

Supplementary material. I. Dataset of 2-cell library identified transcripts

The dataset, currently in Microsoft Excel format, may be downloaded from:
<ftp://ftp.informatics.jax.org/pub/informatics/datasets/index.html>

Supplementary material. II. Distribution of BLAST hits in the mouse genome assembly^a for the 2-cell and somatically expressed genes

Genes expressed only in somatic tissues

Gene symbol	CDS length	Match to a gene	Match to a paralog	Contiguous sequence match, >80% aligned	Contiguous sequence match, <80% aligned
<i>Ahsg</i>	1038	1	0	0	0
<i>Alb1</i>	1827	1	2	0	0
<i>Amy2</i>	1527	1	1	0	0
<i>Apoa4</i>	1185	1	0	3	36
<i>Car3</i>	1067	1	3	0	0
<i>Cryaa</i>	591	1	1	1	1
<i>Csnb</i>	696	1	0	0	3
<i>Eno1</i>	1305	1	9	9	8
<i>Fgb</i>	1446	1	3	0	0
<i>Fgfr2</i>	2466	1	23	0	5
<i>Gnat1</i>	1053	1	9	1	3
<i>Gpx1</i>	606	1	2	1	0
<i>Hbb-b2</i>	444	1	5	0	0
<i>Hp</i>	1044	1	0	0	0
<i>Hpxn</i>	1383	1	0	0	1
<i>Kap</i>	366	1	0	0	5
<i>Klk6</i>	786	1	28	0	0
<i>Knng</i>	1299	1	1	0	2
<i>Lalba</i>	432	1	0	0	0
<i>Lao1</i>	1572	1	2	0	0
<i>MGC37245</i>	1728	1	4	0	1
<i>Mup1</i>	543	1	7	0	1
<i>Pah</i>	1362	1	3	0	1
<i>Pou2af1</i>	771	1	0	1	1
<i>Rbp4</i>	606	1	0	0	0
<i>Rom1</i>	1056	1	0	0	0
<i>Sag</i>	1212	1	2	0	1
<i>Scp2</i>	432	1	1	3	0
<i>Sema4a</i>	2283	1	3	0	0
<i>Sepp1</i>	1143	1	0	0	1
<i>Slc13a3</i>	1803	1	5	2	33
<i>Slc2a4</i>	1533	1	4	0	1
<i>Spi1-1</i>	1242	1	27	0	2
<i>Syp</i>	927	1	1	0	0
<i>Thrsp</i>	453	1	0	0	0
<i>Tiarp-pending</i>	1413	1	1	0	0
<i>Trfr2</i>	2397	1	1	0	1
<i>Tr</i>	444	1	0	0	0
<i>Umod</i>	1929	1	3	0	1
<i>Vtn</i>	1437	1	0	0	0

2-cell-stage expressed genes

Gene symbol	CDS length	Match to a gene	Match to a paralog	Contiguous sequence match, >80% aligned	Contiguous sequence match, <80% aligned
<i>Ask-pending</i>	1992	1	0	0	4
<i>Bmi1</i>	750	1	6	0	0
<i>Btg4</i>	753	1	1	0	1
<i>Catnb</i>	2346	1	1	0	0
<i>Ccne1</i>	1476	1	1	0	0
<i>Cdr2</i>	1368	1	1	4	32
<i>Clptm1</i>	1548	1	0	0	1
<i>Cnih</i>	435	1	3	0	0
<i>Ctsl</i>	1005	1	6	0	3
<i>Dnajb9</i>	669	1	2	1	4
<i>Ebp2</i>	921	1	0	0	1
<i>Ei24</i>	1077	1	0	0	2
<i>Eif4g2</i>	2724	1	1	0	1
<i>Gtl6</i>	3312	1	0	0	1
<i>Hspa5</i>	1968	1	6	9	29
<i>Ingl</i>	840	1	3	0	7
<i>Maid</i>	933	1	0	0	0
<i>Map1lc3</i>	378	1	1	7	1
<i>Nsep1</i>	966	1	2	11	24
<i>Oas1d</i>	1086	1	9	2	1
<i>Odc</i>	1386	1	5	9	6
<i>Pa2g4</i>	1185	1	0	1	4
<i>Ppm1a</i>	1149	1	1	1	4
<i>Ppp1r2</i>	621	1	2	6	12
<i>Prps1</i>	957	1	3	2	3
<i>Psmc5</i>	1515	1	0	0	1
<i>Ptbp1</i>	1584	1	2	1	9
<i>Sat</i>	516	1	1	2	1
<i>Set</i>	870	1	3	19	30
<i>Siah2</i>	978	1	2	2	1
<i>Slc34a2</i>	2094	1	2	0	26
<i>Spin</i>	789	1	0	6	1
<i>Sui1-rs1</i>	342	1	1	26	23
<i>Timm17a</i>	516	1	1	4	2
<i>Tpt1</i>	519	1	0	11	16
<i>Ttrap</i>	1113	1	0	0	2
<i>Ube2d2</i>	444	1	4	23	15
<i>Vcp</i>	2421	1	8	1	5
<i>Xpo1</i>	1695	1	0	0	8
<i>Yap</i>	1419	1	3	0	15

