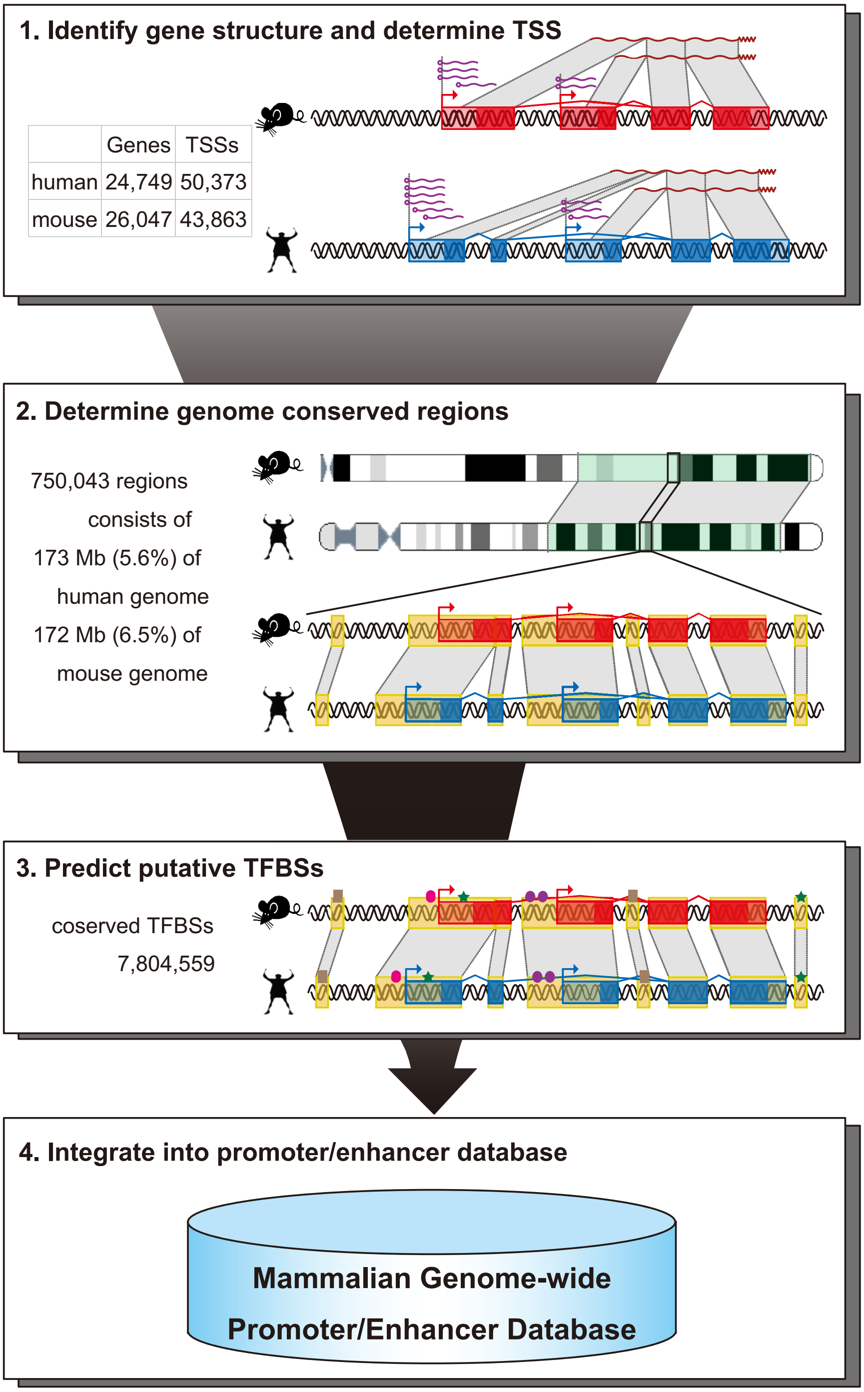
Supporting Information

Kumaki *et al.* 10.1073/pnas.0802636105



	Genes	TSSs
human	24,749	50,373
mouse	26,047	43,863

Fig. S1. Construction of the mammalian promoter/enhancer database. Mammalian full-length cDNA and EST sequences were mapped onto mammalian genome sequences to identify 24,749 human genes with 50,373 TSSs and 26,047 mouse genes with 43,863 TSSs. These mammalian genes were then compared to identify 16,268 human-mouse orthologs (65.7% of human genes and 62.5% of mouse genes). The positional information of adjacent orthologs was used to determine 434 human-mouse syntenic regions, which contained 750,043 human-mouse conserved genomic regions. The 862 consensus TFBSs from TRANSFAC were then mapped on these conserved genomic regions to identify the 7,804,559 sites conserved between human and mouse in non-coding regions. Human-rat comparisons were performed using the same procedure. Finally, visualization of putative promoter/enhancer and TFBSs data, and curation of current genes were integrated into a free and publicly accessible website (Mammalian Promoter/Enhancer Database; http://promoter.cdb.riken.jp/).

