Cover Page

Universiteit Leiden

Leiden University
Repository

The handle http://hdl.handle.net/1887/39295 holds various files of this Leiden University dissertation

Author: Polman, J.A.E.
Title: Glucocorticoid signature in a neuronal genomic context
Issue Date: 2016-05-10

Chapter Two

Specific regulatory motifs predict glucorticorticoid responsiveness of hippocampal gene expression

N.A. Datson[1], J.A.E. Polman[1], R.T. de Jonge[1], P.T.M. Van Boheemen[1],
E.M.T. Van Maanen[1], Jennifer Welten[1], B. McEwen[2], H.C. Meiland[3],
O.C. Meijer[1]

[1] Division of Medical Pharmacology, Leiden/Amsterdam Center for Drug Research &
    Leiden University Medical Center, Leiden, the Netherlands
[2] Laboratory of Neuroendocrinology, the Rockefeller University,
    1230 York Avenue, New York, NY 10065, USA
[3] ICT shared service center, Leiden University, Leiden, The Netherlands

T HE glucocorticoid receptor (GR) is an ubiquitously expressed ligand-activated transcription factor that mediates effects of cortisol in relation to adaptation to stress. In the brain, GR affects the hippocampus to modulate memory processes through direct binding to glucocorticoid response elements (GREs) in theDNA. However, its effects are to a high degree cell specific, and its target genes in different cell types as well as the mechanisms conferring this specificity are largely unknown. To gain insight in hippocampal GR signaling, we characterized to which GRE GR binds in the rat hippocampus. Using a position-specific scoring matrix, we identified evolutionary-conserved putative GREs from a microarray based set of hippocampal target genes. Using chromatin immunoprecipitation, we were able to confirm GR binding to 15 out of a selection of 32 predicted sites (47 %). The majority of these 15 GREs are previously undescribed and thus represent novel GREs that bind GR and therefore may be functional in the rat hippocampus. GRE nucleotide composition was not predictive for binding of GR to a GRE. A search for conserved flanking sequences that may predict GR-GRE interaction resulted in the identification of GC-box associated motifs, such as Myc-associated zinc finger protein 1, within 2 kb of GREs with GR binding in the hippocampus. This enrichment was not present around nonbinding GRE sequences nor around proven GR-binding sites from a mesenchymal stem-like cell dataset that we analyzed. GC-binding transcription factors therefore may be unique partners for DNA-bound GR and may in part explain cell-specific transcriptional regulation by glucocorticoids in the context of the hippocampus.

2.1 Introduction

Glucocorticoid hormones, i.e. cortisol in humans and corticosterone in rodents (both abbreviated as CORT), released by the adrenal gland in response to stress, are important mediators of the stress response throughout the body and the brain. Cellular adaptation to stress is highly tissue dependent, but mechanisms responsible for the high degree of cell specificity of CORT target genes are largely unknown. The brain is a major target of CORT, which readily passes the blood-brain barrier to affect a wide variety of processes, both in neurons and glia cells. CORT has profound effects on neuronal plasticity and neuronal survival, with consequences for behavior, learning, and memory. These effects are mediated by the coordinate action of high-affinity mineralocorticoid (MR) and low-affinity glucocorticoid receptors (GR), colocalized in neurons of the limbic brain, in particular the hippocampus, and in control of gene expression networks (Datson et al., 2008).

Part of the CORT effects on neuronal function and viability depends on genomic mechanisms involving binding of GR and/or MR to glucocorticoid response elements (GREs) regulating expression of target genes. GRE-dependent processes are important in the brain, because modulation of hippocampal excitability and spatial memory were impaired in $GR^{dim/dim}$ mutant mice, in which the mutation prevented GR homodimerization and therefore binding to most GREs, whereas protein-protein interactions of the receptor with other transcription factors remained undisturbed (Karst et al., 2000; Oitzl et al., 2001).

Several studies have focused on identifying GREs in peripheral tissues (Phuc Le P. et al., 2005) and cell lines, including the A549 human lung epithelial carcinoma cell line and mouse mesenchymal stem-like cells (Reddy et al., 2009; So et al., 2007; So et al., 2008; Wang et al., 2004). However, the GREs responsible for action of GR in vivo in the brain are largely unknown. It is likely that there are brain-specific GREs that selectively function in a neuronal context, given the large diversity of CORT-regulated genes when comparing CORT responses in different tissues. Taking this even a step further, within the brain there are likely to be sequence motifs that determine why CORT induces expression of a particular gene in the dentate gyrus (DG) subregion of the hippocampus, whereas having no effect in cornu ammonis 1, despite the fact that both subregions express GR (Gemert Van et al., 2009; Lee et al., 2003; Schaaf et al., 1998). Understanding the molecular context in which GREs function is necessary for a better understanding of the way in which CORT, via GR, affects the function and morphology of different brain regions and adaptation to stress.

Although in many cases chromatin accessibility is a prerequisite for binding of transcription factors, evolutionary conservation appears to be a major predictor of functionality of a subset of transcription factor-binding sites (Kunarso et

al., 2010), including the GRE (So et al., 2008). We took advantage of this to predict evolutionary-conserved GREs *in silico* using a position-specific scoring matrix from 44 GREs described in literature. Using this matrix, we scanned large genomic regions surrounding CORT-responsive genes in two different expression datasets enriched for CORT-responsive genes: 1) an expression dataset derived from *in vivo* CORT responses in rat hippocampus (Datson, N. A. and B. S. McEwen, unpublished data), and 2) a published expression dataset consisting of genes up-regulated by CORT in mouse C3H10T1/2 mesenchymal stem-like cells (So et al., 2008). Our goals were to 1) identify GREs in the vicinity of GR-induced genes in the hippocampus, 2) analyze how true GREs in the hippocampus differ from nonbinding sequences, and 3) elucidate how primary GR targets in hippocampus differ from those in mesenchymal stem-like cells.

2.2 Materials and Methods

Microarray datasets

Two microarray datasets enriched for CORT-responsive genes were used in this study.

CORT-responsive genes in the rat hippocampus

This *in vivo* dataset was derived from the hippocampal DG subregion of rats injected sc with $5\,\mathrm{mg/kg/ml}$ CORT dissolved in propylene glycol and killed $3\,\mathrm{h}$ after injection. The DG was isolated using laser microdisection and used for microarray analysis on Affymetrix Rat Genome 230 2.0 GeneChips. The microarray experiment lead to the identification of 538 CORT-responsive genes, comprising 183 up-regulated and 118 down-regulated genes that could be linked to a gene symbol. We continued with the 183 up-regulated genes for GRE predictions (Table 2.3).

CORT-responsive genes in mouse C3H10T1/2 mesenchymal stem-like cells (So et al., 2008)

This *in vitro* dataset was derived from C3H10T1/2 mesenchymal stem-like cells treated with $1\,\mu\mathrm{M}$ dexamethasone, a synthetic glucocorticoid, for $90\,\mathrm{min}$. Sixty-nine genes were found to be up-regulated after treatment, and 17 genes were down-regulated. In this set, 50 GRE sites were shown to bind GR, whereas 119 "predicted GR-binding sequences" did not bind GR. The GR-binding and GR-nonbinding GREs in this study (as in ours) did not differ in nucleotide content (So et al., 2007).
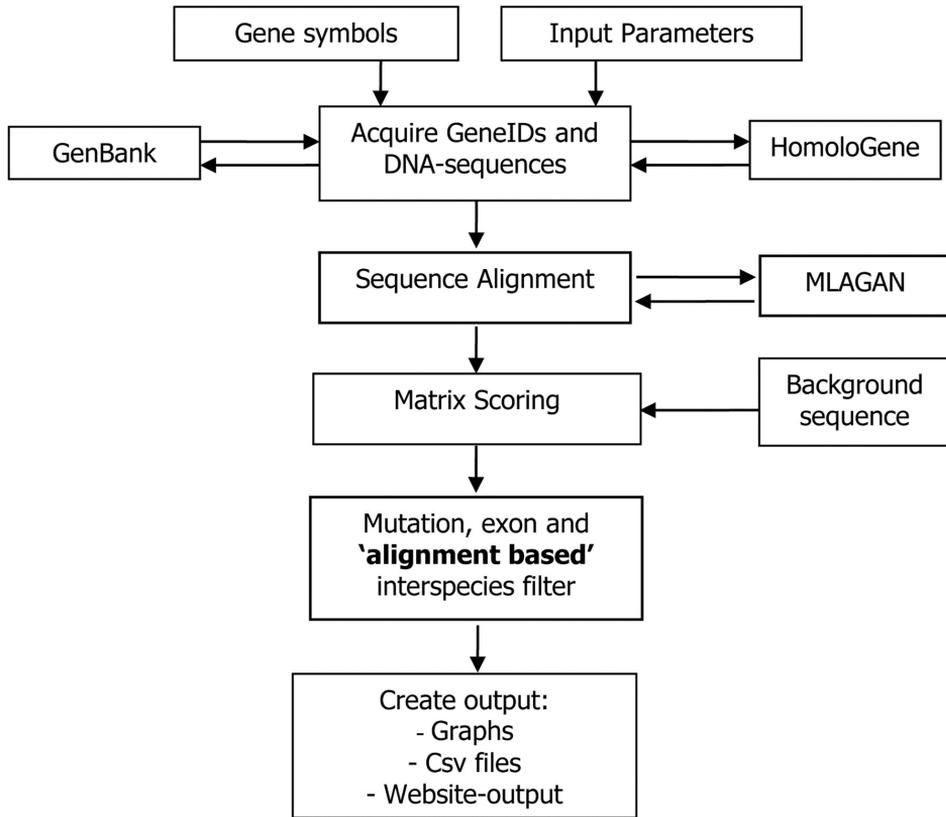
Figure 2.1: Outline of the *in silico* GRE-screening procedure.
Please see text for details.

In silico GRE prediction

For the current study, we constructed a GRE matrix, which is based on 44 GREs described in literature called matrix-44 (Table 2.4). A list of gene symbols representing CORT-responsive genes from the rat hippocampal microarray data was used to score for GRE-like sequences using matrix-44. Only upregulated genes were selected, because these depend on binding to classical GREs. Homologous sequences for multiple species (human, cow, mouse) were retrieved from the homologene database. Mouse and human were chosen for completeness of genomic annotation and supplemented with one additional species that is phylogenetically in between rodent and human (i.e. cow). A genomic region of 50 kb up- and downstream was retrieved per gene per species from the National Center for Biotechnology Information GenBank website. Exonic sequences were excluded. To facilitate identification of positionally and evolutionary-conserved GREs, the sequences of the different species were aligned before scoring using the BioPerl module dedicated to the LAGAN Toolkit, based on the multiple alignment algorithm MLAGAN (Brudno et al., 2003).

After alignment, the sequences of the different species were individually scored, and an interspecies filter was applied, which matches and selects for predicted GREs based on their position in the alignment between multiple species. A maximum of four mismatches between the different species was tolerated. More differences than this resulted in discarding the GRE from further analysis. For each position on the DNA-sequence, a score was computed for the full length of the matrix using a sliding window of 14 nucleotides. Please note that the classical canonical GRE sequence is 15 nucleotides. However, the first position in our matrix-44 did not show any base pair preference and as such did not contribute to the score. Therefore, we decided to omit the first base. Subsequently, a threshold of 0.8 (out of a maximum score of 1) was set, based on a frequency of less than 0.1% of scores of 0.8 or higher in random DNA sequences (data not shown). The criteria for considering a sequence to be a putative GRE were: 1) location within a region spanning 50 kb upstream and 50 kb downstream of a gene upregulated by CORT in our microarray dataset, 2) a GRE score of at least 0.8 in four different species (rat, mouse, cow, and human), and 3) a maximum of four mismatches between the different species. An outline of the approach is depicted in Figure 2.1.

Animals and treatment

Male Sprague Dawley rats (Harlan, Leiden, The Netherlands) weighing approximately 250 g on the day of surgery were group housed on a 12 h light, 12 h dark cycle (lights on at 0700h) in a temperature-controlled facility. Animals were handled daily for a week before the start of the experiment. Food and water were provided *ad libitum*. All experimental manipulations were done in the morning. Experiments were approved by the Local Committee for Animal Health, Ethics, and Research of the University of Leiden (Dierexperimentencommissie no. 07166). Animal care was conducted in accordance with the European Commission Council Directive of November 1986 (86/609/EEC).