^a Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520-562 (2002).

Supplementary material. III. Comparison of the genes from the 2-cell library with the dataset of “stemness” genes expressed in three types of stem cells^a.

A. List of “stemness” genes found in the 2-cell library.

Gene symbol	Description
<i>0610009E22Rik</i>	RIKEN cDNA 0610009E22 gene
<i>0610013E23Rik</i>	RIKEN cDNA 0610013E23 gene
<i>1110007M04Rik</i>	RIKEN cDNA 1110007M04 gene
<i>1110008B24Rik</i>	RIKEN cDNA 1110008B24 gene
<i>1200014I03Rik</i>	RIKEN cDNA 1200014I03 gene
<i>1300007B12Rik</i>	RIKEN cDNA 1300007B12 gene
<i>1500011J06Rik</i>	RIKEN cDNA 1500011J06 gene
<i>1500016H10Rik</i>	RIKEN cDNA 1500016H10 gene
<i>2310006I24Rik</i>	RIKEN cDNA 2310006I24 gene
<i>2410005K20Rik</i>	RIKEN cDNA 2410005K20 gene
<i>2410080P20Rik</i>	RIKEN cDNA 2410080P20 gene
<i>2410141K03Rik</i>	RIKEN cDNA 2410141K03 gene
<i>2610003B19Rik</i>	RIKEN cDNA 2610003B19 gene
<i>2610318G08Rik</i>	RIKEN cDNA 2610318G08 gene
<i>2610511E03Rik</i>	RIKEN cDNA 2610511E03 gene
<i>2700079K05Rik</i>	RIKEN cDNA 2700079K05 gene
<i>3110065C23Rik</i>	RIKEN cDNA 3110065C23 gene; expressed sequence AA408566
<i>4432405K22Rik</i>	RIKEN cDNA 4432405K22 gene
<i>5730409F23Rik</i>	RIKEN cDNA 5730409F23 gene
<i>5730591C18Rik</i>	RIKEN cDNA 5730591C18 gene
<i>5830482G23Rik</i>	RIKEN cDNA 5830482G23 gene
<i>6230400O18Rik</i>	RIKEN cDNA 6230400O18 gene
<i>AA407558</i>	expressed sequence AA407558; gene trap ankyrin repeat (Gtar-pending)
<i>AI327140</i>	Mus musculus, clone IMAGE:3491909, mRNA, partial cds; "expressed sequence AI327140"
<i>AL023001</i>	Mus musculus, clone IMAGE:4504751, mRNA, partial cds; "expressed sequence AL023001"
<i>Arih1</i>	ariadne ubiquitin-conjugating enzyme E2-binding protein homolog 1 Drosophila
<i>Bach1</i>	BTB and CNC homology 1
<i>BC003993</i>	cDNA sequence BC003993
<i>Chd1</i>	chromodomain helicase DNA binding protein 1
<i>Cops4</i>	COP9 (constitutive photomorphogenic) homolog, subunit 4 (Arabidopsis thaliana)
<i>Cops7a</i>	COP9 (constitutive photomorphogenic) homolog, subunit 7a (Arabidopsis thaliana)
<i>Crsp3</i>	cofactor required for Sp1 transcriptional activation, subunit 3 (130kD)
<i>D5Ertd363e</i>	RIKEN cDNA 2700038B03 gene; "D5Ertd363e"
<i>D5Ertd689e</i>	DNA segment, Chr 5, ERATO Doi 689, expressed
<i>D7Ertd373e</i>	DNA segment Chr 7 ERATO Doi 373 expressed
<i>Dnajb6</i>	DnaJ (Hsp40) homolog, subfamily B, member 6
<i>Eif4g2</i>	eukaryotic translation initiation factor 4 gamma 2
<i>Etl1</i>	enhancer trap locus 1
<i>Fbxo8</i>	f-box only protein 8

