsing in a water bath. The DNA was detected using UV light.

Pyrosequencing®

5-10 µl PCR product was immobilized to 2 µl Streptavidin Sepharose™ HP beads (GE Healthcare) in 1XBinding buffer (Biotage), on a rocking board at 1400 rpm in 25°C for 10 min. The biotinylated DNA strands were isolated using a vacuum prep tool (VPT). The DNA immobilized to Sepharose was caught on a filter, transferred to 70% ethanol for 5 s, Denaturation solution (Biotage) for 5 s and Wash buffer (Biotage) for 10 s. The isolated DNA was transferred to a 96-well Pyrosequencing plate containing sequence primer (0.33 µM) in 12 µl 1XAnneal Buffer (Biotage) per well. The sequence primers annealed to the DNA at 80°C for 2 min. Ready to use enzyme mixture, substrate mixture and dNTP's (Biotage) were added to the nucleotide dispensing tips (NTD's). The PSQ HSA System was used with the Pyro Q-CpG 1.0.9 Software.

Western blot

Harvest

Adhesion growing cells were harvested when approx. 95% confluent and suspension growing cells when having reached a number of 10-15 million cells. Before lysis cells were washed twice with PBS. Adhesion growing cells were lysed in 13 µl/cm^2 RIPA lysis buffer (20 mM Tris, pH 7.4, 1% Triton X-100, 0.5% DOC, 0.1% SDS, 5mM EDTA and 150 mM NaCl) and suspension growing cells in 1 ml RIPA buffer/10-15 million cells for 20-40 min on ice. Lysates were collected and cell debris was removed by centrifugation for 10 min at 13000 rpm in 4°C. Supernatants were transferred to new tubes and stored in -20°C until use.

Relative protein concentration measurement

Protein aggregates that might have formed during freezing and thawing of the cell lysates were moved to the bottom of the tubes by centrifugation for 10 min at 13000 rpm in 4°C. A protein concentration measurement reagent was mixed consisting of BCA™ Protein Assay Reagent B (Pierce) diluted 1:50 in BCA™ Protein Assay Reagent A (Pierce). Of each cell lysate 20 µl was mixed with 1 ml of the reagent mixture and incubated for 30 min in 37°C. The absorbance was measured of each sample with WPA Biowave S2100 Diode Array Spectrophotometer (Biochrom). As a blank 1 ml of the reagent mixture was used. Relative protein concentrations of the samples were obtained from the spectrophotometer.

SDS-PAGE

Equal amounts of proteins from each cell lysate were mixed with 10 µl loading dye and incubated in 100°C for 5 min. The samples used for detection of SHP1, SHP2 and STAT1 were loaded onto 8% polyacrylamide gels and samples used for detection of SOCS1 and SOCS3 onto 10% polyacrylamide gels (Running gel: 8% respectively 10% acrylamide, 0.4 M Tris-HCl pH 8.8, 0.1% N,N,N',N'-Tetramethylethylenediamin (TEMED) and 0.04% APS in dH_2O, Stacking gel: 4% acrylamide, 0.1 M Tris-HCl pH 6.8, 12% glycerol, 0.1% TEMED and 0.05% APS in dH_2O). The gels were run for 1-1.5 h at 175 V in 1xEB (1.9 M glycine, 0.25 M Tris, pH 8.7) buffer containing 0.5% SDS.

Transfer

The gels were incubated in transfer buffer (5.82 g Tris, 2.93 g Glycine, 0.038% SDS and 20% methanol in dH_2O) for 10 min. The membranes (Immobilon P Transfer Membrane, Millipore) were first wet in methanol and thereafter in transfer buffer. Sandwiches were made for the transfer consisting of two filter papers, moistened with transfer buffer, on each side with a membrane and the gel in the middle. Excess fluid was wiped off and a semi-dry transfer was run at 15 V for 75 min in RT.

Blotting

The membranes were blocked for unspecific binding in blocking buffer (PBS containing 0.1% Tween and 5% BSA) for 30 min in RT. Incubation with primary antibodies, diluted 1:200 in antibody solution (PBS containing 0.1% Tween and 1% BSA), was done over night at 4°C. Excess primary antibodies were washed away with washing buffer (PBS with 0.1% Tween) 3×5 min. Incubation with secondary antibodies, diluted 1:30000 (anti-rabbit) or 1:10000 (anti-mouse) in antibody solution, was done for 1 h in RT. Excess secondary antibodies were washed away with washing buffer, 3×5 min.

Before blotting the membranes for β-actin they were stripped from antibodies by incubation in 0.4 M NaOH for 10 min in RT followed by washing in wash buffer 3×2 min.

Primary antibodies used for western blotting were SH-PTP1 (C-19), rabbit polyclonal IgG, Santa Cruz Biotechnology, SH-PTP2 (C-18), rabbit polyclonal IgG, Santa Cruz Biotechnology, SOCS1 (H-93), rabbit polyclonal IgG, Santa Cruz Biotechnology, SOCS3 (H-103), rabbit polyclonal IgG, Santa Cruz Biotechnology, STAT1α p91 (C-111) mouse monoclonal, Santa Cruz Biotechnology and β-actin, anti-mouse, Sigma. Secondary antibodies used for western blotting were anti-mouse and anti-rabbit antibodies (Amersham).

ECL detection

Each membrane was incubated in ECL solution (0.1 M Tris-HCl pH 8.5, 0.05% Luminol, 0.02% P. Coumaric acid and 0.03% H_2O_2) for 1 min in RT. Chemiluminescence from the membranes was detected with a CCD camera (Fuji LAS-1000plus).

Results

Gene regions for CpG methylation analysis

The regions of the SHP1, SOCS1, SOCS3 and STAT1 genes determined to be analyzed for CpG methylation were regions that have shown aberrant methylation patterns in other tumors [12-17] and psoriasis [18], and the region of the SHP2 gene determined to be analyzed was an arbitrarily CpG rich promoter region. In the SHP1 gene, two promoter regions were analyzed, in three Pyrosequencing assays (fig 1A), extending over the regions -4674 to -4636 (assay A), -373 to -328 (assay B) and -262 to -209 (assay C). In the SHP2 gene one region, flanking the promoter and exon 1, was analyzed (fig 1B), extending over the region -92 to +8. In the SOCS1 gene two regions were analyzed (fig 1C); one region, situated in intron 1, extending over the region +515 – +598 (assay A) and another region, situated in exon 2, extending over the region +1013 – +1067 (assay B). In the SOCS3 gene two regions were analyzed, in three Pyrosequencing assays (fig 1D); one region, flanking the promoter and exon 1, extending over the region -74 to +7 (assay A) and another region, situated in intron 1, extending over the region +647 to +768 (assay B and C). In the STAT1 gene one region, situated in exon 1, was analyzed (fig 1E), extending over the region +114 to +189. Assay info can be found in table 1.

Table 1. The analyzed bisulfite converted sequence and number of CpG sites in each Pyrosequencing assay. The analyzed sequence of STAT1 is on the complementary strand compared to the genetic sequence. CpG sites are highlighted in yellow. Positions used as controls for bisulfite treatment are highlighted in grey.

Assay	Region	No. of CpG sites	Analyzed bisulfite converted sequence
SHP1 A	Promoter 1	4	GTYGTTGGTTTAGTTTYGTTTTTTGYGGTTTTTTGTYGT
SHP1 B	Promoter 2	4	YGTTYGGTATTTAGTAGGATTTATTYGATGATAGTTGTTATYGTTA
SHP1 C	Promoter 2	4	YGTGGGATYGTTTGGGTTYGTATGYGTGAAGTATTATTTGGGTTTGGA GTGTGT
SHP2	Promoter	16	YGGYGATTTGTGGAAYGAAATGAATGAAATYGATGTGGTAGYGGGTTY GGAYGGGTYGGTGGYGTAGAYGYGGAGYGYGTAGTTTATATTTGGYG GTYGYGGTTTTTAGGAGGAAGTAAGGATGTTTTGGATATTGTG
SOCS1 A	Intron 1	18	YGGATTYGYGYGGATTTGGTGTTTYGTGTTYGTTTTTTAGGGTYGGGTT YGTYGGGAGYGTYGTTTTTYGGAGTTGTTYGGTYGGTGTATATTTGTTY GGTTTYGTAGYGTTTTAGTTTATYGTTTTG
SOCS1 B	Exon 2	7	GYGYGATAGTYGTTAGYGGAATTGTTTTTTYGTTTTTAGYGTGAAGATG GTTTYGGGATT
SOCS3 A	Exon 1	18	YGGTYGYGTAGTTTTAGGAATYGGGGGGYGGGGYGYGGYGGTYGTTT ATATATTYGYGAGYGYGGTTTTYGYGGYGGTTTYGATTTGGATTTTTTG TTTYGTTGT
SOCS3 B	Intron 1	9	YGYGYGYGAGTTTTTTAYGTTGYGTTTTTGTAGTGYGYGTTTGGGAAGG GGTTGTTYGGGGTTA
SOCS3 C	Intron 1	18	TYGGTAGGGGYGGGAGTYGTGYGGGTTTYGTGAGGYGTTTGGATYGG AGYGYGGGTTTAGGAGAGGGTTTT
STAT1	Exon 1	12	TTTGYGYGTAGGATTYGGAAGGGTTAGGYGGGGGYGYGGYGGTGTAG TTTTTTTYGAGYGYGTTGGGTYGTTTTTGTTYGGTTTGGGGT

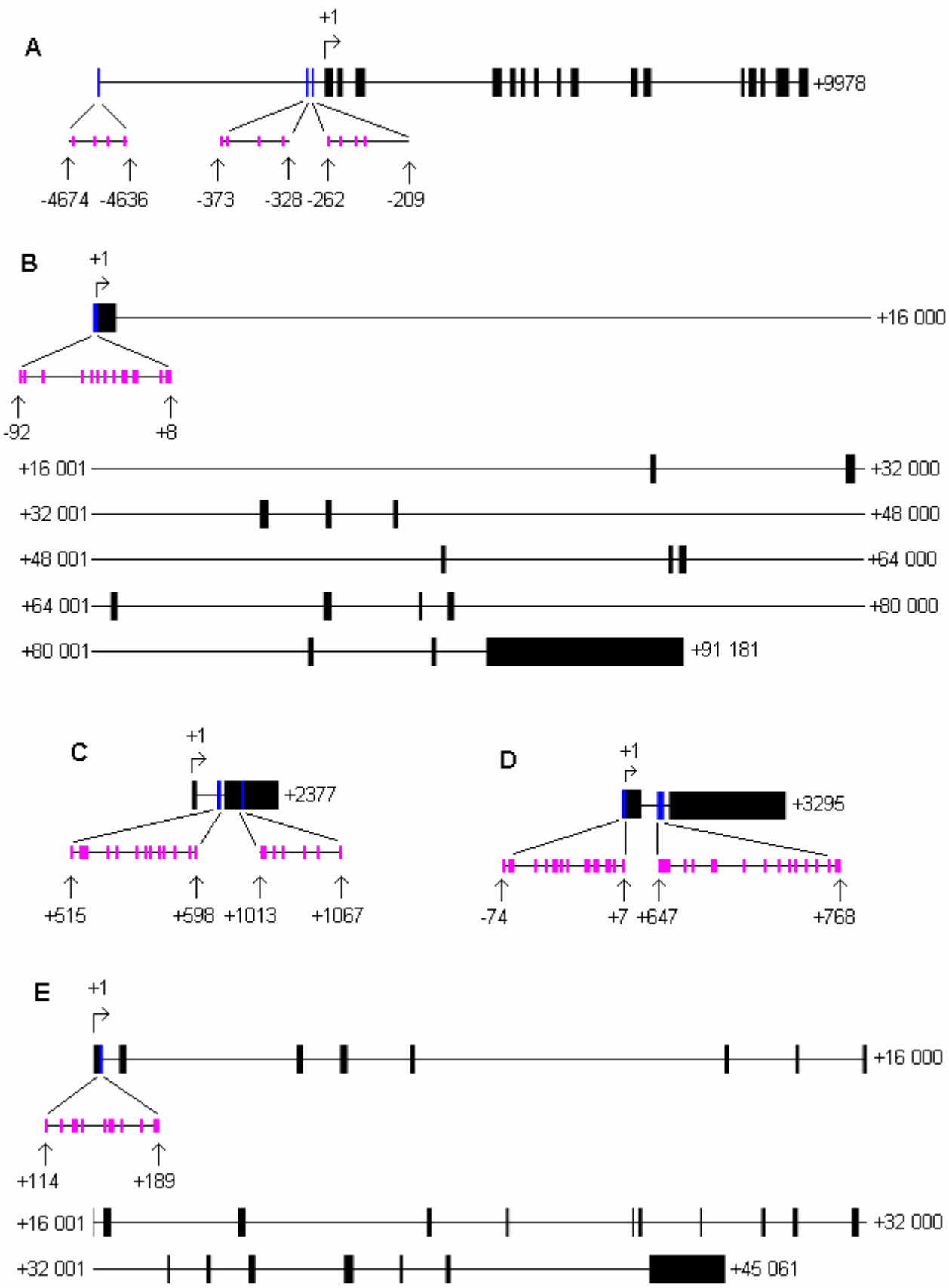


Fig 1. Schematic presentations of the genes SHP1 (A), SHP2 (B), SOCS1 (C), SOCS3 (D) and STAT1 (E) drawn to scale, showing the gene regions analyzed for CpG methylation enlarged. Transcription start sites are defined as +1. Black boxes represent exons. Blue boxes represent analyzed gene regions. Purple boxes represent analyzed CpG sites.

Primer design and PCR optimization

PCR and sequencing primers were designed, using the Assay Design Software (Biotage), to amplify and analyze the gene regions determined to be analyzed for CpG methylation indicated in fig 1. Primer info can be found in table 2.

The PCR was optimized for each PCR primer pair using DNA from normal blood donors, obtained from Monica Pettersson (Biotage AB, Kungsgatan 76, 753 18 Uppsala, Sweden). Five different annealing temperatures, between 46°C and 60°C, and two different $MgCl_2$ concentrations; 1.5 and 3.0 mM, were tested. Each PCR product was analyzed by Pyrosequencing to see which temperature interval and $MgCl_2$ conc. would give the best results, i.e. the Pyrogram most similar to the theoretical outcome and with the highest peaks. The PCR conditions can be seen in table 3. The PCR optimization was also confirmed with agarose gel electrophoresis for SHP1 A, B and C (fig. 2a), SOCS3 A and STAT1 (fig. 2b). It showed that the sizes of the amplified fragments were correct, between 100 and 200 bp, and also showed a correlation between stronger intensity of the bands in the gels and higher peak heights in the Pyrograms.

Table 2. The PCR and sequencing primers used to amplify and analyze CpG methylation in each assay. Two sequence primers were required to cover all the CpG sites in SOCS3 intron 1 (assay B and C), and two sequence primers and two forward primers were required to cover all the CpG sites in the promoter region 2 of SHP1 (assay B and C). Abbreviations: forw/F= forward, rev/R=reverse, P=PCR primer, Seq/S=sequencing primer, B=biotin.

