Table S2. Mapped transcription start sites (TSS)

Gene[a]	Score (bit)[b]	BRE/TATA box sequence	TSS[c]	Coord[d]	Dir[e]	Dist (nt)[f]	PE[g]	RACE[h]	Protein detected[i]	Div[j]
MJ0007	10.71	AAGATAAGATTTTTATG	ACTCTTTTAaT**g**AATT	8805	+	82	+	+	+	+
MJ0035	10.77	ATGGTAATCTTTTTATT	AAAAATATATC**c**GCTT	37321	−	52	ND	+	+	+
MJ0068	12.25	AACAATAACTTTAAATA	CTAAGGTAATT**a**ATAA	65915	−	54	ND	+	+	+
MJ0084	12.06	TTAGAAAAGATTTATA	TAGGTTTTAAt**t**AAAT	78718	+	28	ND	+	+	−
MJ0094	7.92	CTGGTAAAAAATAAAAA	AGATAGTTTTT**a**AGAA	89505	−	24	ND	+	+	−
MJ0113	14.65	ATAATAACGTTTATATA	AACTTAATCTT**a**ATCT	111553	−	38	ND	+	+	+
MJ0136A	13.92	GCAGTAAGCATTTTATA	AATATTAAAAT**a**CTAC	131332	+	135	+	+	+	+
MJ0136B	6.93	TAAAGAATATTAAAATA	TATAGTTTTTT**g**CAAA	131355	+	112	+	+	+	+
MJ0161A	16.30	TCCCAAAAATTTATATA	TAAATCCTATT**a**CCCA	165485	−	326	ND	+	+	+
MJ0161B	10.26	ATCATAATTATTATATG	ATATCTATTTT**a**TTAT	165200	−	41	ND	+	+	+
MJ0168	19.63	ACCGAAAAGTTTAAATA	CAGTATTTATT**g**GAAA	171034	−	24	ND	+	+	+
MJ0176	18.83	ACCAAAAATTTTATATA	TATTTGTTTAT**a**CACT	178134	+	286	ND	+	+	+
MJ0199	14.52	TCCGATAATTATATATA	CTTAATCAATT**a**TTAA	192346	+	36	ND	+	+	−
MJ0201	11.88	AAGATAAGTTATAAATA	TCTTAAACAAT**a**CGGC	193495	−	12	−	+	−	+
MJ0202	8.94	TAAGAATTCTTTAAATA	TAATCATATTT**a**TC**g**C	193488	+	88	−	+	+	+
MJ0205	13.72	AGCAAAAGATATTAATA	AAAATTAAAAG**a**AGGG	198495	−	45	ND	+	+	−
MJ0217	6.25	AGCTGAAAGTATAAAAG	AAAATATAAAT**a**AGGT	209106	−	159	ND	+	+	−
MJ0223	17.07	ATAAAAAATTTTATATA	AAAATATTACC**a**ATAA	214490	−	19	ND	+	+	−
MJ0224A	8.82	TGCAAAGATAAATATA	TTCCCATAATT**a**TAAA	215116	−	47	ND	+	−	+
MJ0224B	10.36	AAAGATAAATATATATT	TAATTATAAAA**a**CCTA	215110	−	41	ND	+	−	+
MJ0237	8.99	CCCATTGGGTTTTTATA	TAAATAAGTTT**a**TTTG	227871	+	149	ND	+	+	+
MJ0252	8.84	ACATCAACTTTTTTATA	AAACTATAACG**c**AAAA	238826	+	20	ND	+	+	+
MJ0269	16.17	CTCCAAAACTTTATATA	AATTGTTAATG**a**TTGA	255926	−	28	ND	+	+	−
MJ0299	15.97	AAAGTTAAATTTATATA	TAAATATAAAT**a**TTAA	282956	−	41	ND	+	+	+
MJ0307	20.41	TTGGAAAACTATTTATA	AAATTTTTATT**a**TAAA	288310	−	30	ND	+	+	−
MJ0313	16.26	ATGCAAAAATTTTTATA	TAATTAACATT**a**CTCC	293490	−	55	ND	+	+	+
MJ0318	16.18	ATGATAAGATATTAATA	AAACTATAATT**g**TTAA	297462	−	25	+	+	+	+
MJ0319	11.73	AACAATAATTTTAAATA	AATATAACAAT**t**AAAA	297586	+	23	+	+	+	+
MJ0369	9.05	CCAAATAATTATATATG	AGTATTCCCTT**a**CATG	336195	−	158	+	+	+	+
MJ0370A	15.81	ACCAATAGATTTTTATA	AATTGAATAAT**a**ATAA	336157	+	154	+	+	+	+

