

SUPPLEMENTARY MATERIALS:

PromH: Promoters Identification Using Orthologous Genomic Sequences V.V.Solovyev¹, I.A.Shahmuradov^{2*}

¹ Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, NY, 10549, USA; Email: victor@softberry.com

² Institute of Botany, Azerbaijan National Academy of Sciences, 370073, Baku, Azerbaijan.

* **Present address:** Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK; E-mail: ilham@cs.rhul.ac.uk

Table s1. Predicted TSS for genes with annotated promoters: TATA or unknown type*

Abbreviated name of gene	Annotated start position of pre-mRNA ¹	Annotated start position of CDS	Sequence source localization of analyzed 5'-region of gene	First positions of homology with ESTs ²	Predicted TSS position	Minimal distance ³	LDF	CONS (TSS) ⁴ %	CONS (TATA) ⁵ %	CONS-R ⁶ %	CONS (RES) ⁷ %	CONS (-100..+41) ⁸ %	CONS (total) ⁹ %
1	2	4	4	5	6	7	8	9	10	11	12	13	14
<i>h-HBB</i>	62137 ^P	62187	60137...62186 (-2000:+50) ^P	-23, -2, +1	62133, (-4)	4 (2)	4.45	66	100	92	82	78	47
<i>ol-HBB</i>	?	51033	49033...51032 (-2000:-1) ^c	-42, -34,	50981, (-10)	?	4.09	77	100	92	85	78	47
<i>h-HBD</i>	54740 ^P	54790	52740...54789 (-2000:+50) ^P	+1, +22	54736, (-4)	4 (4)	4.33	77	81	87	80	67	30
<i>ol-HBD</i>	?	46328	44328...46327 (-2000:-1) ^c	-42, -34	46277, (-9)	?	3.87	88	81	87	51	70	30
<i>h-HBE</i>	19486 ^P , 19289 ^P , 19488 ^P	19541	17289...19540 (-2000:+261) ^P	+1, +14, +16	19482, (+1, +6)¹⁰	5 (5)	4.07	88	75	71	88	79	48
<i>ol-HBE</i>	?	20461	18461...20460 (-2000:-1) ^c	-193 ^H , -28 ^K	20406, (-27)	?	3.43	66	71	75	88	79	48
<i>h-HBGA</i>	39414 ^P	?	37414...39913 (-2000:+500) ^P	-30, -3, -2, +1, 14	39415, (+1)	0 (0)	4.20	66	71	84	81	76	48
<i>h-HBGG</i>	34478 ^P	34531	32478...34530 (-2000:+52) ^P	-30, -6, -3, -2, +1, +3	34478, (+1)	0 (0)	4.16	66	71	82	81	53	50
<i>ol-HBGG</i>	?	28022	26022...28021 (-2000:-1) ^c	?	27969 (-53)¹¹	?	3.26	66	71	82	75	75	50
<i>h-MYL3</i>	1952 ^u , +1	2006	1...2005 (-1951:+54) ^u	-12, +1, +4, +13, +15	1947, (-5)	5 (5)	3.77	100	87	56	89	74	61
<i>r-MLC1V</i>	741 ^P , +1	808	1...807 (-740:+67) ^P	+11, +34	744, (+4)	3	2.94	41	75	66	79	82	61
<i>h-MLC1emb</i>	<2024 ^P , +1	2099	25...2098 (-2000:+74) ^P	+14, +17	2023, (-1)	1	3.13	66	25	50	56	43	28
<i>m-MLC1F</i>	1624 ^P , +1	1749	1...1748 (-1623:+125) ^c	-113, -101, +12, +13	1624, (+1)	0	2.55	25	62	40	45	43	28
<i>h-MYF4</i>	<1177 ^c	1177	1...1176 (-1176:-1) ^c	-152, -24	1126, (-27)	?	4.10	100	100	89	90	94	43
<i>m-MYOG</i>	1570 ^m , +1	1619	1...1618 (-1569:+49) ^m	+6, +10, +17	1568, (-2)	2 (27)	4.05	100	100	89	90	94	43
<i>h-PGAM-M</i>	921 ^P , +1	963	1...962 (-920:+42) ^P	-10, -6, +5, +8, +10, +18,	921, (+1)	0	4.17	88	81	56	71	66	40
<i>r-PGAM2</i>	1182 ^c ,	1182	1...1181 (-1181:-1) ^c	-5, +2, +53	1148, (-1)	1 (from TATA)	3.12	77	81	52	82	68	40