Assay	Region	Type of primer	Primer name	Primer sequence 5'->3'
SHP1 A	Promoter 1	PCR forw	C027FP	TTTGGTTTGGGTTATTGTGTATAG
		PCR rev	C028RPB	CCTCCCTCCAAAACTAACAA
		Seq	C029FS	GGGTTATTGTGTATAGTTGT
SHP1 B	Promoter 2	PCR forw	C025FP	AGAAATTAATTAGATAAGGTATGTGAA
		PCR rev	C023RPB	ACACACTCCAAACCCAAATAATAC
		Seq 1	C026FS	TGTGAAAGTTATTATAGTATAG
SHP1 C	Promoter 2	PCR forw 2	C022FP	TAGTTGGTGGAGGAGGGAGAGAT
		PCR rev	C023RPB	ACACACTCCAAACCCAAATAATAC
		Seq 2	C024FS	GAGGAGGGAGAGATGT
SHP2	Promoter	PCR forw	C046FP	AAGGTTTTATAGTTAATGAGTGGA
		PCR rev	C047RPB	CACAATATCCAAAACATCCTTAC
		Seq	C049FS	TTTATAGTTAATGAGTGGAG
SOCS1 A	Intron 1	PCR forw	C033FP	TTAGTTGTGTTTATTGAGGTTGA
		PCR rev	C034RPB	AAAAAACAAAACCATAAACTAAAAC
		Seq	C035FS	TGTGTTTATTGAGGTTGAA
SOCS1 B	Exon 2	PCR forw	C051FP	GAGTTAGTGGGTATTTTTTTGG
		PCR rev	C052RPB	AATCCCCAAACCATCTTCAC
		Seq	C053FS	GTGGGTATTTTTTTGGT
SOCS3 A	Exon 1	PCR forw	C040FP	TGTTGAGAGTAGTGATTAAATATTATAAG
		PCR rev	C041RPB	CAACAACCAAACAAAAAATCC
		Seq	C042FS	GATTAAATATTATAAGAAGGT
SOCS3 B	Intron 1	PCR forw	C036FP	GGTTATATTTTTGGAGATTTAATTT
		PCR rev	C037RPB	AAAACCCTCTCCTAAACCC
		Seq	C038FS	ATTTTTGGAGATTTAATTTT
SOCS3 C	Intron 1	PCR forw	C036FP	GGTTATATTTTTGGAGATTTAATTT
		PCR rev	C037RPB	AAAACCCTCTCCTAAACCC
		Seq	C039FS	GTTGTTAGGGGTTATTTTG
STAT1	Exon 1	PCR forw	C043FP	GGTTAGAGGATTTTGTTTTTG
		PCR rev	C044RPB	CCTAACAAACCCCAAACC
		Seq	C045FS	GGAATTTTAAGGTTATTTAT

Table 3. The annealing temperatures and $MgCl_2$ concentration used in the PCR for the PCR primer pairs. Abbreviations: F= forward, R=reverse, P=PCR primer, B=biotin.

Primer pair		Annealing temp. [°C]	$MgCl_2$ conc. [mM]
C022FP	C023RPB	55	1.5
C025FP	C023RPB	50	1.5
C027FP	C028RPB	55	1.5
C033FP	C034RPB	50	3.0
C036FP	C037RPB	50	1.5
C040FP	C041RPB	50	3.0
C043FP	C044RPB	55	1.5
C046FP	C047RPB	50	1.5
C051FP	C052RPB	52	1.5

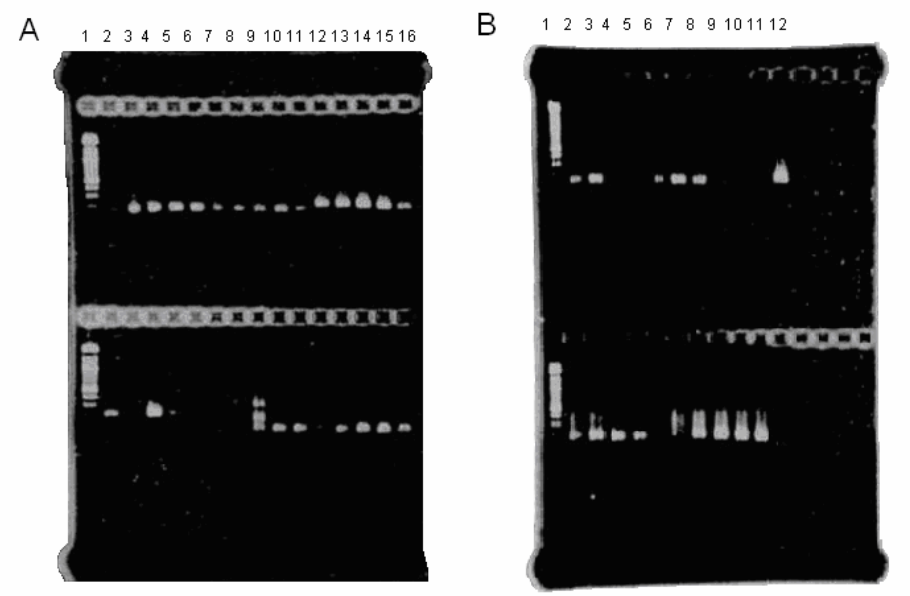


Fig 2. Agarose gel electrophoresis of PCR products from the PCR optimization. Gel A. Well 1: 100 bp ladder. Wells 2-6 and 12-16 upper panel and 7-11 lower panel: 3 mM $MgCl_2$. Wells 7-11 upper panel and 2-6 and 12-16 lower panel: 1.5 mM $MgCl_2$. Wells 2, 7 and 12: 46°C. Wells 3, 8 and 13: 48°C. Wells 4, 9 and 14: 50°C. Wells 5, 10 and 1 : 55°C. Wells 6, 11 and 16 : 60°C. Wells 2-11 upper panel SHP1 A. Wells 12-16 upper panel and 2-6 lower panel: SHP1 B. Wells 7-16 lower panel: SHP1 C. Gel B. Well 1: 100 bp ladder. Wells 2-5 and 11 upper panel and 2-6 lower panel: 1.5 mM $MgCl_2$. Wells 6-9 and 12 upper panel and 7-11 lower panel: 3 mM $MgCl_2$. Wells 2 and 6 upper panel and 2 and 7 lower panel: 46°C. Wells 3 and 7 upper panel and 3 and 8 lower panel: 50°C. Wells 4 and 8 upper panel and 4 and 9 lower panel: 53°C. Wells 11 and 12 upper panel and 6 and 11 lower panel: 58°C. Wells 5 and 9 upper panel and 5 and 10 lower panel: 60°C. Wells 11 upper panel and 12 lower panel: Control PCR without template DNA. Wells 2-12 upper panel: SOCS3 A. Wells 2-12 lower panel: STAT1.

PCR bias analysis

To investigate if the PCR primers amplify methylated and non-methylated DNA equally well in all the analyzed DNA regions, PCR reactions, in duplicates, were run for all primer pairs on DNA samples containing *in vitro* methylated and DNA from normal blood donors in the following ratios; 0:100, 15:85, 25:75, 50:50, 75:25, 85:15 and 100:0. Methylation percentages in the CpG sites in each DNA sample were obtained from Pyrosequencing reactions of all assays of the PCR-products. The mean CpG site methylation in the sample duplicates was plotted against the theoretical CpG methylation percentage. A linear regression curve was made and the R^2 value was calculated for each plot. A representative methylation plot can be seen in fig 3, and the R^2 values for the regression curve for each Pyrosequencing assay can be seen in table 4. Assays SHP1 B, SOCS1 B, SOCS3 A and SOCS3 C showed low levels of methylation (approx. 10 %) in samples with only DNA from normal blood donors.

Primer pairs with a non-linear curve (R^2 < approx. 0.8) were considered to have a bias for either methylated or non-methylated DNA and were not used for further CpG methylation analysis. Therefore, the PCR primer pair for SOCS1 B had to be redesigned once to obtain a linear regression curve with R^2 > 0.8.

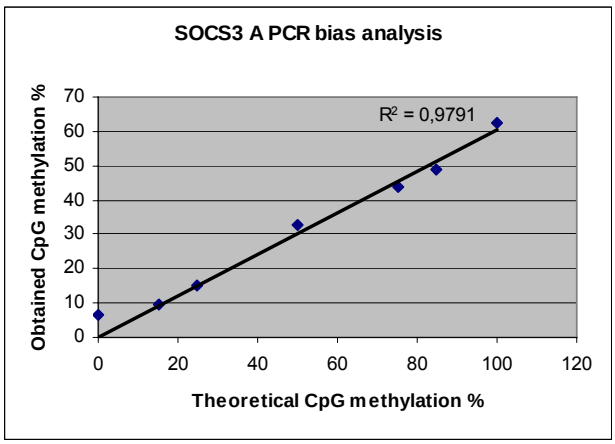


Fig 3. Methylation plot, with regression curve and R^2 value indicated, for assay SOCS3 A. The sample with 0% methylated DNA diverged from the regression curve because of low basal methylation in the DNA from normal blood donors. The methylation plot did not reach 100% methylation because of limited methylation capacity of the Sss I Methylase.

Table 4. R^2 values for the regression curves of the methylation plots, for each Pyrosequencing assay.

Assay	R^2
SHP1 A	0.88
SHP1 B	0.98
SHP1 C	0.97
SHP2	0.98
SOCS1 A	0.95
SOCS1 B	0.98
SOCS3 A	0.98
SOCS3 B	0.96
SOCS3 C	0.95
STAT1	0.86

CpG methylation analysis

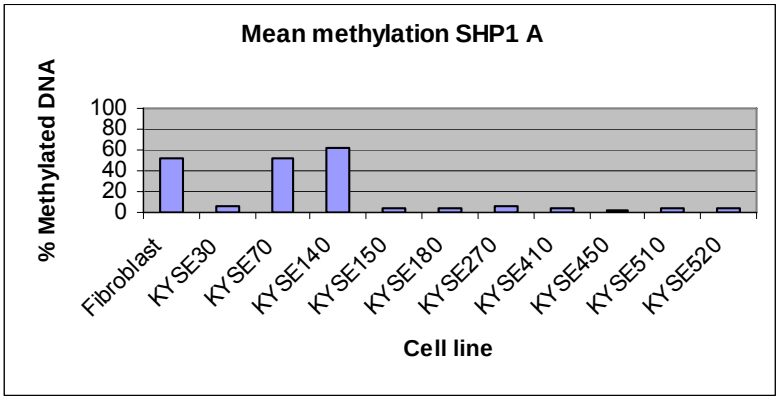
To analyze the levels of CpG methylation in the genes SHP1, SHP2, SOCS1, SOCS3 and STAT1 in EC and LC, DNA from ten different EC cancer cell lines and ten LC cell lines was purified and treated with bisulfite. The analyzed gene regions were amplified by PCR, in duplicates, and analyzed for CpG methylation by Pyrosequencing analysis. Theoretical outcomes together with examples of Pyrograms for each assay can be seen in supplementary fig. S1. Mean values of methylation in the duplicates were used (which all had low standard deviations). For each cell line the mean percent of methylated DNA for all CpG sites in each assay was plotted (in bar charts) and the percent of methylated DNA in each CpG site (in line charts). CpG sites that generated a warning in the Pyro Q-CpG 1.0.9 Software of having a high CpG sum deviation were excluded from the results.

Esophageal cancer

SHP1

The analyzed promoter region 1 of SHP1 (assay A) showed moderate mean degrees of methylation (50%-60%) in two (KYSE70 and KYSE140) of ten EC cell lines, and also in the control cell line (Fibroblast) (fig. 4A). The remaining eight cell lines did not show any significant methylation. The variation in methylation between the CpG sites in this promoter region was fairly low (fig. 4B). Only CpG site 4 diverged from the rest with an increase in methylation for KYSE140 (to 75%) and a decrease in methylation for KYSE70 (to 30%).

A



B

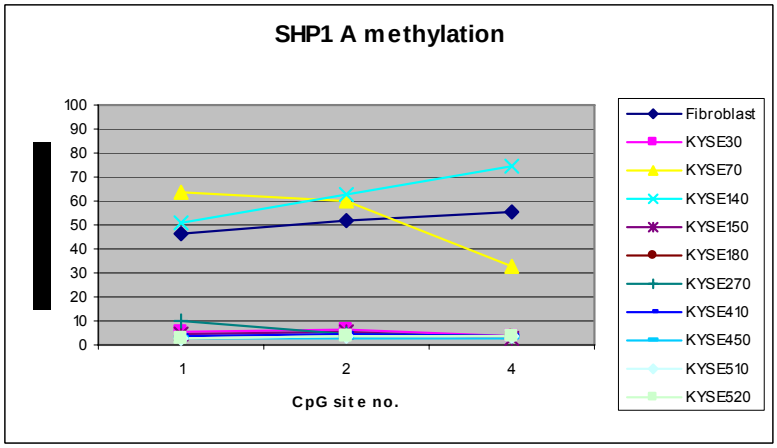


Fig 4. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1, 2 and 4 in the promoter region 1 of SHP1, in ten EC cell lines.

The analyzed promoter region 2 of SHP1 (assay B and C) showed very high mean degrees of methylation (approx. 90%) in nine (KYSE70, KYSE140, KYSE150, KYSE180, KYSE270, KYSE410, KYSE450, KYSE510 and KYSE520) of ten EC cell lines, and also in the fibroblast cell line (supplementary fig. S2A). The remaining cell line (KYSE30) showed a moderate mean degree of methylation (approx. 50%). This was due to the fact that KYSE30 had a large variation in methylation between CpG sites in this region, with very high degrees of methylation in CpG site 1 to 4 (85-100%) but low in CpG site 7 to 10 (10-20%) (supplementary fig. S2B).

SHP2

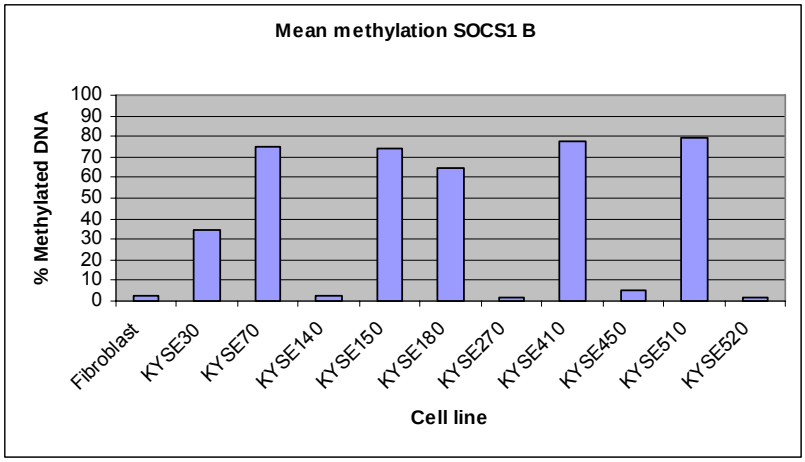
The analyzed promoter region of SHP2 did not show any significant degree of methylation in any of the EC cell lines (supplementary fig. S3A). However, there were low degrees of methylation in CpG site 1 for KYSE270 (20%) and CpG site 2 for the fibroblast cell line (15%) (supplementary fig. S3B).

SOCS1

The analyzed region of intron 1 in SOCS1 (assay A) showed low mean degrees of methylation (10-15%) in five (KYSE30, KYSE70, KYSE150, KYSE270 and KYSE410) of ten EC cell lines (supplementary fig. S4A). The remaining five EC cell lines did not show any significant degree of methylation. CpG site 2, 5 and 13 showed slightly increased degrees of methylation compared to the other CpG sites, in all EC cell lines (supplementary fig. S4B).

The analyzed region of exon 2 in SOCS1 (assay B) showed high mean degrees of methylation (65-80%) in five (KYSE70, KYSE150, KYSE180, KYSE410 and KYSE510) of ten LC cell lines (fig. 5A). One cell line (KYSE30) showed a moderate mean degree of methylation (35%). The remaining four LC cell lines did not show any significant degree of methylation in this region. There was a steep decrease in methylation in CpG sites 6 and 7 in all cell lines (towards 10%) (fig. 5B).

A



B

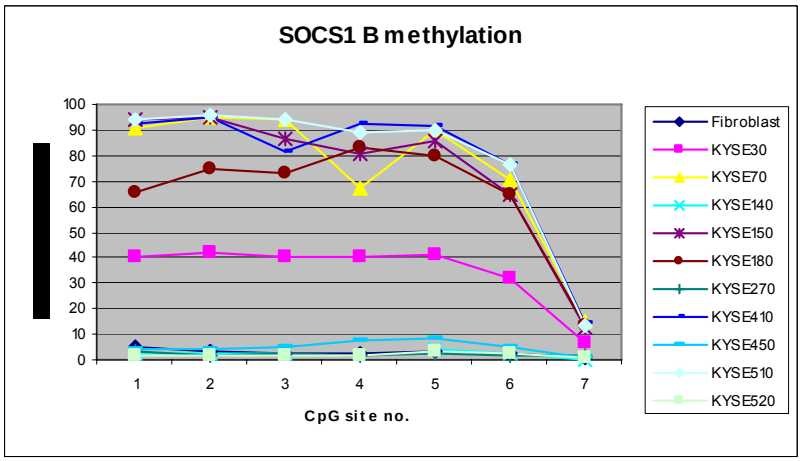


Fig 5. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 7 in the analyzed region of exon 2 in SOCS1, in ten EC cell lines.

