

Table S5 - 21 protein-DNA pairs
Test set composed of 42 known binding sites.

Protein	Gene	Binding position	Strand	Binding site	Refs
SP1 (HS)	DHFR (MM)	-93	+	GGGGCGGGG	(Kriwacki et al. 1992)
SP1 (HS)	DHFR (MM)	-45	+	GGGGCGGAG	(Kriwacki et al. 1992)
SP1 (HS)	DHFR (MM)	-141	+	GGGGCGGAG	(Kriwacki et al. 1992)
SP1 (HS)	MT2A (HS)	-139	+	GGGGCGGGG	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-467	+	GGGGCGGGA	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-488	+	GGGGCGGGA	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-521	+	GGGGCGGAG	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-500	+	TGGGCGGAG	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-479	+	TGGGCGGAG	(Kriwacki et al. 1992)
SP1 (HS)	SV40	-510	+	TGGGCGGAA	(Kriwacki et al. 1992)
EGR1 (MM)	JunD (MM)	-63	+	GCGGGGGCG	(Kriwacki et al. 1992)
EGR1 (MM)	JunD (MM)	-43	+	GCGGGGGCC	(Kriwacki et al. 1992)
EKLF (MM)	β-Globin(MM)	-87	-	AGGGTGTGG	(Perkins 1999; Lee et
Krox20 (RN)	PN1 (RN)	-44	+	GCGGGGGCG	(Erno et al. 1996)
Krox20 (RN)	PN1 (RN)	-63	+	GCGGGGGAG	(Erno et al. 1996)
Krox20 (RN)	PN1 (RN)	-69	+	CCGGGGGCG	(Erno et al. 1996)
EGR1 (HS)	IL2 (HS)	-305	-	GGGGGTGGG	(Skerka et al. 1995)
SP1 (HS)	IL2 (HS)	-308	-	GGTGGGGAT	(Skerka et al. 1995)
Krox20 (HS)	IL3 (HS)	-118	-	GCGGGGGGA	(Koyano-Nakagawa et
DB1 (HS)	IL3 (HS)	-127	-	GGGGGGAGG	(Koyano-Nakagawa et
SP1 (HS)	Dnmt1 (MM)	-155	+	GGGGAGGTG	(Kishikawa et al. 2002)
SP1 (HS)	HBV-LSA	-45	+	AAGGCGGGT	(Raney and McLachlan
EGR1 (HS)	FGF2 (HS)	-182	-	TCGGGGGCG	(Biesiada et al. 1996)
EGR1 (HS)	FGF2 (HS)	-281	+	GCGGGGGTG	(Biesiada et al. 1996)
SP1 (HS)	FGF2 (HS)	-185	-	GGGGCGGGG	(Biesiada et al. 1996)
SP1 (HS)	FGF2 (HS)	-284	+	TGGGCGGGG	(Biesiada et al. 1996)
SP1 (HS)	MDR1 (HS)	-52	+	GGGGCGTGG	(Bartsevich and Juliano
EGR1 (HS)	MDR1 (HS)	-49	+	GCGTGGGCT	(Bartsevich and Juliano
WT1 (HS)	MDR1 (HS)	-55	+	GCCGGGGCG	(Bartsevich and Juliano
WT1 (HS)	MDR1 (HS)	-49	+	GCGTGGGCT	(Bartsevich and Juliano
SP1 (RN)	Cx40 (RN)	-114	+	GGGGAGGGA	(Teunissen et al. 2002)
SP1 (RN)	Cx40 (RN)	-17	+	TGGGCGGGA	(Teunissen et al. 2002)
SP1 (RN)	Cx40 (RN)	-58	+	TGGGAGGAG	(Teunissen et al. 2002)
SP1 (HS)	hTERT (HS)	-91	-	GGGGCGGGG	(Kyo et al. 2000)
SP1 (HS)	hTERT (HS)	-32	-	GGGGAGGGG	(Kyo et al. 2000)
SP1 (HS)	hTERT (HS)	-10	-	AGGGCGGGG	(Kyo et al. 2000)
SP1 (HS)	hTERT (HS)	-57	-	GGGGCTGGG	(Kyo et al. 2000)
SP1 (HS)	hTERT (HS)	-111	-	GAGGCGGAG	(Kyo et al. 2000)
Krox20 (MM)	HOXa2 (MM)	-722	-	GCGTGGGTG	(Nonchev et al. 1996)
Krox20 (MM)	HOXa2 (MM)	-693	+	CTGTGGGCA	(Nonchev et al. 1996)
Krox20 (MM)	HOXb2 (MM)	-102	+	GAGGGGGAG	(Sham et al. 1993)
Krox20 (MM)	HOXb2 (MM)	-213	-	GCGTGGGTG	(Sham et al. 1993)
Krox20 (MM)	HOXb2 (MM)	-81	+	CCGTGGGAG	(Sham et al. 1993)

(HS: Homo sapiens; MM: Mus musculus; RN: Ratus norvegicus; HBV LSA: HBV large surface antigen)