MJ0370B	7.72	AGGGAATACTATATAAA	ATTATTTGGTT**a**ATAG	336238	+	73	+	+	+	+
MJ0400	15.70	AACACAAATTTTAAATA	ATTTCTATATT**a**AAAC	361536	+	54	ND	+	+	+
MJ0405A	10.78	TTTGAACCCTTTATATA	AAATTGTTTAT**a**TAGA	365443	–	61	+	+	+	+
MJ0405B	13.06	CAAAAATTGTTTATATA	AAGTTAAAATT**a**TTGA	365418	–	36	+	+	+	+
MJ0406	9.37	CTAAAAATCTTTAAGTA	AATTTTAACTT**t**AAAG	365430	+	128	–	+	+	+
MJ0428A	12.76	CACAAAATATTTTTATA	AATTATAAATC**c**TTAA	385102	–	119	–	+	+	+
MJ0428B	6.83	TCCTTAAACTTTTTGTA	AGTTTTGGAGT**a**TATA	385063	–	80	–	+	+	+
MJ0429	7.39	TAGGTGAGTTATATATA	TAAGCAATAA**a**T**a**CA	385087	+	64	–	+	+	+
MJ0445A	14.82	ATCAACACATATATATA	AAGTAGATATT**a**CAAA	399406	–	226	ND	+	+	+
MJ0445B	7.82	AATCTAAGGTTTATTAA	CACTTTATATT**g**GACA	399347	–	167	ND	+	+	+
MJ0450	15.46	ACTATAAGATATAAATA	TAAAAATAATT**a**AGAA	403211	+	63	ND	+	+	+
MJ0507A	8.75	AAAAGTATTTATATATA	ATGGGGTTTTT**a**TATA	448306	–	38	ND	+	+	–
MJ0507B	12.39	TATATAAAGTATTTATA	ATATAATATCT**aa**GTT	448295	–	27	ND	+	+	–
MJ0510A	17.11	ACCGAAAATTTTATATA	ATCTTAGCAAT**a**TAAA	450696	–	140	ND	+	+	–
MJ0510B	12.06	TAAGTAAAATATTTATA	ATAATATGGAT**g**CTTT	450593	–	37	ND	+	+	+
MJ0533A	6.05	AATATTAATTATAAGAA	AAGATTACAAC**a**TTAA	469072	–	120	ND	+	+	+
MJ0533B	7.01	GATTACAACATTAAATA	AGTAATATATA**a**TAAA	469043	–	91	ND	+	+	+
MJ0533C	9.12	GGCGATAGTAATATATA	AATGTTAATTT**a**AGTA	469022	–	70	ND	+	+	+
MJ0547	9.05	ACCGAAGAGTATATATT	ACTTTAGTGTT**g**TAAA	484446	–	39	ND	+	+	–
MJ0555	9.36	AACATAAATCTAAAATA	ATACATTAACT**a**CTAT	490758	+	24	ND	+	+	+
MJ0561	12.72	TTAATTAACTTTAAATA	AAAAAAATATA**a**GGTT	495573	+	30	ND	+	+	+
MJ0622	15.46	AAGTTAAAGTATATATA	AAGTAAATTTT**g**TATA	552424	–	24	ND	+	+	+
MJ0660	14.23	ATCCATAAGTATTTATA	AGGTTGTTATT**g**AGCG	586407	–	42	ND	+	+	–
MJ0666A	2.18	CCACTAATTTCTTTATC	TTATATTTTCT**a**AACT	592655	–	45	–	+	–	+
MJ0666B	8.98	TCTATTAAATTTTTATA	ATATTTTCTTT**a**ATAT	592635	–	25	–	+	–	+
MJ0666C	12.47	TTTCTAAACTTTATATA	TACTTCTCATT**g**GTGA	592619	–	9	–	+	–	+
MJ0667A	4.59	ATGAGAAGTAATATTAA	AAGTTTAGAAA**a**TATA	592661	+	41	+	+	+	+
MJ0667B	14.73	TAAAGAAATATATAAA	AAAAATTTAAT**ag**ATA	592676	+	26	+	+	+	+
MJ0673	12.91	ATAGTTAAGTTTTTATA	ATTCATTATTC**a**TAGA	599083	+	171	+	+	+	+
MJ0696	8.91	CCAAAAAATAATAAATC	AATAATAACTT**a**TATG	622498	–	147	–	+	–	+
MJ0697A	0.64	CCATATATGCATATAAG	TATATTGCATT**a**GATT	622521	+	25	–	+	+	+
MJ0697B	10.60	ATAAGTTATTATTTATA	TATTATTTTTT**g**GTGA	622537	+	9	+	+	+	+
MJ0720	14.51	AACAAAACATTTTTATG	TTATTAATATT**g**CT**a**T	654955	–	28	ND	+	+	+