To study GR dynamics, animals were challenged with a high dose of CORT (3 mg/kg ip CORT-hydroxypropyl-cyclodextrin; Sigma-Aldrich, St. Louis, MO). Tail blood samples were taken before and during the challenge to monitor CORT levels in blood. Animals were decapitated ($n = 8$ per time point per treatment group) 0, 60, and 180 min after injection. Brain tissue was collected, snap frozen in isopentane on dry ice, and stored at $-80\,°C$ until further processing. Of each animal, one hippocampus was isolated for chromatin immunoprecipitation (ChIP).

Chromatin immunoprecipitation

ChIP to study binding of GR to predicted GREs was performed as published ChIP to study binding of GR to predicted GREs was performed as published (Sarabdjitsingh et al., 2010a). Briefly, fixed chromatin derived from the hippocampi of three animals was pooled and sheared, yielding fragments of 100–500 bp (20 pulses of 30 sec;

Bioruptor; Diagenode, Liege, Belgium). Immunoprecipitation was performed with either $6\,\mu g$ of GR-specific H300 or normal rabbit IgG (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) overnight at $4\,°C$. Immunoprecipitation with a nonspecific antibody (normal IgG) did not result in increased DNA recovery after treatment and was used to correct the GR immunoprecipitated samples for nonspecific binding. The criterion for binding was a more than 2-fold increase in the yield of the real-time quantitative PCR (RT-qPCR) reaction, compared with the no-hormone condition, and a total recovery of more than $0.1\,\%$ of input material.

Selection of GREs for validation

Out of 183 up-regulated genes, 156 were annotated in all four species that we used for alignment (human, mouse, rat, and cow). GRE predictions were made for these genes, including $50\,kb$ of up- and downstream sequence. Thirty-two predicted GREs from up-regulated genes were selected for validation using RT-qPCR on ChIP material, of which all fitted the criteria of a score above 0.8 in all four species except for two (Adra1d_2 and Slc15a13_3), which were taken along to test how stringent these criteria are. Binding to the metallothionein and myoglobin locus was used as positive and negative control, respectively.

Primer design and RT-qPCR

After DNA recovery (Nucleospin; Macherey-Nagel, Düren, Germany), RT-qPCR was performed *in duplo* to study enrichment of GR-immunoprecipitated DNA fragments harboring the predicted GREs in the different treatment groups. Primers were designed around the *in silico*-predicted GREs using National Center for Biotechnology Information's PrimerBlast and were tested for absence of hairpin formation using Oligo 7. A list of all primers is available in the Table 2.5. RT-qPCR was performed using the LightCycler FastStart DNA Master PLUS SYBR Green I kit (Roche, Indianapolis, IN), according to the manufacturer's instructions.

Motif finding

The flanking sequences of the experimentally tested GREs were screened for additional transcription factor binding motifs. These analyses were done using the motif finding tools MEME (Bailey et al., 2006),MDScan (Liu et al., 2004), and F-Match (Kel et al., 2003). After motif identification, TOMTOM v4.3.0 was used to find corresponding transcription factors in the TRANSFAC database. For MDScan, default settings were used with the following changes: motif width 12 and 5000 nucleotides of mouse intronic sequence (available on the website) was used as background. For both MEME and MDScan, 250 nucleotides left and right of the predicted GRE were used for motif finding. F-Match was used on the BioBase website and is part of the eXPlain v3.0 package. Validation of motif overrepresentation was then determined using 500 nucleotides left and right of predicted GREs.

2.3 Results

Matrix derived from 44 GREs in literature

We used a position-specific scoring matrix based on 44 known GREs from literature (Table 2.4). The resulting sequence logo of these 44 GREs is shown in Figure 2.2.

Prediction of GREs is improved by aligning genomic sequences of multiple species

Because conservation analysis has been shown to predict *in vivo* occupancy of GR-binding sequences at CORT-induced genes (So et al., 2008), we applied an inter-species filter to identify evolutionary-conserved high-scoring GREs in rat, mouse, human, and cow (Figure 2.3). However, the success of this approach is highly de-pendent on a positionally conserved gene structure, in which conserved GREs are present at exactly the same location and not shifted. Because this is often not the case when comparing multiple species, we applied a multiple sequence alignment before scanning for GREs. The effect of this alignment and interspecies filter is ev-ident from the α-1D-adrenergic receptor (*Adra1d*) gene, which shows large inter-species genomic insertions/deletions. Before alignment, the two high-scoring GREs in the different species are highly dispersed, with distances between species differ-ing up to over 20 kb (Table 2.1 and Figure 2.3). However, after alignment, the pre-dicted GREs are located at exactly the same position, thus facilitating their recogni-tion as conserved sites (Table 2.1 and Figure 2.3).

B Relative frequency matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	0.09	0.54	0.63	0.02	0.70	0.20	0.27	0.31	0.06	0.00	0.06	0.11	0.02	0.13
C	0.20	0.18	0.15	0.70	0.11	0.27	0.27	0.34	0.02	0.02	0.02	0.25	0.90	0.25
G	0.65	0.09	0.09	0.18	0.02	0.31	0.18	0.15	0.04	0.86	0.11	0.13	0.02	0.02
T	0.04	0.18	0.11	0.09	0.15	0.20	0.27	0.18	0.86	0.11	0.79	0.50	0.04	0.59

C Log likelihood matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	-1.2630	1.0588	1.2730	-2.5849	1.4150	-0.2630	0.1154	0.3219	-1.5849	-3.5849	-1.5849	-1	-2.5849	-0.7776
C	-0.2630	-0.4150	-0.5849	1.4150	-1	0.1154	0.1154	0.4150	-2.5849	-2.5849	-2.5849	0.0000	1.7725	0.0000
G	1.3219	-1.2630	-1.2630	-0.4150	-2.5849	0.3219	-0.4150	-0.5849	-2	1.7004	-1	-0.7776	-2.5849	-2.5849
T	-2	-0.4150	-1	-1.2630	-0.5849	-0.2630	0.1154	-0.4150	1.7004	-1	1.5849	0.9385	-2	1.1699

Figure 2.2: Graphical representation of the GRE matrix based on 44 proven GREs from literature.
A, Logo representation, in which *letter size* corresponds to the frequency of occurrence of nucleotides at each position (`http://weblogo.berkeley.edu/`). B, The frequency matrix. C, The log-transformed likelihood matrix that was used in the scoring procedure, in which scores were expressed relative to the maximal outcome of the matrix, which was set at the value of 1.

	GRE sequence	Position of GRE before alignment	Position of GRE after alignment	GRE score
Adra1d_1				
Homo Sapiens	gaacaccctgtact	77,873	170,141	0.93
Mus musculus	gaacgccctgtact	53,496	170,141	0.83
Rattus norvegicus	gaacggcctgtacc	58,681	170,141	0.81
Bos taurus	gaacaccctgtact	56,374	170,141	0.93
Adra1d_2				
Homo Sapiens	gaacaggacgtcct	−31,585	−25,038	0.84
Mus musculus	ggacaggatgtcct	−37,462	−25,038	0.89
Rattus norvegicus	ggacaggacgtcct	−43,891	−25,038	0.78
Bos taurus	gaacaagatgcctt	−46,743	−25,038	0.74

Table 2.1: Location of GREs before and after alignment in vicinity of *Adra1d* gene.
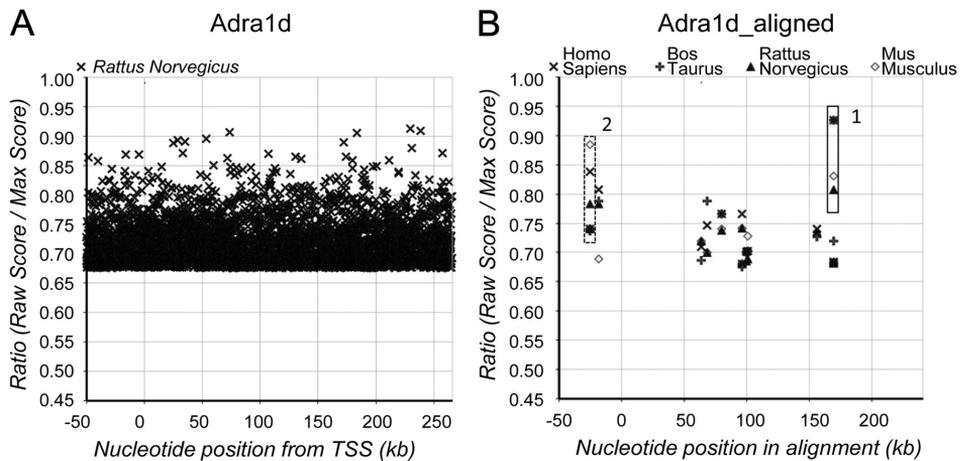
Figure 2.3: Predicted GREs with and without applying an interspecies filter to select for evolutionary-conserved GREs in the *Adra1d* gene.
Boxed scores are the ones aligned to each other. A, GRE matrix scores with a value over 0.65 in rat, plotted relative to the transcription start site of the gene. There are many GREs with a score above 0.8. B, After alignment of the sequences of multiple species and selection for evolutionary-conserved GREs, most of the predicted sites in the rat are lost, leaving two conserved GREs.

Prediction of evolutionary-conserved GREs in CORT-responsive genes in the hippocampus

The 20 selected genes up-regulated by CORT in the hippocampus contained a total of 1614 GREs with a matrix score above 0.8. Adding the demand of conservation of the score being more than 0.8 in all four species strongly reduced the amount of predictions to 32 (Table 2.2). The number of predicted GREs also decreased dramatically if the threshold was raised to 0.85 or 0.9. The 32 evolutionary-conserved GREs with a score above 0.8 were validated experimentally for GR binding. The GRE sequences for the different species are listed in Table 2.6.

Highest score	No. of predicted GREs			No. of conserved GREs		
	>0.9	>0.85	>0.8	>0.9	>0.85	>0.8
Abhd14a	4	15	56	0	1	1
Akap7	3	20	120	0	0	1
Arhgef3	3	15	93	1	1	1
Daam1	2	25	135	0	0	1
Ddit4	2	18	74	1	1	2
Dgat2	5	14	87	0	0	1
Errfi1	0	11	53	0	0	1
Fam55c	1	10	66	0	1	1
Fkbp5	6	19	122	1	1	2
Kcnj11	1	15	63	0	0	2
Klf9	1	16	70	0	1	2
Lyve1	1	8	49	0	0	1
Mfsd2	1	7	63	0	2	2
Msx1	1	9	60	1	1	1
Slc25a13	3	20	143	0	1	1
Srxn1	3	10	58	0	1	3
Tiparp	1	17	62	0	1	4
Tle3	2	12	82	0	0	4
Zfyve28	1	17	111	0	0	1
Znf592	1	9	47	0	1	2
Total	**42**	**287**	**1,614**	**4**	**13**	**34**
Average per gene	2.1	14.4	80.7	0.2	0.7	1.7

Table 2.2: Number of predicted GREs in 20 selected genes before and after selection for evolutionary conserved sequences.

In vivo GR occupancy of predicted GREs in hippocampus

RT-qPCR analysis on ChIP material enriched for GR binding allowed confirmation which of the predicted GREs were bound by GR *in vivo* in the rat hippocampus. Of the 32 tested GREs, GR binding could be shown for 15. Interestingly, the two GREs that did not fully fit the criteria could not be validated. For example, of the two predicted GREs in *Adra1d*, only Adra1d_1, which fitted the criteria, showed GR binding in the tissue and under the conditions we tested in this study (Figure 2.4). Validation of the predicted GREs using ChIP/RT-qPCR *in vivo* in hippocampal rat neurons resulted in a success rate of nearly 50 %. In other words, by applying specific criteria (up-regulation of the gene, evolutionary conservation, and score at least 0.80 in four different species) we can predict with 50 % accuracy GR-binding sites in the hippocampus (Figure 2.5).