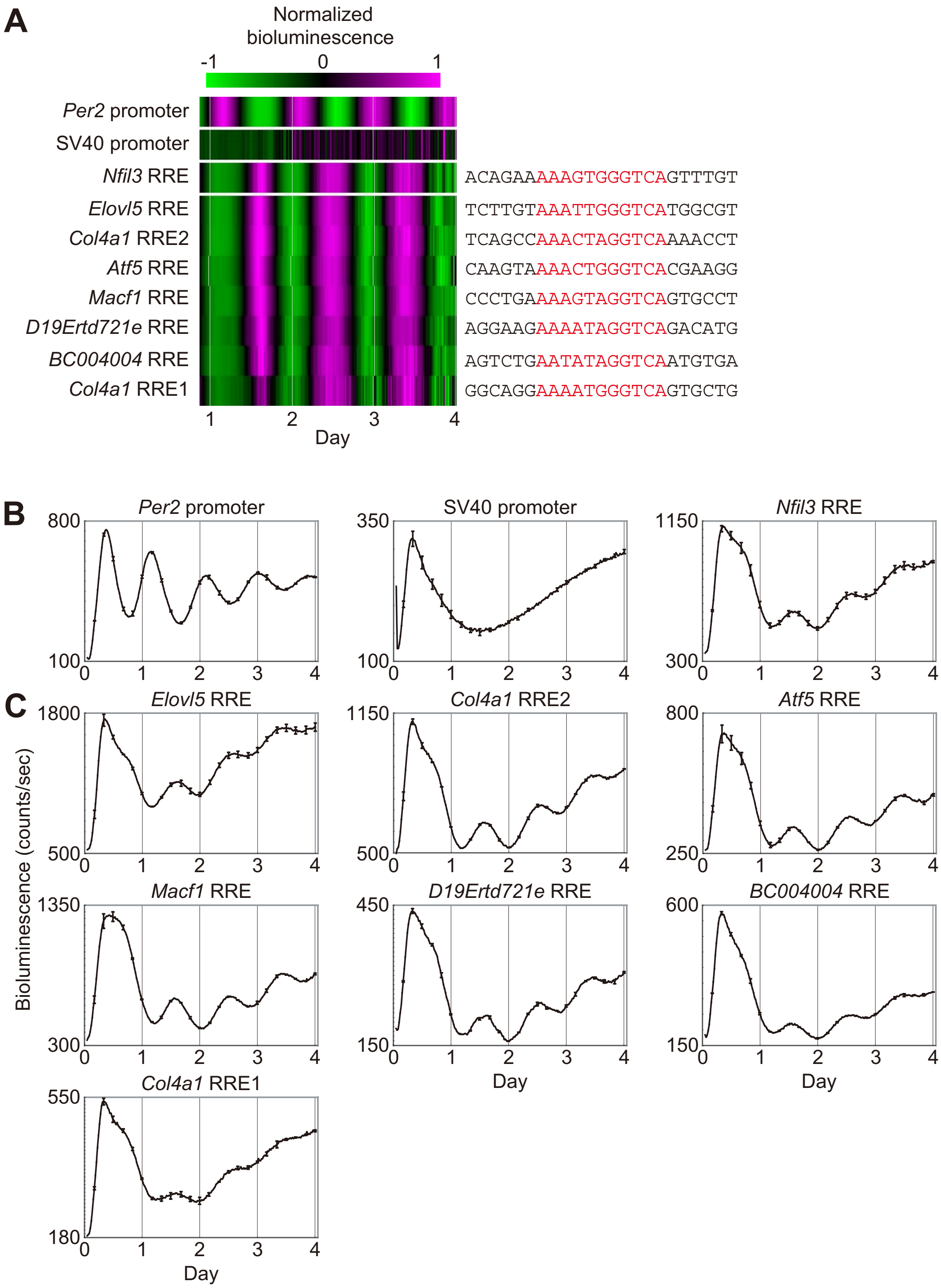


Fig. S2. Functional clock-controlled elements used to train HMMs. (*A*) Bioluminescence from functional clock-controlled elements fused to the SV40 basic promoter driving a dLuc reporter (SV40-dLuc) in NIH 3T3 fibroblasts. Known clock-controlled promoter (*Per2* promoter), known clock-controlled element fused to the SV40 basic promoter (*Nfil3* RRE), or the SV40 basic promoter alone (SV40 promoter) driving luciferase were used as controls. The colors in descending order from magenta to black to green represent the detrended bioluminescence. Columns represent time points, and rows represent the predicted elements on the designated genes. (*B* and *C*) Raw bioluminescence data of positive and negative controls (*B*), or functional clock-controlled elements inserted into SV40-dLuc reporters (*C*). Error bars indicate the standard error of mean (SEM) determined from independent experimental duplicates for each condition.

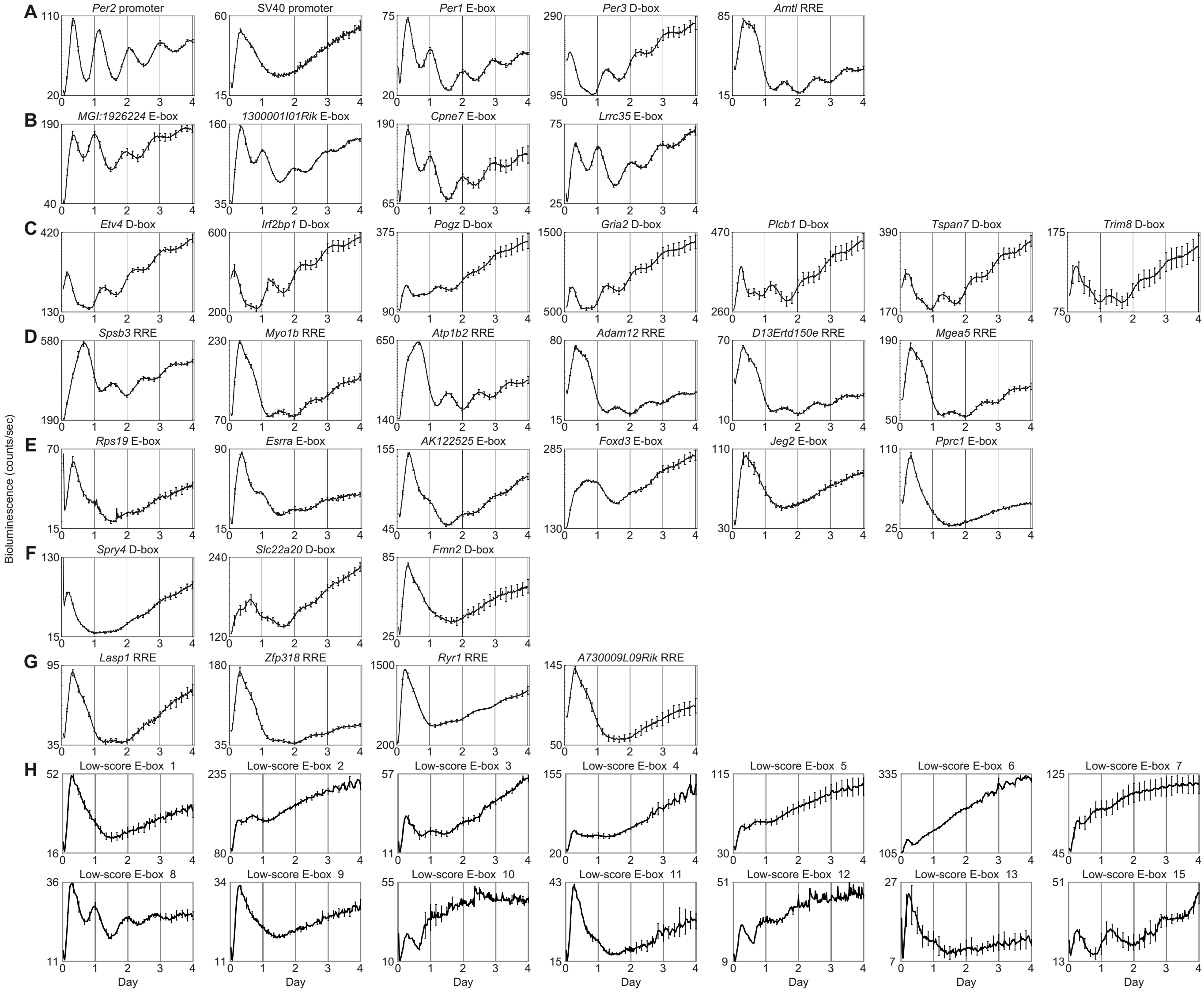


Fig. S3. Detailed information on experimental validation of HMM-based predictions at the cellular level. Raw bioluminescence data from positive and negative controls (*A*), and predicted clock-controlled elements inserted into the SV40-dLuc reporter that generate strong circadian transcriptional activity with high amplitude of the (*B*) E-box, (*C*) D-box, and (*D*) RRE, and weak circadian transcriptional activity with low amplitude or arrhythmic transcriptional activity of the (*E*) E-box, (*F*) D-box, and (*G*) RRE. Raw bioluminescence data from 14 low-score E-boxes inserted into the SV40-dLuc reporter (*H*). Only the low-score E-box No. 8 generated strong circadian transcriptional activity with high amplitude. The low-score E-box No. 14 was not measured due to the difficulty of reporter construction. Error bars indicate the standard error of mean (SEM) determined from independent experimental replicates (n = 2 or 3) for each element tested.