Table s1 (continued)

1	2	4	4	5	6	7	8	9	10	11	12	13	14
<i>h-NPPA</i>	>c(33843) ^c	33198/c	33199...35843/c (-2000:+645) ^p	+251, +313	34063, (-<220)	?	2.82	77	47	51	51	54	31
<i>r-NPPA</i>	3638 ^p , +1	3723	1638..3722 (-2000.+85) ^c	+46, +53	3637, (-1)	1	6.74	100	100	78	87	78	31
<i>h-GLUT4</i>	2033 ^m , +1	2225	33...2224 (-2000:+192) ^m	?	1928, (-105)	105 (?)	4.01	88	25	97	52	65	30
<i>m-GLUT4</i>	786 ^m , +1	?	1...955 (-785:+170) ^m	+15, +17, +57	740, (-46)	46	10.23	88	TATA ⁻	57	86	71	30

^a Promoter type for *r-MLC1V*, *h-MYF4*, *h-PGAM-M*, *h-NPPA*, *r-NPPA* and *hGLUT4* genes are not known. Other 15 genes were annotated to have TATA-promoters.

h-HBB, human beta-hemoglobin: gb AC U01317.1 (and references therein); *o1-HBB*, otolemur beta-hemoglobin: U60902; *h-HBD*, human delta-hemoglobin : U01317.1 (and references therein); *o1-HBD*, otolemur delta-hemoglobin: U60902; *h-HBE*, human epsilon-hemoglobin: U01317.1 (and references therein); *o1-HBE*, otolemur epsilon-hemoglobin: U60902; *h-HBG4*, human hemoglobin Gamma A: U01317.1 (and references therein); *h-HBG*, human hemoglobin Gamma-G: U01317.1 (and references therein); *o1-HBG*, otolemur hemoglobin gamma-G: U60902; *h-MYL3*, human ventricular myosin light chain gene: M76408 (s1); *r-MLC1V*, rat MLC1V gene encoding alkali myosin ventricle light chain: X16325; *h-MLC1emb*, human MLC1emb gene for embryonic myosin alkaline light chain: X58851, X55000 (s2); *m-MLC1F*, mouse MLC1F/MLC3F gene for myosin alkali light chain, exon 1: X12973 (s3); *h-MYF4*, human myogenin (MYF4) gene: AF050501; *m-MYOG*, mouse myogenin: M95800 (s4); *h-PGAM-M*, human phosphoglycerate mutase, PGAM-M, gene: J05073 (s5); *r-PGAM2*, rat gene phosphoglycerate mutase: NM_017328; *h-NPPA*, human atrial natriuretic factor ANF precursor (atrial natriuretic peptide ANP /prepronatriuretic/isoform 2) gene : AL021155; *r-NPPA*, rat atrial natriuretic factor (ANF) gene: J03267; *h-GLUT4*, human glucose transporter, GLUT4, gene: M91463; *m-GLUT4*, mouse glucose transporter gene: M29660.

1 “?” means that no annotation is available. ² Positions of ESTs corresponding to Alu or other repetitive sequences are not shown. ³ Minimal distance from annotated pre-mRNA/CDS or EST start; it is given if the corresponding distance from pre-mRNA start is not known and no other information is available.

⁴ **CONS (TSS)**, interspecies conservation level around TSS (-3...TSS...+5). ⁵ **CONS (TATA)**, interspecies conservation level of TATA box (-2...TATA...+6). ⁶ Interspecies conservation level right to TSS (+1...+40). ⁷ **CONS (REs)**, average interspecies conservation level of regulatory motifs left to TSS. ⁸ **CONS (-100..+41)** : interspecies conservation level of the region -100..+41. ⁹ **CONS (total)**: full-length interspecies conservation level of analyzed orthologous sequences. These abbreviations are also used in the Table s3.

¹⁰ For *h-HBE* gene, three alternative pre-mRNA start positions (19289, 19486, 19488) have been annotated, and the predicted TSS position is given relative to 19486 position. ¹¹ Predicted TSS position is given relative to CDS start position (information about TSS or corresponding EST position is not available).

