Supplementary Figures and Tables

Identification and characterization of DNA sequences that prevent glucocorticoid receptor binding to nearby response elements

Jonas Telorac¹, Sergey V. Prykhozhij^{1,2}, Stefanie Schöne¹, David Meierhofer¹, Sascha Sauer³, Morgane Thomas-Chollier^{4#}, Sebastiaan H. Meijsing^{1#}

¹ Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, D-14195, Berlin, Germany.

² Dalhousie University, Halifax, Canada.

³ CU Systems Medicine, University of Würzburg, Josef-Schneider-Strasse 2, D-97080 Würzburg, Germany.

⁴ Computational Systems Biology, Institut de Biologie de l'Ecole Normale Supérieure (IBENS), CNRS, Inserm, Ecole Normale Supérieure, PSL Research University, F-75005 Paris, France

[#] To whom correspondence should be addressed:

Sebastiaan H. Meijsing Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73 14195, Berlin, Germany Tel: +49-30-84131176 Email: meijsing@molgen.mpg.de

Morgane Thomas-Chollier Institut de Biologie de l'Ecole Normale Supérieure 46 rue d'Ulm 75230, Paris, France Email: mthomas@biologie.ens.fr

Additional footnotes: # co-corresponding author

This PDF includes:

- Supplemental Figures
- Supplemental Tables

Supplemental figures:



Figure S1. Motif occurrence and functional analysis of depleted sequences around MyoDbound regions. (a) Occurrence of motif around MyoD-bound regions was analyzed as described for Fig. 1a. (b) Cells were transfected with luciferase reporter constructs with 3 MyoD binding sites, flanked by a single NRS or control sequence as indicated. Along with reporter, U2OS cells were transfected with either an empty, or a MyoD-encoding expression construct. Reporter activity, normalized to control reporter + MyoD, is shown ±SEM (n=3).



Figure S2. Depletion of NRS sequences identified in previous studies. DNA sequences from GR ChIP-Seq peaks in U2OS cells stably expressing GR were aligned at the peak summit and flanking genomic DNA +/- 4000 bp was sub-divided into 50 bp bins. For each bin, the relative frequency distribution of sequence motifs for (a) Cux1 (M00102) and (b) SATB1 (M01232) was determined by scanning for alignment to these motifs. The normalized number of occurrences for each motif per bin is shown.



Figure S3. Analysis of motif occurrence around HoxD13-bound regions. (a) Occurrence of motif around HoxD13-bound regions was analyzed as described for Fig. 1a. Examples of motifs that are either (b) enriched, or (c) depleted around the peak summit of HoxD13-bound regions.

Table S1: Oligos used for cloning.

Transient NRS-GBS reporters (Fig. 2a, 2b, 2c)

GBS (FKBP5): CCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCT

Control: GTACGCAAGCCTACCTCG CTAGCGAGGTAGGCTTGC NRS1: GTACGAGGTTAATTAACG CTAGCGTTAATTAACCTC Control +5: GTACGCGAGGTAGGCTTGGCTGA CTAGTCAGCCAAGCCTACCTCGC NRS1 +5: GTACGAGAGGTTAATTAAGCTGA CTAGTCAGCTTAATTAACCTCTC AAAATT: GTACGAGGAAAAATTGCG CTAGCGCAATTTTTCCTC AAAAAAAA: GTACGAGGAAAAAAAACG CTAGCGTTTTTTTTCCTC TTTTTAA: GTACGAGGTTTTTAAGCG CTAGCGCTTAAAAACCTC TATATATA: GTACGAGGTATATATACG CTAGCGTATATATACCTC Control +10: GTACAGAGGTAGGCTTGGAGCTGCTGACTAGCCCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCTCCGGGCTAGTCAGCAGCTCCAAGCCTACCTCT NRS2 +10:

GTACTTAATTCAATTAAGAGCTGCTGACTAGCCCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCTCCGGGCTAGTCAGCAGCTCTTAATTGAATTAA Control +20:

GTACAGAGGTAGGCTTGCAGTTGGCGAGAGCTGCTGACTAGCCCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCTCCGGGCTAGTCAGCAGCTCTCGCCAACTGCAAGCCTACCTCT NRS2 +20:

GTACTTAATTCAATTAACAGTTGGCGAGAGCTGCTGACTAGCCCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCTCCGGGCTAGTCAGCAGCTCTCGCCAACTGTTAATTGAATTAA

GTACGAGGTTTGTTTGCG CTAGCGCAAACAAACCTC NRS2: **GTACGTTAATTCAATTAA** CTAGTTAATTGAATTAAC Control #2 +5: GTACGAGAGGTTTGTTTGGCTGA CTAGTCAGCCAAACAAACCTCTC NRS2 +5: GTACGTTAATTCAATTAAGCTGA CTAGTCAGCTTAATTGAATTAAC AAAAAAA: GTACGAGGTTTTTTTGCG CTAGCGCAAAAAAACCTC AAAATTTT: GTACGAGGAAAATTTTCG CTAGCGAAAATTTTCCTC ATTTTTA: GTACGAGGATTTTTAGCG CTAGCGCTAAAAATCCTC TTTTAAAA: GTACGAGGTTTTAAAACG CTAGCGTTTTAAAACCTC

Control #2:

Integrated NRS-GBS reporters (Fig. 3b)