Pattern1	
Per1 E-box2	TAGAGCCACGTGAGGGCG
Per1 E-box3	TTTAGCCACGTGACAGTG
Per1 E-box5	CAAGTCCACGTGCAGGGA
Nr1d1 E-box1	CGGGCCCACGTGCTGCAT
Nr1d1 E-box2	GGGTGCCACGTGCGAGGG
Nr1d2 E-box1	ACTGGCCACGTGCACGGT
Dbp E-box	CCTCGCCACGTGAGTCCG
Bhlhb2 E-box	CCGGGCCACGTGAAGCGT
Rorc E-box	AGGTGCCACGTGCACCAG
High-scoring E-box	CGGGGCCACGTGCAGGCG

Pattern2	
Per1 E-box1	AGGGAACACGTGCAGGCT
Per1 E-box4	TAACGACACGTGGGCCCT
Nr1d2 E-box2	CGGAGACACGTGAGGCCG
MGI:1926224 E-box	AAGAGACACGTGCTGGGG
Lrrc35 E-box	ACGTGACACGTGCGGCGG
1300001I01Rik E-box	CGGGGACACGTGCGCGCA
Cpne7 E-box	CCGAGACACGTGTGCCCG

Fig. S4. Two patterns revealed in the adjacent sequence of high-amplitude E-boxes. High-amplitude E-box sequences classified into two patterns were shown. Information contents of the relative frequency of each nucleic acid at the position of the pattern were shown by sequence logos.

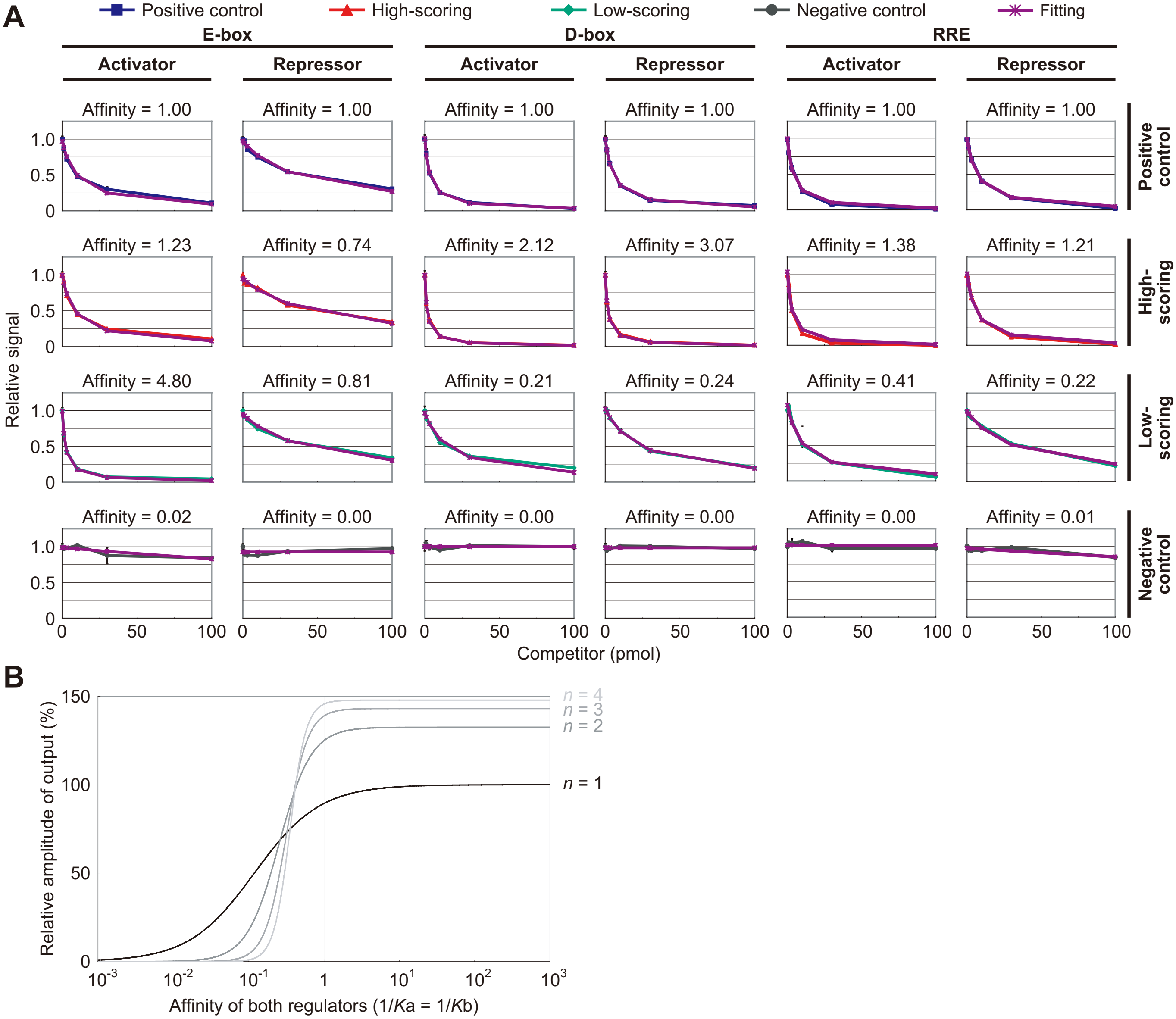


Fig. S5. Analysis of affinity to amplitude mechanism. (*A*) Affinity analysis of competitive binding data. The relative affinities of regulators for known clock controlled elements vs. synthetic elements were determined from competitive binding assay data shown in Fig. 4. A series of signal of binding between labeled oligonucleotide of positive control element and regulators, which were challenged with unlabeled oligonucleotides of positive control element (blue), high-scoring element (red), low-scoring element (green), or negative control element (black) in competition assays. This data were fitted by the model data (purple) using the least-squares method and the affinities relative to that of positive control were determined from fitting data. The value of relative affinities (positive control is 1.0) are indicated. See also *SI Appendix* for more detail. (*B*) *In silico* analysis of affinity to amplitude mechanism when affinity of activator and repressor are balanced. The relative amplitudes of oscillation of output plotted against strength of regulators binding affinity when the affinity of an activator and a repressor are the same; *n* indicates the Hill coefficient at competitive regulation. Amplitude was normalized so that the maximum value at $n = 1$ is 100%.