Analyzing GRE-flanking regions for conserved motifs

We next compared the sequences and flanking regions of the *bona fide* hippocampal GREs with the predictions that could not be validated. The actual sequence, score, or the extent of conservation was not different. Although the highest fold
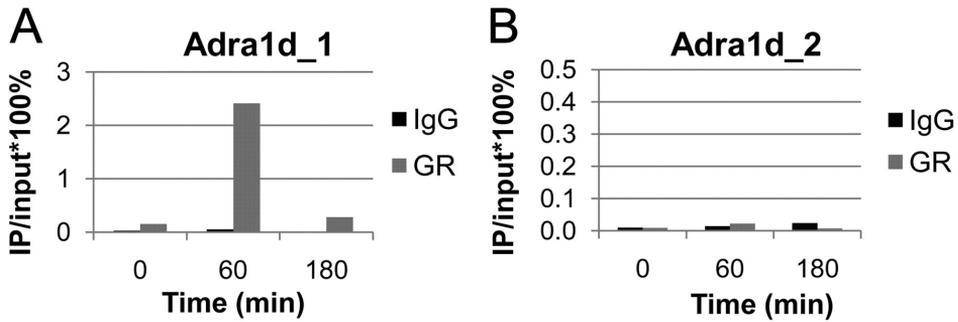
Figure 2.4: ChIP analysis of two predicted GREs indicated in Figure 2.3.
A, Strong enrichment of sequence Adra1d_1 after immune precipitation with a GR antibody of hippocampus material $60\,\text{min}$ after glucocorticoid treatment, compared with $t = 0$ and $t = 180\,\text{min}$ and compared with control IgG. B, Lack of binding to the Adra1d_2 sequence in the same material.
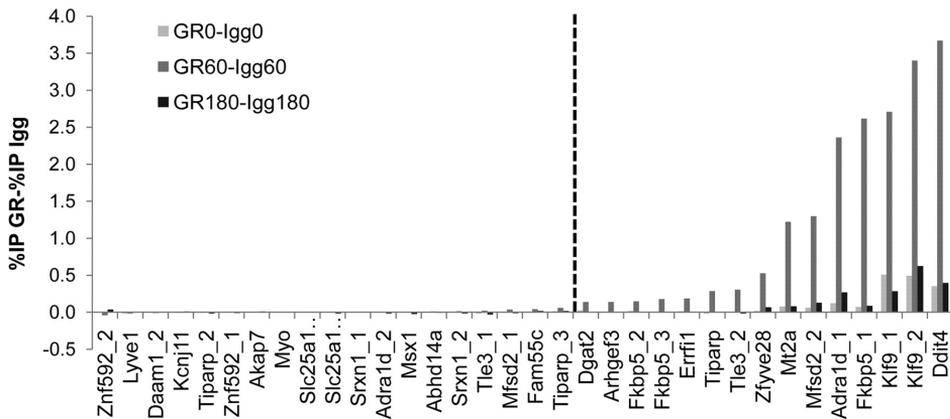
Figure 2.5: Binding profiles for GR on 32 predicted and two control sequences at three time points after CORT injection, expressed as percentage of input material.
Cut-off for enrichment was set at $0.1\,\%$ of input material and enrichment of a factor 2 relative to the time point $0\,\text{min}$. The predicted GREs are ordered by magnitude of GR binding at timepoint $(t) = 60$ (*red bars*), after correction for IgG binding. At $t = 180\,\text{min}$, GR-binding levels are comparable with those before GR activation. Seventeen sequences (*right from dashed line*) were found to be enriched for GR binding at $t = 60$ Myoglobin (Myo) functioned as negative control, the GRE controlling the metallothionein gene (Mt2a) as the positive control.

enrichment tended to be on GREs that were completely conserved (five out of six highest enrichment values), there were also three fully conserved predicted GREs with high scores that showed no enrichment at all. Closer inspection of the flanking sequence of GR-binding and GR-nonbinding GREs showed that they differed strikingly. Scanning 250 nucleotides up- and downstream of the GR-binding GREs showed specific enrichment of a number of predicted binding sites for enriched for the nucleotides cytosine (C) and guanine (G), such as Myc-associated zinc finger protein 1 (MAZ1), Specificity Protein 1 (SP1), Wilms' tumor 1, and zinc finger protein (Znf) 219 (Figure 2.6, A–F). More complete scanning of the sequences around the

GRE showed a higher presence of predicted MAZ1 and SP1 sites up to a distance of about $2{,}000\,\text{bp}$, with a bias at the $5'$ end. These binding sites correspond with a general increase in GC-motifs in these areas. The sequence motifs to which SP1 and MAZ1 bind are shown in Figure 2.7. As a second control for specificity, we checked for general increase in transcription factor-binding sites by scoring nuclear factor κB sites, which were not different between GR-binding and GR-nonbinding GRE sequences (Figure 2.6, G and H). The enrichment was not related to distance to the transcription start site or extent of conservation (data not shown). Most interestingly, the signature was not detected around the bound GREs from the So dataset derived from mesenchymal stem-like cells (Figure 2.8).

2.4 Discussion

The aim of this study was to identify genes regulated by direct GR-GRE binding in the brain based on *in silico* GRE screening of CORT-responsive genes. By doing so, we revealed that GR-binding sequences differ from nonbinding sequences by the nearby presence of predicted GC-rich binding sites for transcription factors such as MAZ1 and SP1. This characteristic of binding was found to be absent in another dataset with GR-binding and GR-nonbinding sites, suggesting a mechanism for tissue-specific CORT signaling that may determine GRE usage in the hippocampus.

Importance and validity of alignment and matrix

Because there are substantial differences in genomic organization of genes and their flanking regions between species, proper alignment facilitates screening for evolutionary conservation of GREs. Although in most cases alignment works well, we cannot exclude that some conserved sequences were missed due to suboptimal alignment by the available algorithms.

Previous papers strongly suggested that evolutionary conservation (up to $11/15$ nucleotides) is an important predictor of GRE functionality (Reddy et al., 2009; So et al., 2007). The strength of conservation analysis for this sequence is demonstrated by the striking difference in numbers of predicted GREs with and without screening for evolutionary conservation. In the 20 genes listed in Table 2.2, there is a total of 1614 predicted GREs with a score in rat above 0.8. This number drops dramatically to only 32 GREs that survive the evolutionary filter, which is much more in a realistic range. Almost half of these conserved GREs can be validated, confirming that evolutionary conservation has an important predictive value for GR binding.

The matrix that was used for identifying GREs was based on 44 GREs from literature, with proven GR binding in either EMSA or deoxyribonuclease footprint assays. These GREs represent different species, different responsive tissues, and have a bias for sequences proximal to the transcription start sites. There are minor differences
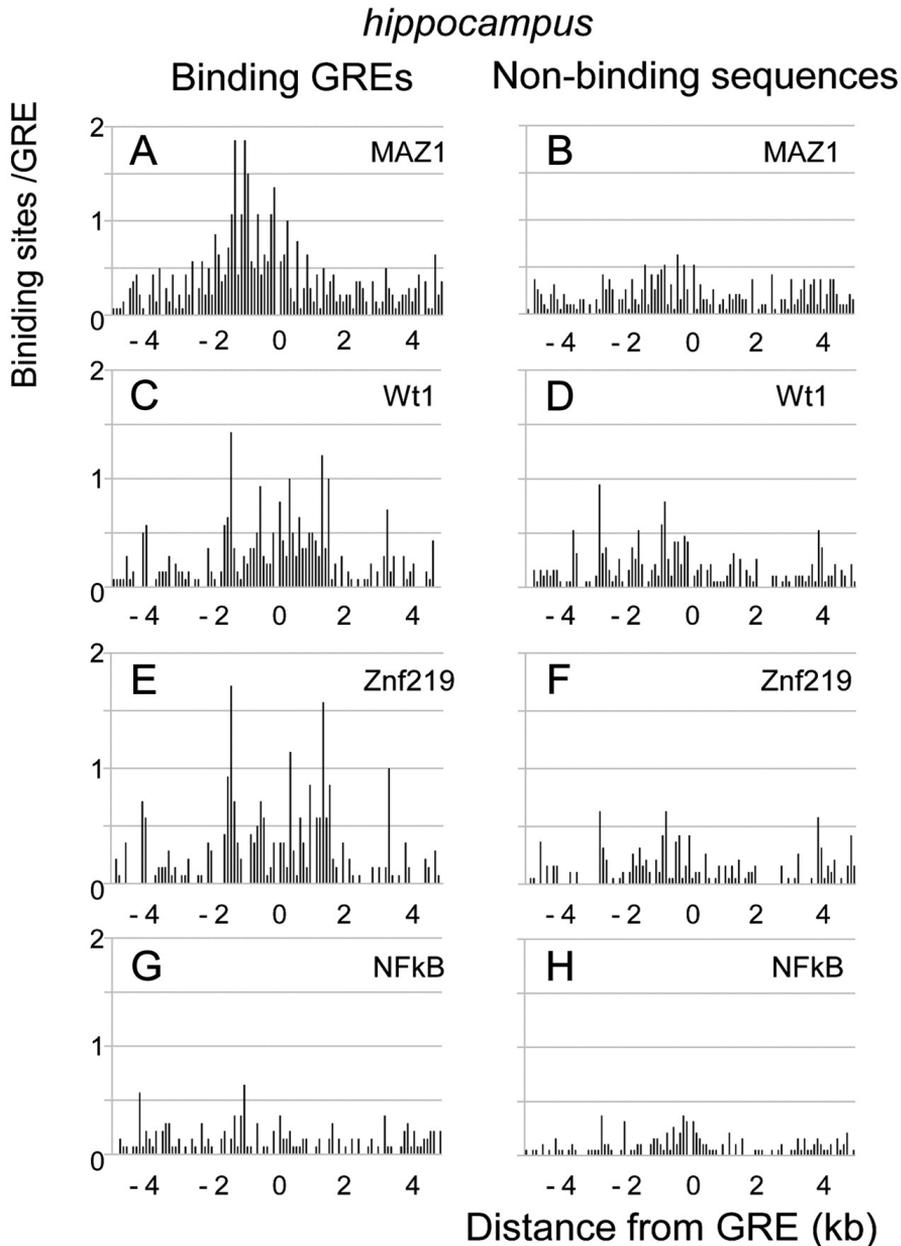
Figure 2.6: Presence of predicted transcription factor-binding sites surrounding true GREs and nonvalidated GREs from rat hippocampus.

Expressed is the occurrence of sites per GRE sequence at particular distances from the GRE sequence. The comparison per transcription factor is between validated GREs (A, C, E, and G) and nonbinding GRE sequences (B, D, F, and H). Binding sites for GC binders, such as MAZ1 (A and B), Wilms' tumor 1 (C and D), and Znf219 (E and F), are enriched up to 2 kb from the GRE, with a tendency for skew on the 5′ site. Nuclear factor κB (NFκB) response element consensus frequency (G and H) did not differ between GR-binding and GR-nonbinding GREs.
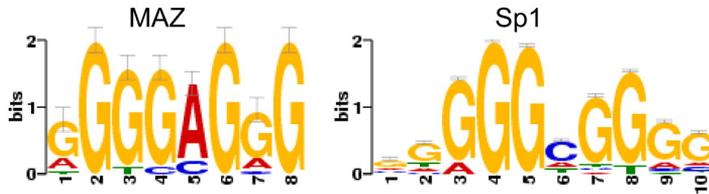
Figure 2.7: Sequence motif of MAZ1 and SP1 transcription factor-binding sites.
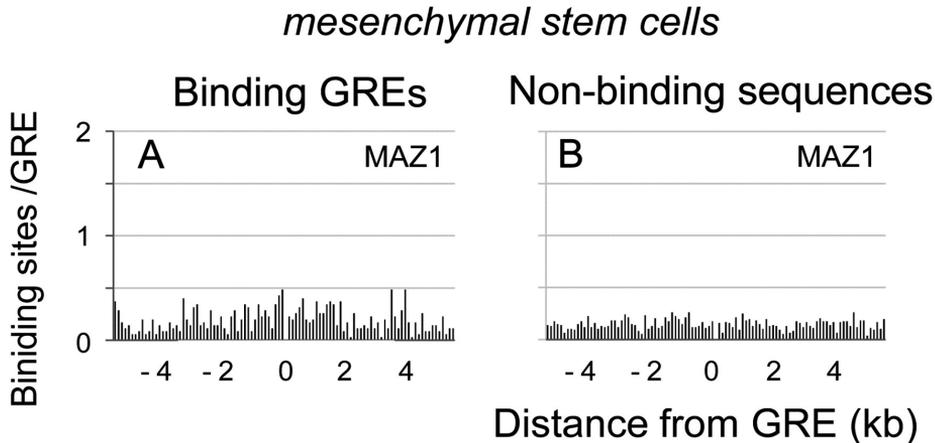
Figure 2.8: MAZ1 sites are not enriched around GR-binding GREs in mesenchymal stem cells. No differences are found for MAZ1 site occurrence between GR-binding (A) and GR-nonbinding (B) GRE sequences.

with matrices derived from large-scale chromatin occupation studies in cell lines (Reddy et al., 2009), which may be due to cell type-specific characteristics of GR binding. However, despite these points, our matrix clearly is suited to predict a substantial number of GREs in the brain, located at considerable distances from the transcription start site of CORT-responsive genes, as demonstrated by the successful validation of 15 GREs consisting of at least 10 novel previously described GREs.

