

Table 1. Protein binding DNA sequences

No.	ID	Sequence	Prediction	S/N prox	ChIP
1	stSG598148	ATCTCCCTGGACTTTGGACCCTCAGC	10	28.44	3.15
2	stSG601538	TGCAAACAGGGCAAAGTCCAAACTAC	10	26.9	2.69
3*	stSG601804B	GGAGATGTGGACTTTGCCCCCATGAG	10	70.01	3.45
4	stSG602885	ATAATCCTGGACCTTGGCCCACTTCA	10	25.1	5.75
5	stSG607281	AGGGCTGTGCACCTTGTACTATTTCA	10	43.1	8.36
6	stSG603946	ATTTAGGGGTCCAAAGTGCAGCTGTG	10	14.02	4.01
7	stSG609153	CACTTCCTGAACTTTGCCCTTCCCCTCC	10	44.93	2.52
8	stSG619706	TTAAAAAGTGGCAAAGTCCCAATTCA	10	50.9	5.04
9	stSG622741	TCGCGCGGGGTCAAAGTCCATGCAAG	10	23.75	5.64
10	stSG630984	GTGTTTTTGTACTTTGCCCTAAGGATT	10	25.99	3.78
11	stSG633449	TAGACCCTGGACTCTGGATATCTCTC	10	13.83	12.86
12	stSG635516	ACAGAACAGTGCAAAGTTCAACCTGT	10	17.15	2.9
13	stSG601777	AACCCCTTCGCAAAGTCCCAGACCC	9	10.7	2.74
14	stSG601795	AGATCACTTCCCTTTGCCCTAATCAG	9	11.08	5.75
15*	stSG601807	AGGGGTCAGTCCAGAGGTCAGAGTCA	9	-	6.68
16	stSG622467	TGGCCCTGTCTCAGAGTCCATGGAGT	9	16.33	4.74
17	stSG623016	GATGTGAAGGACTTTGAGCTTCTTAG	9	6.28	3.4
18	stSG601166	CATGGTAAGGACTTTGACTTTTACT	8	4.08	3.63
19	stSG601963	TGGTGCTTATCAAAGGTCAGCCACT	8	7.16	4.13
20	stSG606999	CAATGTTTTCCAAAGTACAGCTGAG	8	5.13	7.5
21	stSG607585	TTGTCCAGTGCAAAGTACAAATCTCC	8	5.58	2.58
22	stSG604432	AGACTTAGGTACAAGGTGCAATACAA	8	2.23	2.6
23	stSG623108	GTGTTATATGACTCTGACTTCCATGA	8	3.56	6.76
24	stSG597766	TGCAAACAGCCCCCTTGACATGGGTA	7	3.6	12.53
25	stSG602142	TCACAACTGCACCTTGACCAGTTAA	7	1.84	4.43
26	stSG603065	AAAGCTGTGACCCTTGACTTGTGCA	7	1.64	4.16
27	stSG607120	GTAGCCTTGACCTCTGCACACCTTA	7	5.06	2.21
28*	stSG621988	TGGGGCGTGGACTTTGCTCCAGCAGC	7	4.59	4.52
29	stSG624412	ACTGAATACTCAAAGGCCAAGGCTA	7	4.92	3.79
30 [†]	stSG601804A	GGTCTCTTGTCAAGGTACACCCAAGA	6	5.31	3.45
31 [†]	stSG601810A	GGCGCTCCTGGCCTCTGCCCGTAAGC	6	5.39	3.86
32 [†]	stSG601817D	CAAGGCCTGAACCTTGAGCTGGGGAG	6	4	8.17
33	stSG607367	GTATACTTTGACCTTGCACTTTAAAC	6	2.68	3.84
34	stSG609102	GCCCTGTGCCCTCTGCCAGCTGGG	6	8.94	3.14
35	stSG616836	GCCCTCTGGCACTTTGGACAGGCTGG	6	4.99	2.77
36	stSG622520	GGGAGAAGGACCTTGGATTTGGCTT	6	2.71	2.63
37	stSG629094	ATTACCTTATCCAAGGGCAAACCTTT	6	-	3.66
38	stSG633466	AGTGGTGTTCAGAGTGCATGCTGTA	6	1.54	6.25
39	stSG636542	GTATTTTATCCAAAGGAAAGAGACAA	6	5.86	11.74
40	stSG609339	GTTGATTTTTCAAAGAGCAGTATAAA	5	0.77	3.26
41	stSG620120	GGGAACTGTGGCACAGGTCACTCTTG	5	1.73	7.4
42	stSG622527	CCCCGGTAGGACTTCGAACTTTCTGA	5	4.5	6.94
43	stSG626337	GGGCATTTGAACTTTGAACTGAAT	5	3.84	8.45
44	stSG628176	ATCTCATGGGACCCTGCAAGGGTTC	5	8.34	4.38
45	stSG633115	ACTAGGAGGCCCTTTGAACGTGGGA	5	1.45	2.76
46*	stSG601810B	CAGCAGGTGACCTTTGCCCAGCGCCC	4	0.55	3.86
47 [†]	stSG601804C	CAGGGTCACAAAAAGTCCAAGAGGCC	-	1.64	3.45
48 [†]	stSG601817A	CCCCCACTGAACCTTTGACCCCTGCC	-	3.58	8.17
49 [†]	stSG601817B	CTTAAGTTCCACATTTGCCAGGACCAG	-	2.07	8.17
50 [†]	stSG601817C	TGCCCACTCTATTTGCCCAGCCCCAG	-	-	8.17

The sequence column shows the variable part of proximity probe sequence. The prediction column shows Bioprospector predictions of the likelihood of HNF-4 α binding. S/N prox indicates the calculated S/N values of proximity ligation. ChIP indicates the fold enrichment of ChIP-chip arrays.

The colored sequences correspond to sites found in gene promoters, where binding sequences have previously been reported. *, sequences in red are binding sites commonly identified by Rada-Iglesias *et al.* (1 and in other previous studies. †, sequences in green are binding sites identified in this study. ‡, sequences in blue are binding sequences reported in previous studies only. The APOA4 promoter includes the sequences **stSG601804A**, **stSG601804B**, and **stSG601804C**. The APOC3 promoter includes the sequences **stSG601810B** and **stSG601810A**. APOA4/C3 intergenic includes the sequence **stSG601807**. APOA1 promoter includes the sequences **stSG601817A**, **stSG601817B**, **stSG601817C**, and **stSG601817D**. F10 promoter includes the sequence **stSG621988**.

1. Rada-Iglesias A, Wallerman O, Koch C, Ameer A, Enroth S, Clelland G, Wester K, Wilcox S, Dovey OM, Ellis PD, *et al.* (2005) *Hum Mol Genet* 14:3435-3447.