Other Supporting Information Files

SI Appendix

Table S1. Functional RREs from known clock-controlled genes used in this study

<table>
<tr><th>Gene</th><th>GeneID</th><th>Element</th><th>Organism</th><th>chr</th><th>Start</th><th>End</th><th>Sequence</th><th>Evidence</th><th>Affymetrix Probe ID</th></tr>
<tr><td rowspan="2"><i>Elov15</i></td><td rowspan="2">68801</td><td rowspan="2">RRE</td><td>Mouse</td><td>9</td><td>78317303</td><td>78317325</td><td rowspan="2">Mouse TCTTGTAAATTGGGTCATGGCGT
      | || ||| | ||||||| |||
Human TTGTGCAAAGTAGGTCATGCCGT</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">93496_at</td></tr>
<tr><td>Human</td><td>6</td><td>53273932</td><td>53273910</td></tr>
<tr><td rowspan="2"><i>BC004004</i></td><td rowspan="2">80748</td><td rowspan="2">RRE</td><td>Mouse</td><td>17</td><td>27851244</td><td>27851266</td><td rowspan="2">Mouse AGTCTGAATATAGGTCAATGTGA
        | |||| ||||||| | |||
Human TTTGGGAATCTAGGTCATTCTGA</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">95517_i_at</td></tr>
<tr><td>Human</td><td>6</td><td>36927126</td><td>36927148</td></tr>
<tr><td rowspan="2"><i>Macf1</i></td><td rowspan="2">11426</td><td rowspan="2">RRE</td><td>Mouse</td><td>4</td><td>121994185</td><td>121994207</td><td rowspan="2">Mouse CCCTGAAAAGTAGGTCAGTGCCT
      |||||||||||||||||| ||||
Human CCCTGAAAAGTAGGTCAGCGCCT</td><td rowspan="2">ref. 2</td><td rowspan="2">98402_at</td></tr>
<tr><td>Human</td><td>1</td><td>39277738</td><td>39277716</td></tr>
<tr><td rowspan="2"><i>Atf5</i></td><td rowspan="2">107503</td><td rowspan="2">RRE</td><td>Mouse</td><td>7</td><td>32213618</td><td>32213640</td><td rowspan="2">Mouse CAAGTAAAACTGGGTCACGAAGG
        || |||||||||||| |||
Human GGAGGGTAACTGGGTCACGCAGG</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">103006_at</td></tr>
<tr><td>Human</td><td>19</td><td>55118372</td><td>55118350</td></tr>
<tr><td rowspan="2"><i>Col4a1</i></td><td rowspan="2">12826</td><td rowspan="2">RRE1</td><td>Mouse</td><td>8</td><td>11309356</td><td>11309378</td><td rowspan="2">Mouse GGCAGGAAAATGGGTCAGTGCTG
       ||||||||||||||||||| ||
Human AGCAGGAAAATGGGTCAGTGATG</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">101093_at</td></tr>
<tr><td>Human</td><td>13</td><td>109748860</td><td>109748882</td></tr>
<tr><td rowspan="2"><i>Col4a1</i></td><td rowspan="2">12826</td><td rowspan="2">RRE2</td><td>Mouse</td><td>8</td><td>11337102</td><td>11337124</td><td rowspan="2">Mouse TCAGCCAAACTAGGTCAAAACCT
      ||||||||| |||||||||||
Human TCAGCCAAAATAGGTCAAAACAG</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">101093_at</td></tr>
<tr><td>Human</td><td>13</td><td>109781193</td><td>109781215</td></tr>
<tr><td rowspan="2"><i>D19Ertd721e</i></td><td rowspan="2">225896</td><td rowspan="2">RRE</td><td>Mouse</td><td>19</td><td>7972840</td><td>7972862</td><td rowspan="2">Mouse AGGAAGAAAATAGGTCAGACATG
      | ||||||| ||||||||||||
Human AAGAAGAAATTAGGTCAGACATC</td><td rowspan="2">ref. 1, ref. 2</td><td rowspan="2">97240_g_at, 97241_at</td></tr>
<tr><td>Human</td><td>11</td><td>62201916</td><td>62201894</td></tr>
</table>

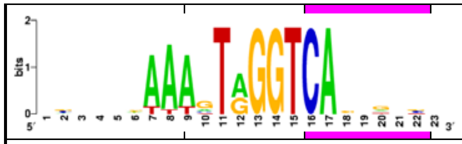
References
1. Ueda HR, *et al*. (2002) A transcription factor response element for gene expression during circadian night. *Nature* 418:534-539.
2. Panda S, *et al*. (2002) Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* 109:307-320.

Element information on the gene symbol ('Gene'), NCBI GeneID ('GeneID'), element type ('Element'), the position ('Start' and 'End') on a chromosome ('chr') of mouse and human ('Organism') and the sequence alignment of human-mouse RREs ('Sequence') are indicated. Experimental evidence on clock-controlled genes ('Evidence') and Affymetrix probe ID detecting circadian expression of the genes ('Affymetrix Probe ID') are also indicated. The canonical consensus sequences for the RRE ([A/T]A[A/T]NT[A/G]GGTCA) is indicated in red. The function of these mouse RREs sequences in the context of a luciferase reporter was experimentally determined in **Fig. S2**.

Table S2. Functional clock-controlled elements used as train HMMs

Gene	GeneID	Element	Organism	chr	Start	End	Score of HMMER	Sequence	Evidence
Per1	18626	E-box 1	Mouse	11	68707450	68707433	15.39	AGGGAACACGTGCAGGCT	ref. 1
Per1	18626	E-box 2	Mouse	11	68707634	68707617	17.87	TAGAGCCACGTGAGGGCG	ref. 1
Per1	18626	E-box 3	Mouse	11	68710224	68710241	11.48	TTTAGCCACGTGACAGTG	ref. 1
Per1	18626	E-box 4	Mouse	11	68710970	68710987	12.10	TAACGACACGTGGGCCCT	ref. 1
Per1	18626	E-box 5	Mouse	11	68711333	68711350	13.45	CAAGTCCACGTGCAGGGA	ref. 2
Nr1d1	217166	E-box 1	Mouse	11	98445288	98445271	14.71	CGGGCCCACGTGCTGCAT	ref. 2
Nr1d1	217166	E-box 2	Mouse	11	98445008	98444991	15.77	GGGTGCCACGTGCGAGGG	ref. 2
Nr1d2	353187	E-box 1	Mouse	14	14904545	14904528	16.38	ACTGGCCACGTGCACGGT	ref. 2
Nr1d2	353187	E-box 2	Mouse	14	14904442	14904425	17.71	CGGAGACACGTGAGGCCG	ref. 2
Dbp	13170	E-box	Mouse	7	33110149	33110166	15.23	CCTCGCCACGTGAGTCCG	ref. 2
Bhlhb2	20893	E-box	Mouse	6	109049680	109049697	18.55	CCGGGCCACGTGAAGCGT	ref. 2
Rorc	19885	E-box	Mouse	3	94357502	94357519	16.45	AGGTGCCACGTGCACCAG	ref. 2
									
Per1	18626	D-box	Mouse	11	68711473	68711496	17.30	GCCTGGCATTATGCAACCCGCCTC	ref. 2
Per2	18627	D-box	Mouse	1	91377818	91377795	16.48	TGTGCGTCTTATGTAAAGAGAGCG	ref. 2
Per3	18628	D-box 1	Mouse	4	148930820	148930797	18.62	CCCGCGCGTTATGTAAGGTACTCG	ref. 2
Per3	18628	D-box 2	Mouse	4	148930773	148930750	16.10	GCCCGCGGTTATGTAACCCCCGCC	ref. 2
Nr1d1	217166	D-box	Mouse	11	98445195	98445218	19.41	GGAGCTCATTATGTAACGAGGCCG	ref. 2

Nr1d2	353187	D-box	Mouse	14	14904456	14904479	17.41	AGCTCGCATTATGTAATGCTGCGT	ref. 2
Rora	19883	D-box 1	Mouse	9	69195245	69195222	11.78	AAGCTGTTTTATGTAATAGCTTTG	ref. 2
Rora	19883	D-box 2	Mouse	9	68926158	68926135	16.84	CACTGCTGTTATGTAACCAAACGT	ref. 2
Rora	19883	D-box 3	Mouse	9	68894279	68894302	17.42	CGAGCGGGTTATGTAACAGGGTTA	ref. 2
Rorb	225998	D-box	Mouse	19	18304970	18304993	14.84	TCCAGTTCTTATGTAATGAATATA	ref. 2
									