Hereafter localizations of analysed 5'-regions are given relative to pre-mRNA (^p), mRNA (^m), CDS (^c) or UTR (^u) start positions. For *H-HBE* gene, three alternative pre-mRNA start positions (19289, 19486, 19488) have been annotated, and we use 5'-region including all of them. “?” is unknown location.

Table s2. Comparing TSS prediction results for 21 genes with promoters of TATA or unknown type by TSSW and PromH

	<i>o1-HBE</i>	<i>r-MLC1V</i>	<i>h-MYF4</i>	<i>m-MYOG</i>	<i>I-PGAM2</i>	<i>o1-HBGG</i>	Total estimation of TSS predictions for 21 genes		
							True (reasonable)	False – positive	False - negative
TSS predicted by TSSW	not found	not found	-1066	-862	not found	not found	15	3	6
TSS predicted by PromoterH	-10	+4	-27	-2	-1	-53¹	21	0	0

Putative TSSs, predicted from -300 to +100 (in relation to the annotated TSS), but always before CDS start, are believed as true or reasonable predictions and are shown in bold.

¹Predicted TSS position is given in relation to CDS start (TSS was not annotated).

Table s3. Predicted TSSs for genes with annotated TATA-less promoters ¹

Abbreviated name of gene	Annotated start position of pre-mRNA	Annotated start position of CDS	Sequence source localization of analyzed 5'-region of gene	Predicted TSS position	Minimal distance	LDF	CONS (TSS) %	CONS-R, %	CONS (REs), %	CONS (-100..+41), %	CONS (total), %
<i>I</i>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>	<u>12</u>
<i>h-NCX</i>	<2170	2170	1...2172 (-2169:+3)	1817 (-353) 1256 ^{TATA}	? (353?) ² 45	13.13	100	87	81	76	65
<i>m-NCX</i>	<u>1301</u>	1599	1...1600 (-1300:+300)	(-45)		4.71	88	71	78	62	65
<i>m-SURI</i>	<u>1200</u>	1346	1...1380 (-1199:+281)	1261 (+61)	61	23.11	69	46	66	50	33
<i>h-RPE65</i>	<u>1919</u>	1973	2001...4008 (-1918:+90)	1915 ^{TATA} (-4)	4	3.92	88	80	68	72	48
<i>h-HHEY1</i>	<u>2860</u>	2999	1001...4008 (-2859:+149)	2663 (-197)	197	22.59	55	64	77	65	45
<i>m-HHEY1</i>	<u>2841</u>	2979	1...2988 (-2840:+148)	2747 (-94)	94	20.51	77	53	72	62	45
<i>h-ABCD2</i>	<3474-3513	<u>3684</u>	3001...6686 (-3683:+3)	3504 (+1)	0	5.75	77	84	88	77	36
<i>m-ABCD2</i>	<3192-3231	<u>3394</u>	1001...3397 (-3393:+4)	3181 (-9)	9	8.15	83	69	91	77	36
<i>h-Nrr1</i>	<u>1516</u>	1867	1...1620 (-1515:+105) ³	1339 (-111)	111	9.42	66	89	65	68	34
<i>h-SP4</i>	2572-2647	<u>2858</u>	16061...19020 (-2858:+102) ⁴	2639 (+1)	0	22.29	100	100	86	84	53
<i>m-SP4</i>	2656-2731	<u>2942</u>	1001...4020 (-2941:+79)	2685 (+1)	0	27.55	77	94	86	82	53
<i>h-B29</i>	<u>998</u>	1090	1...1211 (-997:+214)	993 (+1)	0	5.67	88	69	84	81	44
<i>h-HEB</i>	<u>847</u>	>847	1...1303 (-1313:+214)	950 (+103)	103	7.92	88	59	65	63	34
<i>m-HEB</i>	<u>1351</u>	>1350	1...1586 (-1350:+236)	1271 (-80)	80	16.92	66	71	71	64	34
<i>h-PRNP</i>	<u>1561</u>	>1560	1001...2880 (-1560:+320)	1632 (+71)	71	10.43	50	47	65	51	40
<i>m-PRNP</i>	<u>1151</u>	>1150	1...1440 (-1150:+290)	1141 (-10)	10	16.30	66	65	67	53	40
<i>h-HIC-1</i>	1479 (major), 1480-1500	>1478	1...1631 (-1478:+153)	1504 ^{TATA} (+4)	4	4.68	100	94	91	93	61
<i>h-neu4</i>	<u>1259⁶</u>	1437 ⁶	1...1446 ⁷ (-1258:+188)	1258 ^{TATA} (-1)	1	3.26	88	66	81	79	67
<i>r-neu4</i>	303,348,359, 363 (major)	506	1...650 (-362:+288)	311 (+8)	8	7.99	100	83	71	76	62
<i>h-UTN</i>	<u>897</u>	?	1...1200 (-896:+304)	823 (-74)	74	8.50	100	70	66	64	63