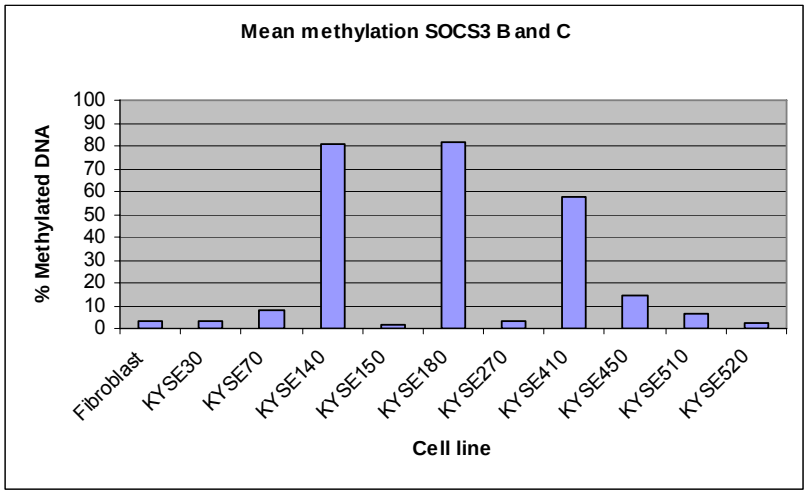
SOCS3

The analyzed promoter region of SOCS3 (assay A) did not show any significant mean degree of methylation in any of the EC cell lines (supplementary fig. S5A). However, there was a trend of slightly increased methylation in CpG site 4 (10-15%), 10 and 14 (up to 10%) in all EC cell lines and in CpG site 1 for KYSE270 and the fibroblast cell line (15%) (supplementary fig. S5B).

The analyzed region of intron 1 in SOCS1 (assay B and C) showed high mean degrees of methylation (60-80%) in three (KYSE140, KYSE180 and KYSE410) of ten EC cell lines (fig. 6A). Two EC cell lines (KYSE70 and KYSE450) showed low mean degrees of methylation (10-15%). The remaining five cell lines did not show any significant degree of methylation. There was a great variation in methylation between CpG sites in this

region for the cell lines KYSE140 and KYSE180 (from 55% to 100%) and for KYSE410 (from 25% to 95%) (fig. 6B).

A



B

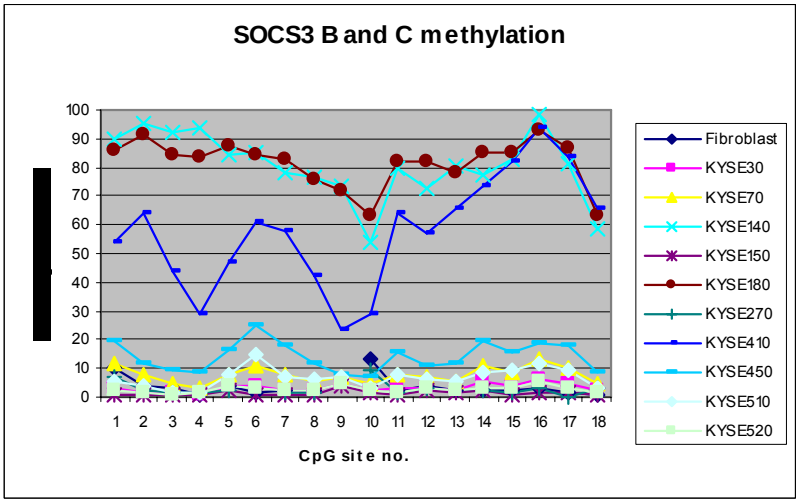


Fig 6. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 18 in the analyzed region of intron 1 in SOCS3, in ten EC cell lines.

STAT1

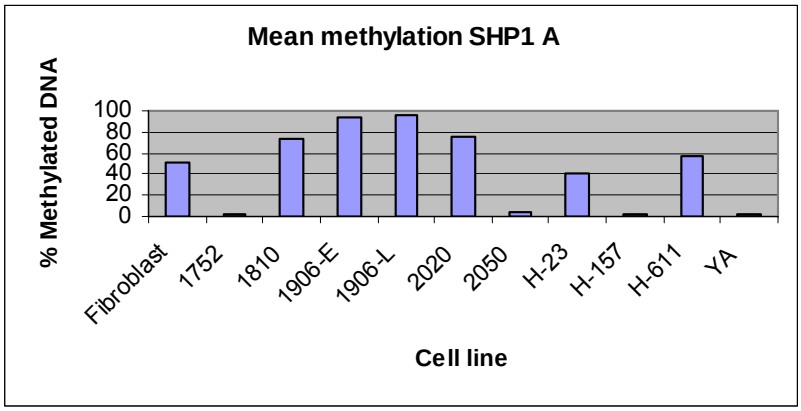
The analyzed region of exon 1 in STAT1 showed no significant mean degree of methylation in any EC cell line (supplementary fig. S6A). However, there was a low degree of methylation in CpG site 1 in the fibroblast cell line (10%) (supplementary fig. S6B).

Lung cancer

SHP1

The analyzed promoter region 1 of SHP1 (assay A) showed high mean degrees of methylation (70%-100%) in four (1810, 1906-E, 1906-L and 2020) of ten LC cell lines (fig 7A). Two cell lines (H-23 and H-611) and the control cell line (Fibroblast) showed moderate mean degrees of methylation (40-60%). The remaining four LC cell lines did not show any significant degree of methylation. There was an increased degree of methylation in CpG site 4 in 1810 and H-611 (to 95% and 75% respectively) and in the fibroblast cell line (to 60%) and a decreased degree of methylation in 2020 (to 50%) (fig. 7B).

A



B

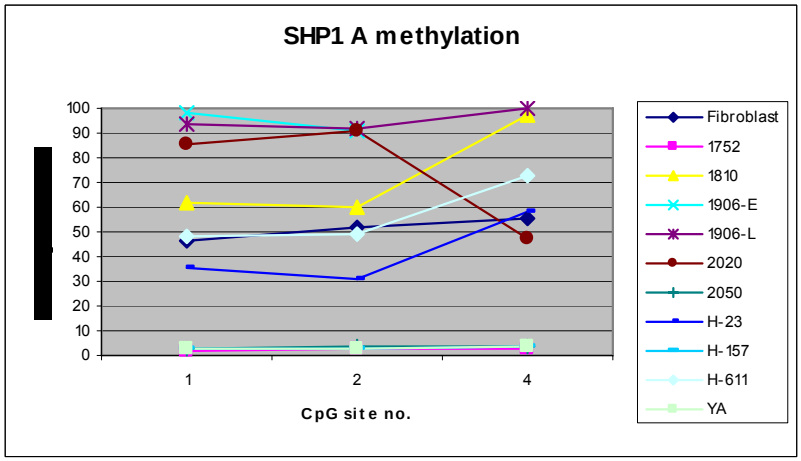


Fig 7. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1, 2 and 4 in the promoter region 1 of SHP1, in ten LC cell lines.

The analyzed promoter region 2 of SHP1 (assay B and C) showed very high mean degrees of methylation (90-95%) in all ten LC cell lines and also in the fibroblast cell line (supplementary fig. S7A). There was no significant variation in methylation between CpG sites in this region, except for slightly decreased degrees of methylation in CpG site 2 for the LC cell lines 1810 (to 60%) and 2050 (to 75%) and the fibroblast cell line (to 70%) (supplementary fig. S7B).

SHP2

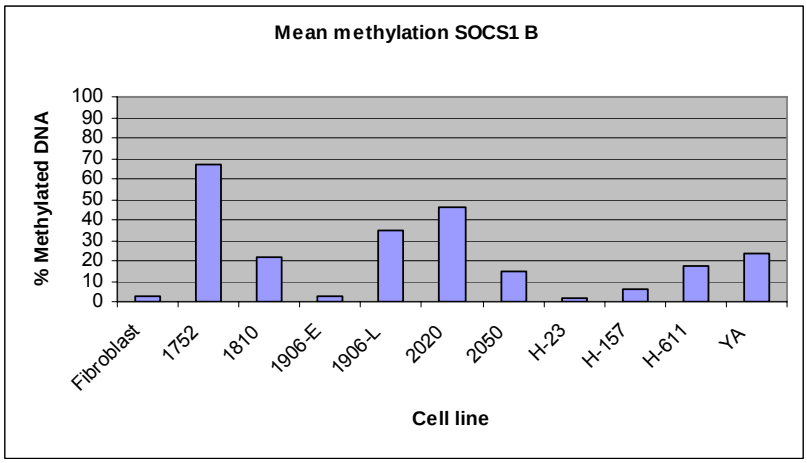
The analyzed promoter region of SHP2 did not show any significant mean degree of methylation in any of the LC cell lines (supplementary fig. S8A). However, there were low degrees of methylation in CpG sites 1, 2 and 3 for 1906-E (10-20%) and in CpG sites 1, 2 and 7 for the fibroblast cell line (10-15%) (supplementary fig. S8B).

SOCS1

The analyzed region of intron 1 in SOCS1 (assay A) did not show any significant degree of methylation (supplementary fig. S9A). However, there were low degrees of methylation in CpG site 1 and 2 for the fibroblast cell line and 1906-E (10-25%) and in CpG site 11 and 13 for 1752 (10-15%) (supplementary fig. S9B).

The analyzed region of exon 2 in SOCS1 (assay B) showed a high mean degree of methylation (70%) in one (1752) of ten LC cell lines (fig. 8A). Two LC cell lines (1906-L and 2020) showed moderate degrees of methylation (35-45%). Four cell lines (1810, 2050, H-611 and YA) showed low mean degrees of methylation (10-20%). The remaining two cell lines did not show any significant degree of methylation. There was a trend of increasing methylation in CpG site 2 to 6 in the eight methylated cell lines (fig. 8B).

A



B

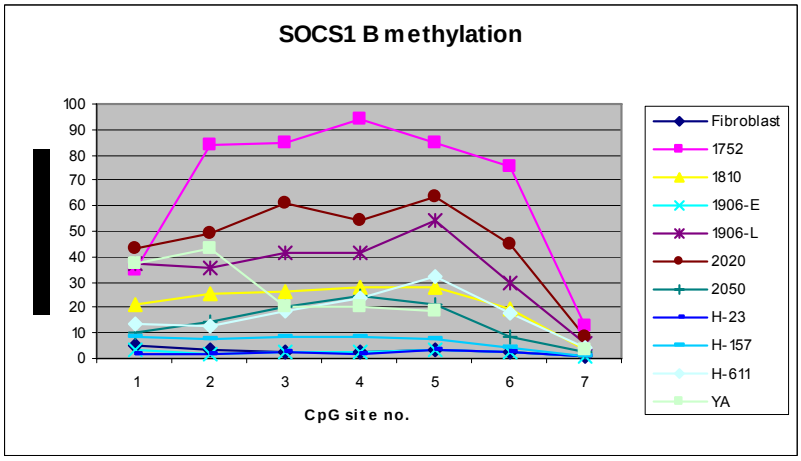


Fig 8. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 7 in the analyzed region of exon 2 in SOCS1, in ten LC cell lines.

SOCS3

Neither the analyzed region of the promoter nor of intron 1 in SOCS3 (assay A) showed any significant mean degree of methylation (supplementary fig. S10 A and C). However, there were low degrees of methylation in the promoter region in CpG site 1 for 1906-E and the fibroblast cell line (10-15%) and in CpG site 4 in all LC cell lines (10%) (supplementary fig. S10B). In the intron 1 region there were low degrees of methylation in CpG site 1 (10%) and 10 (15%) for the fibroblast cell line (supplementary fig. S10D).

STAT1

The analyzed region of exon 1 in STAT1 did not show any significant mean degree of methylation in any LC cell line (supplementary fig. S11A). There was a low degree of methylation in CpG site 1 in the fibroblast cell line (10%) (supplementary fig. S11B).

Protein expression analysis

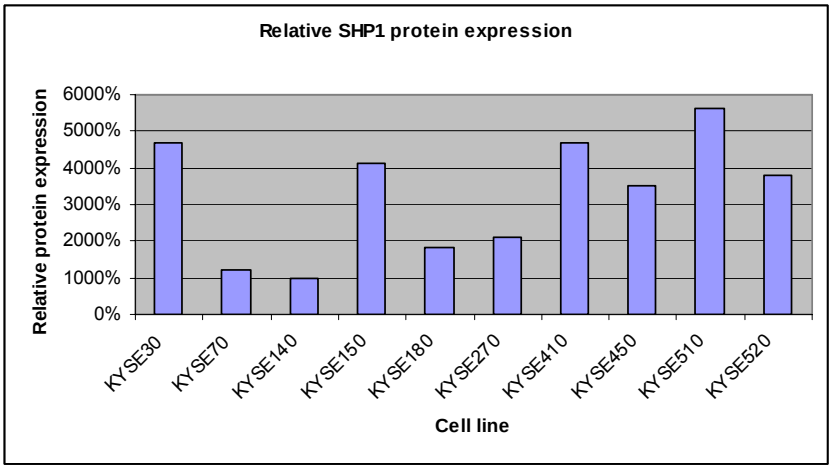
To analyze whether the levels of CpG methylation in the genes SHP1, SHP2, SOCS1, SOCS3 and STAT1 could be correlated to protein expression in EC and LC, cell lysates from ten different EC cell lines and nine LC cell lines were analyzed in western blot analyses. The SCLC cell line 2050 was not included in the protein expression studies due to poor growth. The band intensities of the EC and LC cell lines were normalized against the amount of β-actin in each lane (as a loading control) and correlated to the band intensity of the control cell line (fibroblast) on each membrane, which gave a measurement of relative protein expression. All western blots were carried out once. The results can be seen in supplementary fig S12. The western blots for SOCS1 and SOCS3 showed no distinct bands, possibly due to poor quality of the antibodies. Western blots with new antibodies are still left to be done for these two genes.

Esophageal cancer

SHP1

The protein expression of SHP1 in the EC cell lines was low in KYSE70 and KYSE140, moderate in KYSE180, KYSE270, KYSE450 and KYSE520 and high in KYSE30, KYSE150, KYSE410 and KYSE510 (fig. 9A). The differences in protein expression levels could not be correlated to the levels of methylation in promoter region 2, since all cell lines showed approximately equally high levels of methylation in this region. But interestingly, there seemed to be a correlation between methylation in promoter region 1 and a reduction in protein expression, since the two cell lines that showed highest levels of methylation in this region (KYSE70 and KYSE140) were the ones with lowest protein expression. Thus, there was a slight trend of decreasing protein expression with increasing methylation in this promoter region, even though the linear coefficient of the regression curve was fairly low (fig 9B).