False negatives

There is undoubtedly a number of false negative findings in our GRE scoring. First, the score threshold that we used, requiring a score above 0.8, may be too stringent, thus missing some *bona fide* GREs. Conversely, lowering the threshold below 0.8 results in a number of GRE predictions that likely includes many false positives. Similarly, the requirement for a GRE to be conserved in four different species may also result in missing some GREs, because evolutionary conservation may not be a good predictor for all transcription factor-binding sites (Schmidt et al., 2010). Furthermore, the matrix that we used is not particularly suited for identifying nontypical GREs that deviate from the consensus, despite evidence for direct GR binding

(Costeas and Chinsky, 2000; Kooi van der et al., 2005). Although we included a number of these nontypical GREs in our matrix, their contribution to the matrix is too small to adequately identify such sequences in the scoring procedure. In addition, for those responsive genes in which no GRE could be identified, we simply may have not scanned the relevant DNA region. GREs have been shown to occur at distances up- and downstream of transcription start sites that are even further than the $50\,\mathrm{kb}$ that we used here (Reddy et al., 2009; So et al., 2007).

False positives

Overall, $47\,\%$ (15 out of 32) of the predicted and selected GREs showed GR binding in the hippocampus of rats after administering CORT. Because the GR-GRE interaction and consequently GR-driven gene expression occurs in a cell type-specific way, it is likely that several of the 17 GREs for which we did not observe any GR binding could very well bind GR in other tissues (*e.g.* the perfectly conserved predicted GRE in the *msx* gene, which has an almost maximal matrix score). Chromatin organization controls GRE availability in a number of ways (Biddie et al., 2010), and work on the estrogen receptor and GR has indeed shown considerable cell-specific variation in response to element use (Krum et al., 2008). Future work should elucidate whether nonbinding sequences lack the necessary accessory sites or whether those GRE sequences are inaccessible due to epigenetic regulation. An additional issue is that some of the GREs that we selected for validation may not functionally be associated with the responsive genes in the hippocampus but rather to another gene.

In addition to cell specificity, the kinetics of glucocorticoid signaling may be a basis for elements that came up as false positive, because responses range from immediate early responses (Gemert van et al., 2006) to slower continuous induction (John et al., 2009). Lastly, there may also be "true false positives": sequences that we assign as GREs but that may not bind GR in any tissue or circumstance but rather related nuclear receptors, such as MR, androgen, and progesterone receptors (Nelson et al., 1999), or that have different reasons for evolutionary conservation.

Analyzing GRE-flanking regions

Although the GRE sequence itself may contain information relevant to epigenetic mechanisms (Biddie et al., 2010), we found no relation with responsiveness. Because the exact sequence of the GRE did not distinguish binding from nonbinding, the context of the surrounding sequence may be of relevance (So et al., 2008). Indeed, the binding of GREs could be linked to the presence of MAZ1 and SP1-binding sites. The presence of additional motifs in the flanking sequences of hormone response elements has been reported before (Carroll et al., 2006; Phuc Le P. et al., 2005), but its tissue specificity is less commonly reported. In the current study, we identified an overrepresentation of consensus sites for several GC-box binders, indicating the

presence of a GC-rich area in the flanking region of a substantial part of the GR-binding GREs. The presence of SP1-binding sites surrounding GREs was previously also reported in human A549 lung carcinoma cells (Reddy et al., 2009). Because we did not find an overrepresentation of GC-box transcription factor motifs in the non-validated genes in this study or in the vicinity of the GR-binding GREs identified by So *et al.* (So et al., 2008) in mesenchymal stem cells, we suggest that GC-boxes may play a role in determining tissue specificity of GR binding to a defined group of GREs. Interestingly, a screen on GR-binding sites in mouse liver pointed to enrichment of CCAAT-enhancer-binding protein (C/EBP), rather than GC binders (Phuc Le P. et al., 2005). A recent screen in two mouse cell lines found different motifs associated with GR binding, which were however partly exclusive rather than accessory to GRE (John et al., 2011). Whether such transcription factors determine binding site availability, or the nature of transcriptional responses once GR has bound, remains to be determined. Although both GR- and GC-binding transcription factors are ubiquitously expressed, the combined action in particular target genes may be part of a combinatorial code for specific responses to stress. Irrespective of the exact binding factors, the GC-rich area could be used as an extra criterion in predicting to which GRE GR binds in specific tissues, such as for example the hippocampus.

As a start to further investigate candidate binding factors to the recognized motifs, we queried expression data from the Allen Mouse Brain Atlas (Lein et al., 2007). *MAZ1* had the highest hippocampal expression level compared with the other identified motif-associated proteins. Other factors were expressed at lower levels (*e.g.* SP1 and Znf219) or nondetectable in the brain [Zic family member 3 (Zic3) and zinc finger and BTB domain containing 7B (Zbtb7b)]. MAZ protein is a broadly expressed Cys2His2 zinc finger protein that can interact with SP1 at the same GC-rich binding sites (Song et al., 2001; Song et al., 2003). Among their common target genes are the *N*-methyl-D-aspartic acid (*NMDA*) (Okamoto et al., 2002) and the adrenal medulla glucocorticoid responsive phenylethanolamine *N*-methyltransferase (*PNMT*) gene (Her et al., 2003). Interestingly, the SP family of proteins has been implicated as integratory factors in gene regulation associated with other hormonal signaling pathways (Solomon et al., 2008).

2.5   Conclusion

Using a matrix of 44 published GREs, we have successfully identified 15 GREs that are bound by GR in the rat hippocampus, of which at least 10 are novel. Furthermore, we have identified a signature that distinguishes GR-binding from GR-nonbinding GRE sites in the hippocampus but not in mesenchymal stem cells. This signature is a GC-box, to which transcription factors such as SP1 and MAZ1 can bind. Analysis of additional datasets is essential to further elucidate whether this motif plays a role in determining tissue specificity of GR-responsive transactivated genes. In addition, ChIP analysis with antibodies directed at members of the SP1 family and MAZ proteins could help to further identify exactly which cross talk partners are active in conjunction with GR. We view our current finding as a first step toward understanding the direct downstream pathways of GR signaling in the brain.

Table 2.3: List of 183 genes upregulated by CORT in the dentate gyrus region of the hippocampus.
Please note: some genes are represented by multiple probe sets.

Probe set	Gene Symbol	Gene Title	Parametric p-value	Direction of regulation by CORT
1396113_at	Abhd14a	abhydrolase domain containing 14A	0.0051565	up
1368534_at	Adra1d	adrenergic receptor, alpha 1d	0.0001817	up
1382272_at	Agtrap	angiotensin II, type I receptor-associated protein	2E−006	up
1373078_at	Ahcyl2	S-adenosylhomocysteine hydrolase-like 2	0.0043009	up
1389496_at	Akap7	A kinase (PRKA) anchor protein 7	0.0008697	up
1368365_at	Aldh3a2	aldehyde dehydrogenase family 3, subfamily A2	2.2E−006	up
1373250_at	Anln	anillin, actin binding protein (scraps homolog, Drosophila)	3E−007	up
1391673_at	Arhgap20	Rho GTPase activating protein 20	0.00243	up
1377750_at	Arhgef3_predicted	Rho guanine nucleotide exchange factor (GEF) 3 (predicted)	5.36E−005	up
1368563_at	Aspa	aspartoacylase	2.29E−005	up
1374539_at	Atp10d	ATPase, class V, type 10D	0.0037805	up
1375030_at	B3galt5_predicted	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 5 (predicted)	< 1E−07	up
1374323_at	Bccip_predicted	BRCA2 and CDKN1A interacting protein (predicted)	0.0065831	up
1379368_at	Bcl6_predicted	B-cell leukemia/lymphoma 6 (predicted)	1.1E−006	up
1394375_x_at	Bcl6b	B-cell CLL/lymphoma 6, member B	0.0007852	up
1381804_at	Bcl6b	B-cell CLL/lymphoma 6, member B	1.3E−006	up
1386833_at	Bcl6b	B-cell CLL/lymphoma 6, member B	3.3E−006	up
1373733_at	Bok	Bcl-2-related ovarian killer protein	0.0092297	up
1372855_at	Brd4	Bromodomain containing 4	0.0084915	up
1367657_at	Btg1	B-cell translocation gene 1, anti-proliferative	0.001966	up
1368393_at	C1qr1	complement component 1, q subcomponent, receptor 1	0.0012737	up
1375353_at	Cables1_predicted	Cdk5 and Abl enzyme substrate 1 (predicted)	0.0086606	up
1381637_at	Camk2a	Calcium/calmodulin-dependent protein kinase II, alpha	0.0071659	up
1388736_at	Ccdc43	coiled-coil domain containing 43	0.0031497	up
1384192_at	Chst1	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	0.0052078	up
1396150_at	Cldn1	claudin 1	0.003693	up
1372774_a t	Coq6	Coenzyme Q6 homolog (yeast)	0.0011843	up
1372629_at	Coro2b	coronin, actin binding protein, 2B	4.93E−005	up
1384454_at	Cpa6_predicted	carboxypeptidase A6 (predicted)	0.0025525	up
1398611_at	Cul4b_predicted	cullin 4B (predicted)	0.0073757	up
1367940_at	Cxcr7	chemokine (C-X-C motif) receptor 7	3.3E−005	up
1386904_a_at	Cyb5	cytochrome b-5	0.0004326	up
1389294_at	Cyfip1_predicted	cytoplasmic FMR1 interacting protein 1 (predicted)	0.0027769	up
1389318_at	Daam1_predicted	dishevelled associated activator of morphogenesis 1 (predicted)	0.0002217	up
1374480_at	Daam1_predicted	dishevelled associated activator of morphogenesis 1 (predicted)	0.0017818	up
1384788_at	Daglb	diacylglycerol lipase, beta	0.0010505	up
1368025_at	Ddit4	DNA-damage-inducible transcript 4	2.21E−005	up
1380817_at	Depdc2_predicted	DEP domain containing 2 (predicted)	8.56E−005	up
1389615_at	Derl1	Der1-like domain family, member 1	0.0033244	up