Table s3 (continued)

<i>I</i>	2	3	4	5	6	7	8	9	10	11	12
<i>m-UTN</i>	<u>839</u>	?	1...1200 (-838:+362)	795 (-44)	44	4.71	66	76	81	65	63
<i>h-KDR/fk1</i>	<u>781</u>	1084	1...1020 (-503:+37)	775 (-5)	5	24.32	55	68	57	60	60
<i>m-KDR/fk1</i>	<u>213⁷</u>	<u>504</u>	1...540 (-503:+37)	207 (-6)	6	18.92	88	58	75	57	60
<i>h-UROS(hp)</i>	<u>1558</u>	?	1...1757 (-1557:+200)	1675 (-1117)	117	6.38	25	45	32	35	42
<i>m-UROS(hp)</i>	<u>1773 (major)</u> , 1834	?	1...2040 (-1772:+268)	1523 (-250)	250	5.95	69	60	64	68	42
<i>h-Smad7</i>	<u>1336⁸</u>	>1377	1...1377 (-1335:+42)	1196 (-140)	140	25.03	88	84	90	88	77
<i>r-Smad7</i>	<1320, <u>1316</u>	2741	1...1440 (-1315:+125)	1085 (-231)	231	17.29	88	88	84	79	75

h-NCX, human NCX gene: gb B008502 (s6); *m-NCX*, mouse NCX gene: AB008500 (s6); *m-SURL*, mouse sulfonylurea receptor-1 gene: AF037274(s7); *h-RPE65*, human retinal pigment epithelium-specific 61 kDa protein gene: U20476 (s8); *h-HEY1*, human partial *HEY1* gene for bHLH transcription factor: AJ277506 (s9); *m-HEY1*, mouse partial *HEY1* gene for bHLH transcription factor: AJ277507 (s9); *h-ABCD2*, human adrenoleukodystrophy-related gene: AF302501 (s10); *m-ABCD2*, mouse adrenoleukodystrophy-related gene: AF302500 (s10); *h-Ntr1*, human neurotensin receptor gene: U69634 (s11, s12); *h-Sp4*, human *Sp4* gene: AC004595 (s13); *m-Sp4*, mouse *Sp4* gene: AB019147 (s13); *h-B29*, human immunoglobulin-associated (*B29*) gene: U22954 (s14); *h-HEB*, human helix-loop-helix protein (*HEB*) gene: U35052 (s15); *m-HEB*, mouse E-box binding protein (*ME1*) gene: U19863 (s15); *h-PRNP*, human partial *PRNP* gene: AJ289875 (s16); *m-PRNP*, mouse prion (*PrP*) gene: U52821 (s16); *h-HIC-1*, human *HIC-1* gene: AJ404688 (s17); *h-neu*, human *c-erb B2/neu* protein (*HER2*) gene: J05264 (s18); *r-neu*, rat transforming oncogene neu: M61004 (s18); *h-UTN*, human utrophin gene: X95523 (s19); *m-UTN*, mouse utrophin gene: X95523 (s19); *h-KDR/fk1*, human KDR/flk-1 gene: X89776 (s20), *m-KDR/fk1* mouse KDR/flk-1 gene: X89777 (s20); *h-UROS(hp)*, human uroporphyrinogen III synthase (*UROS*) gene, housekeeping promoter: AF230663 (s21); mouse uroporphyrinogen III synthase (*UROS*) gene, housekeeping promoter: AF156731, AF254791 (s22); *r-Smad7*, rat *Smad7* gene: AF156727 (s22).

