Published online 12 November 2008 *Nucleic Acids Research, 2008, Vol. 36, No. 22* ***7207–7218***
doi:10.1093/nar/gkn897

Highly individual methylation patterns of alternative glucocorticoid receptor promoters suggest individualized epigenetic regulatory mechanisms

Jonathan D. Turner[1,2], Laetitia P. L. Pelascini[1], Joana A. Macedo[1,2] and Claude P. Muller[1,2,*]

[1]Institute of Immunology, Laboratoire National de Santé, 20A rue Auguste Lumière, L-1011, Luxembourg and [2]Department of Immunology, Graduate School of Psychobiology, University of Trier, D-54290, Germany

Received August 22, 2008; Revised October 22, 2008; Accepted October 23, 2008

ABSTRACT

The transcription start sites (TSS) and promoters of many genes are located in upstream CpG islands. Methylation within such islands is known for both imprinted and oncogenes, although poorly studied for other genes, especially those with complex CpG islands containing multiple first exons and promoters. The glucocorticoid receptor (*GR*) CpG island contains seven alternative first exons and their promoters. Here we show for the five *GR* promoters activated in PBMCs that methylation patterns are highly variable between individuals. The majority of positions were methylated at levels >25% in at least one donor affecting each promoter and TSS. We also examined the evolutionarily conserved transcription factor binding sites (TFBS) using an improved *in silico* phylogenetic footprinting technique. The majority of these contain methylatable CpG sites, suggesting that methylation may orchestrates alternative first exon usage, silencing and controlling tissue-specific expression. The heterogeneity observed may reflect epigenetic mechanisms of *GR* fine tuning, programmed by early life environment and events. With 78% of evolutionarily conserved alternative first exons falling into such complex CpG islands, their internal structure and epigenetic modifications are bound to be biologically important, and may be a common transcriptional control mechanism used throughout many phyla.

INTRODUCTION

Methylation of cytosine residues is a major epigenetic mechanism in mammalian cells, and the only known endogenous covalent DNA modification (1). Mammalian DNA methylation occurs on both DNA strands at the 5′ cytosine in a CpG pair. CpG dinucleotides are the least frequent nucleotide pairs and they are distributed in clusters, or CpG islands, throughout the genome. Using the classical CpG island definition (2), there are approximately 29 000 CpG islands (each greater than 200 bp long, containing >50% GC and with an observed/expected CpG ratio of >0.6) representing ~1% of the human genome. Overall 60–90% of genomic CpGs are thought to be methylated (3–5), and this is largely determined by their location within the genome. CpG pairs outside of CpG islands are usually methylated, those within such islands have long been thought to be protected against methylation. More recent studies have shown that CpG islands, in particular of tumor specific genes (e.g. *SPHK*1) can also be methylated (6).

DNA methylation plays an important role in gene regulation and differential gene silencing. In all organisms investigated, methylation patterns are maintained during mitosis. During meiosis parental genomes are demethylated and individual foetal methylation patterns are established during embryogenesis. Evidence from monozygotic twins suggests that the foetal environment tightly regulates methylation patterns: differences between monozygotic twins only develop after birth, increasing with age when global methylation levels fall (7). X-chromosome inactivation, and gene imprinting, both methylation dependent, occur at this early stage of development (8). DNA methylation studies have, largely, been limited to less than 100 genes known to be imprinted (9). However, DNA methylation is not limited to imprinted regions. More recently it has been suggested that DNA methylation outside these imprinted regions may underlie long-term programming effects, especially perinatally (10,11).

Methylation patterns established during embryogenesis vary significantly between tissues. Although a role for DNA methylation in tissue-specific gene expression was suggested already in 1975 (12,13), it was not until recently

*To whom correspondence should be addressed. Tel: +352 490 604 220; Fax: +352 490 686; Email: claude.muller@lns.etat.lu

that this has been experimentally confirmed, by Song *et al*. (14) who identified some 150 such regions. *In vivo*, embryonic stem cells differentiating for instance to neuronal precursor cells and astrocytes undergo gradual tissue-specific methylation, particularly in CpG islands. *In vitro*, these cells become hypermethylated after extended proliferation (15,16). Also, astroglial differentiation is associated with demethylation of JAK-STAT pathway genes (17,18).

CpG islands frequently contain transcription start sites (TSS) and 5′ promoter regions of housekeeping genes as well as tissue-specific genes. Although it is now recognised that methylation also occurs in CpG islands, the role of methylation of promoter elements in CpG islands, especially those associated with non-imprinted or non-oncogenes, has received little attention. Detailed examination of mRNA transcripts has recently shown that the vast majority of human genes show significant diversity in their 5′ region (19). Currently, it is estimated that more than 50% of genes possess alternative 5′ TSS each with its own promoter (20). On average genes have 3.1 promoters, but they can have as many as 10 alternative promoters and TSS (21). Evolutionarily conserved alternative promoters, identified by inter-species sequence homology, were predominantly (78%) located within CpG islands upstream of housekeeping or ubiquitously expressed genes. This raises the question as to the relations between CpG methylation and transcript variability within CpG islands and their roles, especially in the context of ubiquitously expressed, but tightly controlled genes. Although differential methylation of these sensitive regions may play a critical role in the regulation of many genes, there are very few studies where methylation patterns have been analysed within a single complex CpG island containing multiple first exons and alternative promoters.

The type II glucocorticoid receptor (*GR*, OMIM +138040; *NR3C1*) is an ubiquitously transcribed nuclear hormone receptor with an unusually complex promoter structure. *GR* transcription is controlled through nine promoters each associated with an alternative TSS, seven of which are found in an upstream CpG island (22,23). This transcript variability does not alter the protein produced since the ATG translation start codon is in exon 2. We and others have hypothesised that TSS usage, resulting in differential *GR* tissue expression profiles depends on short, immediately upstream proximal promoter regions within the CpG island. The *GR* and glucocorticoids (GC) are involved in many physiological processes, and play a major role in the hypothalamus-pituitary-adrenal (HPA) response to stress. This is a highly regulated system, controlled by GC feedback to the hippocampus and hypothalamus. This system is also known to be susceptible to differential methylation (24). Variability in HPA axis feedback sensitivity in the rat has been explained by changes in methylation of the transcription inducing nerve growth factor inducible-A (NGFI-A, also called KROX, EGR1, or ZIF286) binding site in the promoter of exon 1_7. As most other *GR* promoters, 1_7 is located in an upstream CpG island (25,26). We investigate here the positions and levels of CpG methylation of *GR* alternative exons and their promoters in order to understand the role and pattern of methylation in the CpG island upstream of a gene containing a complex promoter structure.

For this purpose, we have analysed differential methylation of alternative *GR* promoters in peripheral blood mononuclear cells (PBMC) by bisulphate sequencing. We have previously shown that PBMC express almost all alternative first exons (e.g. 1B, C, D, E, F and H) (22). We also examine the evolutionarily conserved TFBS pattern using an improved *in silico* phylogenetic footprinting technique. The majority of CpG positions in conserved TFBS were methylatable, and methylation patterns were highly individualised among donors. The heterogeneity of methylation patterns observed in our study may reflect differences in early life environment and events.

MATERIALS AND METHODS

Genomic alignments

Genomic DNA of the human chromosome 5 (July 2004, UCSC) was aligned with rat (AJ271870, NCBI), mouse (NCBIM35:18:39863309:39869358:1, Ensembl), chimpanzee (CHIMP1A:5:149403812:149409861:1, Ensembl) and cow (ChrUn.313:334239:339538:1, Ensembl) sequences upstream of the *GR* exon 2 using Vector NTi 10.1.1 (Invitrogen, Paisley, UK) as previously reported (27).

In silico phylogenetic footprinting (*IS*PF)

Genomic sequences predicted by homology to contain the promoter regions of exons 1D, 1E, 1F and 1H were screened for potential transcription factor binding sites (TFBS) using the MATCH algorithm (28) from TRANSFAC® Professional (version 10.1). The cut-off values for the core and matrix similarity scores of acceptable binding site predictions were set to 0.98 and 0.95, respectively (1.0 = exact match). TRANSFAC-predicted binding sites were transposed onto the multi-species genomic alignment to identify coincident sites. The results were expressed as the number of species where the first nucleotide of identical transcription factor matrices coincide for each position within the alignment.

Subjects

Twenty-six healthy subjects (22 female and 4 male, age range 35–67 years, mean 51 ± 8 years) were recruited by the Department of Psychobiology, University of Trier, Germany. The study protocol was approved by the ethics committee of the Rheinland-Pfalz State Medical Association and written informed consent was given by all participating subjects.