Probe set	Gene Symbol	Gene Title	p-value	Direction
1371615_at	Dgat2	diacylglycerol O-acyltransferase 2	3.86E−005	up
1368189_at	Dhcr7	7-dehydrocholesterol reductase	0.0010634	up
1367516_at	Dtnbp1	distrobrevin binding protein 1	0.0020268	up
1370830_at	Egfr	epidermal growth factor receptor	0.0005467	up
1391442_at	Ehd3	EH-domain containing 3	0.0033817	up
1373093_at	Errfi1	ERBB receptor feedback inhibitor 1	1.27E−005	up
1389146_at	Fam107b	family with sequence similarity 107, member B	0.0004397	up
1398425_at	Fam110b	family with sequence similarity 110, member B	0.0008749	up
1385046_at	Fam55c /// LOC682630	family with sequence similarity 55, member C /// hypothetical protein LOC682630	0.0012999	up
1374255_at	Farsla	Phenylalanine-tRNA synthetase-like, alpha subunit	0.0073605	up
1387351_at	Fbn1	fibrillin 1	0.000529	up
1368829_at	Fbn1	fibrillin 1	0.0035091	up
1387606_at	Fgf2	fibroblast growth factor 2	0.0007805	up
1388901_at	Fkbp5	FK506 binding protein 5	2E−007	up
1380611_at	Fkbp5	FK506 binding protein 5	7.51E−005	up
1372016_at	Gadd45b	growth arrest and DNA-damage-inducible 45 beta	0.0040849	up
1368577_at	Gjb6	gap junction membrane channel protein beta 6	0.0013882	up
1371363_at	Gpd1	glycerol-3-phosphate dehydrogenase 1 (soluble)	2.03E−005	up
1374648_at	Gpr155_predicted	G protein-coupled receptor 155 (predicted)	0.0003912	up
1388243_at	Gpr176	G protein-coupled receptor 176	0.0004745	up
1374043_at	Gramd3	GRAM domain containing 3	0.0027761	up
1367900_at	Gyg1	glycogenin 1	0.0047381	up
1370491_a_at	Hdc	histidine decarboxylase	0.0001866	up
1373963_at	Hdhd3	haloacid dehalogenase-like hydrolase domain containing 3	0.0073321	up
1374440_at	Hsd17b11	hydroxysteroid (17-beta) dehydrogenase 11	0.0004592	up
1370912_at	Hspa1b	heat shock 70kD protein 1B (mapped)	0.0053864	up
1382220_at	Igf2bp2	insulin-like growth factor 2 mRNA binding protein 2	0.0045324	up
1376895_at	Il16	interleukin 16	0.0048023	up
1373970_at	Il33	interleukin 33	0.0002442	up
1386987_at	Il6ra	interleukin 6 receptor, alpha	1E−006	up
1371091_at	Irs2	insulin receptor substrate 2	0.0039053	up
1383082_at	Jarid1b	jumonji, AT rich interactive domain 1B (Rbp2 like)	0.0029768	up
1390473_at	Kcng2	potassium voltage-gated channel, subfamily G, member 2	0.0007367	up
1387698_at	Kcnj11	potassium inwardly rectifying channel, subfamily J, member 11	0.0013338	up
1391007_s_at	Kcnj11	potassium inwardly rectifying channel, subfamily J, member 11	0.0052578	up
1370209_at	Klf9	Kruppel-like factor 9	0.0003826	up
1373210_at	Lamb1	laminin, beta 1	0.0025957	up
1368006_at	Laptm5	lysosomal-associated protein transmembrane 5	0.0005548	up
1383863_at	Lmo2	LIM domain only 2	7.91E−005	up
1397439_at	LOC497978	similar to diacylglycerol kinase epsilon	0.0035734	up
1372973_at	Lss	Lanosterol synthase	0.0074712	up
1367832_at	Lypla1	lysophospholipase 1	0.0068274	up
1382192_at	Lyve1	lymphatic vessel endothelial hyaluronan receptor 1	0.0004275	up
1371875_at	Manba	mannosidase, beta A, lysosomal	0.0073347	up
1390905_at	Mast4	microtubule associated serine/threonine kinase family member 4	0.0005155	up
1388774_at	Mbd2	methyl-CpG binding domain protein 2	8.17E−005	up

Probe set	Gene Symbol	Gene Title	p-value	Direction
1372966_at	Mfsd2	major facilitator superfamily domain containing 2	0.0009973	up
1372599_at	Mgst2_predicted	microsomal glutathione S-transferase 2 (predicted)	0.0085716	up
1383952_at	Mical1_predicted	microtubule associated monoxygenase, calponin and LIM domain containing 1 (predicted)	3.89E−005	up
1389433_at	Mkks	McKusick-Kaufman syndrome protein	0.0090644	up
1373189_at	Mkl1	megakaryoblastic leukemia (translocation) 1	0.0013132	up
1376410_at	Mmp17_predicted	matrix metallopeptidase 17 (predicted)	4.56E−005	up
1382363_at	Mpp5_predicted	membrane protein, palmitoylated 5 (MAGUK p55 subfamily member 5) (predicted)	0.0065158	up
1368302_at	Msx1	homeo box, msh-like 1	0.0044827	up
1371237_a_at	Mt1a	metallothionein 1a	0.0040924	up
1397644_at	Mtap_predicted	Methylthioadenosine phosphorylase (predicted)	0.00275	up
1371543_at	Mtmr2_predicted	myotubularin related protein 2 (predicted)	0.0020394	up
1394182_at	Mtmr4_predicted	myotubularin related protein 4 (predicted)	0.0065778	up
1372093_at	Mxi1	Max interacting protein 1	0.001482	up
1388139_at	Myh2	myosin, heavy polypeptide 2, skeletal muscle, adult	0.0034219	up
1387004_at	Nbl1	neuroblastoma, suppression of tumorigenicity 1	0.0002314	up
1389507_at	Nedd4l	neural precursor cell expressed, developmentally down-regulated gene 4-like	1.73E−005	up
1370408_at	Nid67	putative small membrane protein NID67	0.0080132	up
1395408_at	Nostrin	nitric oxide synthase trafficker	0.0027857	up
1390828_at	Npy1r	neuropeptide Y receptor Y1	0.0002798	up
1387497_at	Npy5r	neuropeptide Y receptor Y5	0.0053514	up
1373577_at	Nrp1	Neuropilin 1	0.0064527	up
1384112_at	Nt5e	5′ nucleotidase, ecto	0.0007823	up
1369969_at	Parp1	poly (ADP-ribose) polymerase family, member 1	0.0023734	up
1393454_at	Pcdh17_predicted	protocadherin 17 (predicted)	0.0028394	up
1384509_s_at	Pcdh17_predicted	protocadherin 17 (predicted)	0.0028995	up
1368262_at	Phlpp	PH domain and leucine rich repeat protein phosphatase	3.2E−006	up
1368119_at	Pib5pa	phosphatidylinositol (4,5) bisphosphate 5-phosphatase, A	0.0010451	up
1384741_at	Pla2g3_predicted	phospholipase A2, group III (predicted)	0.0002752	up
1368700_at	Plcl1	phospholipase C-like 1	0.0052902	up
1380661_at	Pld3	phospholipase D family, member 3	0.0040563	up
1384355_at	Plxna2_predicted	plexin A2 (predicted)	0.0021167	up
1392157_at	Plxna2_predicted	plexin A2 (predicted)	0.0061316	up
1382604_at	Polr3g	polymerase (RNA) III (DNA directed) polypeptide G	0.0046518	up
1381386_at	Pop5_predicted	Processing of precursor 5, ribonuclease P/MRP family (S. cerevisiae) (predicted)	0.000567	up
1391187_at	Ppl_predicted	periplakin (predicted)	0.002245	up
1373465_at	Pqlc1	PQ loop repeat containing 1	7.7E−006	up
1372135_at	Ptk9l_predicted /// LOC684352	protein tyrosine kinase 9-like (A6-related protein) (predicted) /// similar to twinfilin-like protein	0.0033	up
1378541_at	Pus7l_predicted	pseudouridylate synthase 7 homolog (S. cerevisiae) like (predicted)	0.0013239	up
1383232_at	Rab33b_predicted	RAB33B, member of RAS oncogene family (predicted)	0.0012736	up
1395326_at	Rbm9_predicted	RNA binding motif protein 9 (predicted)	0.0007839	up
1393502_at	RGD1306153	similar to predicted CDS, putative protein of bilate-rial origin (4J193)	0.0006468	up
1391239_at	RGD1306926 _predicted	similar to hypothetical protein FLJ22175 (predicted)	0.0015965	up

Probe set	Gene Symbol	Gene Title	p-value	Direction
1377524_at	RGD1307155	similar to CG18661-PA	0.0096282	up
1374176_at	RGD1308059	similar to DNA segment, Chr 4, Brigham & Womens Genetics 0951 expressed	0.0053442	up
1372420_at	RGD1308064_predicted	similar to FKSG24 (predicted)	0.0006842	up
1372843_at	RGD1309410_predicted	LOC363020 (predicted)	0.0032345	up
1374596_at	RGD1309594	similar to RIKEN cDNA 1810043G02; DNA segment, Chr 10, Johns Hopkins University 13, expressed	0.0096133	up
1372805_at	RGD1310444_predicted	LOC363015 (predicted)	0.0011833	up
1382097_at	RGD1310754_predicted	similar to G2 (predicted)	0.0006971	up
1388945_at	RGD1311307	similar to 1300014I06Rik protein	1E−007	up
1383874_at	RGD1560812_predicted	RGD1560812 (predicted)	0.0024907	up
1373075_at	RGD1560888_predicted	similar to Cell division protein kinase 8 (Protein kinase K35) (predicted)	0.0019401	up
1390964_at	RGD1561115_predicted	similar to Gene model 1568 (predicted)	0.0088555	up
1381924_at	RGD1561507_predicted	similar to hypothetical protein FLJ31606 (predicted)	0.0052705	up
1378310_at	RGD1562710_predicted	similar to neuromedin B precursor - rat (predicted)	4.29E−005	up
1379816_at	RGD1563342_predicted	similar to RIKEN cDNA 2410025L10 (predicted)	0.000553	up
1376809_at	RGD1563342_predicted	similar to RIKEN cDNA 2410025L10 (predicted)	0.0016411	up
1379077_at	RGD1564695_predicted	similar to A830059I20Rik protein (predicted)	0.0041516	up
1375151_at	RGD1565168_predicted	Similar to RAP2A, member of RAS oncogene family (predicted)	0.0047119	up
1390942_at	RGD1565884_predicted	Similar to Pellino protein homolog 2 (Pellino 2) (predicted)	1.03E−005	up
1391075_at	Rgs17_predicted	regulator of G-protein signaling 17 (predicted)	0.0065764	up
1388937_at	Rnf19a	ring finger protein 19A	0.0001664	up
1378524_at	Rnf19a	ring finger protein 19A	3.35E−005	up
1368662_at	Rnf39	ring finger protein 39	0.0032395	up
1389202_at	Rpe	ribulose-5-phosphate-3-epimerase	0.003193	up
1371774_at	Sat1	spermidine/spermine N1-acetyl transferase 1	0.0064071	up
1389367_at	Schip1	schwannomin interacting protein 1	0.0010009	up
1388334_a t	Sec14l1	SEC14-like 1 (S. cerevisiae)	0.0001474	up
1373610_at	Sec24d_predicted	SEC24 related gene family, member D (S. cerevisiae) (predicted)	0.0042078	up
1387294_at	Sh3bp5	SH3-domain binding protein 5 (BTK-associated)	0.0015789	up
1376040_at	Sipa1l2	signal-induced proliferation-associated 1 like 2	0.0010188	up
1378356_at	Slc24a4_predicted	solute carrier family 24 (sodium/potassium /calcium exchanger), member 4 (predicted)	0.0056586	up
1389622_at	Slc25a13	solute carrier family 25 (mitochondrial carrier, adenine nucleotide translocator), member 13	0.0001236	up
1392978_at	Slc25a28	solute carrier family 25, member 28	0.0050879	up
1370848_at	Slc2a1	solute carrier family 2 (facilitated glucose transporter), member 1	0.008105	up
1382136_at	Slc2a9	solute carrier family 2 (facilitated glucose transporter), member 9	0.001057	up