¹ Annotated promoter for human *c-erbB2/neu* gene is TATA-promoter. Other 37 genes have been annotated to have TATA-less promoters. ² Distance is relative to CDS start position. ³ Nucleotides from 1 to 1515 and 1506 to 1620 were taken from GenBank (AC U69634) and from the literature (10), respectively. ⁴ Human gene TSS was predicted by homology with mouse orthologous gene promoter. ⁵ gb AC J05264. ⁶ gb AC M16789. ⁷ Mouse TSS was predicted by homology with human orthologous gene promoter. ⁸ gb AC AF156731. ^{TATA} Predicted promoters are TATA-promoters.

TSS or start position of CDS used for numbering of analyzed 5' -region sequences are underlined. No TSS was predicted for the following genes: *m-RPE65*, mouse retinal pigment epithelium-specific 61 kDa protein gene: AF271297, interspecies conservation level – 48% (s9); *h-Ldc*, human lumican gene: AF239660, 51% (s23); *m-Ldc*, mouse lumican gene: AF186467, AF013262, 51% (s23); *m-NTR1*, mouse neurotensin receptor gene: AF172326, 34% (s11, s12); *m-B29*, mouse immunoglobulin-associated *B29* gene: AF002279, 44% (s14); *h-Slc10a1*, human sodium/bile acid cotransporter gene: AF184235, 37% (s24); *m-Slc10a1*, mouse sodium/bile acid cotransporter gene: AF190698, 37% (s24); *h-PITG1*, human pituitary tumor-transforming 1 gene: AF167560, 34% (s25); *m-PITG1*, mouse pituitary tumor transforming gene: AF060887, 34% (s25); *h-SURL*, human potassium channel subunit: AF053478, 33% (s7); *m-HIC-1*, mouse *HIC-1* gene: AF133277, 61% (s17).

Table s4. Comparing TSS prediction results for 40 genes with annotated TATA-less promoters by TSSW and PromH

	<i>h-ABCD2</i> (15)	<i>h-neu4</i> (74)	<i>m-UTN</i> (76)	<i>h-B29</i> (43)	<i>h-HIC-1</i> (61)	Total estimation of TSS predictions for 40 genes		
						True (reasonable)	False - positive	False - negative
TSS predicted by TSSW	-2550	-1104	-458	not found	-798	22	6	10
TSS predicted by PromH	+1	-1	-44	+1	+4	27	2	9

Putative TSSs, predicted from -300 to +150 (in relation to the annotated TSS), but always before CDS start, are believed as true or reasonable predictions (are given in bold).