Exclusion criteria were based on medical history and physical examination. They were designed to minimize the influence of disease and drugs, and included: steroid use (except for oral contraception or hormone replacement therapy due to menopause); infections during the preceding 2 weeks; dietary weight loss of 5 kg or more within 6 weeks before study entry; pregnancy or breast-feeding; alcohol or drug dependence; severe allergies; hematological, endocrine, cardiovascular, pulmonary,

gastrointestinal, renal, hepatic, autoimmune or psychiatric disorders. The subjects were also evaluated with the German version of the Structured Clinical Interview for DSM-IV (29) to exclude major depression disorder. From each subject, one blood sample was collected between 08:30 and 09:00 hours under basal conditions.

Isolation of genomic DNA from PBMCs

Human PBMCs were purified from heparinised blood by Ficoll-Isopaque (Amersham, UK) using Leucosep tubes (Greiner Bio-one, Germany). After three washes in HBSS (Biowhittaker, Verviers, Belgium), cells were resuspended in RPMI, 10% FBS, 10% DMSO and stored in liquid nitrogen until further analysis. For DNA isolation, cells were washed and resuspended in 2 ml of PBS. DNA was purified using the QIAamp® DNA Blood Midi kit (Qiagen, Venlo, The Netherlands) and stored at –20°C.

Bisulphite treatment

The EpiTect Bisulfite kit (Qiagen) converts unmethylated cytosine residues to uracil, whereas methylated cytosines remain unmodified. According to the manufacturers' protocol, the first two steps of bisulphite treatment were performed on 400 ng of genomic DNA. The thermal cycler (Mastercycler gradient, Eppendorf, Hamburg, Germany) program comprised several incubation steps for DNA denaturation, sulphonation of unmethylated cytosine, and deamination of cytosine sulphonate to uracil sulphonate. Cycling conditions were as follows: first denaturation for 5 min at 99°C, first incubation for 25 min at 60°C, second denaturation for 5 min at 99°C, second incubation for 85 min at 60°C, third denaturation for 5 min at 99°C, third incubation for 175 min at 60°C and hold at 20°C. Alkaline uracil desulphonation and sample purification were performed using a spin column. Modified DNA was eluted and stored at –20°C until analysis.

PCRs and methylation quantification

The bisulphite-modified DNA was used to amplify promoters of exons 1D, 1E, 1F and 1H with the primers and under the conditions shown in Table 1.

PCR was performed using 20 mM Tris–HCl (pH 8.4), 50 mM KCl, 200 mM deoxynucleoside triphosphates, 1× SyBR Green and 2.5 U Platinum *Taq* DNA polymerase (Invitrogen, Merelbeke, Belgium) on an Opticon 2 thermal cycler (MJ Research). Cycling conditions were as follows: 95°C for 2 min, 40 cycles, at 95°C for 20 s, annealing temperature (Table 1) for 20 s and 72°C for 25 s. Products of D_B, E, F_A, F_B, F_C, H_A, H_B, H_C and H_D PCRs were used as a template (20-fold dilution) for nested amplifications using the primers and conditions in Table 1.

Methylation analysis of promoter 1D, 1E, 1H and the 215 first nucleotides of promoter 1F were determined manually. Such peak intensity-based quantification has previously been validated for methylation quantification (30). The nested PCR products (D_{BII}, E_{II}, F_{AII}, F_{BII}, F_{CII}, H_{AII}, H_{BII}, H_{CII}, H_{DII} and H_{DIII}) and the D_A PCR products were directly sequenced after spin column purification (Genomed, Löhne, Germany).

Methylation analysis of the promoter 1F proximal area (F_{CII}) was determined by multiple trace single population quantification. The F_{CII} semi-nested PCR products were cloned into the pCR4-TOPO vector (Invitrogen) and used to transform Top10 electrocompetent *E. coli* (Invitrogen). For each subject, 10–12 clones were sequenced using 100 nM M13 primers and the BigDye 3.1 Terminator cycle sequencing reagent (Applied Biosystems, Nieuwerkerk, The Netherlands) on the ABI 3130 sequencer (Applied Biosystems). Methylation percentage at each position is calculated from the ratio of the numbers of methylated to unmethylated colonies.

RESULTS

Identification of known TFBS by *IS*PF

*IS*PF was performed over all genomic regions containing *GR* promoters with experimentally identified transcription factors (Table 2). The genomic DNA of the complete CpG islands of the five species, the mouse, rat, cow, chimpanzee and man were aligned. The coincidental predictions of the TRANSFAC analyses and genomic alignment identified the evolutionarily conserved transcription factors. Transcription factors have been experimentally identified for promoters 1B (31), 1_7 [1F homologue, (25)] and probably 1D (32).

*IS*PF analysis identifies two of the three known Yin Yang-1 (YY1) sites in promoter 1D, at position -4819 and -4603. The third known YY1 site (-4642 from ATG) was missed because of a discrepancy between the aligned genomic sequences and both the complete and core TRANSFAC matrices. The observed core was ATGGT rather than the expected ATGGC. Reducing the stringency of the TRANSFAC search correctly identifies this YY1 binding site, although with a much increased risk of false positive predictions (data not shown). Similarly, *IS*PF analysis of promoter 1F identifies the NGFI-A-binding site as being conserved between all species, after its initial identification in the homologous rat promoter 1_7. *IS*PF analysis of the constitutive promoter 1B, successfully predicts the three known SP1 sites (31).

Epigenetic sequence coverage

We previously showed that exons 1B–1H are differentially expressed in immune cells (22). Most of the five promoter regions were successfully bisulphite-sequenced in PBMCs. Multiple amplifications (nine nested and one simple reaction, Figure 1) were required to cover 78% of the relevant promoters (1274 nucleotides) representing 40% of the total CpG island of 3.2 kbp.

For some promoters it was not possible to get a complete sequence because of difficulties in amplifying long stretches of identical nucleotides after bisulphite modification. However, all conserved TFBS within the promoters of interest except YY1 (-4814, 1D) were completely sequenced.

Table 1. PCR primers and reaction conditions for bisulphite sequencing

	Primer sequence[a]	Amplified region[b]	Length (bp)	T_m (°C)[c]	[$MgCl_2$] (mM)	[Primers] (μM)
Promoter 1D						
D_A	Fwd: 5′-ATATTAGATAATGTATAGGGAATYGTTTAT-3′ Rev: 5′-CCRCCRCCATCTTAATAAAATACAATATC-3′	-4724 to -4584	140	49	2	0.5
D_B	Fwd: 5′-TGTTAAGATGGTGGTYGYGGGGAYGG-3′ Rev: 5′-AACTACCRCAACTCCACCTAATCC-3′	-4644 to -4438	206	59	2	0.1
D_{BII}	Fwd: 5′-GGTYGYGGGGAYGGGTTGGYGATATTGT-3′ Rev: 5′-CCTACTCRAACRCTCRACCACAACC-3′	-4632 to -4460	172	59	1	0.1
Promoter 1E						
E	Fwd: 5′-GGGAGTTGAAYGTTGGTATTTTAAAGTTG-3′ Rev: 5′-CTCRAAAAAAATTACACRCCAAATAC-3′	-4129 to -3781	348	54	2	0.1
E_{II}	Fwd: 5′-GTATTTTGTTTTATTTGTAGGGGTAGG-3′ Rev: 5′-TACAAAACCTCCAACRAACTAAATT-3′	-4097 to -3804	293	55	1	0.1
Promoter 1F						
F_A	Fwd: 5′-TGAAGATTYGGTYGTTTAGATGAT-3′ Rev: 5′-ACCRAATTACRTAAAATATATCACTTCRAA-3′	-3605 to -3378	227	50	2	1
F_{AII}	Fwd: 5′-TGGTGGGGGATTTGTYGGTAYGYGA-3′ Rev: 5′-TCACTTCRAAAAAAACTACRAAATTACA-3′	-3577 to -3398	179	52	2	0.5
F_B	Fwd: 5′-GGTYGAGAYGTTGYGGTATYGTTTTYGTG-3′ Rev: 5′-CCTTAACRACAAACRCCRCCAATAC-3′	-3452 to -3268	184	50	4	1
F_{BII}	Fwd: 5′-GTTGYGGTATYGTTTTYGTGTAATTT-3′ Rev: 5′-ACRAATAACAACRAACRAACCACAA-3′	-3443 to -3297	146	53	2	0.1
F_C	Fwd: 5′-TTGTGGTTYGTTYGTTGTTATTYGTAGG-3′ Rev: 5′-CACCRAATTTCTCCAATTTCTTTTCTC-3′	-3321 to -3177	144	54	2	1
F_{CII}	Fwd: 5′-TTGTGGTTYGTTYGTTGTTATTYGTAGG-3′ Rev: 5′-CAATTTCTTTTCTCRCTACCTCCTTCC-3′	-3321 to -3190	131	52	2	0.5
Promoter 1H						
H_A	Fwd: 5′-TTYGGTTGYGGYGGGAATTGYGGAYGGTG-3′ Rev: 5′-AAACTAATAAAAATTTATAAACTCC-3′	-2422 to -2258	164	56	2	0.5
H_{AII}	Fwd: 5′-GGTGGYGGGYGAGYGGTTTTTTTGTTAGAG-3′ Rev: 5′-ACTCCCRCRACRACCCCCRAATTATCTC-3′	-2397 to -2276	121	58	1	0.1
H_B	Fwd: 5′-GGGYGYGTTYGTTTTTTYGAGGTGTYGTTG-3′ Rev: 5′-CTCCCCCTCRACCCRACCAAA-3′	-2341 to -2135	206	55	1	0.1
H_{BII}	Fwd: 5′-GAGATAATTYGGGGGTYGTYGYGGGAG-3′ Rev: 5′-ACCCRACCAAAAAACRCCTAC-3′	-2305 to -2145	160	56	1	0.1
H_C	Fwd: 5′-TTTYGTAGGYGTTTTTTGGTYGGGTYGAG-3′ Rev: 5′-AATTCAAACRCRACTTAACRTTCACCACRAA-3′	-2169 to -1967	202	51	2	1
H_{CII}	Fwd: 5′-TTTYGTAGGYGTTTTTTGGTYGGGTYGAG-3′ Rev: 5′-CCAAAATTCCCRCRAAAAATAAAAAACTC-3′	-2169 to -2055	114	55	3	0.1
H_D	Fwd: 5′-GGTYGTTYGATATTYGTTTTYGTGGTG-3′ Rev: 5′-CCCRCTTATACACCCTCAC-3′	-2015 to -1844	171	52	1	0.1
H_{DII}	Fwd: 5′-GTGGTGAAYGTTAAGTYGYGTTTG-3′ Rev: 5′-CCRCACRCCCTCCTCAAACCA-3′	-1994 to -1868	126	53	2	0.1
H_{DIII}	Fwd: 5′-GGTYGTTYGATATTYGTTTTYGTGGTG-3′ Rev: 5′-CCRCACRCCCTCCTCAAACCA-3′	-2015 to -1868	147	52	1	1