Chapter 2

Probe set	Gene Symbol	Gene Title	p-value	Direction
1373565_at	Smarca4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	0.0021364	up
1370159_at	Smarcd2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 2	0.0006773	up
1370049_at	Smpd2	sphingomyelin phosphodiesterase 2, neutral	0.00024	up
1376649_at	Snf1lk2_predicted	SNF1-like kinase 2 (predicted)	0.0001025	up
1394627_at	Snx19_predicted	sorting nexin 19 (predicted)	0.0097343	up
1372633_at	Spg20	spastic paraplegia 20, spartin (Troyer syndrome) homolog (human)	0.0006206	up
1383839_at	Spg20	spastic paraplegia 20, spartin (Troyer syndrome) homolog (human)	0.0079955	up
1372510_at	Srxn1	sulfiredoxin 1 homolog (S. cerevisiae)	2.88E−005	up
1387705_at	Sstr4	somatostatin receptor 4	0.0016267	up
1376572_a_at	Svil_predicted	supervillin (predicted)	0.002046	up
1388679_at	Tbc1d14	TBC1 domain family, member 14	3.8E−006	up
1375074_at	Tbkbp1	TBK1 binding protein 1	0.0078674	up
1367859_at	Tgfb3	transforming growth factor, beta 3	7.96E−005	up
1374446_at	Tiparp_predicted	TCDD-inducible poly(ADP-ribose) polymerase (predicted)	0.0022931	up
1385407_at	Tiparp_predicted	TCDD-inducible poly(ADP-ribose) polymerase (predicted)	0.0086341	up
1387169_at	Tle3	transducin-like enhancer of split 3, homolog of Drosophila E(spl)	0.0039518	up
1368136_at	Tmpo	thymopoietin	5.79E−005	up
1372664_at	Traf2_predicted	Tnf receptor-associated factor 2 (predicted)	0.0054778	up
1397596_at	Trim2	tripartite motif protein 2	0.0012484	up
1375278_a t	Trim2	tripartite motif protein 2	0.0013462	up
1373578_at	Trim2	tripartite motif protein 2	7.47E−005	up
1392972_at	Trio	triple functional domain (PTPRF interacting)	0.0005398	up
1390709_at	Trio	triple functional domain (PTPRF interacting)	7.56E−005	up
1369164_a_at	Trpc4	transient receptor potential cation channel, subfamily C, member 4	0.0082294	up
1376262_at	Uxs1	UDP-glucuronate decarboxylase 1	1.03E−005	up
1370648_a_at	Wipf3	WAS/WASL interacting protein family, member 3	0.0024273	up
1385275_at	Wnt16	wingless-related MMTV integration site 16	0.003533	up
1368641_at	Wnt4	wingless-related MMTV integration site 4	0.0044844	up
1370537_at	Xrcc6	X-ray repair complementing defective repair in Chinese hamster cells 6	0.0081922	up
1372989_at	Zdhhc14	zinc finger, DHHC domain containing 14	0.0001701	up
1376628_at	Zfp189_predicted	zinc finger protein 189 (predicted)	5E−007	up
1391216_at	Zfp509_predicted	zinc finger protein 509 (predicted)	0.0068222	up
1393572_at	Zfp592_predicted	zinc finger protein 592 (predicted)	0.0040185	up
1393556_at	Zfyve28_predicted	zinc finger, FYVE domain containing 28 (predicted)	0.0030633	up
1391478_at	Znf532_predicted	zinc finger protein 532 (predicted)	0.0033106	up

Table 2.4: Proven GREs from literature used to construct a GRE position weight matrix.

#	ID	GRE sequence	Aligned sequence	BP pos.	Symbol	Gene
1	1_1	AGAACAGA-GTGTCCTC	gaacagagtgtcct	−525	pnmt	Phenylethanolamine N-Methyl-transferase PNMT [1]
2	1_2	GGAACATC-CTGAACTA	gaacatcctgaact	−714	pnmt	Phenylethanolamine N-Methyl-transferase PNMT [1]
3	1_4	AGCACATT-ATGTGCCA	gcacattatgtgcc	−950	pnmt	Phenylethanolamine N-Methyl-transferase PNMT [1]
4	2_1	GAACCCA-ATGTTCT	gaacccaatgttct	−2609	gilz	GILZ Human [2]
5	2_2	TTAACAG-AATGTCCT	taacagaatgtcct	−3070	gilz	GILZ Human [2]
6	3_2	GGACTTG-TTTGTTCT	gacttgtttgttct	−2452	tat	Rat Tyrosine aminotransferase (TAT) [3]
7	4	AGAAGAA-ATTGTCCT	gaagaaattgtcct	−660	trhr	Human TRHR Thyrotropin-releasing hormone receptor gene [4]
8	5	GGCACAG-TGTGGTCT	gcacagtgtggtct	−2421	th	Mouse TH Tyrosine hydroxylase gene [5]
9	6	TTATTTTGA-ACACGGGG-ATCCTA	gaacacggggatcc	−75	igfb1	Rat IGFBP-1 Insulin like growth factor binding protein-1 [6]
10	7_1	CGATCAG-GCTGTTTT	gatcaggctgtttt	−183	g6pc	Glucose-6-phosphatase [7]
11	7_2	TGTGCCT-GTTTTGCT	gtgcctgttttgct	−166	g6pc	Glucose-6-phosphatase [7]
12	7_3	AAATCAC-CCTGAACA	aatcaccctgaaca	−142	g6pc	Glucose-6-phosphatase [7]
13	8_1	CACACAA-AATGTGCA	acacaaaatgtgca	−374	pepck	Rat PEPCK Phosphoenolpyru-vate carboxykinase [8]
14	8_2	AGCATATG-AAGTCCA	gcatatgaagtcca	−353	pepck	Rat PEPCK Phosphoenolpyru-vate carboxykinase [8]
15	9	TGTTCAC-TTTGTTAT	gttcactttgttat	−1102	fgg	Human gamma chain fibrio-gen [9]
16	10_1	CTTCCAT-GCTGTTCC	ttccatgctgttcc	−1432	eln	Human elastin gene [10]
17	10_2	ACCCTCC-CCTGTTCC	ccctcccctgttcc	−1310	eln	Human elastin gene [10]
18	10_3	CCACCTC-CCTGTTCC	cacctccctgttcc	−1018	eln	Human elastin gene [10]
19	11	GGAACAA-TGTGTACC	gaacaatgtgtacc	∼2.3 kb dstr. of poly(A)	dexras1	Human dexras1 gene [11]
20	12_1	AGGACAG-CCTGTCCT	ggacagcctgtcct	∼1 kb ustr. of MT II	mt2	Mouse metallothionein [12]
21	12_2	GAAACAC-CATGTACC	aaacaccatgtacc	∼7 kb ustr. of MT-I	mt1	Mouse metallothionein [12]
22	13	GGACATG-ATGTTCC	ggacatgatgttcc	−229	il6	Interleukin-6 Responsive Ele-ment Type2 [13]
23	14_1	CCAAATCA-CTGGACCT	caaatcactggacc	+191	gr	Human glucocorticoid receptor (hGR) protein [14]
24	15	GGAACAA-CAAGGGCA	gaacaacaagggca	−4432	hcar	Human constitutive androstane receptor [15]
25	16	AGAACAG-CCTGTCCT	gaacagcctgtcct	−5042	cdkn1c	Human cyclin dependent ki-nase inhibitor p57Kip2 [16]
26	17	GGGTGAG-CTTGTTCT	ggtgagcttgttct	−365	adrbk2	Rat Beta2-adrenergic receptor gene [17]

#	ID	GRE sequence	Aligned sequence	BP pos.	Symbol	Gene
27	18_1	GTACCAAG-AATGTGTT-CTGCA	caagaatgtgttct	−759	pnmt	Rat Phenyl ethanolamine N-Methyltransferase (PNMT) gene [18]
28	18_2	TTCTGCAC-TCTCTGTT-CTTAC	gcactctctgttct	−773	pnmt	Rat Phenyl ethanolamine N-Methyltransferase (PNMT) gene
29	19	CCCTGGCAC-ATTTCGTGC	ctggcacatttcgt	−150	alpha 1I3	Rat Liver alpha inhibitor III gene [19]
30	20	CGGACAA-ATGTTCT	ggacaaaatgttct	−1159	sgk1	Human sgk1 gene [20]
31	21	TGAACTG-AATGTTTT	gaactgaatgtttt	−1662	cyp2c9	Cytochrome P450 2C9 [21]
32	22	CTGTACAG-GATGTTCT	gtacaggatgttct	−2590	tat	Rat Tyrosine aminotransferase (TAT) [22]
33	23	ACATGAG-TGTGTCCT	catgagtgtgtcct	−583	chga	Rat chromogranin A [23]
34	24	AGCACAC-ACTGTTCT	gcacacactgttct	−1212	serpine1	Rat type1 plasminogen activator [24]
35	25_1	GACACCA-CCCCTCCC	acaccacccctccc	−139	dbt	Alpha-ketoacid dehydrogenase E2 subunit [25]
36	25_2	GCTCGTT-CCTTCTCT	ctcgttccttctct	−110	dbt	Alpha-ketoacid dehydrogenase E2 subunit
37	26	AGAGCAG-TTTGTTCT	gagcagtttgttct	−6300	cps1	Carbamoylphosphate synthetase [26]
38	27	AGAACTA-TCTGTTCC	gaactatctgttcc	1st intron	pfkfb3	6-phosphofructo-2 kinase [27]
39	28	GGAACAT-TTTGTGCA	gaacattttgtgca	−104	agp	Rat Alpha 1-acid glycoprotein [28]
40	29	TGGGACTAC-AGTGTCCTG	gactacagtgtcct	−1193	sult1a3	Human Sulfotransferase 1a3 (SULT1A3) [29]
41	30	TGTCCTGC-TCGAGGTG-GTTCA	ctcgaggtggttca	−630	atp1b1	Human Na/K-ATPase beta1 gene [30]
42	31	AGAACAG-AATGTCCT	gaacagaatgtcct	−1306	scnn1a	alpha-subunit epithelial Na$^+$ channel alpha-ENaC gene [31]
43	32	CAGGGTAC-ATGGCGTA-TGTGTG	cagggtacatggcg	−447	myc	Murine c-myc [32]
44	33	TGTACAC-TATTGTCT	gtacactattgtct	−756	agtr1a	Rat Angiotensin II Type 1A receptor gene [33]

1. Adams, M., Meijer, O. C., Wang, J., Bhargava, A. & Pearce, D. Homodimerization of the glucocorticoid receptor is not essential for response element binding: activation of the phenylethanolamine N-methyltransferase gene by dimerization-defective mutants. Mol Endocrinol 17, 2583–2592 (2003).
2. Wang, J. C., Derynck, M. K., Nonaka, D. F., Khodabakhsh, D. B., Haqq, C. & Yamamoto, K. R. Chromatin immunoprecipitation (ChIP) scanning identifies primary glucocorticoid receptor target genes. Proc Natl Acad Sci U S A 101, 15603–15608 (2004).
3. Jantzen, H. M., Strähle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R. & Schütz, G. Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell 49, 29–38 (1987).
4. Høvring, P. I., Matre, V., Fjeldheim, A. K., Loseth, O. P. & Gautvik, K. M. Transcription of the human thyrotropin-releasing hormone receptor gene-analysis of basal promoter elements and glucocorticoid response elements. Biochem Biophys Res Commun 257, 829–834 (1999).