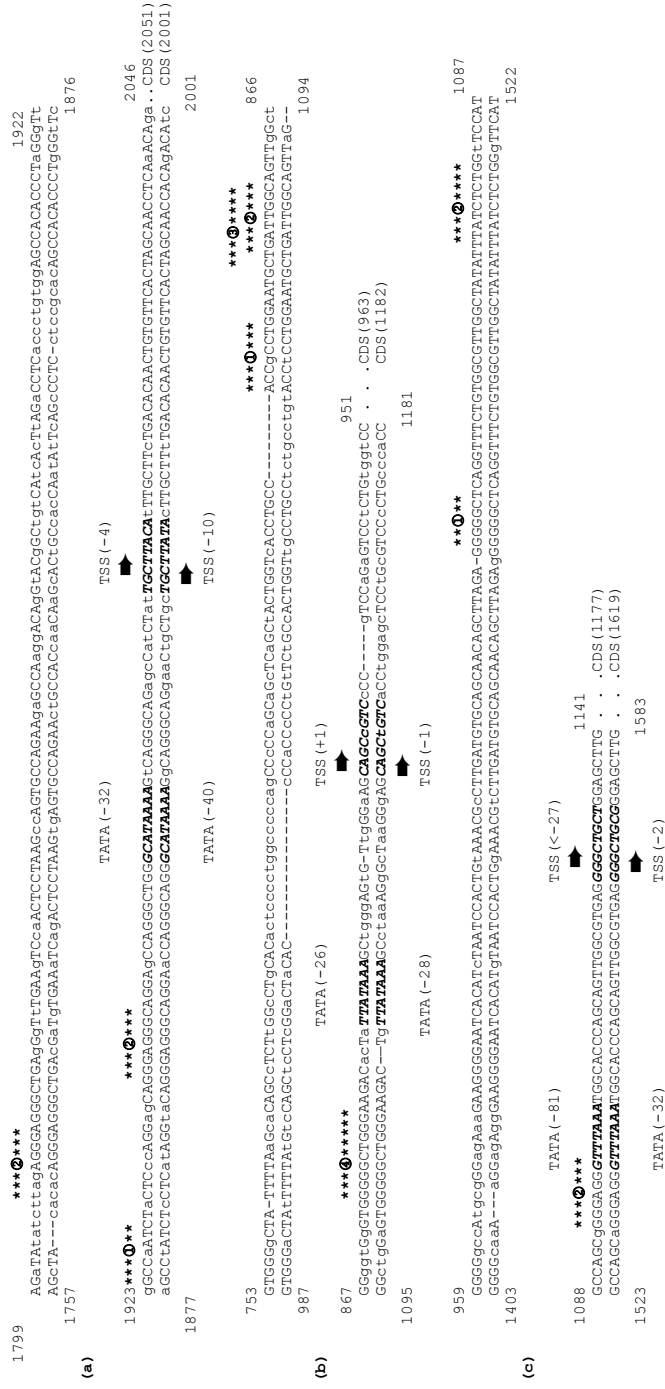


Fig. s1. Location of predicted TSS and TATA boxes in aligned sequences of *h-HBB* and *ol-HBB* (a), *h-PGAM-M* and *r-PGAM2* (b), *h-MYF4* and *m-MYOG* (c) orthologous gene pairs. Human sequences are presented in the corresponding top lines. TSS and TATA boxes are given in bold italic. Annotated start positions of pre-mRNA are denoted by horizontal arrow. When annotation about exact location of TATA-boxes and TSS was not available, positions for these elements are given relative to corresponding ESTs (*ol-HBB* and *h-MYF4*) or annotated TATA-signal (*r-PGAM2*). Some of conservative regulatory motifs from Transfac database (s26), denoted by encircled numbers, are also shown. For human-otolemur genes pair: 1 - MOUSE\$AAG, 2 - MOUSE\$CMYC; for human-rat genes pairs: 1 - HS\$TPI_04, 2 - PAsPY_21, 3 - MOUSE\$M2EB, 4 - SP1\$CONS; for human-mouse gene pairs: 1 - PAsPY_23, 2 - HS\$DG_02, 3 - MOUSE\$CMYC.