MJ0746	16.23	ACCATTAATATTATATA	TATTTTATAATgTAAT	672878	−	31	+	+	+	−
MJ0765	15.95	TCAAAAATATATTTATA	AAGATTTTAATgTAAA	688516	−	62	ND	+	+	−
MJ0784	17.11	ATAGATAATTTTATATA	ACAATAATATTaCCAA	706977	+	38	ND	+	+	−
MJ0800	13.27	GCACAAAGGAATATATA	ATGTTAGAATTaATAA	723642	+	131	ND	+	+	−
MJ0822	15.20	ATCGTAAGATTTATATA	AGGTATATACCgTCAA	744550	−	39	+	+	+	+
MJ0823	12.63	GTAGAAACTTATAAAAA	TATAATGAATTaGAAA	744794	+	35	+	+	+	+
MJ0825	13.01	TCAGTTATGTTTATATA	AAATTATGAAAgATTT	747343	−	56	+	+	+	+
MJ0826	17.29	ACGAAAAATTTTAAATA	CATAATTAATTaAGCA	747429	+	32	+	+	−	+
MJ0841A	9.18	CAGGCAAAGTTTTAAAA	TAAAATCCAATaATTT	768302	−	78	−	+	+	+
MJ0841B	5.78	GTTTTAAAATTAAAATA	AATAATTTTTggTGAA	768294	−	70	−	+	+	+
MJ0842	5.38	AAAACTGGATATATATA	TAACTAATATCgAACC	768428	+	25	+	+	+	+
MJ0847	14.77	ATAGAAAACCTTATATA	AACTAAGTCTTaTGTA	773683	+	20	ND	+	+	−
MJ0864	5.05	GTTTCTAACTATATATA	TACCTTAATGTgAGGT	788088	−	11	+	+	+	+
MJ0865	4.26	ATACCTCATTATAAATA	TAATTTAAAGAgAAAT	788198	+	67	+	+	+	+
MJ0882A	7.77	TTAAGAACTGTTTTATT	TATAATAAATTaACTG	807717	+	61	ND	+	−	+
MJ0882B	3.43	ATTGTATTTTTTATAAT	AAGAATTTgTCaTAAA	807736	+	42	ND	+	−	+
MJ0891	10.05	CGCAAAAAGAATATATT	CAGTTTATAGTgAGGT	821616	+	89	+	+	+	+
MJ0923A	13.97	TTAAGAACTATTATATA	AAATATTAAATaGTAA	853279	+	88	−	+	−	+
MJ0923B	3.10	TGGTTAAATATTAAATA	TATAGTCAGAAgGTAA	853303	+	64	−	+	−	+
MJ0923C	12.13	ATAGTAACACATTTATA	TAATAGAAAATaTTTT	853316	+	51	−	+	−	+
MJ0923D	10.84	ATAGAAAATATTTATA	TAAATAAGTCCaATAG	853347	+	20	−	+	−	+
MJ0936	10.27	TAAAAAATTGTTATATA	TAGATAGTGTTaTAAA	865503	+	41	ND	+	+	−
MJ0952	9.08	ACAATAAGTTTTATTAA	TATCTCTTACTaTAGG	884415	−	118	+	+	+	+
MJ0953	13.82	ATAGTAAGAGATATATA	TAAAACTTATTgTAAG	884456	+	31	−	+	−	+
MJ0968	11.82	CACAATAATTCTTTATA	ATTTAATTAatgGGGT	901986	+	9	ND	+	+	+
MJ0986	11.84	GCAAAAACTATATATT	AAATAACTACCaATAA	917748	−	182	−	+	−	+
MJ0987	13.94	AGGAAAGATTTTTATA	ATTTTACTATCaGCAA	917767	+	141	+	+	+	+
MJ0990	11.26	CAGATTAATTTTTTATA	ACCTACTTTTAtTAAA	919937	−	23	+	+	+	+
MJ0991	14.60	AGAGGAAGCTATTTATA	AATTAAGTTTTaTTAG	920151	+	28	+	+	−	+
MJ0999	16.12	ACCGAAAGGTTTATATA	AACATATACATaCCGA	928545	−	31	+	+	+	+
MJ1018	7.77	ACTATCACATATATATA	ATCATCTATCTaCTGC	949195	−	70	+	+	+	+
MJ1019A	8.04	ACCATAAATTATAAAAT	CTAAAAGCAGTaGATA	949196	+	62	−	+	+	+
MJ1019B	8.26	TCCCAATCATATATATA	TAGGTATTTTggTGAA	949249	+	9	+	+	+	+