Arntl	11865	RRE	Mouse	7	100478066	100478088	19.72	AGGCAGAAAGTAGGTCAGGGACG	ref. 2
Npas2	18143	RRE	Mouse	1	39516271	39516293	14.66	GAAAAATATGTAGGTCAGTGGAA	ref. 2
Nfil3	114519	RRE 1	Rat	17	18094293	18094271	15.64	AGTGTGTTAGTAGGTCAGTTCCG	ref. 2
Nfil3	18030	RRE 2	Mouse	13	52011542	52011520	21.79	ACAGAAAAAGTGGGTCAGTTTGT	ref. 2
Clock	12753	RRE	Mouse	5	75037420	75037442	15.91	AGGAATAAAGTGGGTCACAAGGC	ref. 2
Cry1	12952	RRE 1	Mouse	10	84829678	84829656	20.81	GACTAGAAAGTAGGTCATTGTGA	ref. 2
Cry1	12952	RRE 2	Mouse	10	84829601	84829623	16.47	GTTTCTAAAGTAGGTCATCGCTA	ref. 2
Rorc	19885	RRE	Mouse	3	94352279	94352301	19.40	GGAATAAAAGTGGGTCATCTTGT	ref. 2
Elov15	68801	RRE	Mouse	9	78317303	78317325	17.19	TCTTGTAAATTGGGTCATGGCGT	This study
BC004004	80748	RRE	Mouse	17	27851266	27851244	17.53	AGTCTGAATATAGGTCAATGTGA	This study
Macf1	11426	RRE 2	Mouse	4	121994185	121994207	20.05	CCCTGAAAAGTAGGTCAGTGCCT	This study
Atf5	107503	RRE	Mouse	7	32213618	32213640	14.65	CAAGTAAAACTGGGTCACGAAGG	This study
Col4a1	12826	RRE 1	Mouse	8	11309378	11309356	19.92	GGCAGGAAAATGGGTCAGTGCTG	This study
Col4a1	12826	RRE 2	Mouse	8	11337124	11337102	14.10	TCAGCCAAACTAGGTCAAAACCT	This study
D19Ertd721e	225896	RRE	Mouse	19	7972862	7972840	18.30	AGGAAGAAAATAGGTCAGACATG	This study

References

1. Hida A, *et al*. (2000) The human and mouse Period1 genes: five well-conserved E-boxes additively contribute to the enhancement of mPer1 transcription. *Genomics* 65:224-233.
   and unpabsihed data of Ueda, H.R. *et al*.
2. Ueda HR, *et al*. (2005) System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat Genet* 37:187-192.

Element information on the gene symbol ('Gene'), NCBI GeneID ('GeneID'), element type ('Element'), the position ('Start' and 'End') on a chromosome ('chr') of mouse and rat ('Organism') and the sequence ('Sequence') are indicated. Experimental evidence on functional clock-controlled elements ('Evidence') and match score of HMMER search performed using the constructed HMMs ('Score of HMMER') is also indicated. Canonical consensus sequences for the E-box (CACGTG), D-box (TTATG[T/C]AA) and RRE ([A/T]A[A/T]NT[A/G]GGTCA) are indicated in red. Information contents of the relative frequency of each nucleic acid at the position of the pattern for each element were shown by sequence logos.

Table S3. Predicted clock-controlled elements used for experimental validation

<table>
<tr><th>No.</th><th>Gene</th><th>GeneID</th><th>Element</th><th>FDR</th><th>Organism</th><th>chr</th><th>Start</th><th>End</th><th>Score of HMMER</th><th>Sequence</th></tr>
<tr><td rowspan="2">-</td><td rowspan="2">*Per1*</td><td rowspan="2">18626</td><td rowspan="2">E-box</td><td rowspan="2">0.043</td><td>Mouse</td><td>11</td><td>68711333</td><td>68711350</td><td>13.45</td><td rowspan="2">Mouse CAAGTCCACGTGCAGGGA
      || ||||||||||
Human CAGGTCCACGTGCGCCCG</td></tr>
<tr><td>Human</td><td>17</td><td>7996626</td><td>7996609</td><td>16.13</td></tr>
<tr><td rowspan="2">1</td><td rowspan="2">*Esrra*</td><td rowspan="2">26379</td><td rowspan="2">E-box</td><td rowspan="2">0.002</td><td>Mouse</td><td>19</td><td>6635826</td><td>6635809</td><td>18.03</td><td rowspan="2">Mouse CGGACGCACGTGGCCCCG
      ||||||||||||||||||
Human CGGACGCACGTGGCCCCG</td></tr>
<tr><td>Human</td><td>11</td><td>63828852</td><td>63828869</td><td>18.03</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">*Foxd3*</td><td rowspan="2">15221</td><td rowspan="2">E-box</td><td rowspan="2">0.014</td><td>Mouse</td><td>4</td><td>97986368</td><td>97986351</td><td>17.23</td><td rowspan="2">Mouse CGCGGTCACGTGGCCCCG
      | ||||||||||||||||
Human CACGGTCACGTGGCCCCG</td></tr>
<tr><td>Human</td><td>1</td><td>63500197</td><td>64500214</td><td>15.64</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">*MGI:1926224*</td><td rowspan="2">268859</td><td rowspan="2">E-box</td><td rowspan="2">0.017</td><td>Mouse</td><td>16</td><td>6032279</td><td>6032296</td><td>14.92</td><td rowspan="2">Mouse AAGAGACACGTGCTGGGG
      ||| ||||||||| || |
Human AAGGGACACGTGCGGGCG</td></tr>
<tr><td>Human</td><td>16</td><td>6473013</td><td>6473030</td><td>17.35</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">*AK122525*</td><td rowspan="2">331623</td><td rowspan="2">E-box</td><td rowspan="2">0.021</td><td>Mouse</td><td>10</td><td>43388228</td><td>43388211</td><td>17.00</td><td rowspan="2">Mouse CCCCCGCACGTGGCGCCC
      ||||| ||||||||||||
Human CCCCCCCACGTGGCGCCC</td></tr>
<tr><td>Human</td><td>6</td><td>107543034</td><td>107543051</td><td>15.19</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2">*Lrrc35*</td><td rowspan="2">272589</td><td rowspan="2">E-box</td><td rowspan="2">0.021</td><td>Mouse</td><td>9</td><td>42544300</td><td>42544317</td><td>16.09</td><td rowspan="2">Mouse ACGTGACACGTGCGGCGG
      ||||||||||||||||||
Human ACGTGACACGTGCGGCGG</td></tr>
<tr><td>Human</td><td>11</td><td>120399989</td><td>120399972</td><td>16.09</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2">*1300001I01Rik*</td><td rowspan="2">74148</td><td rowspan="2">E-box</td><td rowspan="2">0.021</td><td>Mouse</td><td>11</td><td>74262482</td><td>74262499</td><td>15.91</td><td rowspan="2">Mouse CGGGGACACGTGCGCGCA
      ||||||||||||||||||
Human CGGGGACACGTGCGCGCA</td></tr>
<tr><td>Human</td><td>17</td><td>2561349</td><td>2561332</td><td>15.91</td></tr>
<tr><td rowspan="2">7</td><td rowspan="2">*Cpne7*</td><td rowspan="2">102278</td><td rowspan="2">E-box</td><td rowspan="2">0.023</td><td>Mouse</td><td>8</td><td>122590893</td><td>122590910</td><td>13.74</td><td rowspan="2">Mouse CCGAGACACGTGTGCCCG
      ||||| |||||| |||||
Human CCGAGCCACGTGCGCCCG</td></tr>
<tr><td>Human</td><td>16</td><td>88169660</td><td>88169677</td><td>17.87</td></tr>
<tr><td rowspan="2">8</td><td rowspan="2">*Pprc1*</td><td rowspan="2">226169</td><td rowspan="2">E-box</td><td rowspan="2">0.024</td><td>Mouse</td><td>19</td><td>45397616</td><td>45397633</td><td>15.71</td><td rowspan="2">Mouse CGGGTCCACGTGGGGGCG
      ||||||||||||||||||
Human CGGGTCCACGTGGGGGCG</td></tr>
<tr><td>Human</td><td>10</td><td>103883430</td><td>103883447</td><td>15.71</td></tr>
</table>