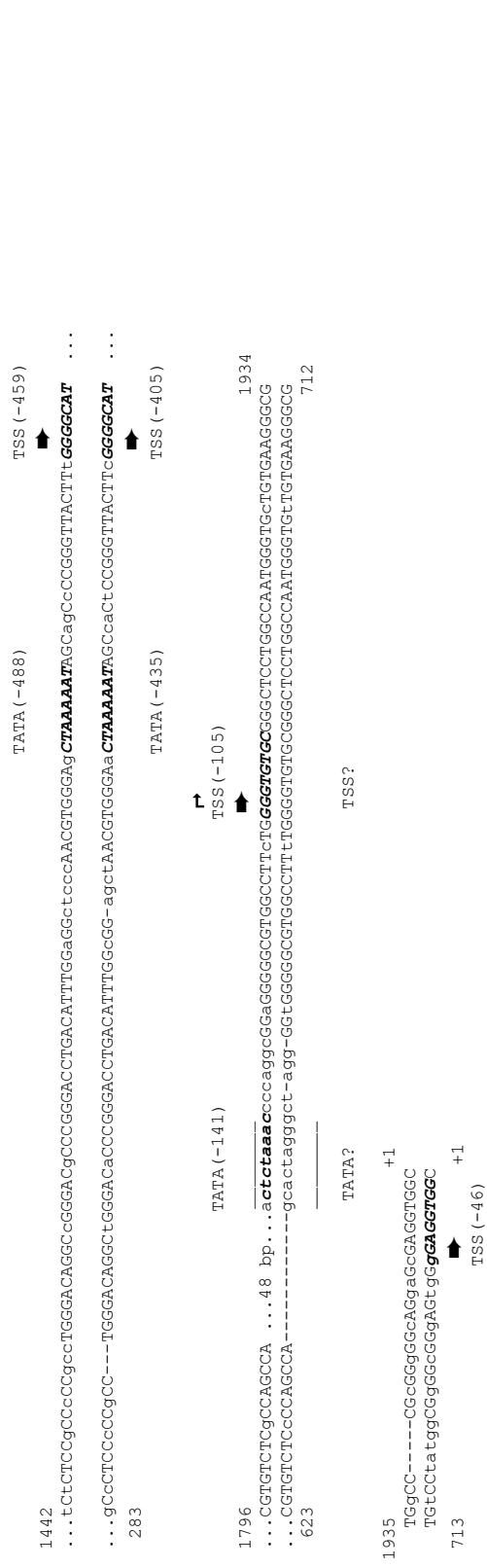


Fig. s2. Localization of predicted TSS and TATA boxes in the aligned sequences of human and mouse *GLUT4* genes (top line: *h-GLUT4*, bottom line: *m-GLUT4*). Numbering of TSS and TATA box positions corresponds to the Table 2. TSSs and TATA boxes are given in bold italic. Annotated start positions of pre-mRNA for both genes (+1) are also shown.

