

Supplemental Methods

Analysis of microarray data

Raw CEL and DAT files were analyzed with dChip software (<http://biosun1.harvard.edu/complab/dchip/>) (Li and Hung Wong, 2001) after normalization. The entire experiment (microinjection, RNA extraction, and microarray analysis) was repeated three times and quantitative RT-PCR used to determine cut-offs (see below). Target genes were determined by comparing microarray and quantitative real-time RT-PCR (qRT-PCR) data. Candidates were preliminarily determined from microarray data based on the following criteria: i) expression, represented by model-based expression indices (MBEI), was increased more than 1.2-fold in DEX-treated groups ($E/B > 1.2$), ii) MBEI differences were larger than 50 ($E-B > 50$), iii) expression was claimed as “present (P)” in at least one sample in each experiment. Genes identified in at least two experiments were considered candidates. Genes induced in DEX-treated, β -galactosidase injected samples compared with the DEX-untreated, Ngnr1-GR or NeuroD-GR injected samples were also determined and excluded from analysis as they were regarded as DEX-induced. From Ngnr1 and NeuroD candidate target lists, we used qRT-PCR to determine fold-change cut-off values. We found that candidate genes showing $E/B > 1.5$ and $E-B > 70$ for Ngnr1 or $E/B > 1.3$ and $E-B > 50$ for NeuroD consistently showed more than 2-fold induction by qRT-PCR. Genes in the final list (Table I) showed at least 1.7-fold induction by Ngnr1 and at least 1.5-fold induction by NeuroD on average in microarray analysis (average fold change for three experiments). In general, fold change values determined by qRT-PCR were greater than those determined by microarray analysis, perhaps reflecting differences in sensitivity intrinsic to the two methods or conservative algorithmic calculation of fold changes by the dChip software. It is also notable that the genetic heterogeneity in *Xenopus laevis* animals, which are not inbred, could contribute to these fold change differences. Polymorphism in *Xenopus* animals may result in mismatches between mRNA sequences and probe sequences on the microarray and increase standard errors, which would lower the average fold change values. We also found that one known NeuroD direct target, Ebf3, was not included in our list, although its expression was highly induced as determined by qRT-PCR (data not shown). In looking at hybridization signals for perfect match (PM) and mismatch (MM) probes for Ebf3 on the array, we found that some MM probes showed higher intensity than PM probes and dChip claimed Ebf3 as ‘absent’. Therefore, some Affymetrix *Xenopus laevis* genome array probe sets may not be ideally designed and we believe that we missed at least some target genes in our study.

Cloning of target promoter/enhancer regions, luciferase constructs, and mutagenesis

To initially screen for putative regulatory elements, Dll1 (Delta-1), Ebf3, Elavl3, Gadd45 γ , HEN1 (Nhlh1), and NeuroD4 (Ath3) genomic loci from various vertebrates were analyzed using the Evolutionary Conserved Region (ECR) Browser (<http://ecrbrowser.dcode.org/>) with default settings (Ovcharenko et al., 2004). For each gene, approximately 10 kb each of 5' and 3' flanking sequences, 5' and 3' untranslated regions (UTRs) and all introns were analyzed and ECRs usually containing 3 or more clustered, conserved E-boxes were chosen as putative enhancers. If multiple ECRs contained clustered E-boxes, 2-3 ECRs closest to the transcription start site (TSS) were selected. Putative enhancers were amplified with KlenTaqLA (DNA Polymerase Technology, Inc.) and mouse genomic DNA as a template. For Elavl3 promo, we replaced the pGL3-Promoter vector's SV40 promoter (Promega) with the Elavl3 gene fragment including its TATA and TSS (Xho I/Hind III). Enhancers lacking a TSS and TATA box were inserted into Nhe I/Hind III sites of the E1X3-TATA luciferase reporter plasmid (Huang et al., 2000) after removing this plasmid's E-boxes. Genomic sequences and TSS followed the Ensembl database and primer sequences (Fig. S1) were defined using Primer3 (Rozen and Skaletsky, 2000).

To determine Ngn and NeuroD consensus sites and Position weight matrix (PWM) files, we used CompareProspector (<http://CompareProspector.stanford.edu>) (Liu et al., 2004) and input 25 ECRs from human and mouse genomic loci for 14 Ngn and NeuroD targets (Ascl1, Dll1 (Delta-1), Ebf2, Ebf3, Elavl3, Gadd45g, HEN1 (Nhlh1), MTGR1, Myo10, MyT1, NeuroD4, Pou3f1, PDK2, and Znf238). For the second enhancer screen, we used this PWM and either the ECR Browser or the Enhancer Element Locator (EEL; available at <http://www.cs.helsinki.fi/u/kpalin/EEL/>) (Hallikas et al., 2006) to locate putative Ngn/NeuroD regulatory enhancers. Input sequences for each gene included 50 kb each of 5' and 3' flanking regions, the 5' and 3' UTRs, and all introns. Two or three putative regulatory elements with the highest EEL scores were PCR amplified and cloned as above. For sequences, see Fig. S2. Site-directed mutagenesis was performed with the QuikChange Site-Directed Mutagenesis protocol (Stratagene) using Pfu Turbo polymerase (Stratagene). All E-box sequences (CANNTG) were changed to CTCGAG, the Xho I restriction site. Mutants were selected by preliminary screening with Xho I digestion and sequence confirmation. For primer sequences, see Fig. S1.

Transgenic Xenopus and morpholino oligonucleotides

REMI transgenesis procedure was performed with minor modifications as described in (Taylor et al., 2006). Plasmids IS-EF1 α -eGFP and ISceI-pBS, which has two I-Sce I sites flanking the pBluescript II SK cloning site, were kindly provided by Drs. Pieler and Grammer.

Transgene constructs were generated by transferring the enhancer-luciferase cassette from luciferase reporter plasmids (see above) to the ISceI-pBS plasmid, using Not I/BamH I or Not I/Sal I sites. For SceI meganuclease transgenesis, we injected 80 pg of I-Sce I digested DNA with 3×10^{-3} units of I-Sce I (NEB Biolabs) in a 4 nl volume into each embryo. Injection was finished within 1 hr after fertilization and injected embryos were kept at 12.5 °C until 2-cell stage. Embryos were raised at 18 °C until st 26-27 and analyzed by in situ hybridization for luciferase expression. Results were consistent for REMI and SceI transgenesis, with the same constructs giving tissue localized or non-localized expression and the frequency of transgenesis was consistent with prior work. REMI gave high frequencies of transgenic embryos that expressed the reporter in the brain and eye, for example, NeuroD4 EEL4, 36%, n=103; MTGR1 EEL1, 29%, n=28; Dll1 EEL1, 57%, n=82; Ebf2 EEL3, 36%, n=42. By contrast, transgenic embryos carrying empty reporter constructs did not show localized transgene expression with either method (REMI transgenic embryos: TATA vector, 1.5%, n=66; pBS vector, 0%, n=31). The SceI method yielded 1-5% transgenic embryos per experiment. For SceI transgenesis, each construct was analyzed by injecting 100-350 embryos per experiment in 3-4 independent experiments.

To reduce *Ngnr1* activity in *Xenopus* embryos, we used a morpholino oligonucleotide (MO) directed against *Ngnr1* 5'-TGGTTAGCCCCAATGTTGCACTGAC, which was previously characterized for specificity and efficacy in suppressing neuronal differentiation (Klisch et al., 2006). As this MO appears to recognize only one *Xenopus* *Ngnr1* gene copy (*Ngnr1a*), we also used a second *Ngnr1* MO (5'-GGTATTCGCACTTCAGCAGCACCAT), which crosses the *Ngnr1* translation start and matches the sequence for both *Ngnr1a* and *1b* gene copies. Injection of either MO comparably reduced expression of *Ngnr1* target genes in *Xenopus* embryos, although the latter was somewhat more effective as expected. 20 or 40ng of each MO plus 30 pg β -galactosidase RNA was injected into 1 blastomere of 2-cell stage albino embryos and embryos were grown to st. 16-17 (Nieuwkoop and Faber, 1967), fixed, X-Gal stained and in situ hybridized as described (Seo et al., 2005).