<table>
<tr><td rowspan="2">9</td><td rowspan="2"><i>Jag2</i></td><td rowspan="2">16450</td><td rowspan="2">E-box</td><td rowspan="2">0.024</td><td>Mouse</td><td>12</td><td>107682267</td><td>107682284</td><td>15.64</td><td rowspan="2">Mouse CCTGGCCACGTGGGCGCG
      ||||||||||||||||||
Human CCTGGCCACGTGGGCGCG</td></tr>
<tr><td>Human</td><td>14</td><td>104704954</td><td>104704971</td><td>15.64</td></tr>
<tr><td rowspan="2">10</td><td rowspan="2"><i>Rps19</i></td><td rowspan="2">20085</td><td rowspan="2">E-box</td><td rowspan="2">0.024</td><td>Mouse</td><td>7</td><td>13623927</td><td>13623944</td><td>15.62</td><td rowspan="2">Mouse CGCGGCCACGTGCGAGCG
      ||||||||||||||||||
Human CGCGGCCACGTGCGAGCG</td></tr>
<tr><td>Human</td><td>19</td><td>47056319</td><td>47056336</td><td>15.62</td></tr>
<tr><td rowspan="2">-</td><td rowspan="2"><i>Per3</i></td><td rowspan="2">18628</td><td rowspan="2">D-box</td><td rowspan="2">0</td><td>Mouse</td><td>4</td><td>148930797</td><td>148930820</td><td>18.62</td><td rowspan="2">Mouse CCCGCGCGTTATGTAAGGTACTCG
      |||||||||||||||| | | |
Human CCCGCGCGTTATGTAACGCGCCCC</td></tr>
<tr><td>Human</td><td>1</td><td>7778455</td><td>7778432</td><td>21.11</td></tr>
<tr><td rowspan="2">1</td><td rowspan="2"><i>Gria2</i></td><td rowspan="2">14800</td><td rowspan="2">D-box</td><td rowspan="2">0.051</td><td>Mouse</td><td>3</td><td>81194005</td><td>81193982</td><td>17.31</td><td rowspan="2">Mouse CGGGGCTGTTACATAATGCCCACC
      |||||||||||||||| |||||||
Human CGGGGCTGTTACATAACGCCCACC</td></tr>
<tr><td>Human</td><td>4</td><td>158501155</td><td>158501178</td><td>17.63</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2"><i>Spry4</i></td><td rowspan="2">24066</td><td rowspan="2">D-box</td><td rowspan="2">0.122</td><td>Mouse</td><td>18</td><td>39043332</td><td>39043309</td><td>16.38</td><td rowspan="2">Mouse AGGTGCGTTTACATAACGCCGGGC
      ||||||||||||||||| || |||
Human AGGTGCGTTTACATAACACCAGGC</td></tr>
<tr><td>Human</td><td>5</td><td>141683774</td><td>141683751</td><td>16.80</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2"><i>Etv4</i></td><td rowspan="2">18612</td><td rowspan="2">D-box</td><td rowspan="2">0.181</td><td>Mouse</td><td>11</td><td>101455967</td><td>101455990</td><td>17.09</td><td rowspan="2">Mouse CACACGTCTTATGTAACCCAGTTC
            |||||||||||| || |
Human ACACGTTCTTATGTAACCGAGCCC</td></tr>
<tr><td>Human</td><td>17</td><td>38978809</td><td>38978832</td><td>15.20</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2"><i>Pogz</i></td><td rowspan="2">229584</td><td rowspan="2">D-box</td><td rowspan="2">0.181</td><td>Mouse</td><td>3</td><td>94832639</td><td>94832662</td><td>16.89</td><td rowspan="2">Mouse CCCCTGTGTTATGTAATCCCGCTC
      ||| |||||||||||||| |
Human CCCTTGTGTTATGTAATCTCTGCT</td></tr>
<tr><td>Human</td><td>1</td><td>148244037</td><td>148244014</td><td>15.15</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2"><i>Fmn2</i></td><td rowspan="2">54418</td><td rowspan="2">D-box</td><td rowspan="2">0.209</td><td>Mouse</td><td>1</td><td>174628879</td><td>174628902</td><td>14.89</td><td rowspan="2">Mouse ACCGCGCATTATGCAAAGCGGCAG
       ||||||||||||||||||||| |
Human GCCGCGCATTATGCAAAGCGGCGG</td></tr>
<tr><td>Human</td><td>1</td><td>236581260</td><td>236581283</td><td>16.89</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2"><i>Plcb1</i></td><td rowspan="2">18795</td><td rowspan="2">D-box</td><td rowspan="2">0.325</td><td>Mouse</td><td>2</td><td>134535724</td><td>134535747</td><td>15.25</td><td rowspan="2">Mouse GGGGCGCGTTATGCAATGGGGCGC
      ||||||||||||||||||||
Human GGGGCGCGTTATGCAATGGGCGCA</td></tr>
<tr><td>Human</td><td>20</td><td>8061473</td><td>8061496</td><td>15.25</td></tr>
<tr><td rowspan="2">7</td><td rowspan="2"><i>Irf2bp1</i></td><td rowspan="2">272359</td><td rowspan="2">D-box</td><td rowspan="2">0.363</td><td>Mouse</td><td>7</td><td>10526613</td><td>10526636</td><td>16.21</td><td rowspan="2">Mouse CCCGCGCGTTATGTAACTTTCCCT
      | ||| |||||||||||||||||
Human CAGGCGTGTTATGTAACTTTCCCT</td></tr>
<tr><td>Human</td><td>19</td><td>51081280</td><td>51081257</td><td>13.67</td></tr>
<tr><td rowspan="2">8</td><td rowspan="2"><i>Slc22a20</i></td><td rowspan="2">381203</td><td rowspan="2">D-box</td><td rowspan="2">0.363</td><td>Mouse</td><td>19</td><td>5775101</td><td>5775078</td><td>14.95</td><td rowspan="2">Mouse CTGCCTTTTTACATAAGGCCTGGG
      | |||| | ||||||| ||| |||
Human CGGCCTCTCTACATAAAGCCGGGG</td></tr>
<tr><td>Human</td><td>11</td><td>64737873</td><td>64737896</td><td>14.81</td></tr>
<tr><td>9</td><td><i>Trim8</i></td><td>93679</td><td>D-box</td><td>0.387</td><td>Mouse</td><td>19</td><td>45844359</td><td>45844336</td><td>14.62</td><td>Mouse GACACTCATTACATAAACAGCAGC
      ||||||||||||||||||||||||
Human GACACTCATTACATAAACAGCAGC</td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td>Human</td><td>10</td><td>104395331</td><td>104395308</td><td>14.62</td><td></td></tr>
<tr><td rowspan="2">10</td><td rowspan="2"><i>Tspan7</i></td><td rowspan="2">21912</td><td rowspan="2">D-box</td><td rowspan="2">0.388</td><td>Mouse</td><td>X</td><td>8817583</td><td>8817560</td><td>13.68</td><td rowspan="2">Mouse AGAGGGTCTTACATAAGCCAGGGG
      ||||||||||||||||||| ||||
Human AGAGGGTCTTACATAAGCCGGGGG</td></tr>
<tr><td>Human</td><td>X</td><td>38177725</td><td>38177702</td><td>15.27</td></tr>
<tr><td rowspan="2">-</td><td rowspan="2"><i>Arntl</i></td><td rowspan="2">11865</td><td rowspan="2">RRE</td><td rowspan="2">0.066</td><td>Mouse</td><td>7</td><td>100478066</td><td>100478088</td><td>20.38</td><td rowspan="2">Mouse AGGCAGAAAGTAGGTCAGGGACG
      |||||||||||||||||||||||
Human AGGCAGAAAGTAGGTCAGGGACG</td></tr>
<tr><td>Human</td><td>11</td><td>13255933</td><td>13255955</td><td>20.38</td></tr>
<tr><td rowspan="2">1</td><td rowspan="2"><i>Atp1b2</i></td><td rowspan="2">11932</td><td rowspan="2">RRE</td><td rowspan="2">0.066</td><td>Mouse</td><td>11</td><td>69217609</td><td>69217631</td><td>20.61</td><td rowspan="2">Mouse GCGGAGAAAGTAGGTCACTGCCG
      |||||||||||||||||| ||||
Human GCGGAGAAAGTAGGTCACAGCCG</td></tr>
<tr><td>Human</td><td>17</td><td>7495878</td><td>7495856</td><td>19.61</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2"><i>Spsb3</i></td><td rowspan="2">79043</td><td rowspan="2">RRE</td><td rowspan="2">0.083</td><td>Mouse</td><td>17</td><td>23484948</td><td>23484970</td><td>18.81</td><td rowspan="2">Mouse GAACGGAAAGTGGGTCAGCGCCG
      | ||||||||||||||| ||||
Human GGGCGGAAAGTGGGTCAGGGCCG</td></tr>
<tr><td>Human</td><td>16</td><td>1772556</td><td>1772534</td><td>19.64</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2"><i>Zfp318</i></td><td rowspan="2">57908</td><td rowspan="2">RRE</td><td rowspan="2">0.104</td><td>Mouse</td><td>17</td><td>44564473</td><td>44564495</td><td>19.35</td><td rowspan="2">Mouse GAGAGAAAAGTGGGTCATTGAGA
       |||||||||||||||||||||
Human AAGAGAAAAGTGGGTCATTGAGC</td></tr>
<tr><td>Human</td><td>6</td><td>43417992</td><td>43417970</td><td>18.25</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2"><i>Mgea5</i></td><td rowspan="2">76055</td><td rowspan="2">RRE</td><td rowspan="2">0.112</td><td>Mouse</td><td>19</td><td>45123691</td><td>45123713</td><td>17.41</td><td rowspan="2">Mouse ATCTCCAAAGTAGGTCAGTGTCT
         | ||||||||||||| |
Human CCATAGAAAGTAGGTCAGTTCTT</td></tr>
<tr><td>Human</td><td>10</td><td>103568543</td><td>103568565</td><td>20.12</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2"><i>Lasp1</i></td><td rowspan="2">16796</td><td rowspan="2">RRE</td><td rowspan="2">0.139</td><td>Mouse</td><td>11</td><td>97476169</td><td>97476191</td><td>16.54</td><td rowspan="2">Mouse TTCGTGAAAGTGGGTCATGGTCT
      | | ||||||||||||||||||
Human TGTGAGAAAGTGGGTCATGGTCT</td></tr>
<tr><td>Human</td><td>17</td><td>34284865</td><td>34284887</td><td>19.34</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2"><i>Ryr1</i></td><td rowspan="2">20190</td><td rowspan="2">RRE</td><td rowspan="2">0.148</td><td>Mouse</td><td>7</td><td>18096374</td><td>18096352</td><td>17.94</td><td rowspan="2">Mouse AGGCTCTGACCTATTTAAATTCT
      || ||||||||||||||||||||
Human AGACTCTGACCTATTTAAATTCT</td></tr>
<tr><td>Human</td><td>19</td><td>43616749</td><td>43616771</td><td>17.45</td></tr>
<tr><td rowspan="2">7</td><td rowspan="2"><i>Adam12</i></td><td rowspan="2">11489</td><td rowspan="2">RRE</td><td rowspan="2">0.148</td><td>Mouse</td><td>7</td><td>121856769</td><td>121856791</td><td>15.44</td><td rowspan="2">Mouse TACTTAAAAGTAGGTCAGAAAAA
       |||| |||||||||||| || |
Human AACTTGAAAGTAGGTCAGTAAGA</td></tr>
<tr><td>Human</td><td>10</td><td>128077170</td><td>128077148</td><td>19.88</td></tr>
<tr><td rowspan="2">8</td><td rowspan="2"><i>A730009L09Rik</i></td><td rowspan="2">402727</td><td rowspan="2">RRE</td><td rowspan="2">0.168</td><td>Mouse</td><td>16</td><td>84900535</td><td>84900513</td><td>16.17</td><td rowspan="2">Mouse TCTTAATGACCCAATTTCTAAAT
      ||||||||||||| |||||||||
Human TCTTAATGACCCACTTTCTAAAT</td></tr>
<tr><td>Human</td><td>21</td><td>25632353</td><td>25632331</td><td>18.49</td></tr>
<tr><td>9</td><td><i>Myo1b</i></td><td>17912</td><td>RRE</td><td>0.179</td><td>Mouse</td><td>1</td><td>52270010</td><td>52269988</td><td>18.02</td><td>Mouse TGGTGCTGACCCACTTTCCTCTT
       | | ||||| ||||||| |||
Human GGATATTGACCTACTTTCCCCTT</td></tr>
</table>