<i>Gfer</i>	growth factor, erv1 (<i>S. cerevisiae</i>)-like (augmenter of liver regeneration)
<i>Gnb1</i>	guanine nucleotide binding protein, beta 1
<i>Itga6</i>	integrin alpha 6
<i>Kif2</i>	kinesin heavy chain member 2
<i>Kras2</i>	Kirsten rat sarcoma oncogene 2 expressed
<i>Laptm4a</i>	lysosomal-associated protein transmembrane 4A
<i>Mki67ip</i>	Mki67 (FHA domain) interacting nucleolar phosphoprotein
<i>Mpdu1</i>	mannose-P-dolichol utilization defect 1
<i>Ncoa6ip</i>	Nuclear receptor coactivator 6 interacting protein
<i>Nyren18-pending</i>	NY-REN-18 antigen
<i>Pdcd2</i>	programmed cell death 2
<i>Plunc</i>	palate lung and nasal epithelium expressed transcript
<i>Ppp1r2</i>	protein phosphatase 1, regulatory (inhibitor) subunit 2
<i>Psmd11</i>	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11
<i>Ptpn2</i>	protein tyrosine phosphatase non-receptor type 2
<i>Pyp</i>	pyrophosphatase; also: "2010317E03Rik"
<i>Qprs</i>	glutamine-proline-tRNA synthetase
<i>Rabggtb</i>	RAB geranylgeranyl transferase, b subunit
<i>Rad23b</i>	RAD23b homolog (<i>S. cerevisiae</i>)
<i>Rnf4</i>	ring finger protein 4
<i>Ryk</i>	receptor-like tyrosine kinase
<i>Sh3d19</i>	SH3 domain protein D19
<i>Snrp1c</i>	U1 small nuclear ribonucleoprotein 1C
<i>Socs2</i>	suppressor of cytokine signaling 2
<i>Strn3</i>	striatin, calmodulin binding protein 3
<i>Thpp2</i>	tripeptidyl peptidase II
<i>Trif-pending</i>	Trif gene
<i>Ube2d2</i>	ubiquitin-conjugating enzyme E2D 2
<i>Uchrp</i>	ubiquitin c-terminal hydrolase related polypeptide
<i>Umps</i>	uridine monophosphate synthetase
<i>Usp9x</i>	ubiquitin specific protease 9, X chromosome
<i>Wtap</i>	Wilms' tumour 1-associating protein
<i>Xpo1</i>	exportin 1, CRM1 homolog (yeast)
<i>Yap</i>	yes-associated protein, 65 kDa
<i>Ywhah</i>	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide
<i>LOC218397</i>	LOC218397: Ras GTPase-activating protein 1 (GTPase-activating protein) (GAP) (Ras p21 protein activator) (p120GAP) (RasGAP)

B. List of “stemness” genes NOT found in the 2-cell library.

Gene symbol	Description
<i>1110002O23Rik</i>	RIKEN cDNA 1110002O23 gene
<i>1110004L07Rik</i>	RIKEN cDNA 1110004L07 gene
<i>1110054G21Rik</i>	RIKEN cDNA 1110054G21 gene
<i>1210001E11Rik</i>	RIKEN cDNA 1210001E11 gene
<i>1700016A15Rik</i>	RIKEN cDNA 1700016A15 gene
<i>1700030G05Rik</i>	RIKEN cDNA 1700030G05 gene
<i>1810009A15Rik</i>	RIKEN cDNA 1810009A15 gene
<i>1810012N18Rik</i>	RIKEN cDNA 1810012N18 gene