Chromatin immunoprecipitation (ChIP)

Antibodies for ChIP were NeuroD (Santa Cruz sc-1084, after (Calella et al., 2007; Martin et al., 2003; Moates et al., 2003), Myc (9E10; Santa Cruz) or isotype matched IgG (Santa Cruz). For tissue ChIP, telencephalon tissue was dissected from e14.5 mouse embryos and 50-60 mg of material was immunoprecipitated with 5 μ g of NeuroD or isotype-matched IgG antibodies. Our tissue ChIP protocol was based on (Chen et al., 2004; Peng and Chen, 2005) and the Farnham

laboratory (<http://genomics.ucdavis.edu/farnham/protocol.html>) and Upstate protocols, with modifications in washing conditions as needed to minimize background and non-specific binding. To process individual samples, 50-60 mg of dissected brain tissue was minced into small pieces and immersed in 1% formaldehyde (in MEM media plus 5% fetal bovine serum) for 12 minutes to crosslink chromatin-bound proteins. Following crosslinking, tissue was disaggregated in ice cold 1X phosphate buffered saline (PBS) with a Dounce homogenizer, spun down briefly in a microfuge, and the cell pellet was then frozen at -80°C until further use. Subsequent steps for cell lysis, isolation and shearing of chromatin, and immunoprecipitation then followed the same protocol as described below.

For P19 ChIP, cells were transfected with 2 µg US2MT-mNeuroD in 100-mm dishes. Cells from a single 100-mm dish were crosslinked with 1% formaldehyde in MEM media plus 5% fetal bovine serum for 12 minutes, scraped from the dish, and re-suspended in 0.5 ml of Active Motif Lysis Buffer with 0.2% SDS or Missonix Lysis Buffer with 1 % SDS. After 30 min. incubation on ice, cell membrane was disrupted by vortexing. Nuclei were collected by centrifugation, washed with TE (tris-EDTA pH 8.0) and re-suspended in 1ml nuclei lysis solution (Tris pH 8.0 50mM, EDTA 10mM, SDS 0.01%). Chromatin was sonicated using a Branson Sonifier 250 at power 2.0, 20s X 20 times to shear chromatin into 250-500 bp fragments. After centrifugation to remove insoluble material, supernatant was aliquoted and stored at - 80 °C. For immunoprecipitation, 900 µl of lysate (approximately 30 µg of chromatin) was diluted in 100 µl of 10 X ChIP Dilution Buffer (Tris pH 8.0 500mM, EDTA 50mM, NaCl 1.5M, Sodium Deoxycholate 1 %, Triton X 100 10 %) and pre-cleared with 60 µl of Salmon Sperm DNA/Protein G agarose-50% slurry (Upstate) for 1 hr at 4 °C. After removing agarose by brief centrifugation, 5 µg of antibodies or normal mouse serum (Jackson ImmunoResearch Lab. Inc.) were added and incubated at 4 °C overnight. Immune complexes were collected with Protein G agarose for 1.5 hr at 4°C. Washing and elution was performed following Upstate ChIPkit protocol. DNA was purified with QIAquick PCR Purification columns and used for quantitative real-time PCR as above. Enrichment was calculated as the ratio of specific antibody to IgG after normalization with GAPDH. We initially designed quantitative ChIP primers for 10 enhancers, which showed robust induction by Ngn and NeuroD in luciferase assays (MTGR1 EEL1, MTGR1 EEL6, Ebf2 EEL1, Ebf2 EEL3, Ebf3 ECR1, Ebf3 ECR2, NeuroD4 EEL4, Zfp238 EEL1, Dll1 EEL1 and HEN1 promo) and one negative control (HEN1 ECR1), but primers for Dll1 EEL1 and HEN1 promoter region did not work (PCR products with smear or a wrong size). Results from the other 8 loci were shown in Fig.4 and Fig. S8.

Genome-wide NeuroD target predictions

Using the constructed Ngn/NeuroD binding profile, the potential Ngn/NeuroD transcriptional regulatory targets were predicted using a previously developed mammalian regulatory sequence analysis model (Chang et al., 2006). This model identifies all evolutionarily conserved transcription factor binding sites and performs enrichment analysis of these sites in a gene's proximal promoter sequence. Using this model, the genomic sequences of human, mouse, and rat were acquired from the NCBI's Genome Assembly Project. Orthologous gene groups were identified using the NCBI's HomoloGene database, and genomic sequences of the orthologous gene loci were aligned using the program TBA (Blanchette et al., 2004). The program PATSER (Stormo et al., 1982) was used to search for matches of the Ngn/NeuroD binding profile in the genomic sequence of each gene locus. Putative Ngn/NeuroD binding sites were searched in the 15 kb proximal promoter sequence and in the multi-species conserved sequences (MCSs) (Margulies et al., 2003) in the entire gene locus. Only Ngn/NeuroD binding sites that were conserved across human and rodents were considered. These sites were used to calculate the binding probability score for each gene using the previously described scoring formula (Chang et al., 2006). The P-value for observing a probability score or higher given the complete set of conserved Ngn/NeuroD binding sites in the genome was then calculated by permutations as follows. In each permutation, Ngn/NeuroD binding sites were randomly assigned to each gene according to the length of the genomic sequence of each gene. After each permutation, a new binding probability score was calculated according to the new binding site distribution. After all the permutations, the P-value of enrichment was calculated using the number of permutations in which a binding probability score higher than the true score was obtained. In this approach, genes with a very low P-value will have over-represented Ngn/NeuroD binding sites and are predicted as potential Ngn/NeuroD regulatory targets. In this study, 10,000 permutations were computed, and 0.01 was used as the P-value cutoff.

A similar analysis was performed using both the Ngn/NeuroD Position Weight Matrix (defined in this manuscript) as well as two divergent E-box binding profiles in the TRANSFAC database, c-Myc/Max (M00118, with a consensus sequence NNACCACGTGGTNN) and E-box (M01034, with a consensus sequence NCACSTGNCN). These two weight matrices are shown below.

c-Myc/Max (M00118)

A	25	28	70	10	0	100	0	0	0	0	2	6	18	31
C	31	7	3	69	100	0	100	0	0	0	19	21	47	13
G	13	47	21	19	0	0	0	100	0	100	69	3	7	31
T	31	18	6	2	0	0	0	0	100	0	10	70	28	25

E-box (M01034)

A	16	0	106	4	23	0	1	47	5	16
C	59	118	7	94	44	0	4	21	71	47
G	21	0	1	15	48	1	112	21	14	26
T	23	1	5	6	4	118	2	30	29	30

GO term enrichment analysis

The Gene Ontology (GO) term enrichment analysis was used to predict the function of a set of potential regulatory targets. For each Gene Ontology (GO) category, the number of potential regulatory target genes annotated by that GO category was counted, and the enrichment of that GO category in the predicted target genes was calculated using the hypergeometric distribution as shown below:

$$P - value = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

In the above formula, M is the total number of genes annotated by a GO category, N is the total number of annotated genes in the human genome, n is the total number of predicted target genes, and k is the number of predicted target genes that are annotated by that GO category.

