Supplementary Table 2: The top 29 features with $P_{in} > 0.5$ in the BART model for the Sox2 ChIP-chip data (tabulated similarly as in Table 2 in the main text)

Feature	P_{in}	avg.	t	Notes
SoxOct*	1.000	3.85	15.61	CWTTNWTATGYAAAT
AAT	1.000	1.74	0.81	background
Sox2*	1.000	1.63	11.56	CWTTGTT
CA	1.000	1.57	-11.57	background
Length	0.986	1.61	7.78	sequence length
CTA	0.955	1.67	-0.35	background
CGC	0.927	1.29	7.10	background
CAA	0.923	1.48	5.67	background
cs	0.899	1.23	8.33	conservation
TA	0.847	1.39	-2.38	background
CAC	0.778	1.10	-8.11	background
GC	0.734	1.01	2.67	background
GAA	0.731	1.02	6.67	background
Nanog*	0.713	0.97	7.17	RSCGATTAWS
NKX25_Q5*	0.700	0.97	7.59	TSYCACTTSM
UF1H3b_Q6*	0.673	0.91	9.82	GCCCCWCCCCRCC
P53_Q2*	0.641	0.90	2.41	RGRCAWGNCY
SP1_Q4_01	0.625	0.81	14.19	NNGGGGCGGGGNN
T3R_Q6	0.601	0.80	1.88	MNTGWCCTN
ATC	0.599	0.76	-3.78	background
GATA_Q6*	0.594	0.82	1.62	WGATARN
NFAT_Q6	0.563	0.74	3.65	NANWGGAAAANN
KROX_Q6	0.557	0.73	12.14	CCCGCCCCCRCCCC
G/C	0.538	0.74	0.72	GC content
AR_Q2	0.533	0.73	4.81	AGWACATNWTGTTCT
BLIMP1_Q6	0.529	0.68	8.31	AGRAAGKGAAAGKR
SP1_Q6	0.518	0.68	14.34	NGGGGGCGGGGYN
SP1_Q2_01	0.515	0.68	13.90	CCCCGCCCCN
NERF_Q2	0.510	0.66	7.68	YRNCAGGAAGYRNSTBDS

*Motifs with reported functions in ESCs or differentiation.