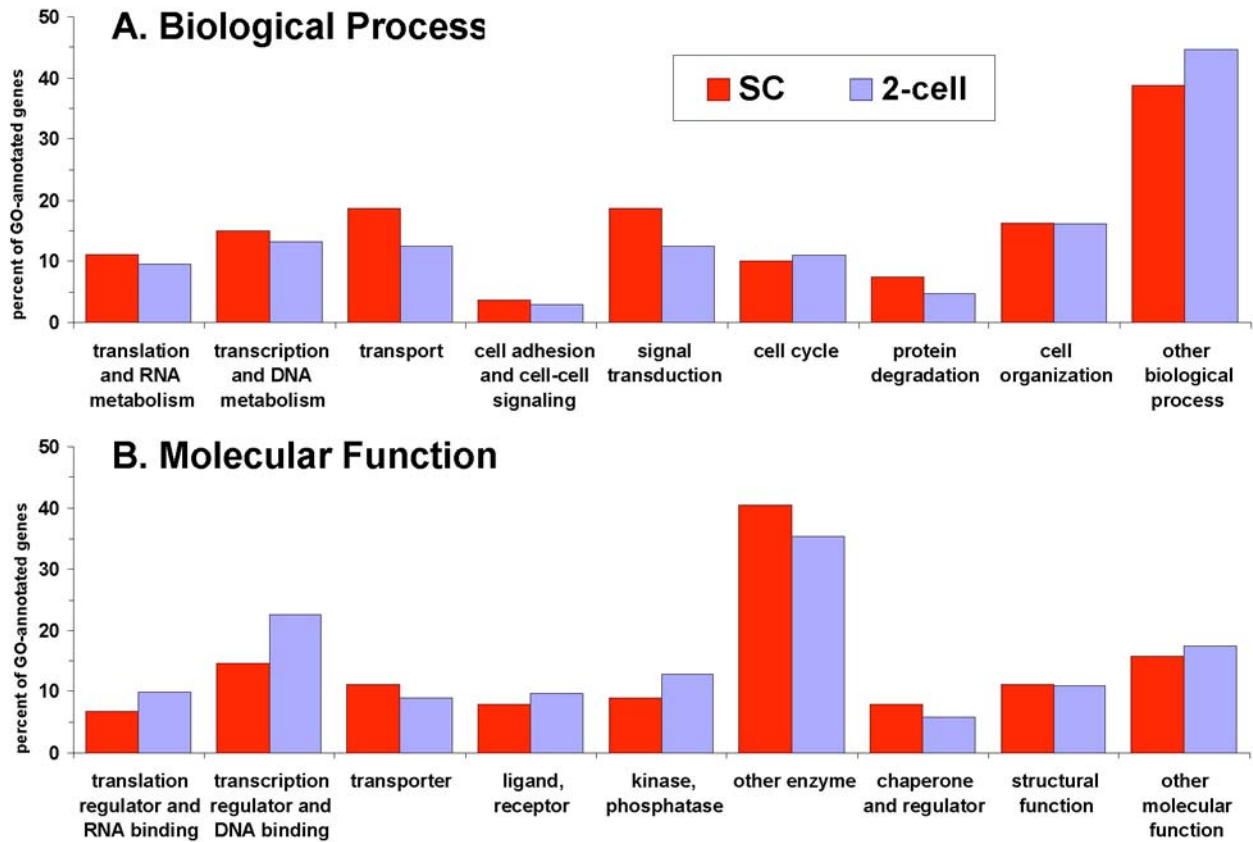
<i>2010002H18Rik</i>	RIKEN cDNA 2010002H18 gene
<i>2010319C14Rik</i>	RIKEN cDNA 2010319C14 gene
<i>2310044F10Rik</i>	RIKEN cDNA 2310044F10 gene
<i>2410001E19Rik</i>	RIKEN cDNA 2410001E19 gene
<i>2410008J01Rik</i>	RIKEN cDNA 2410008J01 gene
<i>2410012M04Rik</i>	RIKEN cDNA 2410012M04 gene
<i>2410015N17Rik</i>	RIKEN cDNA 2410015N17 gene
<i>2410017I18Rik</i>	RIKEN cDNA 2410017I18 gene
<i>2410022L05Rik</i>	RIKEN cDNA 2410022L05 gene
<i>2410091N08Rik</i>	RIKEN cDNA 2410091N08 gene
<i>2500001H09Rik</i>	Mus musculus, clone IMAGE:3481943, mRNA, partial cds; RIKEN cDNA 2500001H09 gene; "expressed sequence C77803"
<i>2610024N01Rik</i>	RIKEN cDNA 2610024N01 gene
<i>2610031L17Rik</i>	RIKEN cDNA 2610031L17 gene
<i>2700084L22Rik</i>	RIKEN cDNA 2700084L22 gene
<i>2810453H10Rik</i>	RIKEN cDNA 2810453H10 gene
<i>3732413I11Rik</i>	RIKEN cDNA 3732413I11 gene
<i>3830408P06Rik</i>	RIKEN cDNA 3830408P06 gene
<i>4933428G09Rik</i>	RIKEN cDNA 4933428G09 gene
<i>5730599O09Rik</i>	RIKEN cDNA 5730599O09 gene
<i>5830427H10Rik</i>	RIKEN cDNA 5830427H10 gene
<i>6330503C17Rik</i>	RIKEN cDNA 6330503C17 gene
<i>AA408880</i>	expressed sequence AA408880
<i>AA409185</i>	expressed sequence AA409185
<i>Abcb1b</i>	ATP-binding cassette, sub-family B (MDR/TAP), member 1B
<i>Acadm</i>	acetyl-Coenzyme A dehydrogenase medium chain
<i>Adam9</i>	a disintegrin and metalloproteinase domain 9 (meltrin gamma)
<i>AI115446</i>	expressed sequence AI115446
<i>AI316867</i>	expressed sequence AI316867
<i>AI504353</i>	expressed sequence AI504353
<i>AI507170</i>	expressed sequence AI507170
<i>AI643885</i>	Mus musculus, Similar to solute carrier family 7 (cationic amino acid transporter, y+ system), member 6, clone MGC:37035 IMAGE:4949962, mRNA, complete cds; "expressed sequence AI643885"
<i>AW111866</i>	expressed sequence AW111866
<i>AW214031</i>	expressed sequence AW214031
<i>AW549877</i>	"expressed sequence AW549877", "expressed sequence AI195826"
<i>Axot</i>	axotrophin
<i>BB128963</i>	<i>Mus musculus</i> , clone IMAGE:3964696, mRNA, partial cds; "expressed sequence BB128963"
<i>Bysl</i>	bystin-like
<i>C78212</i>	expressed sequence C78212
<i>Cbr3</i>	carbonyl reductase 3
<i>Ccnd1</i>	cyclin D1
<i>Cdkap1</i>	CDK2 cyclin-dependent kinase 2 associated protein 1
<i>Cdkn1a</i>	cyclin-dependent kinase inhibitor 1A (P21)
<i>Cenpc</i>	centromere autoantigen C
<i>Cpr2-pending</i>	cell cycle progression 2 protein
<i>Cpxm1</i>	carboxypeptidase X 1 (M14 family)
<i>Crtap</i>	cartilage associated protein
<i>Cttn</i>	cortactin
<i>D10Wsu102e</i>	DNA segment Chr 10 Wayne State University 102 expressed
<i>D11Wsu78e</i>	DNA segment Chr 11 Wayne State University 78 expressed