A



B

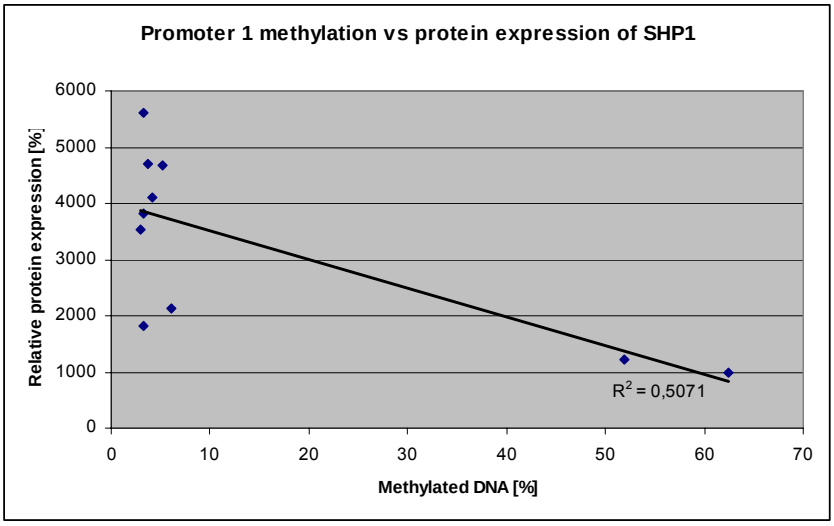


Fig 9. The relative SHP1 protein expression compared to the control cell line (A) and the correlation between promoter region 1 methylation and protein expression of SHP1 (B), in ten EC cell lines.

SHP2

The expression of SHP2 in the EC cell lines was low in KYSE30 and KYSE270, moderate in KYSE70, KYSE140 and KYSE150, KYSE180, KYSE410 and KYSE450 and high in KYSE510 and KYSE520 (fig 10). The variation in protein expression could not be correlated to methylation in the analyzed promoter region, since it did not show any significant methylation.

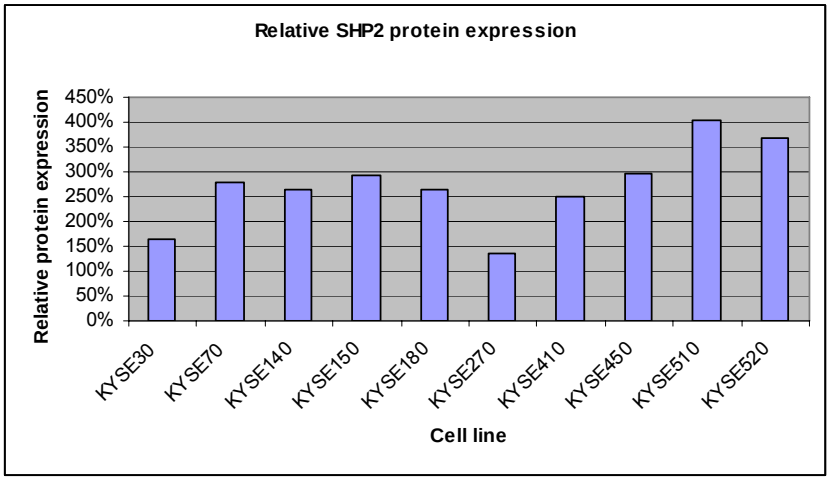


Fig 10. The relative SHP2 protein expression, compared to the control cell line, in ten EC cell lines.

STAT1

The expression of STAT1 was low in KYSE140, KYSE150, KYSE410 and KYSE520, moderate in KYSE30, KYSE270 and KYSE510 and high in KYSE70, KYSE180 and KYSE450 (fig. 11). The variation in protein levels could not be correlated to methylation in the analyzed promoter region since it did not show any significant methylation.

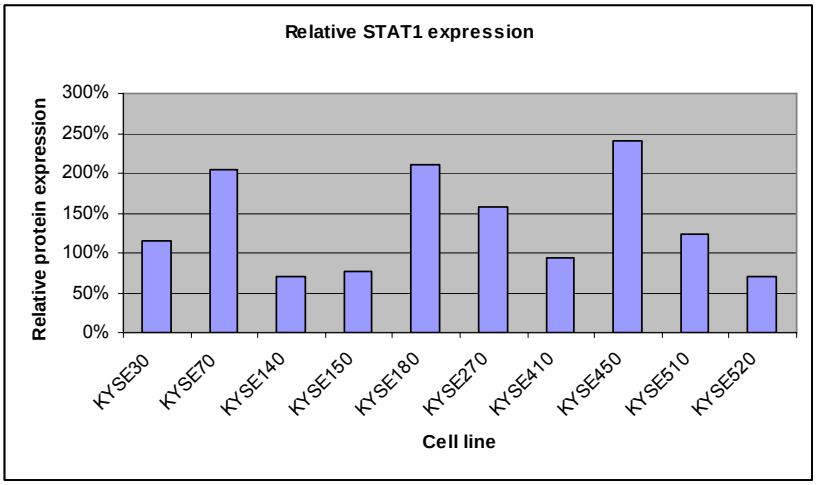


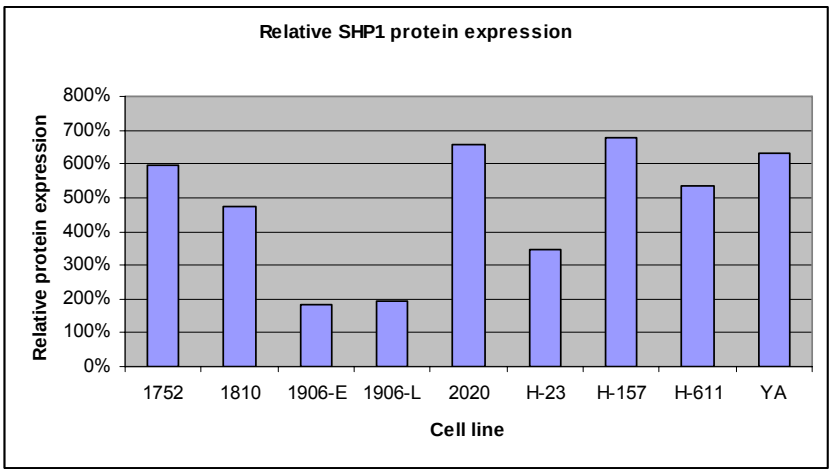
Fig 11. The relative protein expression of STAT1, compared to the control cell line, in ten EC cell lines.

Lung cancer

SHP1

The expression of SHP1 in the LC cell lines was low in 1906-E and 1906-L, moderate in 1810, H-23 and H-611 and high in 1752, 2020, H-157, and YA (fig. 12A). The differences in protein expression levels could not be correlated to the methylation levels in promoter region 2, since all cell lines showed approximately equally high levels of methylation in this region. However, there seemed to be a correlation between methylation in promoter region 1 and a reduction in protein expression. The two cell lines with highest levels of methylation in this region (1906-E and 1906-L) were the ones with lowest protein expression, and the cell lines that did not show any significant methylation (1752, H-157 and YA) were among the ones with highest protein expression. Furthermore, the cell lines, 1810, H-23 and H-611, with high and moderate levels of methylation showed low and moderate levels of protein expression. One exception was 2020, which showed both high level of methylation and protein expression. Thus, there was a trend of decreasing protein expression with increasing methylation in this promoter region, even though the linear coefficient of the regression curve was fairly low (fig 12B).

A



B

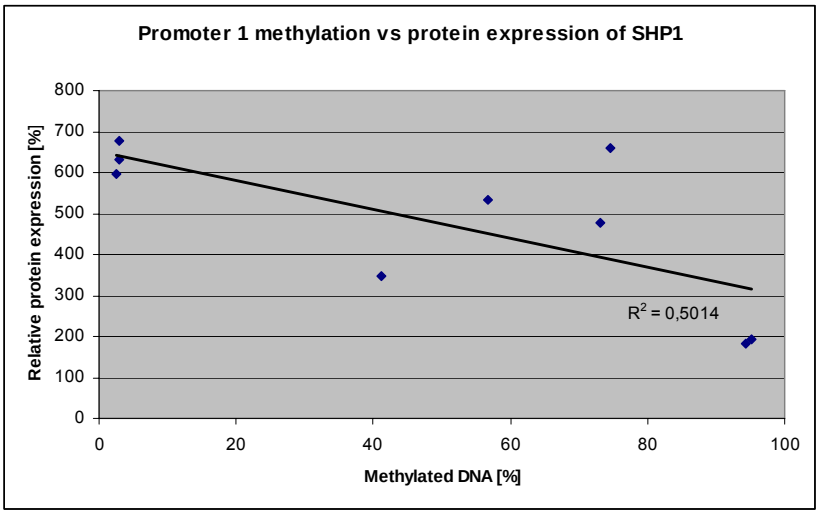


Fig 12. The relative SHP1 protein expression compared to the control cell line (A) and the correlation between promoter region 1 methylation and protein expression of SHP1 (B), in nine LC cell lines.

SHP2

SHP2 was expressed at fairly equal levels throughout the LC cell lines, with exception of H-611 and YA which showed higher expression than the rest (fig. 13). This variation could not be correlated to methylation in the analyzed promoter region, since it did not show any significant methylation.

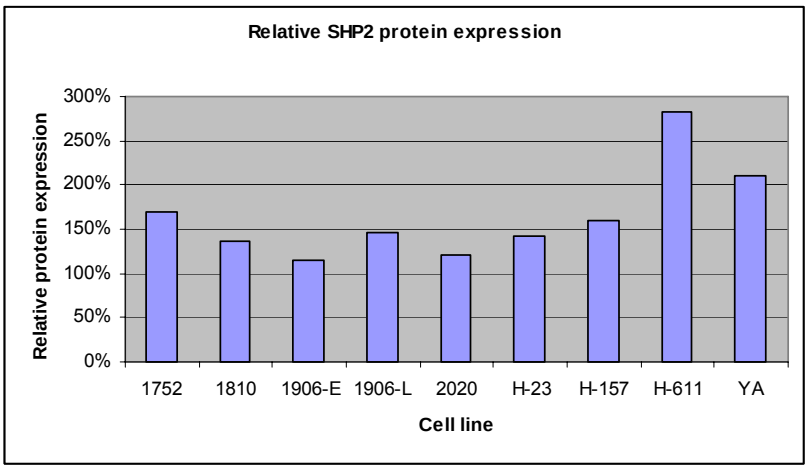


Fig 13. The relative SHP2 protein expression, compared to the control cell line, in nine LC cell lines.

STAT1

The expression of STAT1 in the LC cell lines was low in 1810, 1906-E, 2020, H-23 and H-157, moderate in 1752, 1906-L and YA and high in H-611 (fig. 14). The variation in protein expression could not be correlated to methylation in the analyzed region of exon 1, since it did not show any significant methylation.

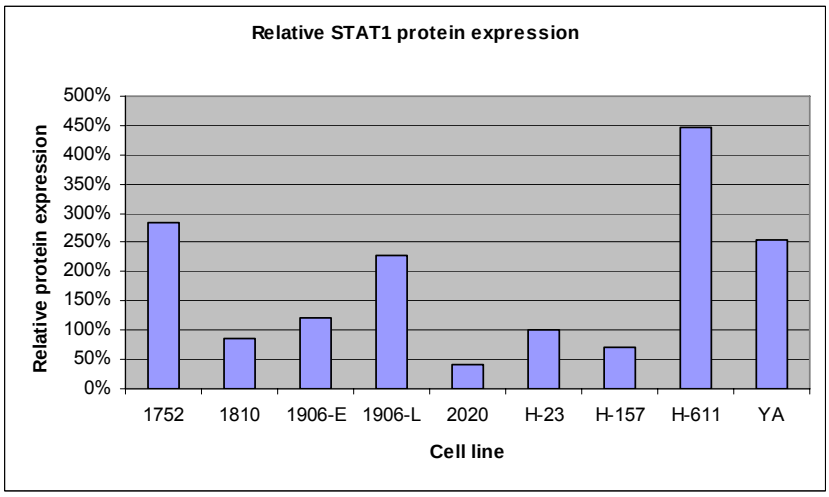


Fig 14. The relative STAT1 protein expression, compared to the control cell lines, in nine LC cell lines.

Discussion

In this study, DNA CpG methylation was examined in the genes SHP1, SHP2, SOCS1, SOCS3 and STAT1 in ten EC cell lines and ten LC cell lines. The effect of DNA methylation on protein expression was further investigated in SHP1, SHP2 and STAT1 in the same EC cell lines as in the methylation studies and in nine of the LC cell lines. The methylation studies showed that the genes SHP1, SOCS1 and SOCS3 are methylated at varying levels throughout the analyzed EC and LC cell lines; whereas SHP2 and STAT1 were not methylated in any cell line. The protein expression studies interestingly showed that the observed methylation levels in SHP1 were associated with reduction of protein expression in both the EC and LC cell lines.

SHP1 has earlier shown methylation-associated inactivation in tumors such as lymphomas and leukemia [24], but to our knowledge no previous studies have been made in EC and LC. This study showed that the promoter region 1 (SHP1 A) of SHP1 has varying methylation levels throughout the EC and LC cell lines, whereas promoter region 2 (SHP1 B and C) is highly methylated in all cell lines. However, the epigenetic regulation of SHP1 has earlier been attributed to promoter region 2 [25, 26]. In this study on the other hand, the reduced protein expression seems to be associated with methylation in promoter region 1. This could imply that, since the two promoter regions direct the expression of two different mRNA isoforms of SHP1 [27], EC and LC cells express isoform I. Surprisingly, in this study SHP1 was also methylated in normal fibroblast cells, in both promoter regions. SHP1 has earlier been reported to be methylated in promoter region 2 in normal epithelial cells [18], but to our knowledge there have been no reports on methylation in promoter region 1 in normal cells. This is an interesting finding and may be the cause of the observed low protein expression of SHP1 in fibroblast cells. However, the levels of protein expression of SHP1 in fibroblasts compared to EC and LC cell lines are contradictory. Although several EC and LC cell lines have higher levels of methylation in SHP1 than the fibroblast cell line, they all express SHP1 at higher levels than the fibroblasts. Thus, there seems to be other factors controlling the protein expression of SHP1 in normal fibroblasts, and not solely methylation.

To our knowledge, no previous methylation studies have been made on SHP2. Therefore, an arbitrarily CpG rich promoter region of SHP2 was chosen for methylation studies. However, it did not show any significant methylation in any of the analyzed EC and LC cell lines. This may be correlated to the high protein expression observed in all EC and LC cell lines. However, the variation in protein levels, in especially the EC cell lines, can not be correlated to the levels of methylation in the analyzed promoter region. Thus, there seems to be other factors controlling the protein expression of SHP2 in EC and LC, and not solely methylation.

SOCS1 was shown in this study to have low levels of methylation in the intron 1 region (SOCS1 A) in all analyzed EC and LC cell lines, and high levels of methylation in the exon 2 region (SOCS1 B and C) in several of the EC and LC cell lines. This agrees with earlier studies which show that SOCS1 exon 2 has heightened levels of methylation in tumor tissues compared to non-timorous tissue in LC [28]. SOCS1 has also shown methylation-associated inactivation in tumors such as chronic myeloid leukemia [12] and hepatocellular carcinoma [13]. Hence, it would be interesting to investigate if the differences in methylation levels observed between the cell lines in this study are related to the protein expression levels in EC and LC (still left to be done). Especially since, as far as we know, no previous methylation studies have been done in EC. It would also be interesting to investigate the methylation status of the adjacent CpG sites in the analyzed exon 2 region. The last CpG site analyzed in this region diverges from the rest with a very low level of methylation in all EC and LC cell lines. If this trend of low methylation continues in the forthcoming CpG sites it could imply that the methylation of exon 2 is restricted to only a few CpG sites.

In this study SOCS3 showed no significant methylation in the promoter region (SOCS3 A) in any of the EC or LC cell lines or in the intron 1 (SOCS3 B and C) region in the LC cell lines. However, the intron 1 region showed high levels of methylation in a few of the EC cell lines. This is contrary to previous methylation studies made in other NSCLC cell lines and tissues, which have reported that the promoter of SOCS3 is methylated [28, 29]. It could imply that the selection of cell lines in this study does not show a representative result for LC cell lines. If more cell lines would have been included there might be a larger proportion of cell lines with SOCS3 methylation. SOCS3 has also shown methylation-associated silencing in LC [29] and in other tumors such as hepatocellular carcinoma [14]. Hence, it would be interesting to investigate if this occurs also in EC; if there is a relation between the observed variation in methylation levels and protein expression levels (still left to be done). It would be especially interesting to correlate the protein expression of KYSE410 with the other cell lines, since it showed such great variation in methylation levels throughout the intron 1 region. This could give information about whether certain CpG cites have greater effect on protein expression than others.