References

- Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., Haussler, D. and Miller, W. (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res*, **14**, 708-715.
- Calella, A.M., Nerlov, C., Lopez, R.G., Sciarretta, C., von Bohlen Und Halbach, O., Bereshchenko, O. and Minichiello, L. (2007) Neurotrophin/Trk receptor signaling mediates C/EBPalpha, -beta and NeuroD recruitment to immediate-early gene promoters in neuronal cells and requires C/EBPs to induce immediate-early gene transcription. *Neural Develop*, **2**, 4.
- Chang, L.W., Nagarajan, R., Magee, J.A., Milbrandt, J. and Stormo, G.D. (2006) A systematic model to predict transcriptional regulatory mechanisms based on overrepresentation of transcription factor binding profiles. *Genome Res*, **16**, 405-413.
- Chen, S., Peng, G.H., Wang, X., Smith, A.C., Grote, S.K., Sopher, B.L. and La Spada, A.R. (2004) Interference of Crx-dependent transcription by ataxin-7 involves interaction between the glutamine regions and requires the ataxin-7 carboxy-terminal region for nuclear localization. *Hum Mol Genet*, **13**, 53-67.
- Hallikas, O., Palin, K., Sinjushina, N., Rautiainen, R., Partanen, J., Ukkonen, E. and Taipale, J. (2006) Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell*, **124**, 47-59.
- Huang, H.P., Liu, M., El-Hodiri, H.M., Chu, K., Jamrich, M. and Tsai, M.J. (2000) Regulation of the pancreatic islet-specific gene BETA2 (neuroD) by neurogenin 3. *Mol Cell Biol*, **20**, 3292-3307.
- Klisch, T.J., Souopgui, J., Juergens, K., Rust, B., Pieler, T. and Henningfeld, K.A. (2006) Mxi1 is essential for neurogenesis in *Xenopus* and acts by bridging the pan-neural and proneural genes. *Dev Biol*, **292**, 470-485.
- Li, C. and Hung Wong, W. (2001) Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol*, **2**, RESEARCH0032.
- Liu, Y., Wei, L., Batzoglu, S., Brutlag, D.L., Liu, J.S. and Liu, X.S. (2004) A suite of web-based programs to search for transcriptional regulatory motifs. *Nucleic Acids Res*, **32**, W204-207.
- Margulies, E.H., Blanchette, M., Haussler, D. and Green, E.D. (2003) Identification and characterization of multi-species conserved sequences. *Genome Res*, **13**, 2507-2518.
- Martin, C.C., Svitek, C.A., Oeser, J.K., Henderson, E., Stein, R. and O'Brien, R.M. (2003) Upstream stimulatory factor (USF) and neurogenic differentiation/beta-cell E box transactivator 2 (NeuroD/BETA2) contribute to islet-specific glucose-6-phosphatase catalytic-subunit-related protein (IGRP) gene expression. *Biochem J*, **371**, 675-686.
- Moates, J.M., Nanda, S., Cissell, M.A., Tsai, M.J. and Stein, R. (2003) BETA2 activates transcription from the upstream glucokinase gene promoter in islet beta-cells and gut endocrine cells. *Diabetes*, **52**, 403-408.
- Nieuwkoop, P.D. and Faber, J. (1967) *Normal Table of *Xenopus laevis* (daudin)*. North-Holland Publishing Company, Amsterdam.
- Ovcharenko, I., Nobrega, M.A., Loots, G.G. and Stubbs, L. (2004) ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res*, **32**, W280-286.
- Peng, G.H. and Chen, S. (2005) Chromatin immunoprecipitation identifies photoreceptor transcription factor targets in mouse models of retinal degeneration: new findings and challenges. *Vis Neurosci*, **22**, 575-586.

- Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*, **132**, 365-386.
- Seo, S., Richardson, G.A. and Kroll, K.L. (2005) The SWI/SNF chromatin remodeling protein Brg1 is required for vertebrate neurogenesis and mediates transactivation of Ngn and NeuroD. *Development*, **132**, 105-115.
- Stormo, G.D., Schneider, T.D., Gold, L. and Ehrenfeucht, A. (1982) Use of the 'Perceptron' algorithm to distinguish translational initiation sites in *E. coli*. *Nucleic Acids Res*, **10**, 2997-3011.
- Taylor, J.J., Wang, T. and Kroll, K.L. (2006) Tcf- and Vent-binding sites regulate neural-specific geminin expression in the gastrula embryo. *Developmental Biology*, **289**, 494-506.

Supplemental Figure Legends

Figure S1. Primer sequences used in this study. (A-C) qRT-PCR primers for (A) *Xenopus* Ngn and NeuroD target genes, (B) their mouse orthologs, and (C) PAP-predicted Ngn/NeuroD target genes (mouse), to test induction by NeuroD. (D-F) Primers to clone enhancer fragments (mouse) for (D) ECR enhancers, (E) EEL enhancers, (F) Primers for site-directed mutagenesis of E-boxes in enhancer fragments, (G) Primers for chromatin immunoprecipitation experiments.

Figure S2. Sequences of enhancers and E-boxes and position weight matrices. (A) Sequences of functional (Ngn/NeuroD responsive) ECR and EEL enhancers, (B) sequences of non-functional ECR/EEL enhancers, (C) Types of E-boxes found in all ECR versus functional ECR, (D) Experimental and computationally-determined Position Weight Matrices, and (E) artificial enhancer sequences are shown.

Figure S3. Ngn and NeuroD targets involved in cell migration and morphology. Many Ngn and NeuroD target genes or highly related genes have known roles in regulating cytoskeletal rearrangement or migration/morphology in cells including neurons. *Xenopus* Unigene numbers, Gene names and symbols and known structural features and roles in regulating cell migration/morphology are listed.

Figure S4. Expression profiles for Ngn/NeuroD target genes. Gene expression profiles were obtained from the NCBI EST Profile viewer for mouse Unigene entries with the gene symbols shown. Transcript abundance is expressed as transcripts per million (TPM). Tissues with enriched (> than average) expression are shown in yellow for each gene. For each set of genes, tissues/regions where >50% of the surveyed genes had enriched expression are indicated in green while tissues/regions where 20-49% of the surveyed genes had enriched expression are indicated in orange. (A-B) Expression profiles of Ngn/NeuroD target genes from microarray. Profiles for 12 strongly Ngn/NeuroD-responsive genes (A) were compared to those of the 8 genes that did not respond to Ngn or NeuroD at all (B). Ngn/NeuroD-responsive genes generally have more neural-restricted expression (compare green/orange tissue locations in A versus B). (C-D) Similar results were obtained when comparing 9 NeuroD-responsive (C) and 9 non-responsive (D) target genes predicted as NeuroD targets by PAP (see text). (E-F) Summary of expression profiling data for 21 NeuroD responsive target genes (E) versus 17 non-responsive target genes (F).

Figure S5. Identification of regulatory elements in Ngn/NeuroD target genes. Putative regulatory elements (sequences in Fig. S2) were cloned into a minimal promoter luciferase construct, cotransfected with Ngn or NeuroD into P19 cells, and responsiveness measured by luciferase assay. Fold change values were normalized to luciferase activity in the absence of Ngn/NeuroD, arbitrarily defined as 1.

Figure S6. Responsiveness of artificial enhancers to Ngn2 and NeuroD. 10 artificial enhancer-luciferase reporter constructs were generated, each of which contained each 3 copies of a different E-box sequence, with the same intervening sequences used for all constructs (sequences in Fig. S2, section E). Upon cotransfection into P19 cells, the same 4 of these 10 artificial enhancer types (those with GA, GC, AA, or TA type E-boxes) could be activated by both Ngn2 and NeuroD.

Figure S7. Analysis of E-box site frequency in Ngn/NeuroD target loci. ECR Browser was used to perform pairwise cross-species comparisons of genomic sequences for 8 shared Ngn and NeuroD direct target loci (Dll1, Ebf2, Ebf3, Cbfa2t2h/MTGR1, MyT1, NeuroD4, Nhlh1, Zfp238). For each locus, sequences analyzed included locus/introns and 50kb each of 5' and 3' flanking sequences. Mouse was used as the base organism, with pairwise comparisons to corresponding human, opossum, dog, chicken, frog (*X. tropicalis*), fugu and zebrafish genomic sequences. In ECR Browser, rVista analysis of each syntenic pairwise alignment identified conserved E-boxes (CANNTG) of different types, where NN is (WA=T/A, A), AA, (GM=G, C/A), or TA.

(A) Loci are preferentially enriched for GM-type sites. Frequency of GM, WA, or TA type E-box sites was calculated as the sum of all pairwise comparisons of mouse genomic sequence to the syntenic genomic regions in other species (calculations and raw data shown in panel E). Frequency is expressed as a percentage of the total number of functional (e.g. Ngn2/NeuroD-responsive WA+GM type E-boxes) E-box sites present in all pairwise comparisons for that locus. Average=average frequency for all pairwise comparisons across the 8 loci. 69% of Ngn2/NeuroD-responsive functional (GM or WA) sites are the GM type, while 31% are the WA type. AA sites account for 22% and TA sites for 9% of the Ngn2/NeuroD-responsive E-boxes present in the loci.

(B) GM type E-boxes are more frequently found in clusters than are WA E-boxes. If multiple sites fell within a 1.5kb window in the >100kb of genomic sequence analyzed, these

sites were counted as being clustered. Frequency of clustered sites is represented relative to the sum of functional (GM+WA) E-boxes for each locus.

(C) Quantitation of E-box frequencies. Raw numerical data and site frequency calculations for pairwise comparisons of each locus are shown.

Figure S8. Chromatin immunoprecipitation of Ngn2 and NeuroD at enhancers. Ngn2 and NeuroD binding to EEL and ECR enhancers was analyzed by ChIP in P19 cells, following transfection of Myc-tagged Ngn2 and NeuroD. Fold-enrichment is expressed as a ratio of specific (Myc) to non-specific (IgG) immunoprecipitation after normalization with GAPDH.