[a]Fwd, forward or sense primer; Rev, reverse or antisense primer. Degenerate nucleotides following IUPAC codes.
[b]Locations given with respect to the ATG start codon.
[c]T_m, annealing temperature in PCR.

Table 2. Regulatory element of the GR promoter region can be predicted by *in silico* phylogenetic footprinting

Promoter	Transcription factor	Identification technique	*IS*PF	Prediction
1B	SP1 (x3)	RG, FP, E	5,4,5	✓
1B–D	YY-1 (x3)[a]	FP, D, E	5[b],4,4	✓
1D	n/d			
1E	n/d			
1F	NGFI-A	ChIP	5	✓
1G	n/d			
1H	n/d			

n/d: Not experimentally determined.
[a]YY-1-binding site initially identified as part of promoter 1B, subsequent identification of alternate first exons would suggest that it is part of promoter 1D.
[b]This YY-1-binding site is outside the CpG island and not considered further.

Promoter 1D

YY1-binding sites at -4814/9 and -4603/8 include dinucleotides CpG[1], CpG[16] and CpG[17] (Figure 2). The CpG dinucleotide in the core binding site (ggCGccatctt) is known to be functionally sensitive to cytosine methylation (33). Among 26 donors none were methylated in CpG[1], while two and one donor, respectively had <25% methylation at CpG[16] or CpG[17]. Similarly, *IS*PF analysis predicts within promoter 1D, 90–94 bp before the transcription start site a core TTCCG matrix (from either ELK-1 or c-Ets-1, both sharing the same core matrix). The predicted complete binding site for these two factors covers three CpG di-nucleotides (CpG[18-20], Figure 2). CpG[19] forms part of the core matrix of ELK-1 and c-Ets-1. Between 25% and 50% methylation

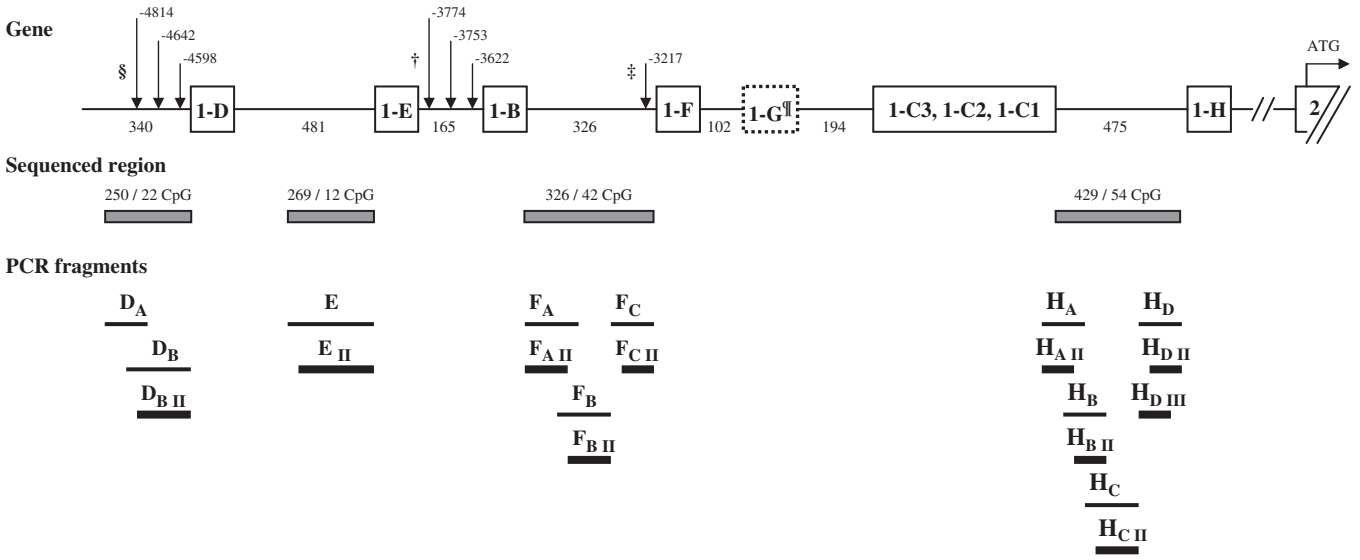


Figure 1. Schematic representation of the Glucocorticoid receptor CpG island internal structure, the bisulphite sequencing strategy and primer locations. Known TFBS are shown as arrows and numbered from the ATG translation start codon in exon 2. § from (32), † from (53), ‡ homologous to the known rat site (25,27), ¶ the ephemeric exon 1G has only been detected in the rat as its homologue 1_8.

of $CpG^{18\text{-}20}$ was observed in 4, 15 and 4 donors, respectively.

Promoters 1J and 1E

Bisulphite sequences of 269 bp covering the complete promoter 1E, exon 1J and part of promoter 1J were obtained (Figure 1). Of the 28 potential CpG methylation sites, 12 were successfully sequenced: four in promoter 1J, four in exon 1J, and four in the promoter region immediately upstream of exon 1E (Figure 3). Each of these sites (except CpG^{27}) was methylated in one to three of the 26 donors. Interestingly, the *in silico* analysis identified nine TFBS that are evolutionarily highly conserved and present in at least four of the five species. However, only two of these nine sites contain CpG di-nucleotides: Pax3 (-4067) in promoter 1J (CpG^{16}), and Sp3 (-4012) in promoter 1E (CpG^{19}). Bisulphite sequencing of these nucleotides showed that whilst methylation was possible it was nevertheless rare (one and three donors, respectively at levels below 50%).