STAT1 has previously shown methylation-associated inactivation in e.g. squamous cell carcinoma of head and neck [30] and also methylation-associated tumor recurrence in non-muscle invasive bladder carcinoma [16]. However, this study shows that STAT1 does not have any significant degree of promoter methylation in any of the analyzed EC and LC cell lines. The observed variation in protein expression can not be correlated to the promoter methylation levels. Thus, there seems to be other factors controlling the protein expression of SHP1 in EC and LC, and not solely methylation.

To conclude the most important findings in this study; SHP1, SOCS1 and SOCS3 are highly methylated in several EC and LC cell lines, which makes them potentially diagnostic and prognostic markers for EC and LC. The relevance from a clinical perspective, e.g. association with tumor grade, stage, growth characteristics and sensitivity toward chemotherapeutics warrants further investigation. Furthermore, the SHP1 promoter region 1 methylation is associated with protein expression reduction, which means its methylation may be a possible contributor to tumor formation in EC and LC, as has been suggested for haematopoietic malignancies previously [31].

Acknowledgements

I want to thank all my supervisors; Monica Pettersson at Biotage, Simon Ekman and Michael Bergqvist at Uppsala University Hospital and Johan Lennartsson at the Ludwig Institute for Cancer Research (LICR), for teaching me new techniques, showing me around in the labs and helping me out with both practical problems and reviewing of my report. I also want to thank my scientific reviewer, Joachim Gullbo, for reviewing and helping me to improve my written report. Furthermore, I like to thank Xuping Wu and Anita Klinga, at the Rudbeck Laboratory, for providing me with cell pellets for the methylation studies and also Anita for showing me the cell cultivation of EC and LC cell lines. Finally, I would like to thank all the people who have taken their time to answer my questions in the labs at Biotage, Rudbeck and LICR and who have been a nice company during my degree project.

References

[1] Fisichella, M.P. and Patti, M. (2006) Esophageal cancer, *eMedicine*, http://www.emedicine.com/med/topic741.htm (18th Jan 2007).

[2] Parkin, D.M., Bray, F., Ferlay, J., and Pisani, P. (2005) Global cancer statistics, 2002, CA Cancer J Clin, **55**, 74-108.

[3] (2006) Esophageal Cancer (PDQ®): Treatment, *National Cancer Institute*, http://www.cancer.gov/cancertopics/pdq/treatment/esophageal/healthprofessional/ (19th Jan 2007).

[4] Line, B.R., Maragh, M.R. and Ahamed, T.B. (2002) Positron Emission Tomography Imaging of Lung and Esophageal Cancer, *Applied Radiology*, **31(6)**, 9-17.

[5] (2007) Lung cancer, *Wikipedia*, http://en.wikipedia.org/wiki/Lung_cancer (22nd Jan 2007).

[6] (2006) Lung Cancer Fact Sheet, *American Lung Association*, http://www.lungusa.org/site/pp.asp?c=dvLUK9O0E&b=669263 (22 Jan 2007).

[7] Granville, C.A. and Dennis, P.A. (2005) An Overview of Lung Cancer Genomics and Proteomics, *American Journal of respiratory cell and molecular biology*, **32**, 169-176.

[8] Baylin, S.B. (2005) DNA methylation and gene silencing in cancer. *Nature Clinical Practice*, **2**, S4-S11.

[9] Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nature Reviews*, **3**, 415-428.

[10] Esteller, M. (2006) The necessity of a human epigenome project, Carcinogenesis, **27(6)**, 1121-1125.

[11] Wu, D.L., Sui, F.Y., Jiang, X.M. and Jiang, X.H. (2006) Methylation in esophageal carcinogenesis, *World J Gastroenterol*, **21;12(43)**, 6933-6940.

[12] Liu, T.C., Lin, S.F., Chang, J.G., Yang, M.Y., Hung, S.Y. and Chang, C.S. (2003) Epigenetic alteration of the SOCS1 gene in chronic myeloid leukemia. *British Journal of Haematology*, **123**, 654-661.

[13] Yoshikawa, H., Matsubara, K., Qian, G.S., Jackson, P., Groopman, J.D., Manning, J.E., Harris, C.C. and Herman, J.G. (2001) SOCS-1, a negative regulator of the JAK/STAT pathway, is silenced by methylation in human hepatocellular carcinoma and shows growth-suppression activity. *Nature Genetics*, **28**, 29-35.

[14] Niwa, Y., Kanda, H., Shikauchi, Y., Saiura, A., Matsubara, K., Kitagawa, T., Yamamoto, J., Kubo, T. and Yoshikawa, H. (2005) Methylation silencing of SOCS-3 promotes cell growth and migration by enhancing JAK/STAT and FAK signalings in human hepatocellular carcinoma. *Oncogene*, **24**, 6406-6417.

[15] Zhang, Q., Wang, H.Y., Marzec, M., Raghunath, P.N., Nagasawa, T. and Wasik, M.A. (2005) STAT3- and DNA methyltransferase 1-mediated epigenetic silencing of SHP-1 tyrosine phosphatase tumor suppressor gene in malignant T lymphocytes. *PNAS*, **102:19**, 6948-6953.

[16] Friedrich, M.G., Chandrasoma, S., Siegmund, K.D., Weisenberger, D.J., Cheng, J.C., Toma, M.I., Huland, H., Jones, P.A. and Liang, G. (2005) Prognostic relevance of methylation markers in patients with non-muscle invasive bladder carcinoma. *European Journal of Cancer*, **41**, 2769-2778.

[17] Johan, M.F., Bowen, D.T., Frew, M.E., Goodeve, A.C. and Reilly, J.T. (2005) Aberrant methylation of the negative regulators RASSFIA, SHP-1 and SOCS-1 in myelodysplastic syndromes and acute myeloid leukaemia. *British Journal of Haematology*, **129**, 60-65.

[18] Ruchusatsawat, K., Wongpiyabovorn, J., Shuangshoti, S., Hirankarn, N. and Mutirangura, A. (2006) SHP-1 promoter 2 methylation in normal epithelial tissues and demethylation in psoriasis, *Journal of Molecular Medicine*, **84,** 175-182.

[19] Uhlén, P., Burch, P.M., Zito, C.I., Estrada, M., Ehrlich, B.E. and Bennett, A.M. (2006) Gain-of-function/Noonan syndrome SHP-2/*Ptpn11* mutants enhance calcium oscillations and impair NFAT signaling, *PNAS*, **103(7)**, 2160-2165.

[20] IPR012152 Protein-tyrosine phosphatase, non-receptor 6/11, *InterPro*, http://www.ebi.ac.uk/interpro/IEntry?ac=IPR012152#PUB00005794 (25th Jan 2007).

[21] You, M., Yu, D. and Feng, G. (1999) Shp-2 Tyrosine Phosphatase Functions as a Negative Regulator of the Interferon-Stimulated Jak/STAT Pathway, *Molecular and Cellular Biology*, **19(3)**, 2416-2424.

[22] Protein CISH_HUMAN, *UniProt*, http://www.ebi.uniprot.org/uniprot-srv/uniProtView.do?proteinId=CISH_HUMAN&pager.offset=0 (25th Jan 2007).

[23] Klampfer, L. (2006) Signal transducers and activators of transcription (STATs): Novel targets of chemopreventive and chemotherapeutic drugs, *Curr Cancer Drug Targets*, **6(2)**, 107-121.

[24] Oka, T., Ouchida, M., Koyama, M., Ogama, Y., Takada, S., Nakatani, Y., Tanaka, T., Yoshino, T., Hayashi, K., Ohara, N., Kondo, E., Takahashi, K., Tsuchiyama, J., Tanimoto, M., Shimizu, K. and Akagi, T. (2002) Gene silencing of the tyrosine phosphatase SHP1 by aberrant methylation in leukemias/lymphomas, *Cancer Research*, **62**, 6390-6394.

[25] Zhang, Q., Raghunath, P.N., Vonderheid, E., Odum, N. and Wasik, M.A. (2000) Lack of phosphotyrosine phosphatase SHP-1 expression in malignant T-cell lymphoma cells results from methylation of the SHP-1 promoter, *Am j Pathol*, **157**, 1137-1146.

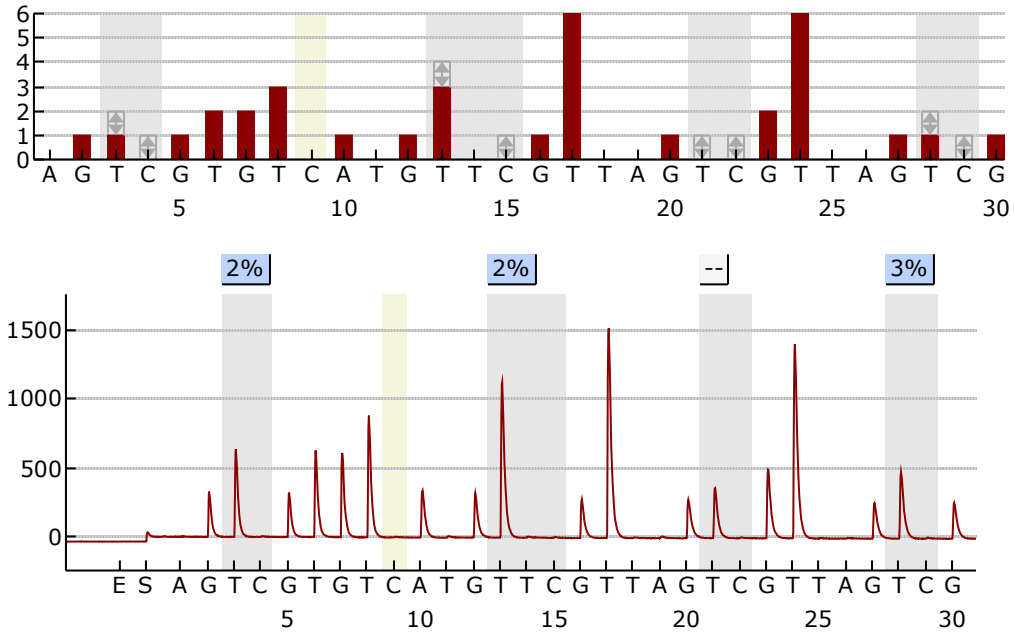
[26] Oka, T., Yoshino, T., Hayashi, K., Ohara, N., Nakanishi, T., Yamaai, Y., Hiraki, A., Sogawa, C.A., Kondo, E., Teramoto, N., Takahashi, K., Tsuchiyama, J. and Akagi, T. (2001) Reduction of hematopoietic cell-specific tyrosine phosphatase SHP-1 gene expression in natural killer cell lymphoma and various types of lymphomas/leukemias: combination analysis with cDNA expression array and tissue microarray, *Am J Pathol*, **159**, 1495-1505.

[27] Banville, D., Stocco, R. and Shen, S.H. (1995) Human protein tyrosine phosphatase 1C (PTPN6) gene structure: alternate promoter usage and exon skipping generate multiple transcripts, *Genomics*, **27**, 165-173.
[28] Shivapurkar, N., Stastny, V., Suzuki, M., Wistuba, I.I., Li, L., Zheng, Y., Feng, Z., Hol, B., Prinsen, C., Thunnissen, F.B. and Gazdar, A.F. (2007) Application of a methylation gene panel by quantitative PCR for lung cancers, *Cancer Letters*, **247**, 56-71.
[29] He, B., You, L., Uematsu, K., Zang, K., Xu, Z., Lee, A.Y., Costello, J.F., McCormick, F. and Jablons, D.M. (2003) SOCS-3 is frequently silenced by hypermethylation and suppresses cell growth in human lung cancer, *PNAS*, **100(24)**, 14133–14138.
[30] Xi, S., Dyer, K.F., Kimak, M., Zhang, Q., Gooding, W.E., Chaillet, J.R., Chai, R.L., Ferrell, R.E., Zamboni, B., Hunt, J. and Grandis, J.R. (2006) Decreased STAT1 expression by promoter methylation in squamous cell carcinogenesis, *J Natl Cancer Ist*, **1:98(3)**, 181-189.
[31] Wu, C., Sun, M., Liu, L. and Zhou, G.W. (2003) The function of the protein tyrosine phosphatase SHP-1 in cancer, *Gene*, **13(306),** 1-12.

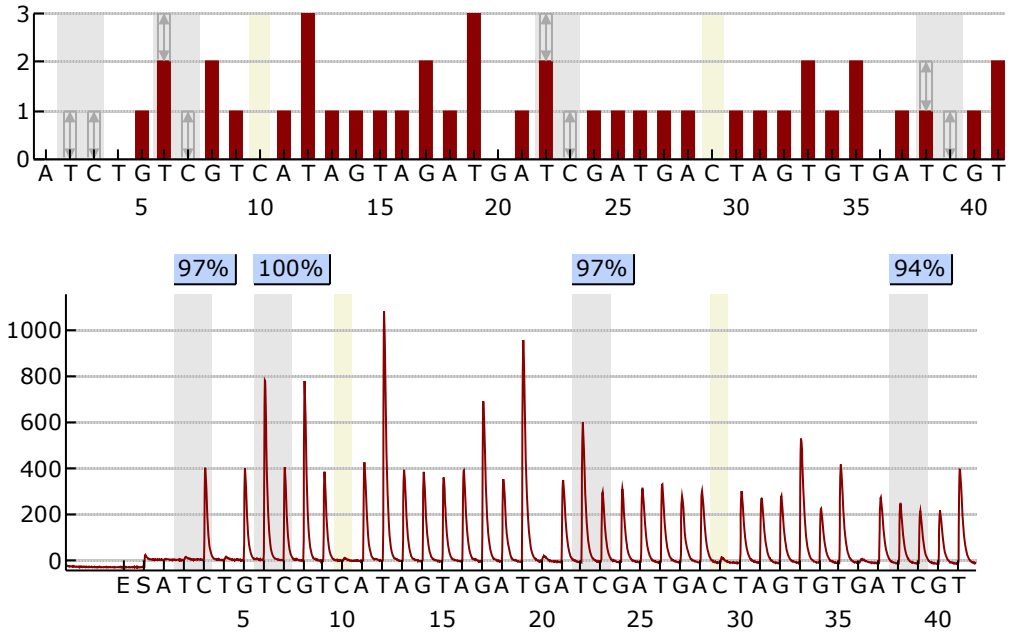
Supplementary information

Figures not included in the report.