GBS (FKBP5): CCGGAGAACAGGGTGTTCT GATCAGAACACCCTGTTCT

Control: GTACGAGGTAGGCTTG CTAGCAAGCCTACCTC NRS1: GTACGAGGTTAATTAA CTAGTTAATTAACCTC Control #2: GTACGAGGTTTGTTTG CTAGCAAACAAACCTC NRS2: GTATTAATTCAATTAA CTAGTTAATTGAATTAA

JT163: CCAGGTCTCAGTACCGTGCCAGAACATTTCTCTATCGATA JT164: CCAGGTCTCATCGACGGATCCTTATCGATTTTACC

6 GBS-NRS-TagRFP reporters (Fig. 4a)

6 GBS (PacI-cons-fkbp5-tat-fkbp5-cons-tat-Ascl): ggaattaattaaAGAACAaaaTGTACCAGAACAgggTGTTCTAGAACAtcccTGTACAAGAACAgggTGT TCTAGAACAaaaTGTACCAGAACAtcccTGTACAggcgcgccttcc

Control:

CGCGCAAGCCTACCTCGGCCAAGCCTACCTCG CGCGCGAGGTAGGCTTGGCCGAGGTAGGCTTG Control #2:

CGCGCAAACAAACCTCGGCCAAACAAACCTCG CGCGCGAGGTTTGTTTGGCCGAGGTTTGTTTG NRS1:

CGCGTTAATTAACCTCGGCTTAATTAACCTCG CGCGCGAGGTTAATTAAGCCGAGGTTAATTAA NRS2:

CGCGTTAATTGAATTAAGCTTAATTGAATTAA CGCGTTAATTCAATTAAGCTTAATTCAATTAA

NRS-3MyoD constructs (Fig. 7)

Control:

CTAGCGAGGTAGGCTTGGGGGCAGCTGCTGAAGGGCAGCTGCTGAAGGGCAGCTGCTGTGCA CAGCAGCTGCCCTTCAGCAGCTGCCCTTCAGCAGCTGCCCCAAGCCTACCTCG Control #2:

 ${\tt CTAGTTAATTCAATTAAGGGCAGCTGCTGAAGGGCAGCTGCTGTGCAGGGCAGCTGCTGTGCAGCAGCTGCCCTTCAGCAGCTGCCCTTCAGCAGCTGCCCTTCAGCAGCTGCCCTTAATTGAATTAA}$

Table S2: Primers used for qPCR analysis.

Gene/Locus:	Fw primer:
h <i>FKBP5</i>	GCATGGTTTAGGGGTTCTTG
h <i>RPL19</i>	ATGTATCACAGCCTGTACCTG
Integr. GBS	GCAGATCGCAGATCAGAACA
h <i>IGFBP1</i>	ACGTCCTGGATACAGTATGTGC
h <i>GAPDH-</i> TSS	AAAAGCGGGGAGAAAGTAGG
h <i>GAPDH</i> +1nuc	CCCCGGTTTCTATAAATTGAGC
h <i>NONO</i>	ACAGCAGGAAGGATTCAAGG
h <i>SFPQ</i>	GAGGAGAAGATCTCGGACTCG
ECFP	ACGTAAACGGCCACAAGTTC
zFKBP5	CAAAAGGGGGAATGCTGTT
TagRFP	GCTGGGAGGCCAACACCGAG

h: human z: zebrafish

Rev. Primer:

TAACCACATCAAGCGAGCTG TTCTTGGTCTCTTCCTCCTTG TATGGTACCGTGCCAGAACA TCATGTTCTTAGGGGGCAAC GGTCTTGAGGCCTGAGCTAC AAAGAAGATGCGGCTGACTG GCATGGCACCTCTGTTGTT CGACATCGCTGTGTGTAAGTTT GCAGATGAACTTCAGGGTCAG TTCTTTTCTGCCCTCTTTGC CAGGGCCATGTCGCTTCTGC Table S3: Identification of NRS-associated proteins by affinity purification and subsequent mass spectrometry analysis. Shown are only proteins with a ratios >2 between NRS experiments versus controls and identified in at least two out of 3 experiments for one or both of the NRS sequences. Amount of unique peptides, sequence coverage in % and a posterior error probability of the identifications are shown.

Protein name:	Gene:	NRS1/control ratio >2	NRS2/control ratio >2	Unique peptides	Sequence coverage [%]	PEP
Splicing factor, proline- and glutamine- rich	SFPQ	3	2	29	43.6	0
Non-POU domain-containing octamer- binding protein	NONO	3	3	20	49.3	0
DNA-3-methyladenine glycosylase	MPG	2	3	14	63.3	1.55E-139
RNA-binding protein 14	RBM14	2	2	18	29.1	5.07E-80
Heterogeneous nuclear ribonucleoprotein D-like	HNRPDL	2	2	7	39.3	1.01E-92
60S ribosomal protein L4	RPL4	2	1	12	34.9	2.78E-75
ATP-dependent RNA helicase A	DHX9	2	1	28	28.4	6.89E-118
Nucleolin	NCL	2	0	21	30.8	7.79E-193
60S ribosomal protein L31	RPL31	2	0	3	26.4	5.97E-31
THO complex subunit 4	ALYREF	1	2	2	48.6	0
PC4 and SFRS1-interacting protein	PSIP1	1	2	13	24.5	5.95E-112

-