5. Hagerty, T., Morgan, W. W., Elango, N. & Strong, R. Identification of a glucocorticoid-responsive element in the promoter region of the mouse tyrosine hydroxylase gene. J Neurochem 76, 825–834 (2001).
6. Goswami, R., Lacson, R., Yang, E., Sam, R. & Unterman, T. Functional analysis of glucocorticoid and insulin response sequences in the rat insulin-like growth factor-binding protein-1 promoter. Endocrinology 134, 736–743 (1994).
7. Vander Kooi, B. T., Onuma, H., Oeser, J. K., Svitek, C. A., Allen, S. R., Vander Kooi, C. W., Chazin, W. J. & O'Brien, R. M. The glucose-6-phosphatase catalytic subunit gene promoter contains both positive and negative glucocorticoid response elements. Mol Endocrinol 19, 3001–3022 (2005).
8. Imai, E., Stromstedt, P. E., Quinn, P. G., Carlstedt-Duke, J., Gustafsson, J. A. & Granner, D. K. Characterization of a complex glucocorticoid response unit in the phosphoenolpyruvate carboxykinase gene. Mol Cell Biol 10, 4712–4719 (1990).
9. Asselta, R., Duga, S., Modugno, M., Malcovati, M. & Tenchini, M. L. Identification of a glucocorticoid response element in the human gamma chain fibrinogen promoter. Thromb Haemost 79, 1144–1150 (1998).
10. Del Monaco, M., Covello, S. P., Kennedy, S. H., Gilinger, G., Litwack, G. & Uitto, J. Identification of novel glucocorticoid-response elements in human elastin promoter and demonstration of nucleotide sequence specificity of the receptor binding. J Invest Dermatol 108, 938–942 (1997).
11. Kemppainen, R. J., Cox, E., Behrend, E. N., Brogan, M. D. & Ammons, J. M. Identification of a glucocorticoid response element in the 3'-flanking region of the human Dexras1 gene. Biochim Biophys Acta 1627, 85–89 (2003).
12. Kelly, E. J., Sandgren, E. P., Brinster, R. L. & Palmiter, R. D. A pair of adjacent glucocorticoid response elements regulate expression of two mouse metallothionein genes. Proc Natl Acad Sci U S A 94, 10045–10050 (1997).
13. Kasutani, K., Itoh, N., Kanekiyo, M., Muto, N. & Tanaka, K. Requirement for cooperative interaction of interleukin-6 responsive element type 2 and glucocorticoid responsive element in the synergistic activation of mouse metallothionein-I gene by interleukin-6 and glucocorticoid. Toxicol Appl Pharmacol 151, 143–151 (1998).
14. Geng, C. D. & Vedeckis, W. V. Steroid-responsive sequences in the human glucocorticoid receptor gene 1A promoter. Mol Endocrinol 18, 912–924 (2004).
15. Pascussi, J. M., Busson-Le Coniat, M., Maurel, P. & Vilarem, M. J. Transcriptional analysis of the orphan nuclear receptor constitutive androstane receptor (NR1I3) gene promoter: identification of a distal glucocorticoid response element. Mol Endocrinol 17, 42–55 (2003).
16. Alheim, K., Corness, J., Samuelsson, M. K., Bladh, L. G., Murata, T., Nilsson, T. & Okret, S. Identification of a functional glucocorticoid response element in the promoter of the cyclin-dependent kinase inhibitor p57Kip2. J Mol Endocrinol 30, 359–368 (2003).
17. Cornett, L. E., Hiller, F. C., Jacobi, S. E., Cao, W. & McGraw, D. W. Identification of a glucocorticoid response element in the rat beta2-adrenergic receptor gene. Mol Pharmacol 54, 1016–1023 (1998).
18. Tai, T. C., Claycomb, R., Her, S., Bloom, A. K. & Wong, D. L. Glucocorticoid responsiveness of the rat phenylethanolamine N-methyltransferase gene. Mol Pharmacol 61, 1385–1392 (2002).
19. Abraham, L. J., Bradshaw, A. D., Northemann, W. & Fey, G. H. Identification of a glucocorticoid response element contributing to the constitutive expression of the rat liver alpha 1-inhibitor III gene. J Biol Chem 266, 18268–18275 (1991).
20. Itani, O. A., Liu, K. Z., Cornish, K. L., Campbell, J. R. & Thomas, C. P. Glucocorticoids stimulate human sgk1 gene expression by activation of a GRE in its 5'-flanking region. Am J Physiol Endocrinol Metab 283, E971-E979 (2002).
21. Gerbal-Chaloin, S., Daujat, M., Pascussi, J. M., Pichard-Garcia, L., Vilarem, M. J. & Maurel, P. Transcriptional regulation of CYP2C9 gene. Role of glucocorticoid receptor and constitutive androstane receptor. J Biol Chem 277, 209–217 (2002).
22. Grange, T., Roux, J., Rigaud, G. & Pictet, R. Cell-type specific activity of two glucocorticoid responsive units of rat tyrosine aminotransferase gene is associated with multiple binding sites for C/EBP and a novel liver-specific nuclear factor. Nucleic Acids Res 19, 131–139 (1991).

23. Rozansky, D. J., Wu, H., Tang, K., Parmer, R. J. & O'Connor, D. T. Glucocorticoid activation of chromogranin A gene expression. Identification and characterization of a novel glucocorticoid response element. J Clin Invest 94, 2357–2368 (1994).

24. Bruzdzinski, C. J., Johnson, M. R., Goble, C. A., Winograd, S. S. & Gelehrter, T. D. Mechanism of glucocorticoid induction of the rat plasminogen activator inhibitor-1 gene in HTC rat hepatoma cells: identification of cis-acting regulatory elements. Mol Endocrinol 7, 1169–1177 (1993).

25. Costeas, P. A. & Chinsky, J. M. Glucocorticoid regulation of branched-chain alpha-ketoacid dehydrogenase E2 subunit gene expression. Biochem J 347, 449–457 (2000).

26. Christoffels, V. M., Grange, T., Kaestner, K. H., Cole, T. J., Darlington, G. J., Croniger, C. M. & Lamers, W. H. Glucocorticoid receptor, C/EBP, HNF3, and protein kinase A coordinately activate the glucocorticoid response unit of the carbamoylphosphate synthetase I gene. Mol Cell Biol 18, 6305–6315 (1998).

27. Pierreux, C. E., Ursø, B., De Meyts, P., Rousseau, G. G. & Lemaigre, F. P. Inhibition by insulin of glucocorticoid-induced gene transcription: involvement of the ligand-binding domain of the glucocorticoid receptor and independence from the phosphatidylinositol 3-kinase and mitogen-activated protein kinase pathways. Mol Endocrinol 12, 1343–1354 (1998).

28. Ratajczak, T., Williams, P. M., DiLorenzo, D. & Ringold, G. M. Multiple elements within the glucocorticoid regulatory unit of the rat alpha 1-acid glycoprotein gene are recognition sites for C/EBP. J Biol Chem 267, 11111–11119 (1992).

29. Bian, H. S., Ngo, S. Y., Tan, W., Wong, C. H., Boelsterli, U. A. & Tan, T. M. Induction of human sulfotransferase 1A3 (SULT1A3) by glucocorticoids. Life Sci 81, 1659–1667 (2007).

30. Derfoul, A., Robertson, N. M., Lingrel, J. B., Hall, D. J. & Litwack, G. Regulation of the human Na/K-ATPase beta1 gene promoter by mineralocorticoid and glucocorticoid receptors. J Biol Chem 273, 20702–20711 (1998).

31. Wang, H. C., Zentner, M. D., Deng, H. T., Kim, K. J., Wu, R., Yang, P. C. & Ann, D. K. Oxidative stress disrupts glucocorticoid hormone-dependent transcription of the amiloride-sensitive epithelial sodium channel alpha-subunit in lung epithelial cells through ERK-dependent and thioredoxin-sensitive pathways. J Biol Chem 275, 8600–8609 (2000).

32. Ma, T., Copland, J. A., Brasier, A. R. & Thompson, E. A. A novel glucocorticoid receptor binding element within the murine c-myc promoter. Mol Endocrinol 14, 1377–1386 (2000).

33. Guo, D. F., Uno, S., Ishihata, A., Nakamura, N. & Inagami, T. Identification of a cis-acting glucocorticoid responsive element in the rat angiotensin II type 1A promoter. Circ Res 77, 249–257 (1995).

Table 2.5: Primers created based on *in silico* predictions.

Gene	GRE pos from TSS	Primer pos from TSS	Fw/Rev	Sequence	Tm	GC	Loop, T, degrees	Length amplicon
Abhd14a	35067	35043	F	CCAGCTCAGGTTACCGTCTT	59.35	55.00	16	88
		35131	R	TGAACTAAGATGGCCAACACC	59.99	47.62		
Adra1d_1	58681	58618	F	TTAAACGGTCCTTGGTGCAT	60.37	45.00		94
		58712	R	TCCTTTATCTGTGGGCTGGTA	59.58	47.62		
Adra1d_2	−43891	−44001	F	TCTGAACCGTGACCAAGGAA	61.64	50.00	17.2	93
		−43908	R	TGACTGAACTGGAAGTGACT	53.03	45.00		
Akap7	154398	154315	F	CATGGGAGATTCTTACAGGCTT	59.61	45.45	19.6	140
		154455	R	GGTGAGGACATGACATTAGCAA	60.00	45.45		
Arhgef3	243344	243319	F	TCCGTCAACATCCTGGATTC	60.87	50.00		90
		243409	R	GAGGTGAAAAGAGGCAGGTG	59.84	55.00		
Daam1	−40707	−40761	F	GAGCAATGGGTTTGTTGGAG	60.50	50.00	2.9	101
		−40660	R	AATCCTCTCTCCATGATGCAC	59.11	47.62		
Ddit4	−20879	−20894	F	CTGTGGGTGAGCTGAGAACA	60.02	55.00		99
		−20866	R	GGCCTGTAGGTCCAGCACTA	60.28	60.00	11.2	
Dgat2	41675	41655	F	CCTGTTTTGTCTGCCTCTCTG	60.04	52.38		116
		41771	R	CACTGAGTCATTTCGCAGGA	59.98	50.00		
Errfi1	−29643	−29723	F	CCTGCATTTCTGGTTTTGAAG	59.73	42.86		105
		−29618	R	TCCTCTCCAGGGGTACACTC	59.10	60.00	23.8	
Fam55c	64366	64283	F	AAATCTTTCACCGGCTCAGA	59.81	45.00		116
		64399	R	CTCACAGCTCCAAACGGAA	59.97	52.63		
Fkbp5_1	62946	62912	F	TCAGCACACCGAGTTCATGT	60.32	50.00		133
		63045	R	CTGGTCACTGCAAAACATCATT	60.04	40.91		
Fkbp5_2	58773	58712	F	GGATGGAGACTGCGTTCTGT	60.27	55.00		99
		58811	R	CTGGAGTTCTGCCTGCACTT	60.59	55.00	30.4	
Fkbp5_3	1097	1026	F	GAACGCGTTGGAAGAAGGTA	60.25	50.00		120
		1146	R	CCGCATGCAGAATTTACTGA	59.83	45.00	7.5	
Kcnj11	−23686	−23707	F	ACCCCTGTCCTTACTCTCCA	53.15	55.00	46.3	79
		−23628	R	ATGGGGCAGGATGTCTATGT	52.16	50.00		
Klf9_1	−6345	−6376	F	ATGATGAAACGTGAGCGCTAT	59.75	42.86	6.8	93
		−6283	R	TTTCCTGTGGTTGTTGTGGA	59.98	45.00	10	
Klf9_2	−5522	−5554	F	ATCTAGGGCAGTTTGTTCAA	54.96	40.00		96
		−5458	R	GGCAGGTTCATCTGAGGACA	61.23	55.00		
Lyve1	−19879	−19951	F	CACCCAGAAAGAAGGCACA	59.81	52.63		104
		−19847	R	CTCTGTAAATGAGGGCCGAG	59.83	55.00	5.1	
Mfsd2_1	−17609	−17675	F	GAGGCATCATACCGGAACTC	59.51	55.00	13	102
		−17573	R	AGAAGATGGGAGATTGGCCT	60.04	50.00		
Mfsd2_2	2297	2215	F	GACCCGTTAGTGACGCTGTT	60.18	55.00	28.9	123
		2338	R	ACAGTGCTCCCATCAGCCTA	60.82	55.00	21.6	
Msx1	30573	30562	F	TGCAAACTCCTGAACAGCCT	60.98	50.00		84
		30646	R	GAGAAGGTGACGCCTGGTTA	60.25	55.00	13.2	
Slc25a13_2	−5588	−5635	F	GGAAAGTCTGCGTCCGTATC	59.7	55.00	9	93
		−5542	R	AGGCAGAAAGCATGAAAGCA	61.05	45.00		
Slc25a13_3	−17970	−18047	F	CTTACCCAGGACCACAAGGA	59.96	55.00		120
		−17927	R	AACAGCCATTAATTTGTGTGGTT	59.7	34.78	7.1	
Srxn1_1	−28091	−28164	F	GATGCTTTTGTGGCCACTCT	60.26	50.00	11.6	100
		−28064	R	GTTGAATGGGAAAGGGACAA	59.77	45.00		
Srxn1_2	−21486	−21509	F	GAATTTCTCATGCACAGCCA	59.81	45.00	16.5	85
		−21424	R	CTCTTTGGACGGGATTCAAG	59.66	50.00		
Tiparp_1	20215	20164	F	GCTAGGATTTCACTCGCACA	59.03	50.00	30.6	107

Gene	GRE pos from TSS	Primer pos from TSS	Fw/ Rev	Sequence	Tm	GC	Loop, T, degrees	Length amplicon
		20271	R	CAAGCTGCTGGTCTCGGTA	60.14	57.89		
Tiparp_2	13493	13481	F	GACCTCCCACATGAACTGC	59.04	57.89		112
		13593	R	CATGTGAACTTAGTTACCAGACCA	58.25	41.67	13.1	
Tiparp_3	1312	1229	F	TTGCCTGGATTGGTGTGATA	59.92	45.00		107
		1336	R	AGGCTCAGTTGGCACAGATT	59.87	50.00		
Tle3_1	70362	70355	F	GTCAAAACAACACCCAGTCC	57.87	50.00	13	116
		70471	R	ATTTGGTGGAGCTGAGCACT	59.87	50.00		
Tle3_2	−725	−782	F	TCGCCGCCTCTGCAGAATCAA	58.85	57.14		119
		−663	R	TGGCGGGAGGGGAGAAAGAGA	58.66	61.90		
Zfyve28	99395	99350	F	CCGGGATTCAGGACTCTAGTT	59.58	52.38		93
		99443	R	CATACAAGCCACTGCAGGAA	59.86	50.00		
Znf592_1	92248	92152	F	CAGCATAGCCCGACTGTGT	59.87	57.89	7.8	102
		92254	R	ATCCCTCTTCCTCCTTCCAG	59.63	55.00		
Znf592_2	90506	90487	F	CCCAGTCTAATCCCTCTTGG	58.59	55.00	8.3	121
		90608	R	ATCCAAGTCCTGCCCTACCT	59.96	55.00		

Table 2.6: Predicted GREs that were selected for validation using ChIP.