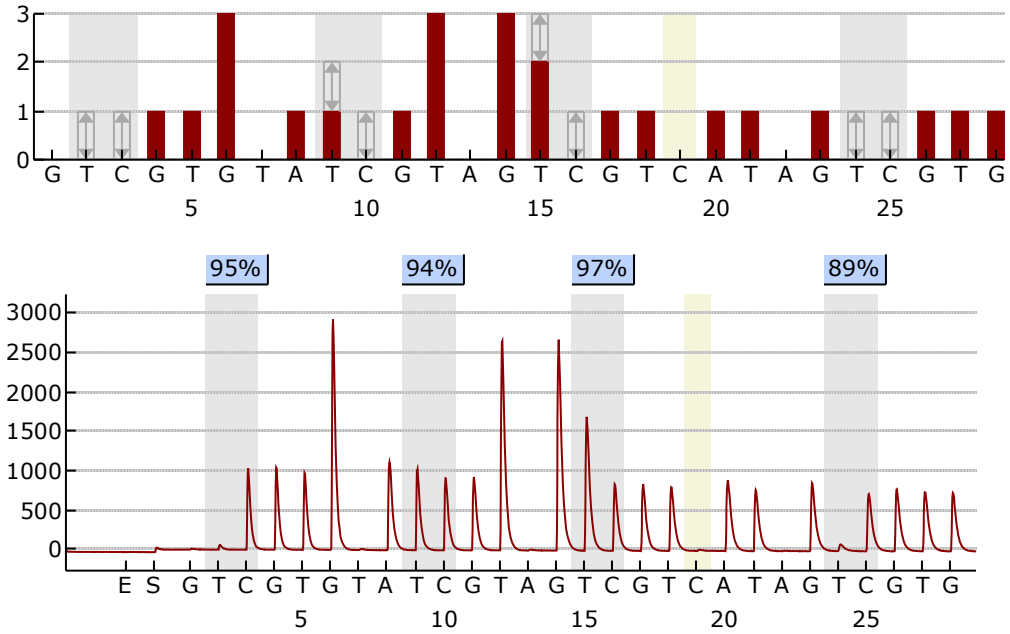
A



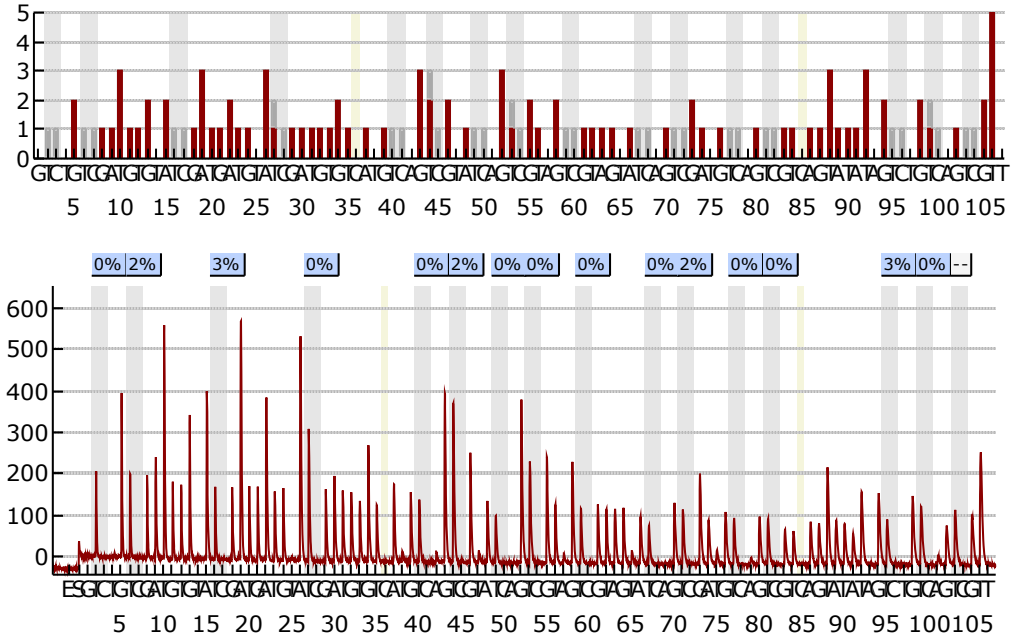
B



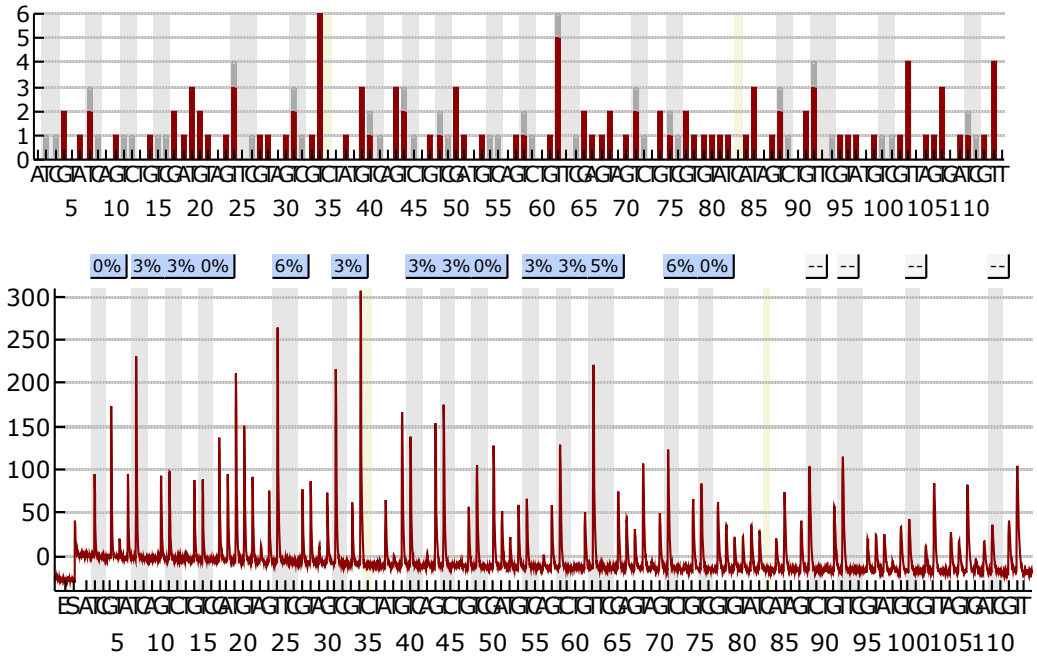
C


3
2
1
0
G T C G T G T A T C G T A G T C G T C A T A G T C G T G
5 10 15 20 25
95%
94%
97%
89%
3000
2500
2000
1500
1000
500
0
E S G T C G T G T A T C G T A G T C G T C A T A G T C G T G
5 10 15 20 25

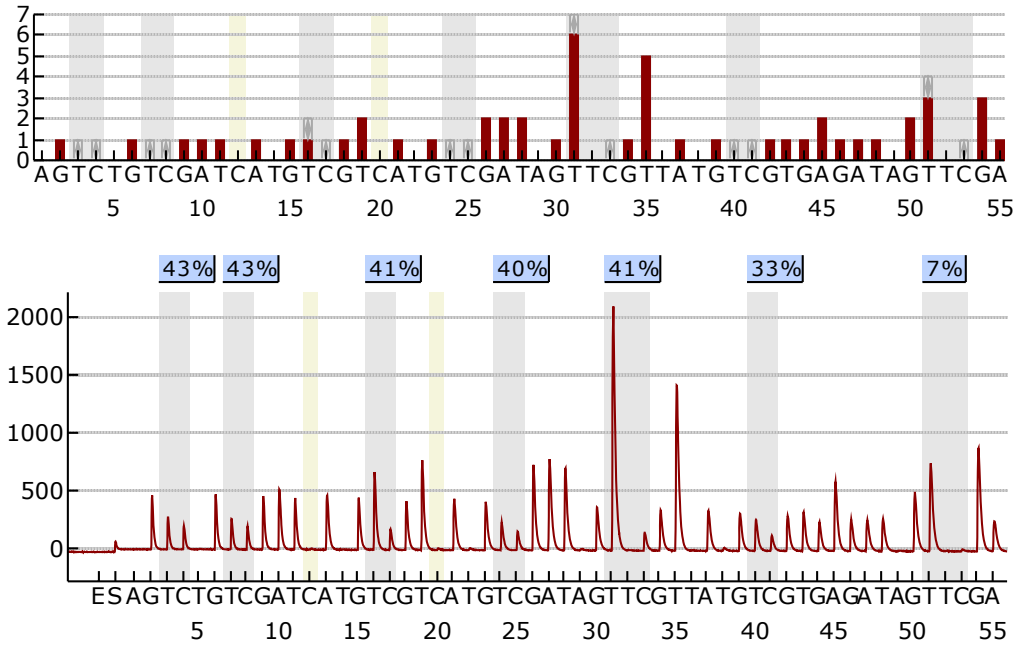
D


5
4
3
2
1
0
GTCTGTCGATGTGATCGATGATGATCGATGTGTCATGTCAGTCGTATCAGTCGTAGTCGTAGTATCAGTCGATGTCAGTCGTCAGTATATAGTCTGTCAGTCGTT
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105
0% 2%
3%
0%
0% 2%
0% 0%
0%
0% 2%
0% 0%
3% 0%
600
500
400
300
200
100
0
ESGTCTGTCGATGTGATCGATGATGATCGATGTGTCATGTCAGTCGTATCAGTCGTAGTCGTAGTATCAGTCGATGTCAGTCGTCAGTATATAGTCTGTCAGTCGTT
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100105

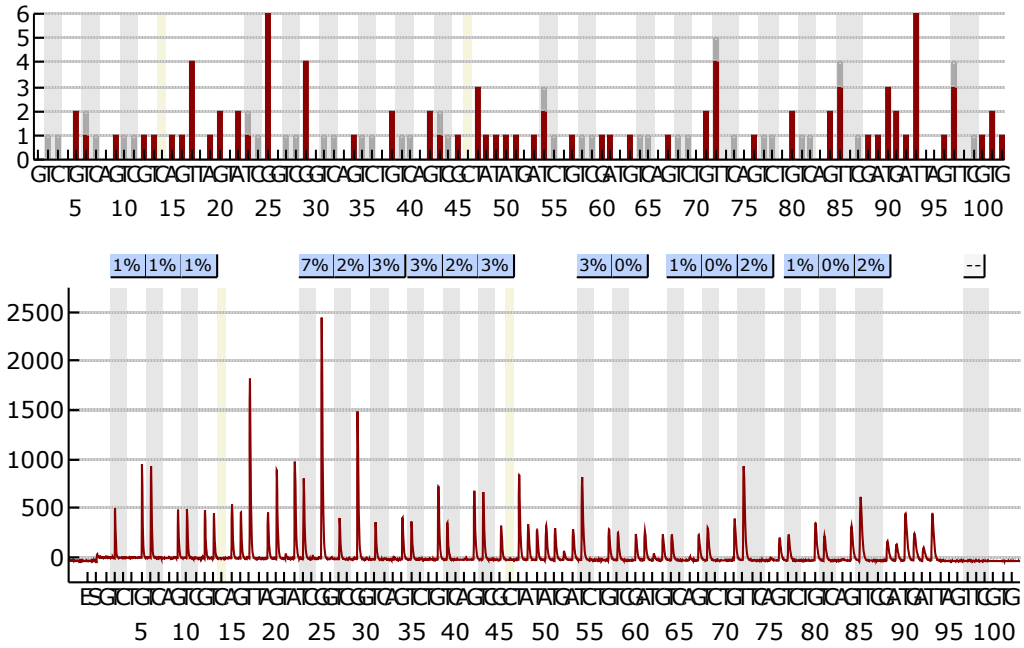
E



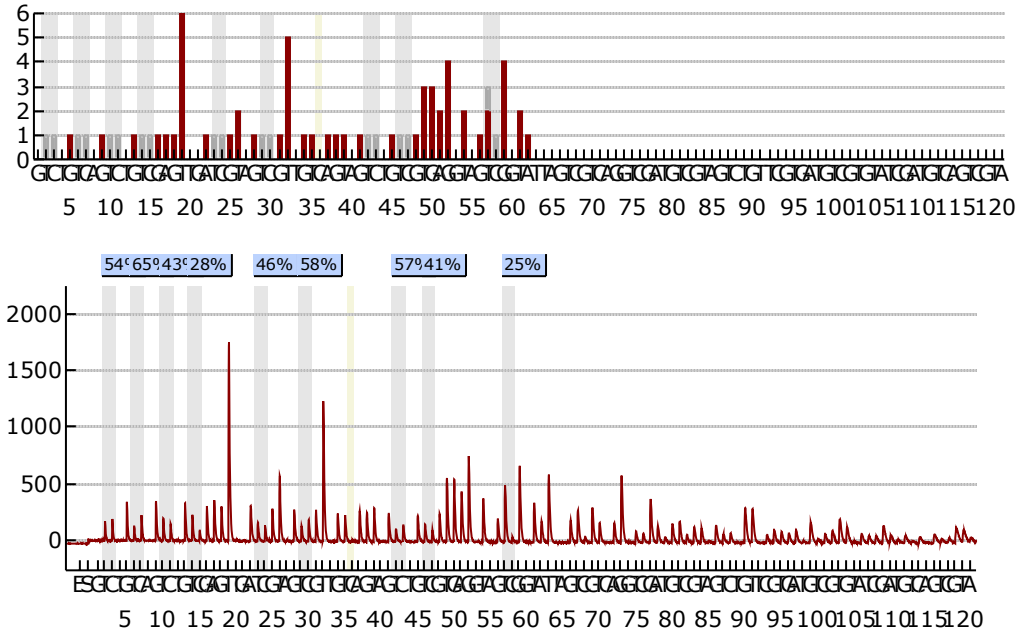
F



G



H



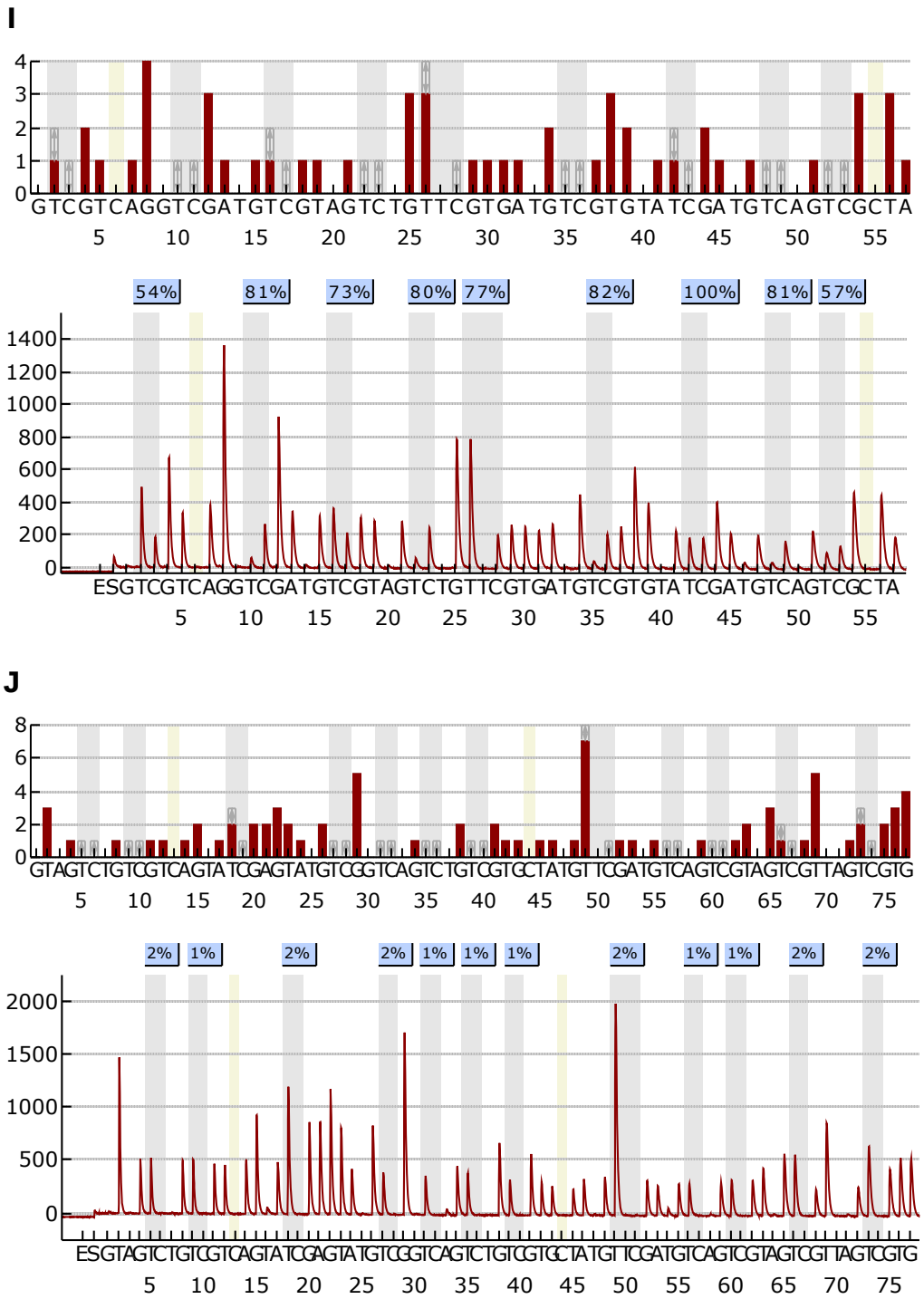
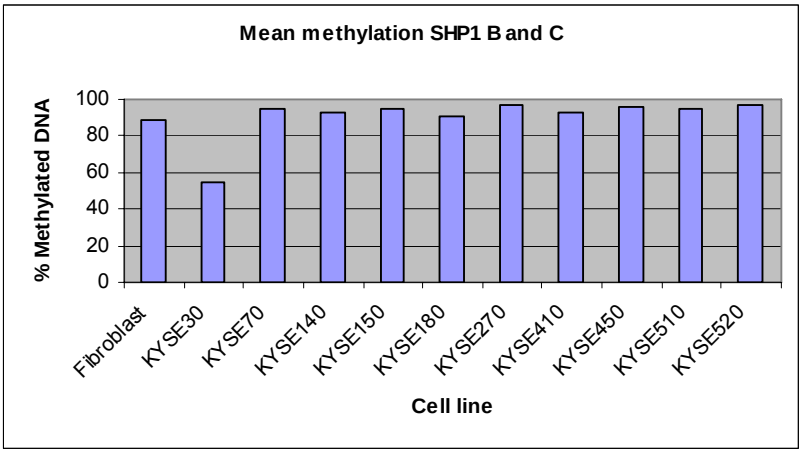


Fig S1. Theoretical outcome (upper panel) and Pyrogram (lower panel) for SHP1 A in H-157 (A), SHP1 B in KYSE150 (B), SHP1 C in 2020 (C), SHP2 in 1752 (D), SOCS1A in 2050 (E), SOCS1 B in KYSE30 (F), SOCS3 A in YA (G), SOCS3 B KYSE410 (H), SOCS3 C in KYSE140 (I) and STAT1 in 1810 (J).