Figure S9. Computational genome-wide prediction of Ngn/NeuroD target genes.

(A) Ngn/NeuroD target genes predicted by PAP in a genome-wide computational screen. The PAP program was used in combination with our position weight matrix for Ngn/NeuroD binding (Fig. 3G) for genome-wide prediction of Ngn/NeuroD targets. Annotated genes were scored based on their enrichment for Ngn/NeuroD binding sites conserved between human and mouse (h/m) or human, mouse and rat (h/m/r) and using a cut-off value of $P < 0.01$ (see Supplemental Methods). 347 Ngn/NeuroD targets were predicted in both h/m and h/m/r comparisons (intersection). PAP predicted many of our Ngn and/or NeuroD targets defined on the microarray, as shown at bottom for the h/m and h/m/r based analyses and intersection between the two.

(B-C) PAP predicted target genes for the M01034 and c-Myc/Max E-boxes. To test the specificity of the PAP predictions for Ngn/NeuroD, PAP analysis was also performed for two E-boxes in TRANSFAC that are diverged from the NeuroD PWM, E-box M01034 (B) and c-Myc/Max, M00118 (C). These CACGTG-type E-boxes did not respond to Ngn or NeuroD in our prior assays. PAP predicted 365 (B) and 127 (C) target genes for these matrices respectively.

Figure S10. Functional clustering of PAP-predicted target genes.

(A-C) PAP predicted target genes for the matrices above were analyzed by functional clustering and top-ranking Gene Ontology (GO) terms are shown. (A) For Ngn/NeuroD targets, nervous system development (green) was the top-ranked GO term. This was not found as a highly ranked GO term for either of the other matrices, which had development (yellow) and cell cycle (pink) as top GO terms (B-C).

(D) Ngn/NeuroD target genes in the nervous system development and development functional clusters are listed. Yellow indicates overlap with development cluster genes predicted as targets using the other E-boxes.

(E-F) M01034 E-box (E) and Myc/Max (F) target genes in top-ranked clusters (development and cell cycle) are listed.

Seo_Supplemental Figure S1. Primer sequences.

A. Primers for qRT-PCR of *Xenopus* genes (*Ngnr1* and *NeuroD* target confirmation)

	gene	direction	sequence	Product size	UniGene number
1	xEf1alpha	F	CACCATGAAGCCCTTACTGA	149	XI.17432
		R	ACCTGTGCCGGTAAAGAACC		
2	xEbf2	F	GTCCTCTCCACCACAGTGAA	148	XI.547
		R	CACTTGACTGATGGCTTTG		
3	xEbf3	F	TACAGCAGGGGTAGTGGATG	103	XI.632
		R	ATTGGATTTCTGAGGTTGG		
4	xDelta-1	F	GCACTACCAGAGCAACGTGT	90	XI.54916
		R	GGAACAACGAAGGAGTTGGT		
5	xMyT1	F	TTGGGATGGTCCCATAGACT	138	XI.57163/XI.31433
		R	TCTTGATCTCCTGCATCTC		
6	xPOU50	F	GGACCTCAGGCTTCTCTTT	138	XI.1274/XI.21449
		R	GCTAGGGGCATCTTCATCTG		
7	xZnf238(RP58)	F	CGGAGAGGCTGAAACACAC	139	XI.23642
		R	CACAGAGAGGTCCAGCACAC		
8	Xash1	F	CCCCAACTATCCACGATA	144	XI.450
		R	TGCTACTCCGCATCTCAGAA		
9	xTwist	F	GAGGCGATCTGCTAGGAAAA	121	XI.879
		R	CCTCTGACTCTGCAGCTCCT		
10	xFRP	F	GTGACCCCTGTGTAAGGAA	140	XI.3150
		R	GAGAAGGAGGAGCCTGGTG		
11	xlro3	F	GAGCCTTCTGCCCTACAG	125	XI.4522
		R	GGGATAAAAAGGCTGGTGTA		
12	xPlx3	F	GACTGCATGACTGCTTCTGG	124	XI.16835
		R	CCCAGACAAAATGCTCAGGT		
13	xMTGR1	F	CAAAACCACTACCCCCATCT	143	XI.4960
		R	CAGCAGTGACAGAAGCAGGT		
14	xGadd45	F	TGACTGGGTTCCCACTATCA	145	XI.12125/XI.16395
		R	GGACCTCAACCAGCGTAATC		
15	xEI-C	F	GCTTTCTATCCTCCCCAGGT	147	XI.1035
		R	TGCCACAGGACACTCTCATC		
16	xAth3	F	CATTCAGGTCCCCCTATCC	123	XI.1263
		R	GGTGTGGGTAAGGCTCGTAA		
17	xDelta-2	F	CTTCCGTGTGTCTCAAGC	139	XI.68
		R	ATGGAAGGGACACGAATTG		
18	xB4galT3	F	GCCAGAGAGCTGGAACATCT	134	XI.71386
		R	GTGCTGCTCGTCTTGTGGAC		
19	MTGR1b	F	CTGTGGCCAAAGCATTGATA	150	XI.16456
		R	GCAGTGACAGAAGCAGGTGT		
20	CIDE-c	F	CCTCTCCCCCAAGTCTCTCT	133	XI.15305
		R	AATGCCTTTCTGAGACTGC		
21	Hes-6	F	TTTCTGGGATCCTCCTCTGA	138	XI.9030
		R	CCATGGTCTCCACATAGCAG		
22	xl.6849	F	GGCAGAGTGTAAGGGACACAC	139 bp	XI.6849
		R	GCTCACCTTTTCTGTTTCT		
23	xScamp2	F	ACTGTCCGTCTCCTCGTCT	123 bp	XI.45029
		R	CATGGTTGCCAGTTGTAATTTG		
24	xl5263	F	AGTGAGCCCAGACCAGAAGAT	127 bp	XI.5263
		R	CCATTTTACCCCATATCCAC		
25	xRgc32	F	TGATATGCTGTGACACCGAAG	146 bp	XI.5886
		R	TATCAGCCGCTTCTCACTCAT		
26	XI.7910	F	AGAACCTGAGCCGACTAGGAG	138 bp	XI.7910
		R	CTTGGGCTGTCCATCAGTAG		
27	xGas6	F	CTTCTTCTCACCTCCTGGTT	141 bp	XI.4807
		R	CCTTTCCAAGTGCCCTGTTT		
28	xLims1	F	TTCGGAAGACTGCTCCAAAT	130 bp	XI.8188
		R	GTAATACTCTCCACCCAAAC		
29	XI.24399	F	GTGCTCGAGTTCTTCTGGT	106	XI.24399
		R	ACACCTGTCCGGCTACATTC		
30	XI.34	F	GTGCAAGCTCAGCAGTATC	151 bp	XI.34
		R	AGCCAAGGATACAAGGATGG		
31	xPrickle	F	CTGGCAGATGAACCATTTCT	138 bp	XI.7556
		R	TCCCCACAGGATTCACAGTA		
32	XI.51207	F	GGTTTTCTGCTGGAGTGGAT	149 bp	XI.51207
		R	GTACAGCAACCCTCCACCTT		
33	xPDK2	F	GCACTGAAGCAGATGATTGG	133 bp	XI.1717

	R	GGTAGTGGTCTGGGAGAAA			
34	x1419	F	GGATTGAGCAAGAGCACTG	149 bp	XI.1419
		R	TCAGGGTTGAGCAAATCAA		
35	xESR7	F	CTACGAAAGCCAGTTGTGGA	124 bp	XI.586
		R	CAGCCTTCTCCAGCTTCC		
36	xNeuroD	F	GAGCAGAGTCAGGACATCCA	139bp	XI.330
		R	GCTTGACGTGGAATACATGG		
37	xgng-13	F	ACGCTACCTTACCGGCCTAT	103 bp	XI.21807
		R	TTTGATGAGGGTGGGCTTAG		
38	xCalponin2	F	GACCCCAAGAACAATCCT	139 bp	XI.554
		R	CAGTCCAGACTTGGTGTCG		
39	xRiddle2	F	ATGTAAGCTGGGATGCCACT	104 bp	XI.6174
		R	TGTTCTCAGGGCAGGAGACT		
40	xAngio1	F	ATTGTCCATTGCTTTCCAG	116 bp	XI.23649
		R	TCACGCAGAATCAGACTTGG		
41	XI.8921	F	GTCACCGGAGTACCAAAGGA	96 bp	XI.8921
		R	TCTTGGTGGTGCAGTCCATA		
42	XI.4183	F	CCAGAGCCTGAGAAGAATGG	123 bp	XI.4183
		R	GGTTTGTGTCCAGCGGTAAT		