Gene	GRE sequence	BP in Aln	BP from TSS	Score	Gene	GRE sequence	BP in Aln	BP from TSS	Score
Adra1d_1					**Lyve1**				
Homo sapiens	gaacaccctgtact	170141	77873	0.93	Homo sapiens	gcacaggctgtgct	596	−18228	0.89
Mus musculus	gaacgccctgtact	170141	53496	0.83	Bos taurus	gcacaggctgtgct	596	−18833	0.89
Rattus norvegicus	*gaacggcctgtacc*	*170141*	*58681*	*0.81*	*Rattus norvegicus*	*tcacagattgtgct*	*596*	*−19879*	*0.80*
Bos taurus	gaacaccctgtact	170141	56374	0.93	Mus musculus	gcacagactgtgct	596	−18909	0.90
Adra1d_2					**Mfsd2_1**				
Homo sapiens	gaacaggacgtcct	−25038	−31585	0.84	Homo sapiens	gaactccatgtcct	4961	−27297	0.90
Bos taurus	gaacaagatgcctt	−25038	−46743	0.74	Bos taurus	gaactccatgtccc	4961	−27984	0.87
Rattus norvegicus	*ggacaggacgtcct*	*−25038*	*−43891*	*0.78*	*Rattus norvegicus*	*gaactccatgtccc*	*4961*	*−17609*	*0.87*
Mus musculus	ggacaggatgtcct	−25038	−37462	0.88	Mus musculus	gaactccatgtccc	4961	−18781	0.87
Abhd14a					**Mfsd2_2**				
Homo sapiens	gaacagcctgtact	86304	37769	0.93	Homo sapiens	gcacactgtgttcc	60613	2402	0.88
Bos taurus	gaacattatgtacc	86304	30132	0.90	Bos taurus	gcacaccatgtccc	60613	2165	0.88
Rattus norvegicus	*gaacagcctgtacc*	*86304*	*35067*	*0.90*	*Rattus norvegicus*	*gcacaccatgtccc*	*60613*	*2297*	*0.88*
Mus musculus	gaacagcctgtacc	86304	33829	0.89	Mus musculus	gcacaccatgtccc	60613	2214	0.88
Akap7					**Msx1**				
Homo sapiens	gcagactttttct	360179	167897	0.81	Homo sapiens	gaacagcctgttct	117908	53528	0.98
Bos taurus	gcacattttgtcct	360179	134767	0.89	Bos taurus	gaacagcctgtccc	117908	40500	0.98
Rattus norvegicus	*gcagatcctgttct*	*360179*	*154398*	*0.88*	*Rattus norvegicus*	*gaacagcctgttcc*	*117908*	*30573*	*0.98*
Mus musculus	gcagacctgttct	360179	150888	0.89	Mus musculus	gaacagcctgttct	117908	26934	0.98
Arhgef3					**Slc25a13_2**				
Homo sapiens	gaacagtctgtcct	302939	16194	0.95	Homo sapiens	taatagtttgttct	38873	−8228	0.81
Bos taurus	gaacacactgtgct	302939	14141	0.93	Bos taurus	taacagattgttct	38873	−12309	0.88
Rattus norvegicus	*gaacaatctgtcct*	*302939*	*243344*	*0.94*	*Rattus norvegicus*	*taacaggctgttct*	*38873*	*−5588*	*0.88*
Mus musculus	gaacactctgtcct	302939	112042	0.95	Mus musculus	taacaggctgttct	38873	−5705	0.88
Ddit4					**Slc25a13_3**				
Homo sapiens	gaacattgtgttct	8595	−24936	0.94	Bos taurus	ccataaaattatct	18803	−21180	0.68
Bos taurus	gaacattgtgttct	8595	−15283	0.94	Homo sapiens	gcataacattagct	18803	−21125	0.68
Rattus norvegicus	*gaacattgtgttct*	*8595*	*−20879*	*0.94*	*Rattus norvegicus*	*ccataaaattttct*	*18803*	*−17970*	*0.76*
Mus musculus	gaacattgtgttct	8595	−22516	0.94	Mus musculus	ccataaaattttct	18803	−19022	0.76
Daam1_2					**Srxn1_1**				
Homo sapiens	ttagattatgttct	−11741	−27887	0.80	Homo sapiens	gaccatcttgtccc	−14268	−28256	0.85
Bos taurus	ttagattatgttct	−11741	−32385	0.80	Bos taurus	gaccaacttgtccc	−14268	−31089	0.85
Rattus norvegicus	*ttagattatgttct*	*−11741*	*−40707*	*0.80*	*Rattus norvegicus*	*gaccatcttgtccc*	*−14268*	*−28091*	*0.85*
Mus musculus	ttagattatgttct	−11741	−31094	0.80	Mus musculus	gaccatcttgtccc	−14268	−28603	0.85
Dgat2					**Srxn1_2**				
Homo sapiens	aaacactatgttct	110257	44912	0.91	Homo sapiens	ctgcaggctgttcc	−3941	−20498	0.80
Bos taurus	aaatactctgttct	110257	49855	0.85	Bos taurus	ctgcagactgttcc	−3941	−23256	0.82
Rattus norvegicus	*gaacactgtgttct*	*110257*	*41675*	*0.95*	*Rattus norvegicus*	*ctgcaggctgttcc*	*−3941*	*−21486*	*0.80*
Mus musculus	gaacactgtgttct	110257	40233	0.95	Mus musculus	ctgcaggctgttcc	−3941	−22059	0.80

Gene	GRE sequence	BP in Aln	BP from TSS	Score	Gene	GRE sequence	BP in Aln	BP from TSS	Score
Errfi1					**Tiparp_1**				
Homo sapiens	gaacgaaatgtact	−16545	−44820	0.82	Homo sapiens	gaatatattgtcct	71852	22639	0.86
Bos taurus	gaacaagatgtact	−16545	−34910	0.90	Bos taurus	gaacatgctgtcct	71852	19024	0.93
Rattus norvegicus	*gaacagagtgtacc*	*−16545*	*−29643*	*0.88*	*Rattus norvegicus*	*gaacatactgtcct*	*71852*	*20215*	*0.94*
Mus musculus	ggacagagtgtgcc	−16545	−29939	0.83	Mus musculus	gaacatgctgtcct	71852	20078	0.93
Fam55c					**Tiparp_2**				
Homo sapiens	gcactttctgttcc	202292	50884	0.85	Homo sapiens	gaactgggtgtgcc	62183	15135	0.82
Bos taurus	gcacttcctgttcc	202292	61589	0.85	Bos taurus	gaactgcatgtgct	62183	12005	0.89
Rattus norvegicus	*gcactttctgttcc*	*202292*	*64366*	*0.85*	*Rattus norvegicus*	*gaactgcatgtgct*	*62183*	*13493*	*0.89*
Mus musculus	gcactttctgttcc	202292	59350	0.85	Mus musculus	gaactgcatgtgct	62183	13312	0.89
Fkbp5_1					**Tiparp_3**				
Homo sapiens	gaacagggtgttct	201653	86842	0.94	Homo sapiens	ccacaatctgtgcc	45867	1537	0.82
Bos taurus	gaacagggtgttct	201653	99485	0.94	Bos taurus	ccacaatctgtgcc	45867	−535	0.82
Rattus norvegicus	*gaacagggtgttct*	*201653*	*62946*	*0.94*	*Rattus norvegicus*	*ccacaatctgtgcc*	*45867*	*1312*	*0.82*
Mus musculus	gaacagggtgttct	201653	20724	0.94	Mus musculus	ccacaatctgtgcc	45867	1308	0.82
Fkbp5_2					**Tle3_1**				
Homo sapiens	gtacacactgttct	187860	77853	0.94	Homo sapiens	caacacccagtccc	130075	73749	0.81
Bos taurus	ctacatactgttct	187860	91074	0.89	Bos taurus	caacacccagtccc	130075	76757	0.81
Rattus norvegicus	*gtacacgctgttct*	*187860*	*58773*	*0.92*	*Rattus norvegicus*	*caacacccagtccc*	*130075*	*70362*	*0.81*
Mus musculus	gtacataccgttct	187860	16445	0.83	Mus musculus	caacacccagtccc	130075	67838	0.81
Fkbp5_3					**Tle3_2**				
Homo sapiens	ggacagtgtgttca	39281	1167	0.85	Homo sapiens	gtacagcttgtctt	26890	−387	0.81
Bos taurus	ggacagagtgtaca	39281	−4322	0.80	Bos taurus	gtacagcttgtctt	26890	−1506	0.81
Rattus norvegicus	*ggacagtgtgtaca*	*39281*	*1097*	*0.80*	*Rattus norvegicus*	*gtacagcttgtcct*	*26890*	*−725*	*0.90*
Mus musculus	ggacagggtgtaca	39281	−43907	0.79	Mus musculus	gtacagcttgtcct	26890	−1692	0.90
Kcnj11					**Zfyve28**				
Homo sapiens	gtacaagatggtca	−11257	−24819	0.80	Homo sapiens	gaacgcagtgttcc	248785	161563	0.85
Bos taurus	gtacaagatggtca	−11257	−24012	0.80	Bos taurus	ggacgccgtgttct	248785	64876	0.80
Rattus norvegicus	*gtacaagatggtca*	*−11257*	*−23686*	*0.80*	*Rattus norvegicus*	*gaacaccatgttcc*	*248785*	*99395*	*0.94*
Mus musculus	gtacaagatggtca	−11257	−22705	0.80	Mus musculus	gaacaccatgttcc	248785	105004	0.94
Klf9_1					**Znf592_1**				
Homo sapiens	ggacaaactgttcc	45287	−5481	0.88	Homo sapiens	gaagataatgttct	218121	92619	0.92
Bos taurus	ggacaaactgttcc	45287	−5884	0.88	Bos taurus	gaggaggatgttct	218121	83347	0.86
Rattus norvegicus	*ggacaaactgttcc*	*45287*	*−6345*	*0.88*	*Rattus norvegicus*	*gaagatactgttcc*	*218121*	*92248*	*0.92*
Mus musculus	ggacaaactgttcc	45287	−6133	0.88	Mus musculus	gaagatactgttct	218121	85402	0.92
Klf9_2					**Znf592_2**				
Homo sapiens	gagcttgatgttcc	46231	−4616	0.81	Homo sapiens	ggacagtatggcct	216299	90890	0.84
Bos taurus	gagcttgatgttcc	46231	−4991	0.81	Bos taurus	gaacagcgtggcct	216299	81672	0.87
Rattus norvegicus	*gagcttgatgttcc*	*46231*	*−5522*	*0.81*	*Rattus norvegicus*	*ggacagcatgacct*	*216299*	*90506*	*0.82*
Mus musculus	gagcttgatgttcc	46231	−5265	0.81	Mus musculus	ggacagcatgacct	216299	83680	0.82

Chapter 2

51