A



B

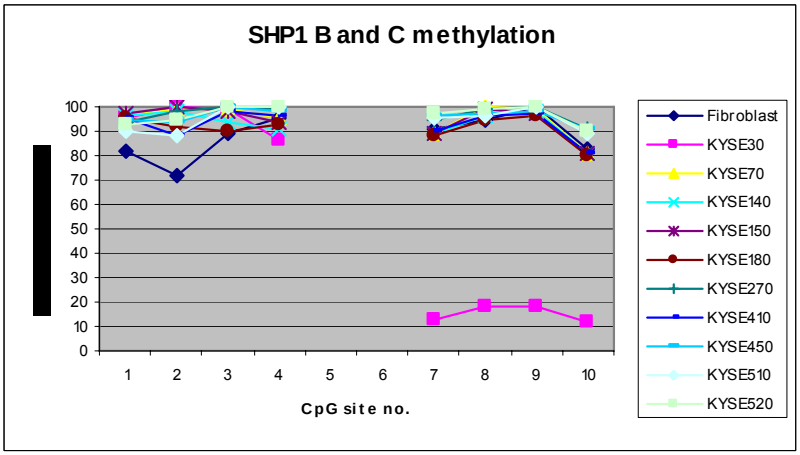
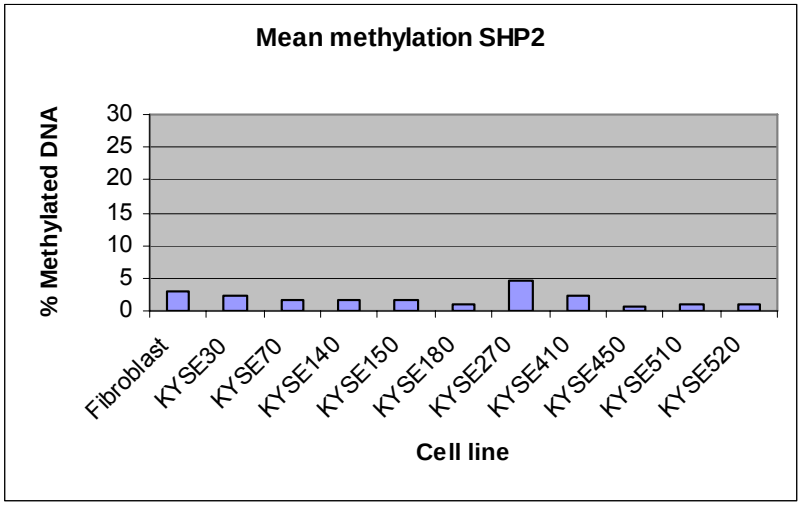


Fig S2. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 4 and 7 to 10 in the analyzed promoter region 2 of SHP1 (B), in ten EC cell lines.

A



B

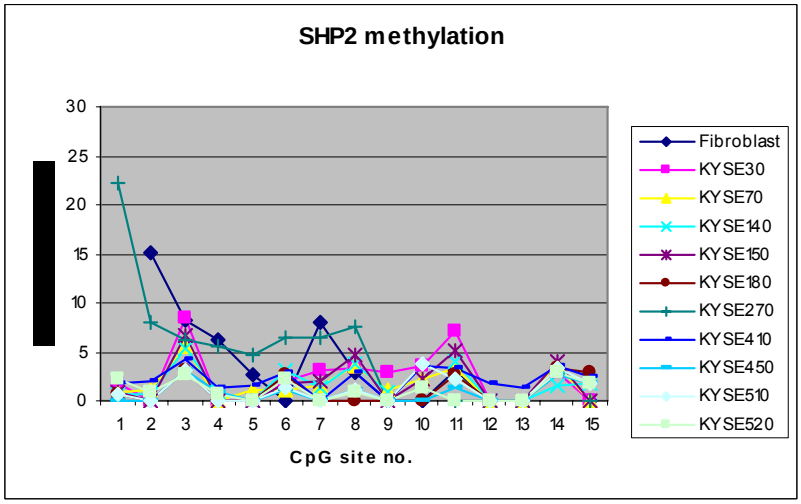
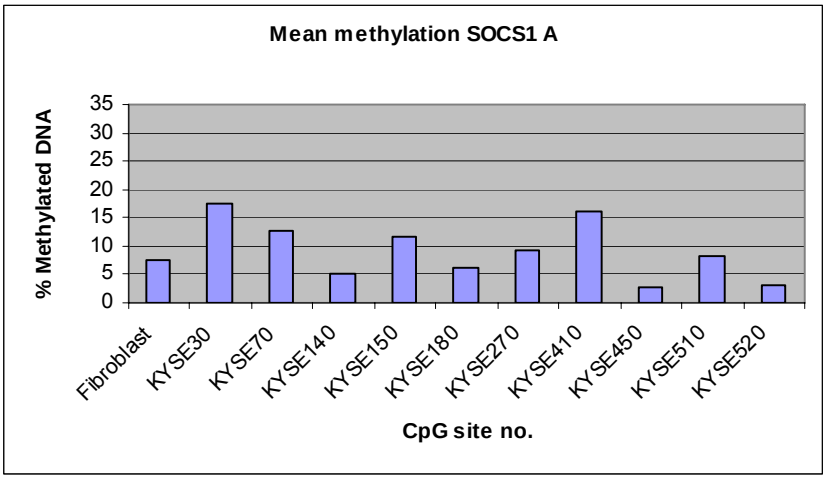


Fig S3. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 15 in the analyzed promoter region of SHP2, in ten EC cell lines.

A



B

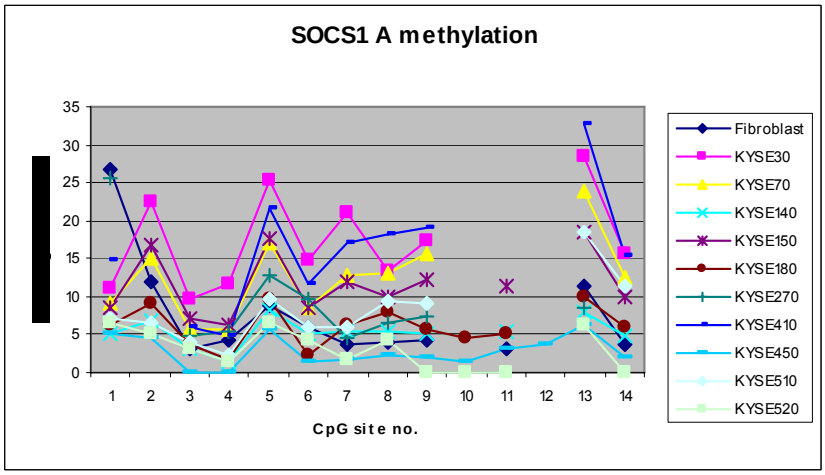
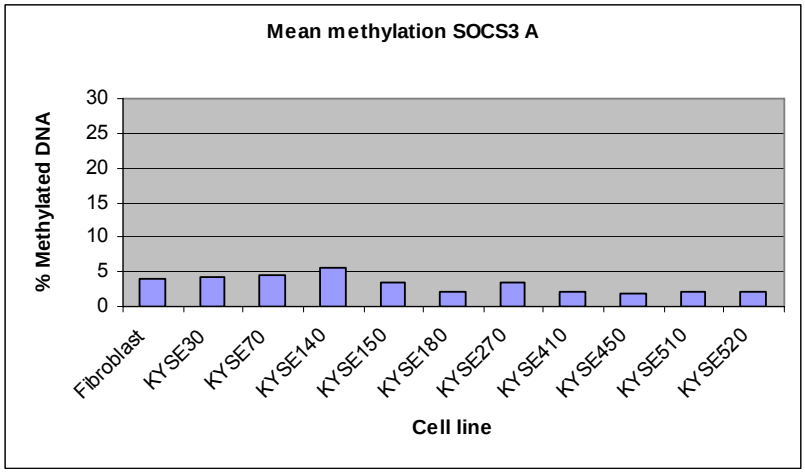


Fig S4. The mean degree of methylation (A) and the degree of methylation in the each CpG site (B) in the CpG sites 1 to 14 in the analyzed region of intron 1 in SOCS1, in ten EC cell lines.

A



B

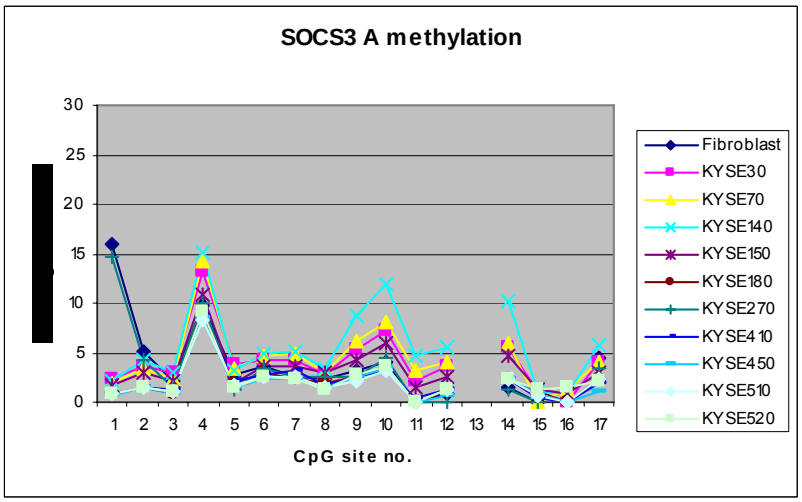
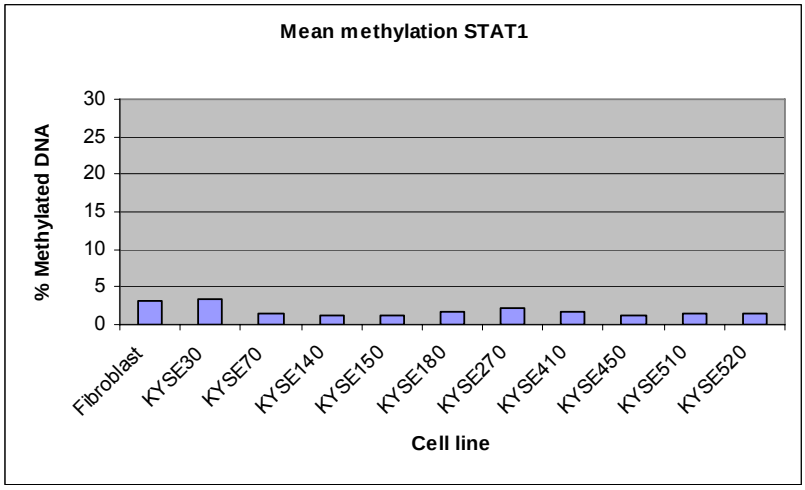


Fig S5. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 17 in the analyzed promoter region of SOCS3, in ten EC cell lines.

A



B

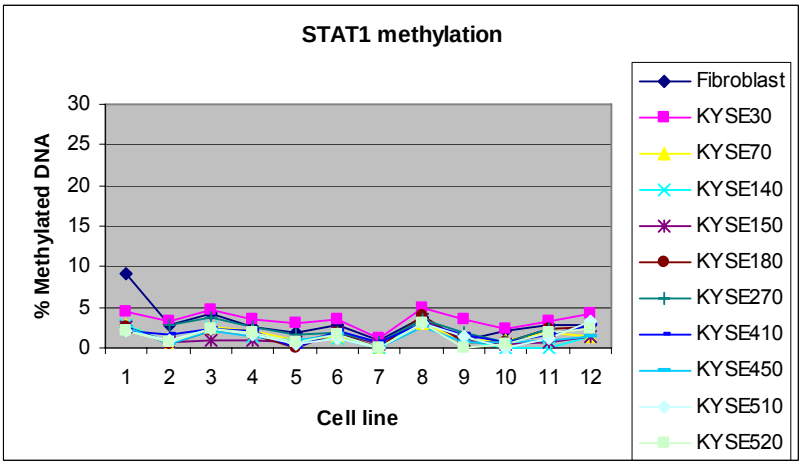
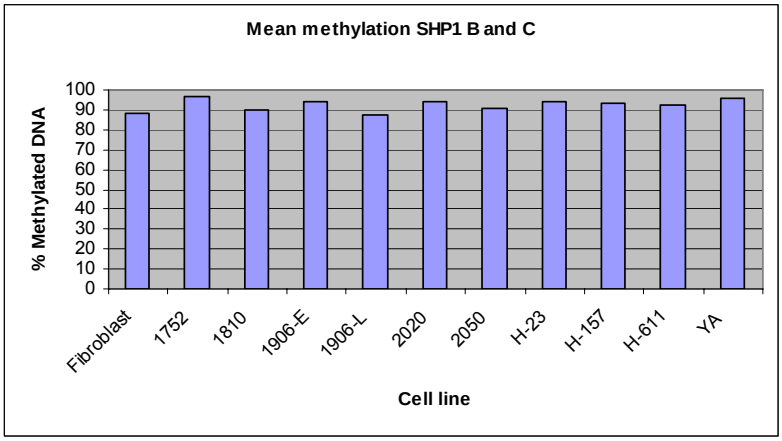


Fig S6. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 12 in the analyzed region of exon 1 in STAT1, in ten EC cell lines.