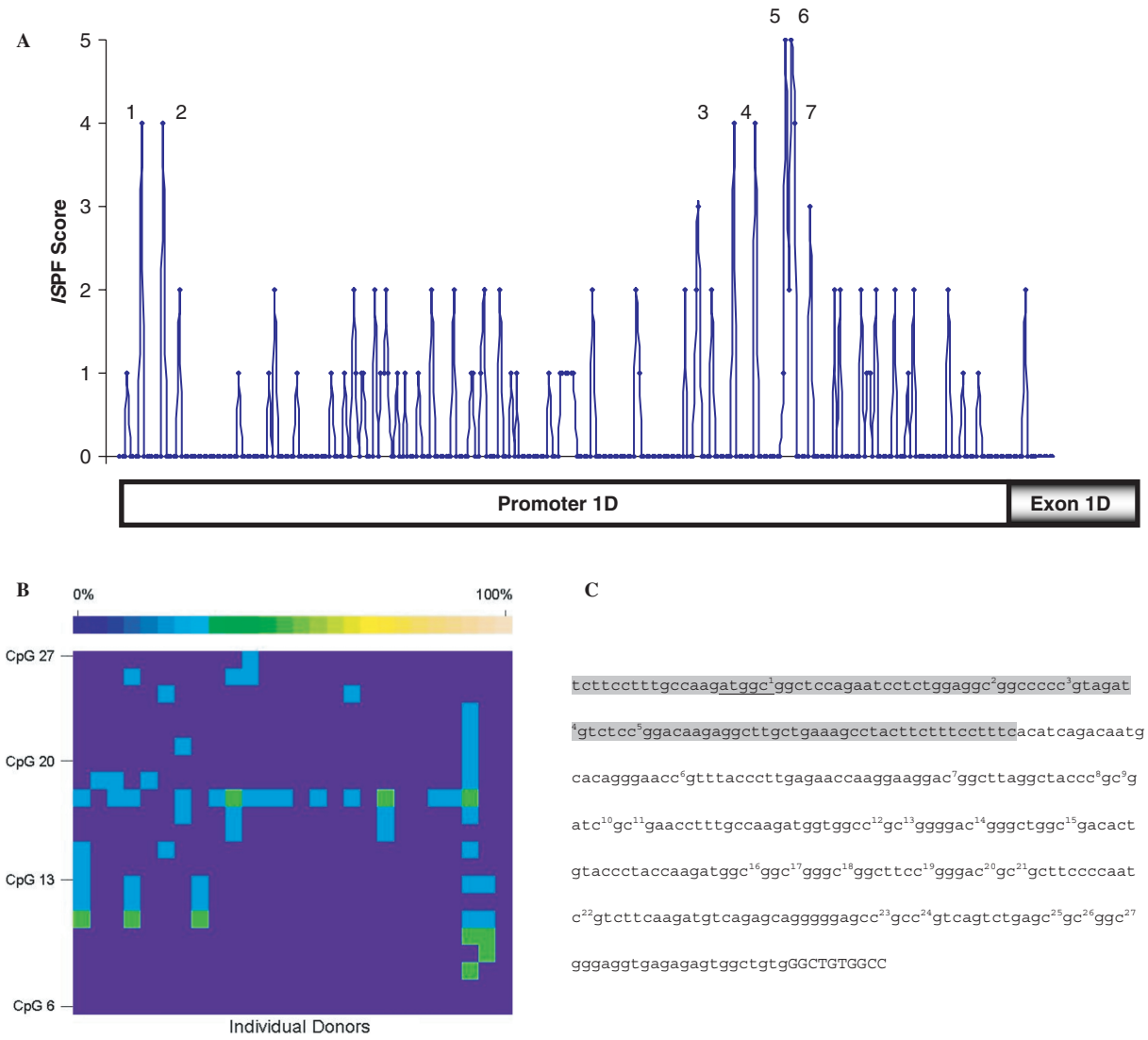
The bisulphite sequenced region included exon 1J, and the first 24 nucleotides of exon 1E. Exon 1J contains four CpG dinucleotides ($CpG^{20–23}$, Figure 3), whilst exon 1E contains only one (CpG^{28}). Within these five exonic CpG dinucleotides, the methylation levels were similar to those in the neighbouring positions.

The methylation of CpG^{27} (six donors, >25% methylation) was much greater than of the other positions within promoter 1E, and this position is the closest to the exon 1E TSS. TRANSFAC analysis of this position identifies an NGFI-A-binding site in the human sequence, but this is not an evolutionarily conserved position beyond human and chimpanzee. However, both the rat and mouse share a CACCCCCTCC sequence, differing by only a single nucleotide from the known active KROX response element (ctccccctc) (34). Thus the NGFI-A response element may be more conserved than predicted by *IS*PF.

Promoter 1F

*IS*PF predicted for promoter 1F several high quality potential TFBS, including the NGFI-A-binding site reported before (25,35). Since methylation of the corresponding NGFI-A-binding site has been well documented in the rat, we tried to quantify the level of methylation by cloning. The F_{CII} PCR products were cloned and 10–12 colonies sequenced. The human NGFI-A core recognition motif, 5′-GCGGGGGCG-3′, includes two CpG sites ($CpG^{41–42}$, Figure 4). Both of these positions were unmethylated in the majority of donors. In only 4 of the 26 donors less than 10% of the colonies screened were methylated at CpG^{41} (5′-end of the NGFI-A site). A similar level of methylation of CpG^{42} was found for only two donors (Figure 4). These results show that methylation of the NGFI-A site is possible in the human, but is neither uniform nor frequent. The level of methylation observed by cloning was in agreement with the low but detectable levels seen in the sequencing electropherogram.

The sequence covering $CpG^{31–33}$ should be interpreted with caution as this region corresponds to the primer binding site of the F_{CII} forward primer. However, this region covers only one predicted human TFBS, HES1, with an *IS*PF value of only 2. Of all the promoters investigated, 1F contains the highest number of confident *IS*PF predictions (10 with *IS*PF > 4), and has the highest percentage (60%) of methylatable CpG dinucleotides. Of the 42 CpG dinucleotides in promoter 1F, *IS*PF predictions cover only $CpG^{14–15}$ and $CpG^{41–42}$. There was no visible methylation, however, for any of the 26 donors at either positions CpG^{14} or CpG^{15}.



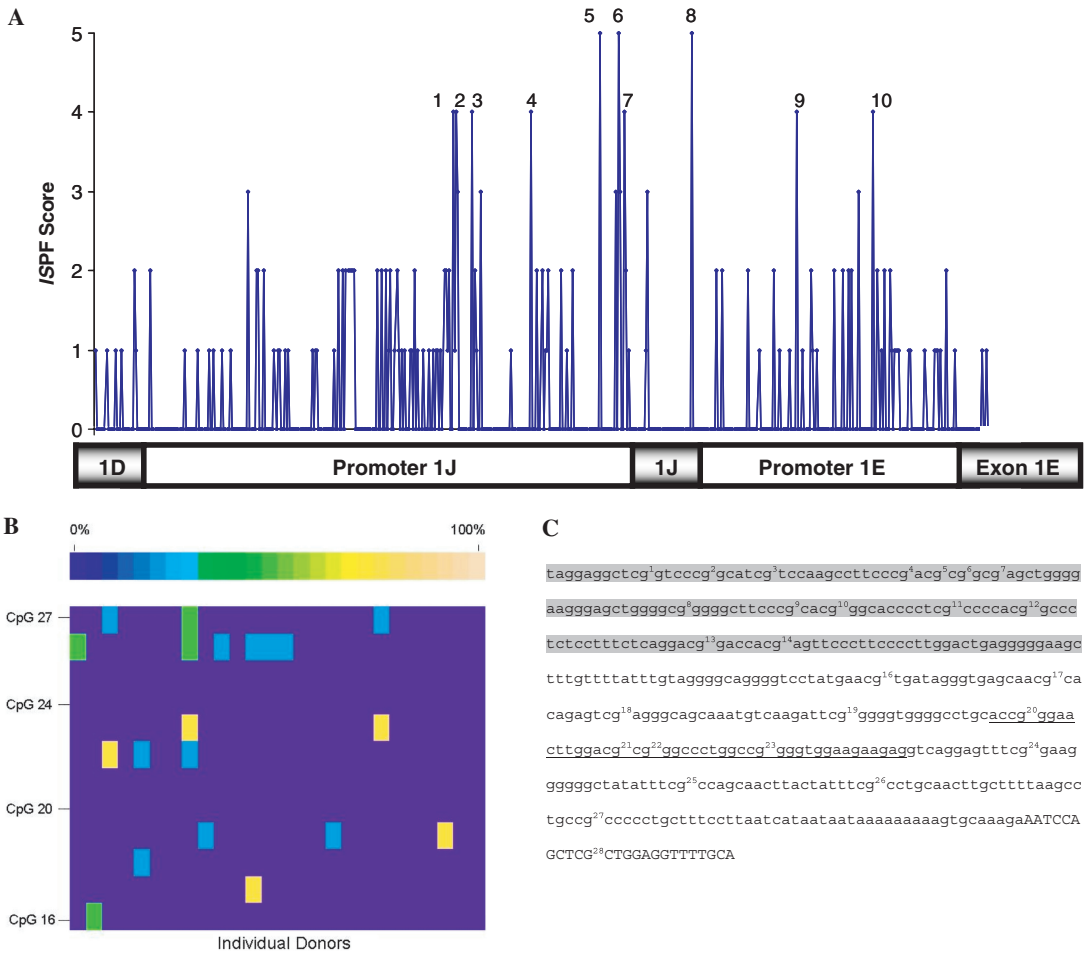
D

TF	*ISPF* peak	Location	+/–	CpG	Core match	Matrix match	Sequence
YY1	1	-4819	–	1	1.000	0.930	ctttgccaagATGGCggctc
NF-muE1	2	-4811	–	1	1.000	0.966	agATGGCgg
YY1	3	-4603	–	16, 17	1.000	0.935	ccctaccaagATGGCggcgg
NF-muE1	4	-4595	–	16	1.000	0.966	agATGGCgg
c-Ets-1 (p54)	5	-4584	–	18 – 20	1.000	0.962	ggcggcTTCCGgacg
Elk-1	6	-4582	–	18 – 20	1.000	0.980	cggcTTCCGgacg
NRF-2	7	-4581	–	18, 19	1.000	0.965	ggCTTCCggg

Figure 2. *In silico* phylogenetic footprinting and bisulphite sequencing of promoter 1D. (**A**) *In silico* phylogenetic footprinting covers the start of the CpG island 340 nucleotides upstream of exon 1D through to exon 1D. (**B**) Percentage methylation was measured by direct electropherogram reading after bisulphite sequencing of 26 donors, covering CpG 6–27. Methylation levels are expressed by colour, levels from the scale at the panel top. (**C**) CpG identifiers and the unsequenced region of promoter 1D (grey background). (**D**) All TRANSFAC predicted human TFBS in promoter 1D are significant *IS*PF predictions.