B. qRT-PCR primers for mouse genes

1	F-mRPL19	TCAGGCTACAGAAGAGGCTTGC	160	Mm.218350
	R-mRPL19	ACAGTCACAGGCTTGCGGATG		
2	F-mAscl1	GCCAACAAGAAGATGAGCAA	148	Mm.136217
	R-mAscl1	TGGAGTAGTTGGGGGAGATG		
3	F-mAth3	TCTGGGCCTTGTCTGAAAGTC	145	Mm.10695
	R-mAth3	TCCAGGAGGGTAGATTGAGG		
4	F-mDelta1	TCCGATACCCAGGTTGTCTC	93	Mm.4875
	R-mDelta1	GGTCTTGGTTGCAGAAAAGG		
5	F-mMyT1	CCCAGCAGCTCTATGACCTC	111	Mm.130005
	R-mMyT1	CTCAGGCTCTTCTCTCGAA		
6	F-mElavl3	AAGCCATCAACACCCTCAAT	148	Mm.347707
	R-mElavl3	AGCTGCTCCATCTCCTTCTG		
7	F-mMTGR1	TTCGTCCATTTGTGATTCCA	142	Mm.29914
	R-mMTGR1	GCGATACTGGTGTGAGCAG		
8	F-mEbf3	CTTAATGGCTCCTCCGCTAA	104	Mm.258708
	R-mEbf3	CCGTGTGTGCTGCTACAGTT		
9	F-mTwist1	GAGCAAGATTCAGACCCTCAA	147	Mm.3280
	R-mTwist1	ATCCTCCAGACGGAGAAGG		
10	F-mHEN1(Nhlh1)	CCCCCTGTCAACATACATTA	150	Mm.2474
	R-mHEN1(Nhlh1)	GCCGGAACAACCTCTCTGAAG		
11	F-mOct6	CAAGCAGTTC AAGCAACGAC	96	Mm.297371
	R-mOct6	GTCTGCCAGAAACAGTTACC		
12	F-mBmn2	AAGTGGTTGGAAGAGGCAGA	148	Mm.129387
	R-mBmn2	GCTTAGGGCATTGAGGAAA		
13	F-mHes5	TGAAACACAGCAAAGCCTTC	112	Mm.137268
	R-mHes5	GTGCAGGGTCAGGAACTGTA		
14	F-mBTBD6	TACCAAGGAGGCTGTGGTCT	134	Mm.18396
	R-mBTBD6	TTGGAATTCGGACCAGGTAG		
15	F-mIrx3	AAACTAGCGACTCGGACGAC	120	Mm.238044
	R-mIrx3	AGGGTGGAGAAGGTGGAGAT		
16	F-mPlk3	AAGCCACTGACACCGAGTCT	134	Mm.259022
	R-mPlk3	GATATGGCGGTGCTGTAGGT		
17	F-mRgc32	AGCACCTAGAGCGCATGAA	149	Mm.29811
	R-mRgc32	AGCTGGAGAGGAGTTGGTTG		
18	F-mSAT1	CGGAAACGAATGAGGAA	147	Mm.379009
	R-mSAT1	GATCAGTCGCAGGATGTCAC		
19	F-mGadd45g	CTGCCTTGAGAAAGCTCAGT	119	Mm.281298
	R-mGadd45g	CACCAAGTCGATCAGACCAA		
20	F-mPDK2	GACGTATCGGGTCAGCTAGG	132	Mm.29768
	R-mPDK2	ATCAGGGAGACCCAGTTTCA		
21	F-mPrickle1	CAAGCCCCACACATACCTCT	109	Mm.150314
	R-mPrickle1	CTCATCGCGTGGACATTCT		
22	F-mSlc7a7	CATGGTCAACAGCACCAAGT	139	Mm.142455
	R-mSlc7a7	GAGACACACGCCATTAAGCA		
23	F-mXL21239	AATGCAGTTGATGGGGTAGC	127	Mm.130063
	R-mXL21239	ACTCCTTGCACTGGTTGGTT		
24	F-mCidec	TGTTTCATGGTCTGCTGAAG	149	Mm.10026
	R-mCidec	GTCCTGAGGGTTCAGCTTGT		

25	F-mB4GalT3	CCAGTACTTTGGCGGAGTTT	115	Mm.274011
	R-mB4GalT3	GACCCCTGGTAGCAATGTCGT		
26	F-mSIP1	GAACCCATTTAGTGCCAAAGC	119	Mm.259595
	R-mSIP1	GTCCTGGGTATGGTCGTAGC		
27	FmRP58(Zfp238)	AGACTGTGTATCGGCAGCAG	83	Mm.330700
	RmRP58(Zfp238)	AACAGTGTGCAAAACGCTTC		
28	FmEbf2	AGGATACATCCGCAACACAA	108	Mm.319947
	RmEbf2	CTGTAGCCGTTTCATGCTGTT		
29	FmSlc43a2	ACACTGGAGCCCATACTTCC	97	Mm.11186
	RmSlc43a2	AGTGGGAGGTTCACTTGTCC		
30	F-mAmotl2	CCAGCTATGGTGTCTGTGG	106	Mm.21145
	R-mAmotl2	CACCTCCAAGGATTCAAGT		
31	F-mMyo10	GCTGGAGAGGTGCAGTATGA	132	Mm.60590
	R-mMyo10	GCTGGAGACGTGTCAAAGA		

C. qRT-PCR primers for PAP-predicted target genes (mouse)

1	FmAtoh8	AGGTGCCGTGCTACTCCTAT	129	Mm.87449
	RmAtoh8	AGCTGAGGTTGCTGTGGTC		
2	FmDCX	ACAGTGCTCAAGCCAGAGAG	134	Mm.12871
	RmDCX	GACCTGCTCGAAAGAGTGG		
3	FmEPHB2	TGCTGAAGAAGTGGATGTGC	144	Mm.250981
	RmEPHB2	AAGGTTCTGATGGACAACC		
4	FmNCAM1	TCCTCCTTGACCCTGAAGAG	150	Mm.4974
	RmNCAM1	CCAGGTGTAGACAGCCACAG		
5	FmROBO1	CACAGGAACAGCCAAATGAC	150	Mm.310772
	RmROBO1	AACTGGGGTCTCTAGGGTTG		
6	FmSema3a	CGGCAATGGAGCTTTCTAC	147	Mm.372039
	RmSema3a	CATCCCAGGCACAGTAAGG		
7	FmSli3	AGGTCAATGGCTACACATGC	115	Mm.202457
	RmSli3	TTCTGGCACTCGTACTGGTC		
8	FmSyn3	AGGCTCTGCTATGCTGGAG	133	Mm.430741
	RmSyn3	ATAATCTCTGCCGTCCTTGC		
9	FmDCC	ACAAGACCTCAACCCACTCC	111	Mm.167882
	RmDCC	GCCGAGCATCTGTAGACTCC		
10	FmSli2	GCTGGGGTTAGTGTGTGCG	150	Mm.289739
	RmSli2	AAATCCAGTCTCTCGGTGTTG		
11	FmEPHA3	GGAATGTCAGGCAGTGTGAG	110	Mm.1977
	RmEPHA3	GTTGGTGTGAGCTAGGAGGTC		
12	FmEPHB1	ACCTCAGACCCACCTACAC	129	Mm.22897
	RmEPHB1	CCACATGACAAATCCCATAGC		
13	FmSynpr	GCTGGCATTCTCTGGAC	171	Mm.317515
	RmSynpr	TGGGACCACTAGGCTGTTG		
14	FmSema3g	ATAACTCTGACCCGGACGAC	154	Mm.27812
	RmSema3g	CTCCATTTGTTCAACCAGCAC		
15	FmEya1	ATCCACAGTCTTCCACACC	150	Mm.250185
	RmEya1	CAGGTCCCAGATGAACACTC		
16	FmSyt1	TTTTGACCGCTTCTCCAAG	170	Mm.289702
	RmSyt1	CGGCAGTAGGGACGTAGC		
17	FmSema6d	GACCTCCATCGACTTCCAG	111	Mm.330536
	RmSema6d	ACCCGTGTCTTTGTGAACC		
18	FmWnt10a	CTCGGAACAAAGTCCCCTAC	133	Mm.5130
	RmWnt10a	GCCTTCAGTTTACCAGAGC		
19	FmST18	AGGGCTCTACCCTCACCAC	149	Mm.234612
	RmST18	GAGTGTTCGATTGCTGTTC		
20	FmPhf17	AAGAAAGCGGAGCTATGAA	140	Mm.286285
	RmPhf17	CTGCTGGGAATGTGAGACAC		
21	FmNrxn1	GTGAACCATACCCAGGCTCT	113	Mm.312068
	RmNrxn1	GGCATAGAGGAGGATGAGGA		
22	FmCrta1	AGCGCAGCTCACCCCTACTAT	127	Mm.313558
	RmCrta1	CCTGAGAAGGCGTTATTGGT		
23	FmOd2	CCTATGCCAACACCACAAAG	144	Mm.339846
	RmOd2	GAGCACTGTTCTCTGGGTTT		
24	FmCacna2d3	GCCCCTATTGAAGCCTATTG	145	Mm.386754
	RmCacna2d3	CCTGATTGGTGAGTTGTTTCG		
25	FmKcnd3	CAGACCTGGAGGACCACTCT	148	Mm.44530
	RmKcnd3	CGGTGCGGTAGAAGTTAAGC		
26	FmProx1	GACAGCACAGACTCCGAAAA	127	Mm.132579
	RmProx1	GGTCCAGCTCACACATCTCA		