A



B

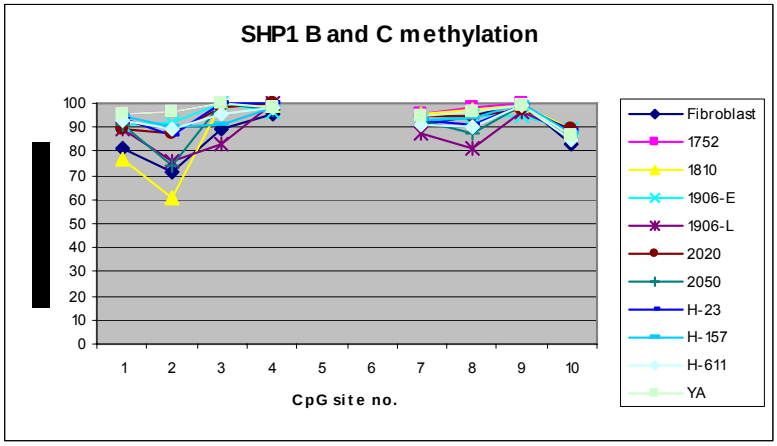
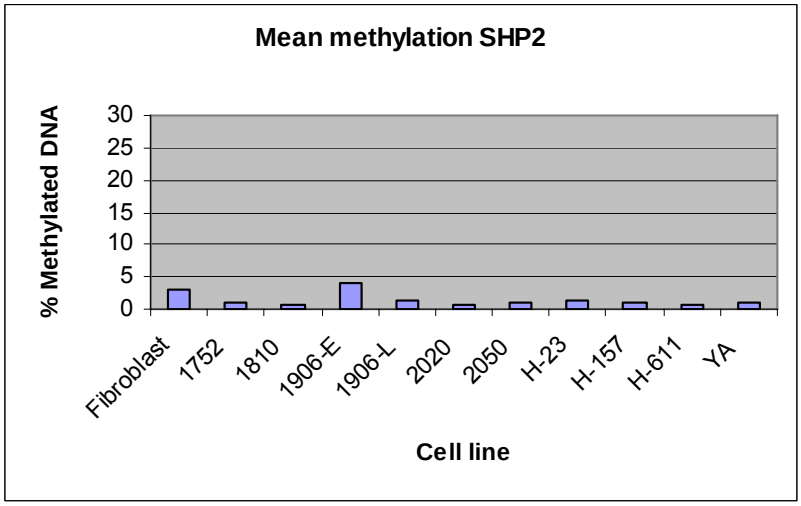


Fig S7. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 4 and 7 to 10 in the analyzed promoter region 2 of SHP1, in ten LC cell lines.

A



B

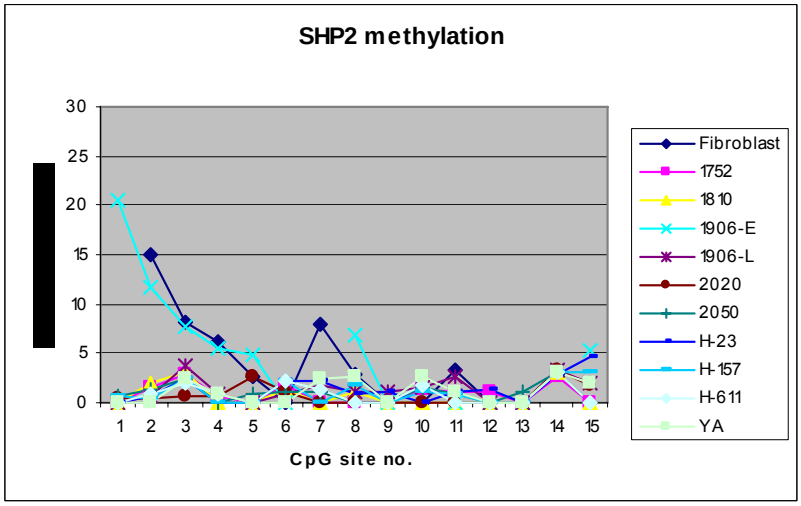
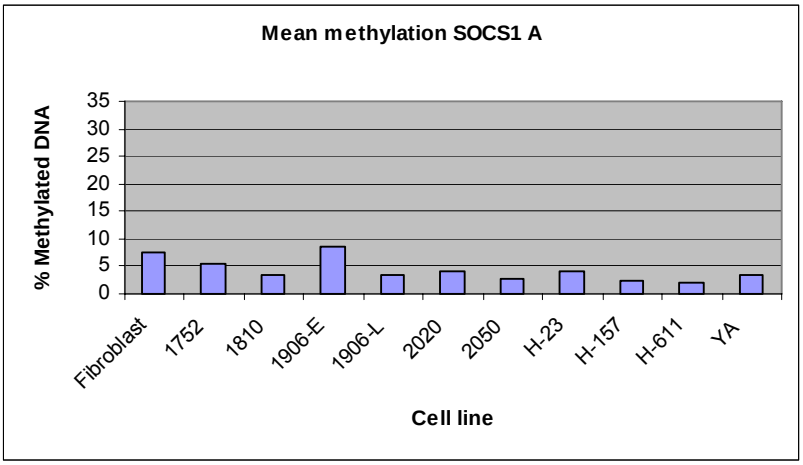


Fig S8. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 15 in the analyzed promoter region of SHP2, in ten LC cell lines

A



B

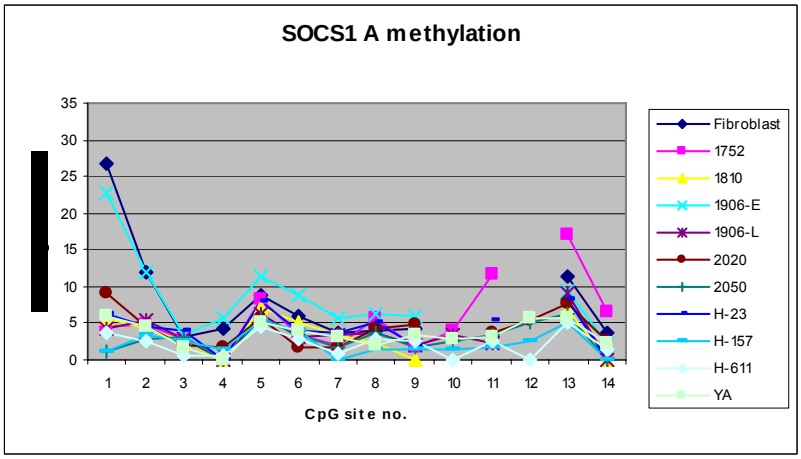
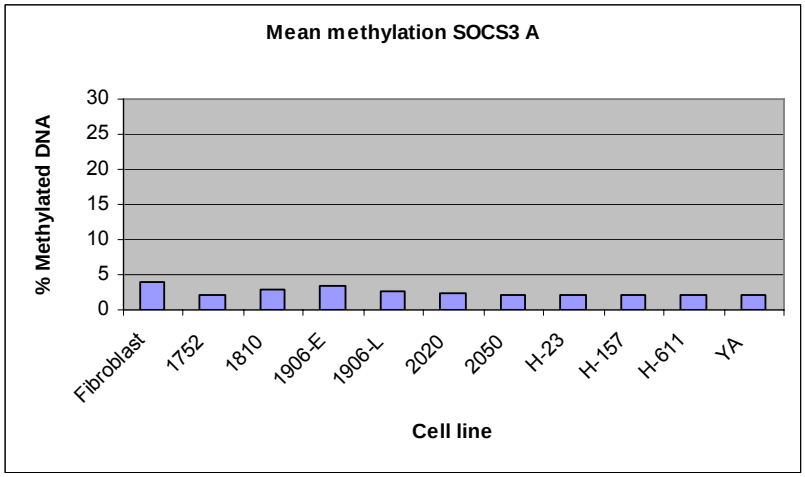
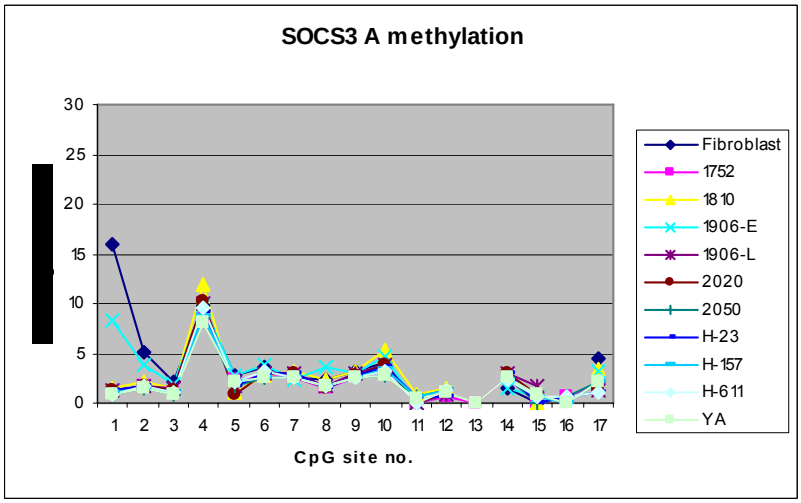


Fig S9. The mean degree of methylation (A) and the degree of meethylation in each CpG site (B) in the CpG sites 1 to 14 in the analyzed region of intron 1 in SOCS1, in ten LC cell lines.

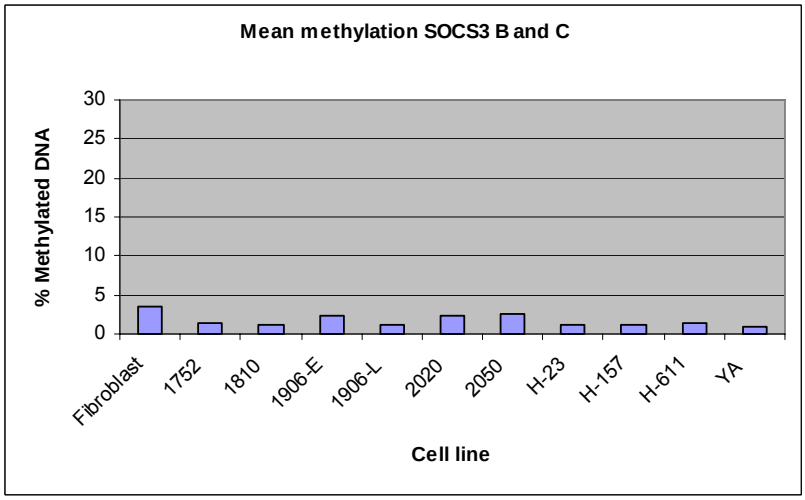
A


Mean methylation SOCS3 A
% Methylated DNA
30
25
20
15
10
5
0
Fibroblast
1752
1810
1906-E
1906-L
2020
2050
H-23
H-157
H-611
YA
Cell line

B


SOCS3 A methylation
30
25
20
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
CpG site no.
Fibroblast
1752
1810
1906-E
1906-L
2020
2050
H-23
H-157
H-611
YA

C


Mean methylation SOCS3 B and C
% Methylated DNA
30
25
20
15
10
5
0
Fibroblast
1752
1810
1906-E
1906-L
2020
2050
H-23
H-157
H-611
YA
Cell line

D

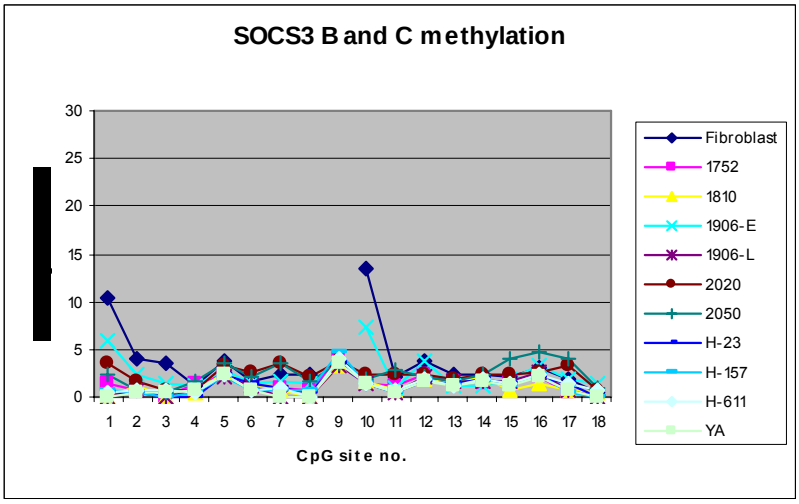
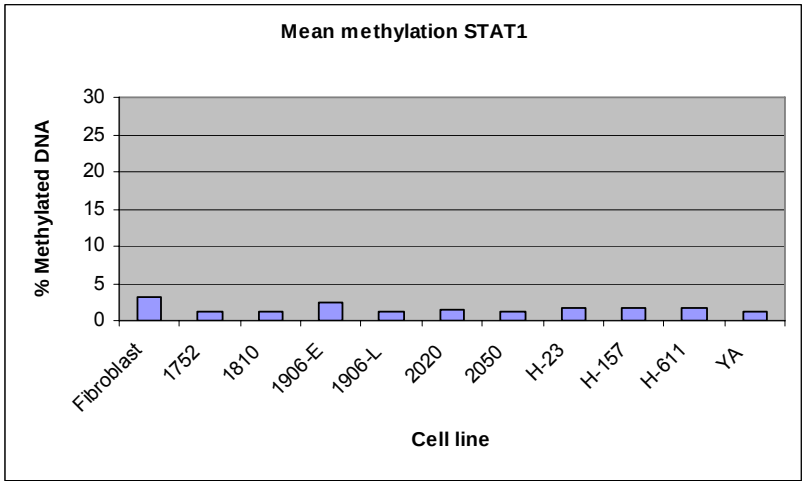


Fig S10. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 17 in the analyzed promoter region of SOCS3, in ten LC cell lines. The mean degree of methylation (C) and the degree of methylation in each CpG site (D) in the CpG sites 1 to 18 in the analyzed region of intron 1 in SOCS3, in ten LC cell lines.

A



B

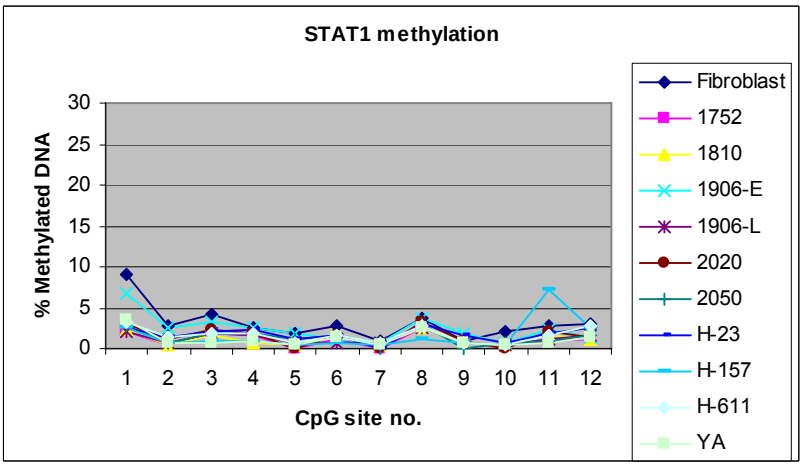


Fig S11. The mean degree of methylation (A) and the degree of methylation in each CpG site (B) in the CpG sites 1 to 12 in the analyzed region of exon 1 in STAT1, in ten LC cell lines.

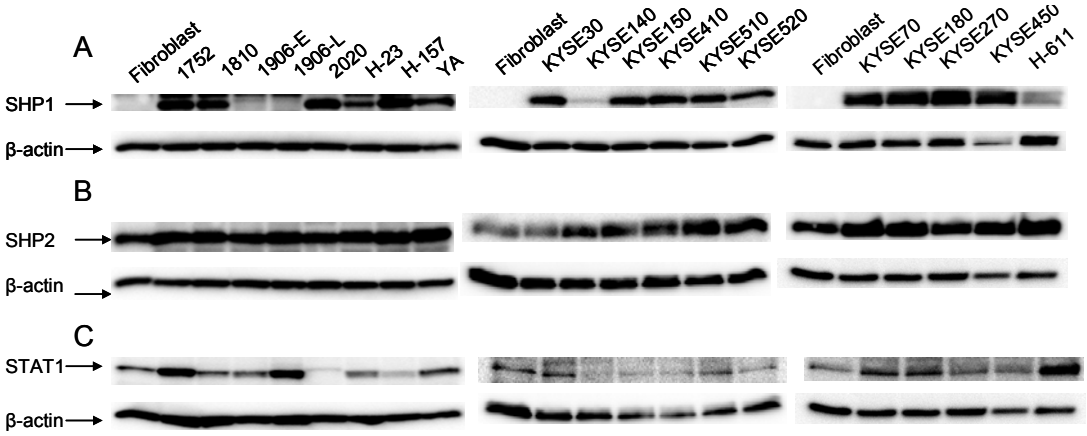


Fig S12. Western blot results for SHP1 (A), SHP2 (B) and STAT1 (C), with corresponding β-actin blots, in ten EC cell lines and nine LC cell lines.