<i>D12Ert7e</i>	DNA segment Chr 12 ERATO Doi 7 expressed
<i>D16Bwg1547e</i>	DNA segment Chr 16 Brigham Women s Genetics 1547 expressed
<i>D17Ert7e</i>	DNA segment Chr 17 ERATO Doi 197 expressed
<i>D7Wsu87e</i>	DNA segment, Chr 7, Wayne State University 87, expressed
<i>Eif4ebp1</i>	eukaryotic translation initiation factor 4E binding protein 1
<i>Eplin-pending</i>	epithelial protein lost in neoplasm
<i>Erc5</i>	excision repair cross-complementing rodent repair deficiency complementation group 5
<i>F2r</i>	coagulation factor II (thrombin) receptor
<i>Fhl1</i>	four and a half LIM domains 1
<i>Fkbp9</i>	FK506 binding protein 9
<i>Gab1</i>	growth factor receptor bound protein 2-associated protein 1
<i>Gas2</i>	growth arrest specific 2
<i>Gcat</i>	glycine C-acetyltransferase 2-amino-3-ketobutyrate-coenzyme A ligase
<i>Gclm</i>	glutamate-cysteine ligase, modifier subunit
<i>Ghr</i>	Growth hormone receptor
<i>Gsta4</i>	glutathione S-transferase, alpha 4
<i>Hrsp12</i>	heat-responsive protein 12
<i>Hspa11</i>	heat shock 70kD protein 1-like
<i>Hspa4</i>	heat shock 70 kDa protein 4
<i>Itgb1</i>	integrin beta 1 (fibronectin receptor beta)
<i>Kcnab3</i>	potassium voltage-gated channel, shaker-related subfamily, beta member 3
<i>Kdt1</i>	kidney cell line derived transcript 1
<i>Lce-pending</i>	long chain fatty acyl elongase
<i>Madh1</i>	MAD homolog 1 (Drosophila)
<i>Madh2</i>	MAD homolog 2 (Drosophila)
<i>Mdfi</i>	MyoD family inhibitor
<i>Mrp17</i>	mitochondrial ribosomal protein L17
<i>Mrp13</i>	mitochondrial ribosomal protein L3
<i>Mrp134</i>	mitochondrial ribosomal protein L34
<i>Mrp145</i>	Mitochondrial ribosomal protein L45
<i>Mrps10</i>	mitochondrial ribosomal protein S10
<i>Mrps2</i>	mitochondrial ribosomal protein S2
<i>Mrps31</i>	mitochondrial ribosomal protein S31
<i>Msh2</i>	mutS homolog 2 (E. coli)
<i>Ndufaf1</i>	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, assembly factor 1
<i>Nol5</i>	nucleolar protein 5
<i>Pex7</i>	peroxisome biogenesis factor 7
<i>Pkd2</i>	polycystic kidney disease 2
<i>Pla2g6</i>	phospholipase A2, group VI
<i>Ppic</i>	peptidylprolyl isomerase C
<i>Psm12</i>	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12
<i>Rab18</i>	RAB18, member RAS oncogene family
<i>Rcn</i>	reticulocalbin
<i>Rmp-pending</i>	RPB5-mediating protein
<i>Rnac-pending</i>	RNA cyclase homolog
<i>Rock2</i>	Rho-associated coiled-coil forming kinase 2
<i>Rpl22</i>	ribosomal protein L22
<i>Sec23a</i>	SEC23A (S. cerevisiae)
<i>Sfrs3</i>	splicing factor, arginine/serine-rich 3 (SRp20)
<i>Snx12</i>	sorting nexin 12
<i>Srcasm</i>	Src activating and signaling molecule