REFERENCES

- s1. Fodor, W.L., Darras, B., Seharaseyon, J., Falkenthal, S., Francke, U. and Vanin, E.F. (1989) Human ventricular/slow twitch myosin alkali light chain gene characterization, sequence, and chromosomal location. *J. Biol. Chem.* 264, 2143-2149.
- s2. Rotter, M., Zimmerman, K., Poustka, A., Soussi-Yanicostas, N. and Starzinski-Powitz, A. (1991) The human embryonic myosin alkali light chain gene: use of alternative promoters and 3' non-coding regions. *Nucleic Acids Res.* 19, 1497-1504.
- s3. Cohen, A., Barton, P.J., Robert, B., Garner, I., Alonso, S. and Buckingham, M.E. (1988) Promoter analysis of myosin alkali light chain genes expressed in mouse striated muscle. *Nucleic Acids Res.* , 16, 10037-10052.
- s4. Edmondson, D.G., Cheng, T.C., Cserjesi, P., Chakraborty, T. and Olson, E.N. (1992) Analysis of the myogenin promoter reveals an indirect pathway for positive autoregulation mediated by the muscle-specific enhancer factor MEF-2. *Mol. Cell. Biol.*, 12, 3665-3677.
- s5. Tsujino, S., Sakoda, S., Mizuno, R., Kobayashi, T., Suzuki, T., Kishimoto, S., Shanske, S., DiMauro, S. and Schon, E.A. Structure of the gene encoding the muscle-specific subunit of human phosphoglycerate mutase. *J. Biol. Chem.*, 264, 15334-15337.
- s6. Iitsuka, Y., Shimizu, H., Kang, M.M., Sasagawa, K., Sekiya, S., Tokuhisa, T. and Hatano, M. (1999) An enhancer element for expression of the *Ncx* (*Enx*, *Hox11L1*) gene in neural crest-derived cells. *J. Biol. Chem.*, 274, 24401-24407.
- s7. Hernandez-Sanchez, C., Ito, Y., Ferrer, J., Reitman, M. and LeRoith D. (1999) Characterization of the mouse sulfonyleurea receptor 1 promoter and its regulation. *J. Biol. Chem.*, 274, 18261-18270.
- s8. Boulanger, A., Liu, s., Henningsgaards, A.A., Yu, S. and Redmond, T.M. (2000) The upstream region of the *Rpe65* gene confers retinal pigment epithelium-specific expression *in vivo* and *in vitro* and contains critical octamer and E-box binding sites. *J. Biol. Chem.*, 275, 31274-31282.
- s9. Maier, M.M. and Gessler, M. (2000) Comparative analysis of the human and mouse *Hey1* promoter: *Hey* genes are new notch target genes. *Biochem. and Biophys. Res. Communic.*, 275, 652-660.
- s10. Pujol, A., Troffer-Charlier, N., Metzger, E., Chimini, G. and Mandel, J.-L. (2000) Characterization of the adrenoleukodystrophy-related (*ALDR*, *ABCD2*) gene promoter: inducibility by retinoic acid and forskolin. *Genomics*, 70, 131-139.
- s11. Tavares, D., Tully, K. and Dobner, P.R. (1999) Sequences required for induction of neurotensin receptor gene expression during neuronal differentiation of N1E-115 neuroblastoma cells. *J. Biol. Chem.*, 274, 30066-30079.
- s12. Le, f., Groshan, K., Zeng, X.P. and Richelson, E. (1997) Characterization of the genomic structure, promoter region, and a tetranucleotide repeat polymorphism of the human neurotensin receptor gene. *J. Biol. Chem.*, 272, 1315-1322.
- s13. Song, J., Mangold, M., Suske, G., Geltinger, C., Kanazawa, I., Sun, K., Yokoyama, K.K. (2001) Characterization and promoter analysis of the mouse gene for transcription factor Sp4. *Gene*, 264, 19-27.
- s14. Akerblad, P., Rosberg, M., Leanderson, T. and Sigvardsson, M. (1999) The *B29* (immunoglobulin β -chain) gene is a genetic target for early B-cell factor. *Mol. And Cell. Biology*, 19, 392-401.
- s15. Shain, D.H., Zuber, M.X., Norris, J., Yoo, J. and Neuman, T. (1998) Selective conservation of an E-protein gene promoter during vertebrate evolution. *FEBS Let.*, 440, 332-336.
- s16. Mahal, S.P., Asante E.A., Antoniou, M. and Collinge, J. (2001) Isolation and functional characterization of the promoter region of the human prion protein gene. *Gene*, 268, 105-114.
- s17. Guerardel, C., Deltour, S., Pinte, S., Monte, D., Begue, A., Godwin, A.K. and Leprince, D. (2001) Identification in the human candidate tumor suppressor gene *Hic-1* of a new major alternative TATA-less promoter positively regulated by p53. *J. Biol. Chem.*, 276, 3078-3089.
- s18. Suen, T.-C. and Goss, P.E. (2000) Characterization of a repressor element and a juxtaposed tissue-restricted activator element located on the distal *neu* gene promoter. *J. Biol. Chem.*, 275, 6600-6607.
- s19. Dennis, C.L., Tinsley, J.M., Deconinck, A.E. and Davies, K.E. (1996) Molecular and functional analysis of the utrophin promoter. *Nucl. Acids Res.*, 24, 1646-1652.
- s20. Wu, Y. and Patterson, C. (1999) The human *KDR/flk-1* gene contains a functional initiator element that is bound and transactivated by TFII-I. *J. Biol. Chem.*, 274, 3207-3214.
- s21. Aizencang, G., Solis, C., Bishop, D.F., Warner, C., Desnick, R.J. (2000) Human uroporphyrinogen-III synthase: genomic organization, alternative promoters, and erythroid-specific expression. *Genomics*, 70, 223-231.

- s22. Stopa, M., Anhof, D., Terstegen, L., Gatsioisi, P., Gressner, A.M. and Dooley, S. (2000) Participation of Smad2, Smad3, and Smad4 in transforming growth factor β (TGF- β)-induced activation of Smad7. *J. Biol. Chem.*, 275, 29308-29317.
- s23. Grover, J., Liu, C.-Y., Kao, W.-Y. and Roughley, P.J. (2000) Analysis of the human laminin gene promoter. *J. Biol. Chem.*, 275, 40967-40973.
- s24. Shiao, T., Iwahashi, M., Fortune, J., Quattrocchi, L., Bowman, S., Wick, M., Qadri, I. and Simon, F.R. (2000) Structural and functional characterization of liver cell-specific activity of the human sodium/taurocholate cotransporter. *Genomics*, 69, 203-213.
- s25. Kakar, S.S. (1999) Molecular cloning, genomic organization, and identification of the promoter for the human pituitary tumor transforming gene (PTTG). *Gene*, 240, 317-324.
- s1. Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., Meinhardt, T., Pr  , M., Reuter, I. and Schacherer, F. (2000) TRANSFAC: an integrated system for gene expression regulation *Nucleic Acids Res.* 28, 316-319.