Promoter 1H

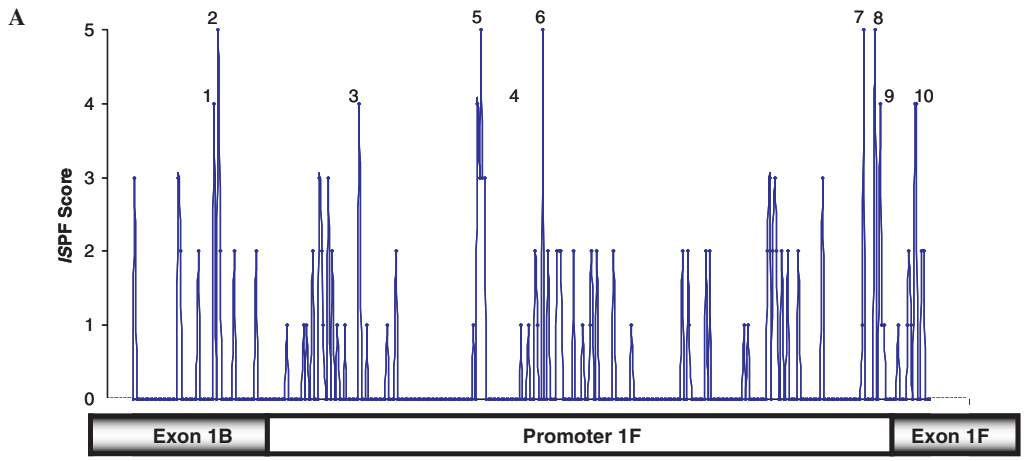
Promoter 1H, in contrast to the others, showed considerable methylation in up to 20 or 21 of the 26 donors (CpG^{41} and CpG^{55} respectively, Figure 5). Methylation levels of individual donors were also much higher than in other *GR* promoters, with up to 75% methylation of CpG^{41} in two individuals. *IS*PF predicts that the region around CpG$^{14–18}$ is evolutionarily the most conserved. Half of the high-quality *IS*PF predictions (*IS*PF > 4) contained CpG dinucleotides. However, methylation of these positions was limited to CpG^{14} and CpG^{15} (9 and 16 donors, respectively, all <50%), whilst CpG$^{16,17 \text{ and } 18}$

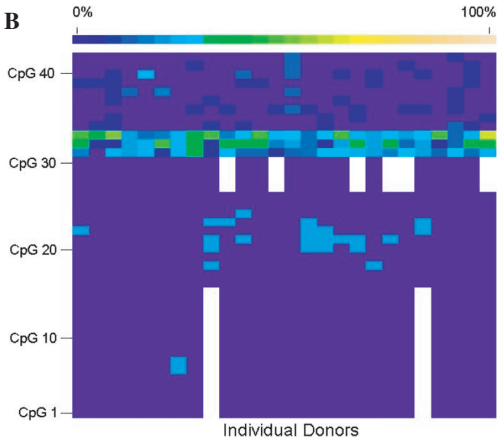


D

TF	*ISPF* Peak	Location	Strand	CpG	Core match	Matrix match	Sequence
MAZ		-4247	+	8	0.888	0.920	gGGGCGgg
NGFI-A		-4247	–	8	1.000	0.895	ggggcGGGGGcttc
NF-κ		-4242	–	9, 10	0.958	0.923	gggggcTTCCCgcacg
E2F		-4237	+	9	0.869	0.838	ctTCCCGc
HES1		-4189	–	13, 14	1.000	0.941	cggacCACGAgttcc
Cdc5	1		–	-	0.911	0.817	gcatTTTAAagc
TATA	2		–	-	1.000	0.950	aTTTTAaagc
Poly A	3		+	-	0.949	0.916	tgccTGTATtttgttt
Pax-3	**4**	**-4067**	**–**	**16**	**1.000**	**0.838**	**cctatgaaCGTGAtagggtga**
GATA-1		-4061	+	16	0.993	0.978	aacgTGATAgggtg
HNF-4		-4049	+	17	0.813	0.809	tgagCAACGcacag
COUPTF		-4042	–	17, 18	0.927	0.820	cgcacagagtcgaGGGCAgcaaa
Evi-1	5		+	-	1	0.732	gcaaatgtcAAGATt
Sp3	**6**	**-4012**	**+**	**19**	**0.925**	**0.882**	**attcggggGTGGGg**
MAZR	7		+	-	0.930	0.949	gggggTGGGGcct
E2F-1		-3992	–	20	0.873	0.907	ccGGGAAc
COUPTF	8		–	-	1.000	0.843	gggtggaagaagaGGTCAggagt
NF-κ		-3948	–	24	0.774	0.796	caggagTTTCGgaagg
Elk-1		-3944	+	24	1.000	0.937	agtttCGGAAgggg
MEF-2		-3930	+	25	0.962	0.876	ggcTATATttcg
RFX1 (EF-C)		-3923	–	25	0.910	0.765	tttcgccaGCAACt
Pax	9		+	-	0.818	0.823	CTGCAacttgc
GATA-1	10		–	-	1	0.967	tcctTAATCataat
HNF-1		-3913	+	26	0.952	0.835	aacTTACTatttcgcctg
NGFI-A		-3877	+	27	1.000	0.894	gccgCCCCCtgct

Figure 3. *In silico* phylogenetic footprinting and bisulphite sequencing of promoters 1J and 1E. (**A**) *In silico* phylogenetic footprinting covers the end of exon 1D through promoter 1J, exon 1J, to exon 1E. (**B**) Percentage methylation was measured by direct electropherogram reading after bisulphite sequencing of 26 donors, covering CpG 16–27. Methylation levels are expressed by colour, levels from the scale at the panel top. (**C**) CpG identifiers and the unsequenced region of promoter 1J (grey background). Exon 1J is underlined. (**D**) TRANSFAC prediction for promoter 1J and E. Significant *IS*PF predictions are numbered corresponding to (**A**). Predictions in bold are high-quality *IS*PF predictions containing CpG dinucleotides.





C

gtacg[1]tatgcg[2]ccg[3]acccccgctatcccg[4]tcccttccctgaagcctccccagagg

gcg[5]tgtcaggccg[6]cccg[7]gccccg[8]agcg[9]cg[10]gccg[11]agacg[12]ctgcg[13]gcaccg[14]

tttccg[15]tgcaaccccg[16]tagccccttccg[17]aagtgacacacttcacg[18]caactcg[19]g

cccg[20]gcg[21]gcg[22]gcg[23]gcg[24]cg[25]ggccactcacg[26]cagctcagccg[27]cg[28]ggaggc

g[29]ccccg[30]gctcttgtggcccg[31]cccg[32]ctgtcacccg[33]caggggcactggcg[34]gcg[35]

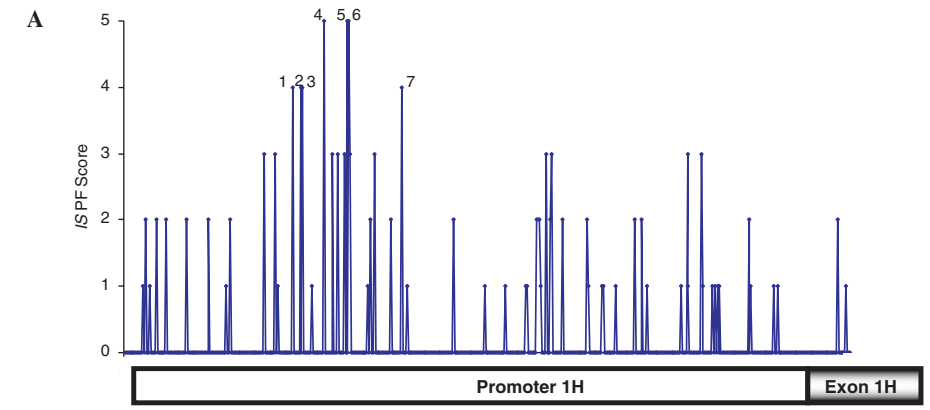
cttgccg[36]ccaaggggcagagcg[37]agctcccg[38]agtgggtctggagccg[39]cg[40]gagct