					Human	2	191953461	191953483	16.47	
10	*D13Ertd150e*	52548	RRE	0.203	Mouse	13	37218548	37218570	16.99	Mouse GATGCGAAAGTGGGTCAGGAATG || ||||||||||||||||||||
					Human	6	6997855	6997877	16.67	Human GACGCGAAAGTGGGTCAGGAATG
1	*Shox2*	20429	low-score E-box	-	Mouse	3	67055808	67055791	0.64	Mouse GATAAACACGTGTGTATC ||||||||||||| ||||
					Human	3	159121615	159121598	0.64	Human GATAAACACGTGTATATC
2	*1700057H15Rik*	78460	low-score E-box	-	Mouse	4	122900755	122900738	1.39	Mouse TATTCTCACGTGATAAAC ||||||||||||||||||
					Human	1	38322547	38322564	1.39	Human TATTCTCACGTGATAAAC
3	-	435253	low-score E-box	-	Mouse	11	35821433	35821450	1.83	Mouse GACAAGCACGTGCCAGAC ||||||||||||||||||
					Human	5	167485618	167485601	1.83	Human GACAAGCACGTGCCAGAC
4	*Tbx15*	21384	low-score E-box	-	Mouse	3	99126851	99126834	1.83	Mouse GTTTTTCACGTGCTTGAC |||||||||||| ||||
					Human	1	119301655	119301672	2.39	Human GTTTTTCACGTGTTTGAT
5	*Stk4*	58231	low-score E-box	-	Mouse	2	164147516	164147533	2.12	Mouse GAAGAGCACGTGATCTGC ||| | |||||| | |||
					Human	20	43041498	43041515	2.28	Human GAAAACCACGTGGTTTGC
6	*Ube1x*	22201	low-score E-box	-	Mouse	X	19009993	19009976	2.61	Mouse GTTACTCACGTGAGGTAC ||||||||||||||||||
					Human	X	46825349	46825332	2.61	Human GTTACTCACGTGAGGTAC
7	-	211823	low-score E-box	-	Mouse	3	129597003	129596986	2.80	Mouse TGCAAACACGTGATTTCC ||||||||||||||||||
					Human	4	112078186	112078203	2.80	Human TGCAAACACGTGATTTCC
8	*Rnf44*	105239	low-score E-box	-	Mouse	13	53770222	53770239	2.80	Mouse GTAAAACACGTGGATTTT ||||||||||||| ||||
					Human	5	175887078	175887095	2.80	Human GTAAAACACGTGGGTTTT
9	*Pabpc1*	18458	low-score E-box	- -	Mouse	15	36663158	36663141	2.93	Mouse AAATACCACGTGTTGAAC ||||||||||||||||||
					Human	8	101802352	101802335	2.93	Human AAATACCACGTGTTGAAC
10	-	-	low-score E-box	-	Mouse	16	74421611	74421628	3.03	Mouse GCTAAGCACGTGGAAGTC ||||||||||||||||||
					Human	3	77581572	77581555	3.03	Human GCTAAGCACGTGGAAGTC