MJ1035A	9.16	TATATAACTTAAATAAA	TATTTTTTaATaACAA	967284	−	116	−	+	+	+
MJ1035B	11.30	GTAGTAAGATATTTATT	ATAATACTAATaTCCC	967199	−	31	+	+	+	+
MJ1053	14.97	TACAAAAACTTTAAATA	AAAATAGATaTaAGGA	990383	−	35	+	+	+	+
MJ1054	14.64	AACGAAAATTATAAGTA	AATTAATTGTTgGAGG	990847	+	117	+	+	−	+
MJ1099	15.18	AACATTAAGTTTTTATA	TAATCATTTTTgAAAT	1040346	−	25	−	+	+	+
MJ1146A	15.04	TAAACAATTATAAAAA	AAGTAATTATTaTTAA	1086364	−	33	+	+	+	+
MJ1146B	8.12	TTATAAAATATATATT	TATTATTAATCcTACA	1086357	−	26	+	+	+	+
MJ1158	14.68	TTAAAAAATAATATATA	TAAATTTTTTAaTGTA	1097639	−	41	−	+	+	+
MJ1185	12.79	AACAAAATCTATAAAAT	AAATAAAAATTaACAT	1125066	−	37	+	+	+	+
MJ1186	9.14	AATATAATATTTAAATA	CTTTACTTTTTgAAAA	1125163	+	28	−	+	−	+
MJ1225	12.75	AGGATAATCATTTTATA	ATATTAATATTaTGTG	1167042	+	58	+	+	+	+
MJ1228	14.47	ATAGTAATATTTATATA	TAGTTATTAAAaGTTT	1171550	−	33	ND	+	+	−
MJ1246	11.39	AACGACACATATAAAAA	ATATAAATTTTaCGGT	1189802	−	30	ND	+	+	+
MJ1249	9.12	TTGAGAAGTTAAATATA	AAATCTAAACaATAA	1191345	+	19	+	+	+	+
MJ1259	16.21	GCAAAAAACTATAAATA	TTGATATTAATaACTC	1201107	+	31	ND	+	+	−
MJ1260	13.38	ACCAAAAATTATTTAAA	CAATAATTATCcTGGC	1202056	+	227	ND	+	+	−
MJ1265	10.72	ATCAAAAGCTTTAAAG	TAATAGTAATTaTATA	1210366	−	25	+	+	+	+
MJ1266	16.91	CTCGAAAAATTTAAATA	AAAATGAAATTgCCAT	1210494	+	73	+	+	+	+
MJ1311	4.04	AACCAAAAAGCTAAAAG	AAATATTAATTtGAAG	1256405	+	47	ND	+	+	+
MJ1333	11.75	GTGTTAAAATATTTATA	TAATGATATTtaAAAT	1282486	+	24	+	+	+	+
MJ1334	11.74	TCACCAAATTATTTATA	ACTATATAATTgTAAC	1284162	−	42	+	+	+	+
MJ1335	9.28	TATGTTACAATTATATA	TAGGTATAAATaATTT	1284195	+	140	−	+	−	+
MJ1338	15.40	ATAATAAAGTATTTATA	AAACTCTAAATgCGTA	1286293	+	31	ND	+	+	−
MJ1403	13.57	ATCGTAAGATTTATATA	ATATACTGCATaAGTA	1365854	−	116	+	+	+	+
MJ1404	13.09	CTGTTAAACTTTAAATA	TAATATCAATCaTCAA	1366077	+	29	+	+	+	+
MJ1486	8.70	ACAAAAAACTATATTAA	AAAGCTAAAGTaTAGA	1458931	−	42	−	+	+	+
MJ1534	5.12	AATAGAAAACTTAATTA	ATATAATAATTgTTGT	1511775	+	31	+	+	+	+
MJ1543	6.20	ACCGAAAACCTTAAAAT	CTAATAAGGTaGCgA	1521373	+	39	ND	+	+	+
MJ1585	15.50	ATAGGAAATTATTAATA	TAACTATTTTTgTAGT	1559201	−	35	+	+	+	+
MJ1586	10.36	CCCACAATAAATATATG	AAAATAGTTATgTGAG	1559214	+	39	+	+	+	+
MJ1592	11.77	ACTATAAATTTTTTATA	GTATAAAGATTaAAAA	1564540	+	57	ND	+	+	−
MJ1597A	15.03	CCAAAAACTTTTTTATA	AAAACATTTTGaaGTA	1568422	+	178	ND	+	+	+
MJ1597B	6.27	AAAGATATTTTAAGTA	ATTTATTGTTTcTaTT	1568566	+	34	ND	+	+	+