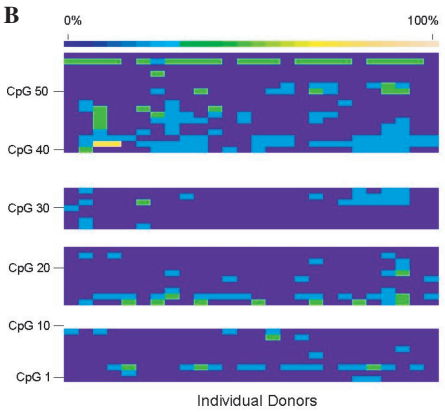
gggcg[41]ggggcg[42]ggaaggAGGTAGCG[43]AGAAAAGAAACTG

D

TF	*ISPF*	Location	Strand	CpG	Core match	Matrix match	Sequence
Evi-1		-3513	–	4	0.757	0.716	tATCCCgtccctcc
EBF	1		+	-	0.976	0.928	cTCCCCagagg
v-Myb	**2**	**-3436**	**–**	**14**	**1.000**	**0.914**	**caCCGTTtcc**
MIF-1		-3435	–	14, 15	1.000	0.940	accgtttccgtGCAACCc
Elk-1	**3**	**-3434**	**–**	**14, 15**	**1.000**	**0.936**	**ccgtTTCCGtgcaa**
RFX1	**4**	**-3434**	**+**	**14, 15**	**1.000**	**0.924**	**ccgtttccgtGCAACcc**
Bach2	5		–	-	0.922	0.915	agtgACACAct
Pax-3		-3399	+	18, 19	1.000	0.797	gacacactTCACGcaactcgg
AhR:Arnt		-3393	–	18, 19	1.000	0.955	cttCACGCaactcggc
CCAAT box		-3363	+	25	0.944	0.925	gcgggCCACTca
HES1		-3326	+	31	0.989	0.970	ggctcTTGTGgcccg
COUPTF		-3281	–	36, 37	0.927	0.816	gcttgccgccaagGGGCAgagcg
E2F		-3278	+	36	1.000	0.903	tgcCGCCAagg
HNF-4 DR1		-3276	–	36	0.893	0.799	ccgcCAAGGggca
NGFI-A	**6**	**-3228**	**–**	**41, 42**	**1.000**	**0.968**	**tgggcGGGGGcggg**
MAZ	**7**	**-3222**	**+**	**42**	**0.888**	**0.920**	**gGGGCGgg**
E2F-1	**8**	**-3220**	**–**	**42**	**0.935**	**0.890**	**gGCGGGaagg**
IRF	9		–	-	0.972	0.970	agaaaAGAAActgga
ICSBP	10		+	-	1.000	0.960	gaaaaGAAACtg

Figure 4. *In silico* phylogenetic footprinting and bisulphite sequencing of promoter 1F. (**A**) *In silico* phylogenetic footprinting covers the end of exon 1B through to exon 1F. (**B**) Percentage methylation was measured by direct electropherogram reading ($CpG^{1–30}$), or sequencing of 10–12 clones per donor ($CpG^{31–43}$) after bisulphite sequencing of 26 donors, covering CpG 1–42. Methylation levels are expressed by colour, levels from the scale at the panel top. Blank squares indicate positions for which no data was obtainable. (**C**) CpG identifiers within the sequence of promoter 1F. (**D**) TRANSFAC prediction for promoter 1F. Significant *IS*PF predictions are numbered corresponding to (A). Predictions in bold are high-quality *IS*PF predictions containing CpG dinucleotides.





C

gtaagaagcg[1]aggcg[2]ggaggggccg[3]gggcg[4]cg[5]ctcg[6]ctcccccg[7]aggtgccg[8]ct

gggaccg[9]gagacaactcg[10]ggggccg[11]ccg[12]cg[13]ggagcctacaaactttattagcc

tcg[14]gggagtgggggtgggggctggcaagggccg[15]ggcg[16]acg[17]gtgacg[18]aaaggg

cagcg[19]cg[20]cg[21]ggtgacagcg[22]ctggcctcttcctctccctccg[23]caggcg[24]tcccc

tggccg[25]ggccg[26]aggggaggaacctgacctcg[27]gacg[28]gcg[29]agcg[30]gagccctgt

cg[31]aactgccg[32]ggggcttcg[33]agcctctcattcctcg[34]cg[35]ggaatcctggcctctt

ttctcccctagtgtccccttccctccaaggggtcg[36]cccg[37]acacccg[38]ttttcg[39]

tggtgaacg[40]ctaagccg[41]cg[42]tctgaattttactcg[43]cccg[44]aatatttgcacg[45]cc

accccg[46]gcg[47]cg[48]cccg[49]agcg[50]cg[51]agcccg[52]ggctccg[53]gggaggccccg[54]gcG[5]

[5]GCG[56]CCTGGCTTGAGGAGGGCG[57]TGCG[58]G

D

TF	*ISPF*	Location	Strand	CpG	Core match	Matrix match	Sequence
MAZ		-2354	+	2	1.000	0.952	cGGGAGgg
MAZ		-2332	–	6	1.000	0.943	cgCTCCCc
AP-2Aα		-2264	+	14	0.989	0.947	ttAGCCTcggggagt
Sp3	**1**	**-2262**	**+**	**14**	**0.765**	**0.870**	**agcctcggGGAGTg**
MAZ	2		+		1.000	0.936	gGGGAGtg
NGFI-A	3		–		1.000	0.898	ggagtGGGGGtggg
CP2	4		–		0.97	0.885	CTGGCaagggc
COUPTF		-2221	–	17 – 20	0.927	0.884	gacggtgacgaaaGGGCAgcgcg
CREB	**5**	**-2220**	**+**	**17, 18**	**1.000**	**0.946**	**acggTGACGaaa**
ATF4		-2218	+	18	1.000	0.903	ggTGACGaaagg
E4F1	**6**	**-2217**	**–**	**18**	**1.000**	**0.930**	**gTGACGaaag**
CREBATF	**6**	**-2217**	**+**	**18**	**1.000**	**0.954**	**gTGACGaaa**
HNF4		-2216	+	18	0.860	0.872	tgacGAAAGggcag
MAZ			–	-	1.000	0.921	ctCTCCCt
Sp3		-2171	–	23, 24	1.000	0.883	cCCTCCgcaggcgt
COUPTF	7	-2134	+	27 – 29	1.000	0.814	gaaccTGACCtcggacggcgagc
Ik-1		-2069	+	34, 35	1.000	0.949	tcgcGGGAAtcct
E2F-1		-2068	–	34, 35	0.935	0.879	cGCGGGaatc
E2F		-2067	–	35	0.869	0.828	gCGGGAat
NF-κ		-2019	–	36, 37	0.752	0.817	aggggTCGCCcgaca
v-Myb		-2004	–	38	1.000	0.937	acCCGTTttc
E2F		-1998	+	39	0.790	0.830	ttTCGTGg
GATA-4		-1962	–	43, 44	0.814	0.792	tcgcccgAATAT
MINI-20		-1951	+	45 – 48	0.992	0.903	tttgcaCGCCAccccggcgcg
NF-κ		-1913	+	53	0.812	0.828	gctccGGGGAggcccc

Figure 5. *In silico* phylogenetic footprinting and bisulphite sequencing of promoter 1H. (**A**) *In silico* phylogenetic footprinting covers promoter 1H and part of exon 1H. (**B**) Percentage methylation was measured by direct electropherogram reading after bisulphite sequencing of 26 donors, covering CpG 1–42. Methylation levels are expressed by colour, levels from the scale at the panel top. Blank squares indicate positions for which no data was obtainable. (**C**) CpG identifiers within the sequence of promoter 1H and the unsequenced region (grey background). (**D**) TRANSFAC prediction for promoter 1H. Significant *IS*PF predictions are numbered corresponding to (A). Predictions in bold are high-quality *IS*PF predictions containing CpG dinucleotides.

were methylated in 2, 0 and 3 donors, respectively (all <25%).

DISCUSSION

The results of this study show that the human *GR* promoter region is extensively methylated, and that levels and positions of methylation are highly diverse in normal donors. Although all of the alternative *GR* promoters investigated here are in a CpG island, most of the conserved TFBS contained methylatable CpG sites, suggesting that epigenetic mechanisms operate throughout the CpG island of this gene. This is the first report of such a comprehensive investigation of a complex CpG island containing multiple alternative first exons each with its own promoter and TSS.

Sequence patterns predicted unrealistically large numbers of TFBS throughout the human *GR* CpG island. By identifying evolutionarily conserved regions through interspecies sequence comparison ('phylogenetic footprinting') (36), and combining it with TRANSFAC TFBS predictions, a technique we term *in silico* phylogenetic footprinting, the number of putative TFBS was reduced by 3- to 4-fold. The successful prediction of the known TFBS within the CpG island (27) validated our approach, at least for the *GR*. It also suggests that mechanisms of transcription of the alternative 5′ mRNA transcript variants of *GR* are conserved in mammalian species.