<i>Stam</i>	signal transducing adaptor molecule (SH3 domain and ITAM motif) 1
<i>Stat1p1</i>	signal transducer and activator of transcription interacting protein 1
<i>Stxbp3</i>	syntaxin binding protein 3
<i>Suclg2</i>	succinate-Coenzyme A ligase, GDP-forming, beta subunit
<i>Tbrg1</i>	transforming growth factor beta regulated gene 1
<i>Tcp1-rs1</i>	t-complex protein 1, related sequence 1
<i>Tead2</i>	TEA domain family member 2
<i>Tex292</i>	testis expressed gene 292
<i>Tjp1</i>	tight junction protein 1
<i>Tmk</i>	thymidylate kinase
<i>Trip6</i>	thyroid hormone receptor interactor 6
<i>Txn1</i>	thioredoxin-like (32kD)
<i>Txnrd1</i>	thioredoxin reductase 1
<i>Upp</i>	uridine phosphorylase
<i>Wbp5</i>	WW domain binding protein 5
<i>Wig1</i>	wild-type p53-induced gene 1
<i>Xrcc5</i>	X-ray repair complementing defective repair in Chinese hamster cells 5
<i>Yes</i>	Yamaguchi sarcoma viral (v-yes) oncogene homolog
<i>Ywhab</i>	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide
<i>Zfp54</i>	zinc finger protein 54
<i>Zfx</i>	zinc finger protein X-linked
<i>LOC230737</i>	LOC230737: similar to Autoantigen NGP-1
<i>MGC27546</i>	hypothetical protein MGC27546
<i>Mm.10509</i>	UniGene: Mus musculus mRNA for zinc finger protein partial cds
<i>Mm.135092</i>	UniGene: ESTs, weakly similar to solute carrier family 4, sodium bicarbonate cotransporter-like, member 10 (Mus musculus)
<i>Mm.22363</i>	UniGene: Mus musculus, clone MGC:28202 IMAGE:3989666, mRNA, complete cds
<i>Mm.32505</i>	UniGene: Mus musculus, Similar to zinc finger protein 281, clone MGC:7737 IMAGE:3498439, mRNA, complete cds
<i>Mm.55092</i>	UniGene: ESTs

^a Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C. & Melton, D.A. "Stemness": transcriptional profiling of embryonic and adult stem cells. *Science* 298:597–600 (2002).

C. Functional annotation of “stemness” genes.



SC – stem cell genes; 2-cell – 2-cell embryo expressed genes. Current GO annotations for “stemness” genes were obtained from the MGI database and clustered using the same set of high-level terms (GO bins) as was used for the 2-cell transcriptome annotation.