27	FmEsrrg	CACCTCTCTACCCCTCTGCT	136	Mm.89989
	RmEsrrg	AGTCTCTTGGGCATGGAGTT		
28	FmLsamp	AGACTGCTCTGCCTTCTTCC	142	Mm.310524
	RmLsamp	ACCAGGCCACTTTTGAGTTC		
29	FmPlxna2	CCTATGTGTCCTGCCTTTT	149	Mm.2251
	RmPlxna2	CTTGAGAAGATGGCAAACA		
30	FmL1rapl2	AGGCACTGTTGCAGAGGATA	121	Mm.328321
	RmL1rapl2	GGGGTAGCTGTCCATTGAGT		
31	FmRunx1t1	TGACAGAGAGCCTTTGCACT	143	Mm.4909
	RmRunx1t1	GGTAATGCTGAGGTGGAGGT		
32	FmRtn4r1	TCTCACACGCTTACCACCTC	134	Mm.82661
	RmRtn4r1	TTCTTGCCTGCCTTCTTGTA		
33	FmPtch1	AGCCAAGGTTGTGGTAATCC	143	Mm.228798
	RmPtch1	TGAACTGGGCAGCTATGAAG		
34	FmPrkcb1	AAGATGACGATGTGGAGTGC	91	Mm.207496
	RmPrkcb1	AAGCAGGAATGGAGCTGAGT		
35	FmCol8a1	GCCAGGTTTTCTTGGTGAAG	149	Mm.130388
	RmCol8a1	CAGGTATGCCAGGTTCTCCT		
36	FmNotch1	CTTCGTGCTCCTGTTCTTTG	120	Mm.290610
	RmNotch1	CTGGCCTCTGACACTTTGAA		
37	FmTshz3	AAACCTTCTCTGGAGCTGGA	132	Mm.44141
	RmTshz3	GTACCGATTATTTGGCGTGA		
38	FmAtbf1	CAATCACAGACACCACGACA	134	Mm.416972
	RmAtbf1	ATACTCGGACATTGCCATCA		
39	FmWnt2b	TATCTCGTCAGCAGGAGTGG	121	Mm.10740
	RmWnt2b	GTCAAAGTCCCTCGTTGAT		
40	FmAlk	ACCCAAAGTCCACTCTCACC	138	Mm.311854
	RmAlk	GGTCACTGTTGCACCTTCAG		
41	FmTcf4	TTTGCCGTCTTCAGTCTACG	107	Mm.4269
	RmTcf4	AGGAGCTAGGAAAGTGCTG		
42	FmPax2	GAGTGGTGTGGACAGTTTGC	112	Mm.192158
	RmPax2	CTGGAAGACATCGGGATAGG		
43	FmNtn1	AGGAGGGCTTCTACCAGAC	111	Mm.39095
	RmNtn1	CAGTGGTTTGATTGCAGGTC		
44	FmDnm2	GCCGAGTTTGAAGTGTGTTG	132	Mm.39292
	RmDnm2	TCCCTGATGTAGGTGGTGAC		
45	FmUnc5c	AGGAACCTACTGGCATTGAC	148	Mm.325205
	RmUnc5c	CAGTCATGGCCTCTTGTTTG		
46	FmEda	CTGACTTTGCCAGCTACGAG	149	Mm.328086
	RmEda	CATCTTCACGGCGATTTC		
47	FmIca1	TGCAGTGTGGAGATGACC	107	Mm.275683
	RmIca1	AAAGCTGTGATGGAAGGAATC		
48	FmGas7	ACAAGACCGAGGAGGACATC	136	Mm.40338
	RmGas7	AGTTCCAGTGTGGTGTGAC		
49	FmDab1	AAGCTCAAGGGTGTGTTGTC	131	Mm.289682
	RmDab1	ACAGCATGGTGTGCTGAAG		
50	FmNrp2	TGAATCTCCAGGTTTCCAG	145	Mm.266341
	RmNrp2	ACAGTCTCCTTCCCCACTT		
51	FmNTRK3	GCCAGAGCCTTTACTGCATC	123	Mm.33496
	RmNTRK3	CTCCTTCTCGACAGTCAAG		

D. Primers for cloning ECR enhancer fragments

name	sequence	size
1 FmGadd45g-prom-min	aacgtagcTCCCCACGTGGCCTTCTG	99
RmGadd45g-prom-min	aacaagcttGCCAGCTGGCGGGAGGCT	
2 FmAth3-ECR1min (NeuroD4)	aacgtagcTCTGAGGCCAGTTTCTCA	166
RmAth3-ECR1min	aacaagcttGGCTGAAAGCCACAGACG	
3 FmAth3-ECR2min	aacgtagcTCTCTGCTTCAATGGCTCC	257
RmAth3-ECR2min	aacaagcttGGCTGTGCAATCAGCAAA	
4 FmAth3-ECR3min	aacgtagcGGAGGAGAATGAATGCAA	144
RmAth3-ECR3min	aacaagcttGTCAAAACCCCTTGACAG	
5 FmDII1-ECR1-min1	aacgtagcCAGACAGAACTCCAGATG	144
RmDII1-ECR1-min1	aacaagcttTGTCTGGTGGTAGTGTG	
6 FmDII1-ECR1-min2	aacgtagcGGCCAACATTACAGGCAAA	173
RmDII1-ECR1-min2	aacaagcttCGCCTTTAGACGCCATCA	
7 FmDII1-ECR2-min1	aacgtagcGGGTGTGAGTATCTTTTG	145
RmDII1-ECR2-min1	aacaagcttCCAACCTCACAGACAATCA	
8 FmDII1-ECR2-min2	aacgtagcTAGGCAGGCTGGGTGATC	118
RmDII1-ECR2-min2	aacaagcttTCGCATGCCAGAGAAAC	
9 FmDII1-ECR2-min4	aacgtagcGGTGACGGTTGTGTAAG	201

RmDII1-ECR2-min4	aacaagcttCTGTTATTGTGCGAGGCT	
10 FmEbf3-ECR1-min	aacgctagcGGTGAGCGCTGTCTCTTT	179
RmEbf3-ECR1-min	aacaagcttGGAGAGGCTTCAATGAGC	
11 FmEbf3-ECR2-min	aacgctagcGCCAGCCGTAGAGTGACA	
RmEbf3-ECR2-min	aacaagcttTAACGATCCTGCCCCAGA	276
12 FmElav3-promo-min	aatctcgagTAGGCCCGCACTCGCTT	
RmElav3-promo-min	aacaagcttATGGGCTCTCTCTCTC	201
13 FmElav3-enh-min	aacgctagcGAAGGAAAAGGGCTCAGA	
RmElav3-enh-min	aacaagcttCTCCTCAGACCCCTTGG	165
14 FmHEN1-ECR1-min (Nhlh1)	aacgctagcTGGCAATGATGATGCAAGA	
RmHEN1-ECR1-min	aacaagcttCCAAGCTCAGCTCTCTCT	179
15 FmHEN1-promo-min	aacgctagcGAGTCTCTCTGCCCTGCT	
RmHEN1-promo-min	aacaagcttGGATCTGGGAACCCCTTG	210