In PBMCs, five of the seven CpG island promoters are known to be used (22). In these five promoters, covering more than 40% of the 3.2 kbp total *GR* CpG island over 70% of all CpG dinucleotides were methylated. Methylation of the *GR* first exon promoters occured throughout the CpG island and affected every promoter. Within our 26 donors, methylation was variable, but the majority of positions were methylated at levels above 25% in at least one donor. Not a single position was methylated in all donors.

Previously, methylation of only a 100 bp section of promoter 1F (37), and 331 bp of the constitutively active promoter 1C (38) have been investigated in humans. In agreement with our results, Lillycrop *et al.* observed variable methylation in both the human promoter 1C in cord blood and its homologue 1_{10} in the rat liver (38,39), but no methylation levels of individual CpG dinucleotides were reported. In human hippocampi of both healthy and depressed individuals methylation of 1F $CpG^{35–43}$ was limited to <10% methylation of CpG^{35} in 1 of 32 donors, similar to our findings in PBMCs (37). In the same study, CpG^{43} was almost 100% methylated in all donors with no methylation in any other site, which is in contrast to our findings in PBMCs where CpG^{43} was almost universally unmethylated. Thus, it is not clear to what extent positions and variability can be extrapolated from PBMCs to the human brain.

This variability in methylation patterns is in stark contrast to the promoter hypo- and hyper-methylation of large regions, and levels that are either undetectable, or nearly 100% in tumours. Because of their consistency CpG island hypermethylation profiles ('hypermethylomes') are now recognised as tumor-specific markers, similar to genetic and cytogenetic markers (40–42), for instance in prostate cancer (43), and in clear cell renal carcinomas (44).

The variability in methylation between donors and positions may reflect an important epigenetic mechanism suggested by studies both in animals and humans. Weaver was able to manipulate the methylation patterns by perinatal interventions. In the rat brain, methylation of the Ngfi-A-binding site in promoter 1_7 was virtually complete in animals that had received poor postnatal maternal care, whilst those receiving better maternal care were virtually unmethylated (25,26,45). By cross fostering pups between dams providing good or poor post-natal care the pups developed the epigenome of the foster mother. Similarly, methylation of promoter 1_{10} was reduced in pups born to dams fed on a protein-restricted diet during pregnancy (38,39). These rat experiments suggest that *Gr* expression and perhaps sensitivity can be fine-tuned as the host adapts to its environment during early life. Recent human data from the *GR* support the above rat data (46). Although in the latter study methylation levels observed were very low, methylation of exon 1F in fetal cord blood was sensitive to maternal mood. Thus the heterogeneity of methylation patterns observed in all the CpG island dinucleotides and in particular the different alternative *GR* promoters reflect a mechanism of epigenetic programming of GR fine tuning programmed by differences in early life environment and events. Methylation of these important evolutionarily conserved TFBS containing methylatable CpG dinucleotides would suggest that such methylation will have considerable effects on the transcription from the TSS immediately downstream. In both these cases, the methylation of single CpG dinucleotides within the promoter that has been implicated. Therefore, the observed variability both in levels and positions of methylation may reflect some specificity. Methylation of different TFBS within a promoter will modulate the promoter response, in a tightly controlled individualised fashion. Data from other genes support the role of the inter-uterine environment in lasting epigenetic conditioning. In a rodent type 2 diabetes model, Park *et al.* (47) showed that retarded interuterine growth resulted in lower *Pdx1* levels, a key correlate of adult β cell function. As for the *GR*, the *Pdx1* promoter is in a CpG island. Methylation of the 14 CpG pairs in the 275 bp promoter region was shown to be responsible for the reduced *Pdx1* levels. Since methylation modulates promoter activity the highly individual methylation patterns of *GR* promoters strongly suggest highly individualised fine tuning of *GR* expression. However, many other factors such as age (7) and drug intake (48–50) may influence and add to this variability.

CpG methylation has also been implicated in tissue-specific regulation. GNAL (51) and PNP22 (52) are such examples, but both genes have a simpler structure with alternative first exons and promoters in independent CpG islands. The tissue-specific expression of Toll-like receptor 2 (TLR2) is, however, the result of strong methylation of the 20 CpG pairs in the ~220 bp proximal promoter in the upstream CpG island. This promoter is very similar in size and number of CpG pairs as each of

the individual *GR* promoters investigated here. Although ubiquitously expressed, in most tissues functional levels of GR are tightly and individually regulated in all tissues. We have previously observed that the multiple alternative *GR* first exons are expressed in a tissue-specific manner each using its own promoter (22,27). Since the complete *GR* CpG island with all the promoters can be methylated we further propose that CpG methylation orchestrates alternative first exon usage and silencing.

Currently an incredible transcriptional variability, especially in the 5′ region unfolds. Already in 2006 the number of genes with multiple alternative first exons has passed 7600 (21), and at least 43 genes have more than 10 first exons. Many of these genes are expressed in a tissue-specific manner. With 78% of the evolutionarily conserved alternative first exons falling into CpG islands epigenetic modulation and tissue-specific expression through methylation of individual promoters within CpG islands is bound to be a biologically important and widely used mechanism throughout many phyla.

ACKNOWLEDGEMENTS

We are grateful to Prof. Dr Jobst Meyer and member of the Functional Genetics group (Trier, Germany) for their excellent technical advice in setting up the bisulphite sequencing. We would like to thank Hartmut Schächinger and Dirk Hellhammer for their initiatives within the Trier-Leiden International Research Training Group (IRTG) and the Graduate School of Psychobiology, respectively.

FUNDING

This work was supported by the Luxembourg Health ministry [SAN-LNSI-20070106]; the Ministry of Culture, Higher Education and Research [REC-LNSI-20070112]; and the Deutsche Forschungsgemeinschaft [GRK1389/1 to C.P.M.]. Funding for open access charge: Luxembourg ministry of culture, higher education and research [REC-LNSI-20070112].

Conflict of interest statement. None declared.

REFERENCES

1. Bird,A. (2007) Perceptions of epigenetics. *Nature*, **447**, 396–398.
2. Gardiner-Garden,M. and Frommer,M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261–282.
3. Ehrlich,M. (2003) Expression of various genes is controlled by DNA methylation during mammalian development. *J. Cell. Biochem.*, **88**, 899–910.
4. Razin,A., Webb,C., Szyf,M., Yisraeli,J., Rosenthal,A., Naveh-Many,T., Sciaky-Gallili,N. and Cedar,H. (1984) Variations in DNA methylation during mouse cell differentiation in vivo and in vitro. *Proc. Natl Acad. Sci. USA*, **81**, 2275–2279.
5. Gruenbaum,Y., Stein,R., Cedar,H. and Razin,A. (1981) Methylation of CpG sequences in eukaryotic DNA. *FEBS Lett.*, **124**, 67–71.
6. Imamura,T., Ohgane,J., Ito,S., Ogawa,T., Hattori,N., Tanaka,S. and Shiota,K. (2001) CpG island of rat sphingosine kinase-1 gene: tissue-dependent DNA methylation status and multiple alternative first exons. *Genomics*, **76**, 117–125.
7. Fraga,M.F., Ballestar,E., Paz,M.F., Ropero,S., Setien,F., Ballestar,M.L., Heine-Suner,D., Cigudosa,J.C., Urioste,M., Benitez,J. *et al.* (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci. USA*, **102**, 10604–10609.
8. Kiefer,J.C. (2007) Epigenetics in development. *Dev. Dyn.*, **236**, 1144–1156.
9. Swales,A.K. and Spears,N. (2005) Genomic imprinting and reproduction. *Reproduction*, **130**, 389–399.
10. Allegrucci,C., Denning,C.N., Burridge,P., Steele,W., Sinclair,K.D. and Young,L.E. (2005) Human embryonic stem cells as a model for nutritional programming: an evaluation. *Reprod. Toxicol.*, **20**, 353–367.
11. Thurston,A., Lucas,E.S., Allegrucci,C., Steele,W. and Young,L.E. (2007) Region-specific DNA methylation in the preimplantation embryo as a target for genomic plasticity. *Theriogenology*, **68(Suppl. 1)**, S98–S106.
12. Riggs,A.D. (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.*, **14**, 9–25.
13. Holliday,R. and Pugh,J.E. (1975) DNA modification mechanisms and gene activity during development. *Science*, **187**, 226–232.
14. Song,F., Smith,J.F., Kimura,M.T., Morrow,A.D., Matsuyama,T., Nagase,H. and Held,W.A. (2005) Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl Acad. Sci. USA*, **102**, 3336–3341.
15. Meissner,A., Mikkelsen,T.S., Gu,H., Wernig,M., Hanna,J., Sivachenko,A., Zhang,X., Bernstein,B.E., Nusbaum,C., Jaffe,D.B. *et al.* (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766–770.
16. Allegrucci,C., Wu,Y.Z., Thurston,A., Denning,C.N., Priddle,H., Mummery,C.L., Ward-van Oostwaard,D., Andrews,P.W., Stojkovic,M., Smith,N. *et al.* (2007) Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. *Hum. Mol. Genet.*, **16**, 1253–1268.
17. Fan,G., Martinowich,K., Chin,M.H., He,F., Fouse,S.D., Hutnick,L., Hattori,D., Ge,W., Shen,Y., Wu,H. *et al.* (2005) DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development*, **132**, 3345–3356.
18. Hamby,M.E., Coskun,V. and Sun,Y.E. (2008) Transcriptional regulation of neuronal differentiation: the epigenetic layer of complexity. *Biochim. Biophys. Acta.*, **1779**, 432–437.
19. Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S. *et al.* (2001) Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. *EMBO Rep.*, **2**, 388–393.
20. Tsuritani,K., Irie,T., Yamashita,R., Sakakibara,Y., Wakaguri,H., Kanai,A., Mizushima-Sugano,J., Sugano,S., Nakai,K. and Suzuki,Y. (2007) Distinct class of putative 'non-conserved' promoters in humans: comparative studies of alternative promoters of human and mouse genes. *Genome Res.*, **17**, 1005–1014.
21. Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H. *et al.* (2006) Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res.*, **16**, 55–65.
22. Turner,J.D. and Muller,C.P. (2005) Structure of the glucocorticoid receptor (NR3C1) gene 5′ untranslated region: identification, and tissue distribution of multiple new human exon 1. *J. Mol. Endocrinol.*, **35**, 283–292.
23. Presul,E., Schmidt,S., Kofler,R. and Helmberg,A. (2007) Identification, tissue expression, and glucocorticoid responsiveness of alternative first exons of the human glucocorticoid receptor. *J. Mol. Endocrinol.*, **38**, 79–90.
24. Weaver,I.C. (2007) Epigenetic programming by maternal behavior and pharmacological intervention. Nature versus nurture: let's call the whole thing off. *Epigenetics*, **2**, 22–28.
25. Weaver,I.C., Cervoni,N., Champagne,F.A., D'Alessio,A.C., Sharma,S., Seckl,J.R., Dymov,S., Szyf,M. and Meaney,M.J. (2004) Epigenetic programming by maternal behavior. *Nat. Neurosci.*, **7**, 847–854.
26. Weaver,I.C., Meaney,M.J. and Szyf,M. (2006) Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors

in the offspring that are reversible in adulthood. *Proc. Natl Acad. Sci. USA*, **103**, 3480–3485.
27. Turner,J.D., Schote,A.B., Macedo,J.A., Pelascini,L.P. and Muller,C.P. (2006) Tissue specific glucocorticoid receptor expression, a role for alternative first exon usage? *Biochem. Pharmacol.*, **72**, 1529–1537.
28. Kel,A.E., Gossling,E., Reuter,I., Cheremushkin,E., Kel-Margoulis,O.V. and Wingender,E. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.*, **31**, 3576–3579.
29. Wittchen,H.-U., Zaudig,M. and Fydrich,T. (1997) *Strukturiertes Klinisches Interview für DSM-IV (SKID)*, Hogrefe, Goettingen.
30. Lewin,J., Schmitt,A.O., Adorjan,P., Hildmann,T. and Piepenbrock,C. (2004) Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplificates. *Bioinformatics*, **20**, 3005–3012.
31. Nunez,B.S. and Vedeckis,W.V. (2002) Characterization of promoter 1B in the human glucocorticoid receptor gene. *Mol. Cell. Endocrinol.*, **189**, 191–199.
32. Breslin,M.B. and Vedeckis,W.V. (1998) The human glucocorticoid receptor promoter upstream sequences contain binding sites for the ubiquitous transcription factor, Yin Yang 1. *J. Steroid Biochem. Mol. Biol.*, **67**, 369–381.
33. Kim,J., Kollhoff,A., Bergmann,A. and Stubbs,L. (2003) Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum. Mol. Genet.*, **12**, 233–245.
34. Swirnoff,A.H. and Milbrandt,J. (1995) DNA-binding specificity of NGFI-A and related zinc finger transcription factors. *Mol. Cell. Biol.*, **15**, 2275–2287.
35. McCormick,J.A., Lyons,V., Jacobson,M.D., Noble,J., Diorio,J., Nyirenda,M., Weaver,S., Ester,W., Yau,J.L., Meaney,M.J. *et al.* (2000) 5′-heterogeneity of glucocorticoid receptor messenger RNA is tissue specific: differential regulation of variant transcripts by early-life events. *Mol. Endocrinol.*, **14**, 506–517.
36. Andersen,M.C., Engstrom,P.G., Lithwick,S., Arenillas,D., Eriksson,P., Lenhard,B., Wasserman,W.W. and Odeberg,J. (2008) In silico detection of sequence variations modifying transcriptional regulation. *PLoS Comput. Biol.*, **4**, e5.
37. Moser,D., Molitor,A., Kumsta,R., Tatschner,T., Riederer,P. and Meyer,J. (2007) The glucocorticoid receptor gene exon 1-F promoter is not methylated at the NGFI-A binding site in human hippocampus. *World J. Biol. Psychiatry*, **8**, 262–268.
38. Lillycrop,K.A., Slater-Jefferies,J.L., Hanson,M.A., Godfrey,K.M., Jackson,A.A. and Burdge,G.C. (2007) Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications. *Br. J. Nutr.*, **97**, 1064–1073.
39. Lillycrop,K.A., Phillips,E.S., Torrens,C., Hanson,M.A., Jackson,A.A. and Burdge,G.C. (2007) Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPARalpha promoter of the offspring. *Br. J. Nutr.*, **97**, 1064–1073.
40. Esteller,M., Corn,P.G., Baylin,S.B. and Herman,J.G. (2001) A gene hypermethylation profile of human cancer. *Cancer Res.*, **61**, 3225–3229.
41. Costello,J.F., Fruhwald,M.C., Smiraglia,D.J., Rush,L.J., Robertson,G.P., Gao,X., Wright,F.A., Feramisco,J.D., Peltomaki,P., Lang,J.C. *et al.* (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.*, **24**, 132–138.
42. Esteller,M., Sanchez-Cespedes,M., Rosell,R., Sidransky,D., Baylin,S.B. and Herman,J.G. (1999) Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.*, **59**, 67–70.
43. Hoque,M.O., Topaloglu,O., Begum,S., Henrique,R., Rosenbaum,E., Van Criekinge,W., Westra,W.H. and Sidransky,D. (2005) Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. *J Clin. Oncol.*, **23**, 6569–6575.
44. Kvasha,S., Gordiyuk,V., Kondratov,A., Ugryn,D., Zgonnyk,Y.M., Rynditch,A.V. and Vozianov,A.F. (2008) Hypermethylation of the 5′CpG island of the FHIT gene in clear cell renal carcinomas. *Cancer Lett.*, **265**, 250–257.
45. Weaver,I.C., Champagne,F.A., Brown,S.E., Dymov,S., Sharma,S., Meaney,M.J. and Szyf,M. (2005) Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. *J. Neurosci.*, **25**, 11045–11054.
46. Oberlander,T.F., Weinberg,J., Papsdorf,M., Grunau,R., Misri,S. and Devlin,A.M. (2008) Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics*, **3**, 97–106.
47. Park,J.H., Stoffers,D.A., Nicholls,R.D. and Simmons,R.A. (2008) Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1. *J. Clin. Invest.*, **118**, 2316–2324.
48. Alonso-Aperte,E., Ubeda,N., Achon,M., Perez-Miguelsanz,J. and Varela-Moreiras,G. (1999) Impaired methionine synthesis and hypomethylation in rats exposed to valproate during gestation. *Neurology*, **52**, 750–756.
49. Manev,H. and Uz,T. (2002) DNA hypomethylating agents 5-aza-2′-deoxycytidine and valproate increase neuronal 5-lipoxygenase mRNA. *Eur. J. Pharmacol.*, **445**, 149–150.
50. Shimabukuro,M., Jinno,Y., Fuke,C. and Okazaki,Y. (2006) Haloperidol treatment induces tissue- and sex-specific changes in DNA methylation: a control study using rats. *Behav. Brain. Funct.*, **2**, 37.
51. Corradi,J.P., Ravyn,V., Robbins,A.K., Hagan,K.W., Peters,M.F., Bostwick,R., Buono,R.J., Berrettini,W.H. and Furlong,S.T. (2005) Alternative transcripts and evidence of imprinting of GNAL on 18p11.2. *Mol. Psychiatry*, **10**, 1017–1025.
52. Huehne,K. and Rautenstrauss,B. (2001) Transcriptional startpoints and methylation patterns in the PMP22 promoters of peripheral nerve, leukocytes and tumor cell lines. *Int. J. Mol. Med.*, **7**, 669–675.
53. Nobukuni,Y., Smith,C.L., Hager,G.L. and Detera-Wadleigh,S.D. (1995) Characterization of the human glucocorticoid receptor promoter. *Biochemistry*, **34**, 8207–8214.