<table>
<tr><td rowspan="2">11</td><td rowspan="2"><i>C330011F03</i></td><td rowspan="2">328837</td><td rowspan="2">low-score E-box</td><td rowspan="2">--</td><td>Mouse</td><td>17</td><td>49541248</td><td>49541231</td><td>3.08</td><td rowspan="2">Mouse GTTTCCCACGTGTTTGGC
||||||||||||||||||
Human GTTTCCCACGTGTTTGGC</td></tr>
<tr><td>Human</td><td>3</td><td>17954468</td><td>17954451</td><td>3.08</td></tr>
<tr><td rowspan="2">12</td><td rowspan="2"><i>Prkce</i></td><td rowspan="2">18754</td><td rowspan="2">low-score E-box</td><td rowspan="2">-</td><td>Mouse</td><td>17</td><td>85008129</td><td>85008146</td><td>3.20</td><td rowspan="2">Mouse ATATAACACGTGCTAAAA
||||||||||||||||||
Human ATATAACACGTGCTAAAA</td></tr>
<tr><td>Human</td><td>2</td><td>46068028</td><td>46068045</td><td>3.20</td></tr>
<tr><td rowspan="2">13</td><td rowspan="2"><i>Map3k9</i></td><td rowspan="2">338372</td><td rowspan="2">low-score E-box</td><td rowspan="2">-</td><td>Mouse</td><td>12</td><td>77032273</td><td>77032290</td><td>2.44</td><td rowspan="2">Mouse GACAAACACGTGTGCGTC
||||||||||||| |
Human GACAAACACGTGTATGCA</td></tr>
<tr><td>Human</td><td>14</td><td>70276648</td><td>70276665</td><td>4.03</td></tr>
<tr><td rowspan="2">14</td><td rowspan="2"><i>Zfhx1b</i></td><td rowspan="2">24136</td><td rowspan="2">low-score E-box</td><td rowspan="2">-</td><td>Mouse</td><td>2</td><td>45057246</td><td>45057229</td><td>3.29</td><td rowspan="2">Mouse GTATTACACGTGAAAAGC
||||||||||||||||||
Human GTATTACACGTGAAAAGC</td></tr>
<tr><td>Human</td><td>2</td><td>145095987</td><td>145095970</td><td>3.29</td></tr>
<tr><td rowspan="2">15</td><td rowspan="2"><i>Ppfibp2</i></td><td rowspan="2">19024</td><td rowspan="2">low-score E-box</td><td rowspan="2">-</td><td>Mouse</td><td>7</td><td>94860226</td><td>94860243</td><td>4.93</td><td rowspan="2">Mouse GTTTCCCACGTGTGTCCC
||||||||||||| | |
Human GTTTCCCACGTGTTTGTC</td></tr>
<tr><td>Human</td><td>11</td><td>7489362</td><td>7489379</td><td>1.76</td></tr>
</table>

The element information including the gene symbol ('Gene'), NCBI GeneID ('GeneID'), element type ('Element'), the position ('Start' and 'End') on a chromosome ('chr') of mouse and human ('Organism'), the sequence alignments of human and mouse ('Sequence') and the match score of HMMER search ('Score of HMMER') are indicated. The estimated accuracy of the HMM-based prediction is indicated as false discovery rate ('FDR'). The mouse elements were used for experimental validation. The element type ‘low-score E-box’ indicate evolutionary conserved 15 most low-score E-boxes with the core consensus sequence ‘CACGTG’ in non-coding regions. The function of these element sequences in the context of a luciferase reporter was experimentally determined in **SI Fig. S3**.

Table S4. Similarity between DNA binding domains of DNA binding activators and repressors

Element	**Gene**	**Alignment**	**Accession**	**Domain position**	**Domain name**	**Type**	**Homology**
E-box	*Arntl*	`-KNAREAHSQIEKRRRDKMNSFIDELASLVPTCNAM--SRKLDKLTVLRMAVQHMKTLR----`	NP_031515.1	71-126	HLH	Activator	Identity 30%, Similarity 57% to *Bhlhb2*
	Clock	`DKAKRVSRNKSEKKRRDQFNVLIKELGSMLPGN-----ARKMDKSTVLQKSIDFLRKHKE---`	NP_031741.1	32-86		Activator	Identity 22%, Similarity 55% to *Bhlhb2*
	Bhlhb2	`KETYKLPHRLIEKKRRDRINECIAQLKDLLPEHLKLTTLGHLEKAVVLELTLKHVKALTNLID`	NP_035628.1	50-112		Repressor	-
		` :  : .:   **:***::*  * :* .::*         :::* .**. ::..::`					
D-box	*Dbp*	`KDEKYWSRRYKNNEAAKRSRDARRLKENQISVRAAFLEKENALLRQEVVAVRQE`	NP_058670.1	254-307	bZIP_2	Activator	Identity 44%, Similarity 69% to *Nfil3*
	Nfil3	`KDAMYWEKRRKNNEAAKRSREKRRLNDLVLENKLIALGEENATLKAELLSLK--`	NP_059069.1	72-123		Repressor	-
		`**  **.:* **********: ***::  :. :   * :*** *: *:::::`					
RRE	*Rora*	`IPCKICGDKSSGIHYGVITCEGCKGFFRRSQQSNATYS-CPRQKNCLIDRTSRNRCQHCRLQKCLAVGMSRDAVKFG`	NP_038674.1	71-146	zf-C4	Activator	Identity 65%, Similarity 81% to *Nr1d1*
	Nr1d1	`-LCKVCGDVASGFHYGVHACEGCKGFFRRSIQQNIQYKRCLKNENCSIVRINRNRCQQCRFKKCLSVGMSRDAVRFG`	NP_663409.2	132-207		Repressor	
		` **:*** :**:**** :*********** *.*  *. * :.:** * * .*****:** ::***:********:**`					

The gene symbol ('Gene') of DNA binding activators and repressors, target element ('Element'), alignment ('Alignment') of DNA binding domain, accession No. ('Accession') of protein sequence, position of DNA binding domain ('Domain position') at the protein sequence and name of the DNA binding domain ('Domain name') are indicated. Homology ('Homology') between DNA binding domains of DNA binding activators and repressors is also indicated.