E. Primers for cloning EEL enhancer fragments

FmCbfa2I2EEL1		
16 (MTGR1)	aacgctagcGGATTCTGGCAAGGTGAAGA	637
RmCbfa2I2EEL1	TGAAAATGAGCCCTGCAGAAGT	
17 FmCbfa2I2EEL6	aacgctagcTGCAGCCATTGTTTAGCTG	
RmCbfa2I2EEL6	aacaagcttTTGTGTCCTTCCCATCTC	566
18 FmDII1EEL1	aacgctagcCGCATACACATGGCTAGG	
RmDII1EEL1	aacaagcttCCCCACTTCCAATTCTAT	202
19 FmDII1EEL2	aacgctagcTCATTGTGCCACGTCTGAAT	
RmDII1EEL2	aacgaattcTGGATGTGGGTACGACTTGA	570
20 FmEbf2EEL1	aacgctagcTGAGGTGAAGTGACGTGCTC	
RmEbf2EEL1	aacgaattcCGCGTAGATCTGCTGTTGAA	383
21 FmEbf2EEL3	aacgctagcTGCTGCGTTAGTTCATTTG	
RmEbf2EEL3	aacaagcttCCAAACTCCAGCCCAAGATA	469
22 FmMyT1EEL	aacgctagcCTCACAGTGCAGGGCATAA	
RmMyT1EEL	aacgaattcTATGACTGCTGTCGGCTCAC	534
23 FmNeuroD4EEL1	aacgctagcACGAAAGCCATCAGAAGGAAT	
RmNeuroD4EEL1	aacaagcttATCTCCCATCCTGAGAGCCTA	340
24 FmNeuroD4EEL2	aacgctagcACTGTTTTCCACAGGGATGTG	
RmNeuroD4EEL2	aacaagcttCCACATGCAATGAGACTGCT	408
25 FmNeuroD4EEL4	aacgctagcGAGAGAAATGCAGCAATCTGG	
RmNeuroD4EEL4	aacaagcttGAAGGTTAAATGTGCCCAAT	408
26 FmZfp238EEL1	aacgctagcTGTAGATGTGAGCTGCCTGAA	
RmZfp238EEL1	aacgaattcTTTCATCCTAACCCCGATACC	323
27 FmZfp238EEL2	ctagcGCACAGCTGGAGGTAGCATTTCATCTGAACATGATGTCAGAGCAGCTGACA	
RmZfp238EEL2	agctTGTCAGCTGCTGACATCATGTTTCAGATGGAAATGCTACCTCCAGCTGTGCg	

F. Primers for site directed mutagenesis of E-boxes in enhancer fragments

name	sequence
1 FmHEN1E2mut (Nhlh1)	AGGCTGAGGGGGTctcgagTGGCCCTGCCCTGG
RmHEN1E2mut	CCAGGGCAGGGCCActcgagAACCCCTCAGCCT
2 FmHEN1E1mut	ATATGCATGGAGAActcgagTCCCTGCTCCCCCA
RmHEN1E1mut	AGGGGGAGCAGGGActcgagTTCTCCATGCATAT
3 FmEbf3ECR1E9mut	CGACATAGGGAGGActcgagGGCACTGGCCATCTG
RmEbf3ECR1E9mut	AGATGGCCAGTGCCctcgagTCTCCCTATGTGCT
4 FmEbf3ECR1E8mut	AGCTGGGCACTGGCctcgagACTGACTCCGGCTTG
RmEbf3ECR1E8mut	AAGCCGGAGTCACTctcgagGCCAGTGCCCACTG
5 FmEbf3ECR1E7mut	TGACTCCGGCTTGActcgagGGAGCCCACTGGTGT
RmEbf3ECR1E7mut	CACCAGTGGGCTCCctcgagTCAAGCCGGAGTCAG
6 FmEbf3ECR1E6mut	CTGCTATTTAAATGctcgagTGAGGGGGCTCATTG
RmEbf3ECR1E6mut	AATGAGCCCCCTCActcgagCATTAAATAGCAGC
7 FmEbf3ECR2E5mut	TTTGGGTGCGGACctcgagATGTAGAAAATGAGC
RmEbf3ECR2E5mut	CTCATTTTCTACATctcgagGTCGCGCACCCAAAG
8 FmEbf3ECR2E4mut	AGCCTGTTGTGAGActcgagATTCCTCCATTTCCGG
RmEbf3ECR2E4mut	CGAAATGGGGGAATctcgagTCTCACAACAGGCTC
9 FmEbf3ECR2E3mut	CCCCATTTCCGGGctcgagGCTGCCGAACACAGA
RmEbf3ECR2E3mut	CTGTGTTCCGGCAGCctcgagGCCGAAATGGGGGA
10 FmEbf3ECR2E2mut	CCAAAATAGCGGGGctcgagCATTTGTTGTCAAAA
RmEbf3ECR2E2mut	TTTGACAACAATGctcgagCCCCGCTATTTTGGGA
11 FmEbf3ECR2E1mut	TAGCGGGGAGCTGctcgagTTGTCAAAAAGGTTT
RmEbf3ECR2E1mut	AACCTTTTTGACAActcgagCAGCTGCCCCGCTAT
12 FmEbf3ECR2POUm	CTGCCGAACACAGaggcctATTCATTTTCGCTT
RmEbf3ECR2POUm	AAGCGAAAATGAATaggcctCTGTGTTCCGGCAG
13 FmMyT1EELmP3	CTGGAGAAAGTGGTaggcctTACAGATGCAGAA
RmMyT1EELmP3	TTCTGCATCTGTAaggcctACCACCTTCTCCAG
14 FmMyT1EELmP2	AAGAGTTTTATGAAaggcctGAACTGACACTTGG

RmMyT1EELmP2	CCAAGTGTCAGTTCAaggcctTTCATAAAAACTCTT
15FmMyT1EELmP1	TGCTGTGTGCGGAaggcctACAAGTAATTACACC
RmMyT1EELmP1	GGTGAATTACTTGTaggcctTCCGCACACAGACA
16F-DII1ECR1m2mE2	GAGTGAAAAACGCctcgagGGTTCGGAGCAGATG
R-DII1ECR1m2mE2	ATCTGCTCCGAACCctcgagGCGTTTTTCACTC
17F-DII1ECR1m2mE1	CATTGGTTCCGAGctcgagGGCTGGCTAGG
R-DII1ECR1m2mE1	CCTAGCCAGCCctcgagCTCCGAACCAAATG

G. Primers for chromatin immunoprecipitation

primer name	primer sequence	Product size
F-mEbf3 ECR1 ChIP	TTAGGTGATGCCACAAACA	291
R-mEbf3 ECR1 ChIP	CCTTGGAGGGTGGGAAGTAT	
F-mEbf3 ECR2 ChIP	CTAGGGTGGAGACCCTTCCT	245
R-mEbf3 ECR2 ChIP	AAGCGAAAAATGAATGCAAATC	
F-mHEN1 ECR1 ChIP (Nhlh1)	ATGGCAATGATGATGCAAGA	282
R-mHEN1ECR1 ChIP	AcaggtgTGCAAAGGATTC	
F-mEbf2 EEL1 ChIP	CCCGACTGTAATGATGCTGA	162
reverse primer	use RmEbf2EEL1	
F-mEbf2 EEL3 ChIP	TTCGCCAATTATTAGCAAGGA	295
R-mEbf2 EEL3 ChIP	CCTGCGATTGCATAACAAAA	
F-mMTGR1 EEL1 ChIP	GGATTCTGGCAAGGTGAAGA	162
R-mMTGR1 EEL1 ChIP	GACAGCCTTCCTCCTCCTTT	
F-mMTGR1 EEL6 ChIP	TGCAGCCATTGTTTTAGCTG	171
R-mMTGR1 EEL6 ChIP	GGAAGTAATCAAGCGCAAGG	
F-mNeuroD4 EEL4 ChIP	CCTTGTCAGCCTTTCAGGAG	270
R-mNeuroD4 EEL4 ChIP	AGAGTTGCTTTCAGGCCAAA	
F-mZnf238 EEL1 ChIP (Zfp238)	GATGTGAGCTGCCTGAATTG	135
R-mZnf238 EEL1 ChIP	AGAGGGACGAAGAAGGAAGC	

Seo_Supplemental Figure S2.

