

Supplementary Table S1. Summary of new vertebrate matrix profiles in the core JASPAR database. Hs: *Homo sapiens*; Mm: *Mus musculus*; Oc: *Oryctolagus cuniculus*; Rn: *Rattus norvegicus*; Rr: *Rattus rattus*

TF name	Species	TF Class	TF Family	Source
ARID3A	Mm	Helix-Turn-Helix	Arid	SELEX (1)
CTCF	Hs	Zinc-coordinating	BetaBetaAlpha-zinc finger	ChIP-seq (2)
EBF1	Mm	Zipper-Type	Helix-Loop-Helix	PAZAR
EGR1	Mm	Zinc-coordinating	BetaBetaAlpha-zinc finger	bacterial 1-hybrid (3)
ESR2	Hs	Zinc-coordinating	Hormone-nuclear Receptor	ChIP-chip (4)
ESRRB	Mm	Zinc-coordinating	Hormone-nuclear Receptor	ChIP-seq (5)
EWSR1-FLI1	Hs	Winged Helix-Turn-Helix	Ets	ChIP-seq (6)
FEV	Hs, Rn	Winged Helix-Turn-Helix	Ets	PAZAR
FOXA1	Hs	Winged Helix-Turn-Helix	Forkhead	ChIP-seq (7)
FOXO3	Hs, Mm	Winged Helix-Turn-Helix	Forkhead	PAZAR
HIF1::ARNT	Hs, Mm, Oc, Rr	Zipper-Type	Helix-Loop-Helix	Compiled (8)
HNF1B	Hs, Mm	Helix-Turn-Helix	Homeo	PAZAR
HOXA5	Hs, Mm	Helix-Turn-Helix	Homeo	PAZAR
INSM1	Hs	Zinc-coordinating	BetaBetaAlpha-zinc finger	PAZAR
MYC	Mm	Zipper-Type	Helix-Loop-Helix	ChIP-seq (5)
NFATC2	Hs, Mm, Rn	Ig-fold	Rel Homology Region	PAZAR
NFE2L2	Hs	Zipper-Type	Leucine Zipper	PAZAR
NFIC	Hs	Other	Nuclear Factor I-CCAAT-binding Transcription Factor (NFI-CTF)	High-throughput SELEX SAGE (9)
Nr2e3	Mm	Zinc-coordinating	Hormone-nuclear Receptor	SELEX (10)
NR4A2	Hs, Mm, Rn	Zinc-coordinating	Hormone-nuclear Receptor	PAZAR
PLAG1	Hs	Zinc-coordinating	BetaBetaAlpha-zinc finger	bacterial 1-hybrid (3)
Pou5f1	Mm	Helix-Turn-Helix	Homeo	ChIP-seq (5)
RXR::RAR_DR5	Hs	Zinc-coordinating	Hormone-nuclear Receptor	PAZAR
SOX2	Mm	Other Alpha-Helix	High Mobility Group (Box)	ChIP-seq (5)
SOX10	Hs, Mm, Rn	Other Alpha-Helix	High Mobility Group (Box)	PAZAR
STAT3	Mm	Ig-fold	Stat Protein	ChIP-seq (5)
TAL1::GATA1	Mm	Zipper-Type	Helix-Loop-Helix	ChIP-seq (submitted)
Tcfcp2l1	Mm	Other	CP2 Transcription Factor	ChIP-seq (5)
Zfx	Mm	Zinc-coordinating	BetaBetaAlpha-zinc finger	ChIP-seq (5)

References

- Herrscher, R.F., Kaplan, M.H., Lelsz, D.L., Das, C., Scheuermann, R. and Tucker, P.W. (1995) The immunoglobulin heavy-chain matrix-associating regions are bound by Bright: a B cell-specific trans-activator that describes a new DNA-binding protein family. *Genes & Development*, **9**, 3067-3082.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823-837.

3. Meng, X., Brodsky, M.H. and Wolfe, S.A. (2005) A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat Biotechnol*, **23**, 988-994.
4. Liu, Y., Gao, H., Marstrand, T.T., Ström, A., Valen, E., Sandelin, A., Gustafsson, J.-A. and Dahlman-Wright, K. (2008) The genome landscape of ERalpha- and ERbeta-binding DNA regions. *Proc Natl Acad Sci USA*, **105**, 2604-2609.
5. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J. *et al.* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106-1117.
6. Guillon, N., Tirode, F., Boeva, V., Zynovyev, A., Barillot, E. and Delattre, O. (2009) The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS ONE*, **4**, e4932.
7. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, **9**, R137.
8. Wenger, R.H., Stiehl, D.P. and Camenisch, G. (2005) Integration of oxygen signaling at the consensus HRE. *Sci STKE*, **2005**, re12.
9. Roulet, E., Busso, S., Camargo, A.A., Simpson, A.J.G., Mermod, N. and Bucher, P. (2002) High-throughput SELEX SAGE method for quantitative modeling of transcription-factor binding sites. *Nat Biotechnol*, **20**, 831-835.
10. Chen, J., Rattner, A. and Nathans, J. (2005) The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes. *J Neurosci*, **25**, 118-129.