Gene		Sequence 1	Sequence 2	Position						
MJ1619	12.94	CACTTAACCCATAAATA	AAATATTATTT**a**GGAT	1594518	−	29	ND	+	+	+
MJ1636	9.36	AACGGAAAATATTTTAA	AAATATCTATA**a**TTGA	1617181	+	54	ND	+	+	−
MJ0722*	8.17	TTAATAATATTAATATA	TATAGTTCaTT**g**CAAA	656690	−	138	+	ND	−	+
MJ0723*	13.07	TCCAAAATTTTTATATA	TATAGTTCaTT**g**CAAA	656604	+	147	+	ND	−	+
MJ0740*	9.34	GTGATTAATGATATATA	AGATAAATAAT**g**ACAA	668883	−	27	+	ND	+	−
MJ0210	NA	NA	processing site	NA	NA	NA	−	−	+	+
MJ0387	NA	NA	processing site	NA	NA	NA	ND	−	+	−
MJ0499	NA	NA	processing site	NA	NA	NA	ND	−	+	+
MJ0006	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ0135	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ0150	NA	NA	NA	NA	NA	NA	ND	−	+	+
MJ0209	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ0537	NA	NA	NA	NA	NA	NA	ND	−	+	−
MJ0672	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ0676	NA	NA	NA	NA	NA	NA	ND	−	+	−
MJ0721	NA	NA	NA	NA	NA	NA	ND	−	+	+
MJ0732	NA	NA	NA	NA	NA	NA	ND	−	+	−
MJ0890	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ0922	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1000	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1036	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1100	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1147	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1157	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1197	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1224	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1248	NA	NA	NA	NA	NA	NA	−	−	−	+
MJ1332	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1423	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1424	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1533	NA	NA	NA	NA	NA	NA	−	−	+	+
MJ1555	NA	NA	NA	NA	NA	NA	ND	−	+	−
MJ1656	NA	NA	NA	NA	NA	NA	ND	−	+	−

[a]Suffix A, B, C, or D was used when more than one promoter was identified for a gene; *, TSS previously determined (1).

[b]Promoter score calculated as the sum of the scores of all the promoter elements (BRE, TATA box, PPE, and Inr) and a spacer penalty score applied for suboptimal spacing between the TATA box and the TSS.
[c]Lower case bold, major start site; lower case but not bold, alternative start site; NA, not available.
[d]Coordinate of the major TSS in the chromosome of *M. jannaschii* (NCBI accession no. NC_000909.1).
[e]Direction of transcription: +, transcription toward a larger genomic coordinate; –, transcription toward a smaller genomic coordinate.
[f]Distance from the major TSS to the start of the nearest downstream gene (translation start locations from David E. Graham, University of Texas).
[g]Primer extension: +, TSS mapped; –, experiment done but TSS not observed; ND, experiment not done.
[h]Rapid amplification of 5′ cDNA ends: +, TSS mapped; –, experiment done but TSS not observed; ND, experiment not done.
[i]+, protein product detected by proteomics (2); –, protein product not detected.
[j]+, divergent gene; –, not divergent gene.

REFERENCES

1. Ouhammouch,M., Dewhurst,R.E., Hausner,W., Thomm,M. and Geiduschek,E.P. (2003) Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc. Natl. Acad. Sci. USA*, **100**, 5097-5102.
2. Zhu,W., Reich,C.I., Olsen,G.J., Giometti,C.S. and Yates,J.R.,3rd. (2004) Shotgun proteomics of *Methanococcus jannaschii* and insights into methanogenesis. *J. Proteome Res.*, **3**, 538-548.