A. Sequences of functional (Ngn/NeuroD-responsive) EEL and ECR enhancers

MTGR1 EEL1 (162 bp)

ggattctggcaaggtgaagagctggtctgtttggcaggacctcc**CATCTG**ggatc**CATCTG**gtttccctggaatgggacaggca
gctctagtcaagtgaatcg**CAGATG**tgagagagctgacacatgctccagctaataaaggaggaggaaggctgc

MTGR1 EEL6 (566 bp)

Tgcagccattgttttagctggtctcttcacccctctgctgttactctgctcttttgggccattgtttctgtaggaaattagcattaatc**CATC**
TGaaacccccatgcccctcaactgcaggtgactgttctggccc**CAGCTG**ccacccttgcgcttgattactccataggggctgtgca
aagaagtagtagcatttattgattgccaccacaagccccacccaataaactgaaatctattaggcctccagtaggctgtagctca
ttggcattacactgtgatgaacgtttacaagcaaaataattatgtctttttcttcccttagcctttccgaatgacagatggagaaaacctga
gttatcatagtgttaattgaaatgctggagctagatgcttctcagtagtgagctctaagaattggatggactgtctgaaaatttagtca
atttaaatatttctgaccacccataggttcaagcaaacatggcctgataagtccttaggatagagatgggaaggacacaaa

DLL1 EEL1 (202 bp)

cgcatacacacatggctagggccgggggtgggggtgggggtggggggcagcattgttgctgggg**CAGCTG**cgggaggggaagac
acaatggag**CAGCTG**tctgccctgcagctgcacac**CAGCTG**tccactcttatctgcacacacactttctgggatagaagaggt
gggctttgtcatagaattgggaagtggggg

DLL1 EEL2 (569 bp)

tcattgtccacgtctgaatgcctgtgaataggcagagcatatcggcctgtcactgcaggtgccca**CATCTG**ttctccagg**CAG**
ATGtcatagaaaaggaggtgggggaggccagggaagcctggagcactaaaagcttcttggcctagcctctctggggcatctg
ggagtggcctttgggtgctgctgtgtccaacagggaacttctggggagggcctgaaccacacgtgctcacacagcaagcatagctcagac
cctagttaagcctgagagcggcaggctttgaggttattgtaagaaactcttgggagcctaactttaacttcccggccaaggccagagcc
tgagcaattagacactggggagagagagagcagggaggggagcagtagggtagcgtggaggcagtgatcttgcacatcgttcga
cagccagctggtcctctgataccaggtgaggtgagagcactgtgccccggccagctgcatggaaccagcagcactctctcaaagc
tttgactcaagctgaccacatcca

Ebf2 EEL1 (383 bp)

tgaggtgaagtacgtgctcaggtccccctcaatggagactcaagcagctctgaagcgcagcttctttcgtcgcgggcaaacgtgc
acccagggcgtgtcctctagcggcgaactctgcagctcgtctgtgagcgtcacctacagctagggggtccggctcaggagcaattga
cactcaaacctgagccctgagccgggaggaaagactcccactgtaatgatgctgagaggtcgggtgagctcgtcgaaccggcg**C**
AGCTGattgctcctccag**CAGATG**gtctttgagcaca**gattgca**tctt**caaat**ctgtttcgcgctggcgttctagattc
aacagcagatctacgcg

Ebf2 EEL3 (469 bp)

tgctgcgttagtcttcccaattattagcaaggagatactaggtgtatgttgcctgccaccctcgcacctcaatgctcagagggtta
aaagatattataattgaacagaagctgctctagcccta**CAGCTG**gctaataacagagttgattgagcgtgagcaaaccca**CAGC**
TGgacttttccacaaggcgcaggttagacagggccctaat**CAGATG**ggccagccgggcagtgggccagtgctttatcggc
CAGATGgtgtgaatttagacacttttattgcaatcgcagggaggagttggggagaagaaaacaaaagccaccatgctgtga
gcttattgagttgaaattagagacagattttgaaagttgaggggtgaaatftaaaaaatgagttggaagggaagacacattagcttattc
ggctggagtttg

Ebf3 ECR1 (179 bp)

ggtgagcgtgtctcttfaaggtcactgcctgctccggcagcagacataggagga**CAGCTG**ggcactggc**CATCTG**actgactcc
ggcttga**CATCTG**ggagcccactggtgtggtggcacggaagcgttgatgctgctattfaaatg**CAAATG**tgagggggctcattg
aagcctctcc

Ebf3 ECR2 (276 bp)

gccagccgtagagtgacaaaaagataaagttggtgaatgataaagaccgtatftccacgcttgggtgcgcgacCAGATGatgtaga
aaatgagctgaaatggattcagccccgagcctgtgtgagaCAGCTGattccccatttcgggcCAGATGgctgccgaacacag
attgcaatcatttcgcttaatatcgtccaaaatagcggggCAGCTGCATTGttgtcaaaaagggttaaaccaccttctttctggg
gcaggatcgta

MyT1 EEL (534 bp)

ctcacagttgcagggcataagtgtgcatgtgtgagggagtctcataaactcattgccaaatgttgcgagtgtgagacaagcacacaaagt
cagcctggagtggtgtaaagtggctcctggtttgttctcCATCTGccaggcctgctggggcagcgcccccagccctCAG
CTGttctatacaatattaatgctgacctctcggctgcgcacacaacgccctaattgtcggcgagcattaaggcaggcgcctCATCT
GcggaagccttcagaaggaaacaagaaaacctcaccctCAGCTGgtcactggagaaggtgtgccaattaCAGATGcagaa
agtttagaacaataacttgaagagttttatgaagcaaatgaactgacacttggaatgtctgtgtgcgattgcaagaataactacaccttatt
tgaatgatcctaccagtaagcaacaattaagtctgctctccagcgcacacaagcacttccgagtgcagccgacagcagtcata

NeuroD4 EEL1 (340 bp)

acgaaagccatcagaaggaatcactgggttacagtctagagtgtgtgactCAGCTGgagcctgaatacaaaaaggctcctgagcagaa
gcatcaggacagacagagggCAGATGaaattgtcctgcagttggcacagtcacaagaatttcaaggataaatgagaagaaaa
tcagccctgggagggagcactgctgacatgtttcctgcttctgagcagCAGCTGcaggCAGATGgagCAGCTGCAGC
TCtcctaggagcctggattcctgggaaagaggcctcaggaaggaaatgggagggagcagcagcataggctctcaggatgggagat

NeuroD4 EEL2 (408 bp)

actgtttaccagggatgtgctccttaagaggctgccatgagctgtctatgtctcactcatgacacttctgcttctgcacatgggctatgt
gcattggcccagtgatgcatagccttttaggaaaaggctCATCTGtgcctgttcggttcatcttccagaccctggagattggcag
agaccctagagaagaactgcttctgggtgtagggCAGATGctgtgCAGCTGagctattaggacaagcctggctctgcagtttta
gaagtcagacagatcagccccattcttactagaggaaattggtgagttctcactctcagtaacctctcaagcccagagagagcgtgta
ttgagaaaaagctgaagaagtagcagctcattgcatgtgg

NeuroD4 EEL4 (408 bp)

gagagaaatgcagcaatctgggcagggtgacctaaactgatctccatagtcctctgtgataatccctgcctctcccctactcctgtcag
ccttcaggaggtcagcagatgacgttactgtgggggtgggagtaggattgagttgccttccctgtCAGCTGgtgtttgtttcatgaCA
GATGgtgggtgatgactgtctttatgtcttccctagccttagaggagttttatgtccagcaaatgtctatacttccagaggaat
ggCATCTGcctctgaggggatgtccagcagcactctgtcttcttcttgcactccttggcctgaaagcaactctgtgacctttagtga
cttaggagacatagattggggcacattaacctt

Zfp238 EEL1 (323 bp)

tgtagatgtgagctgctgaattggccagaCAGCTGtaatcgactcctctattcttaatcttcttcttggcagCAGCTGcCATA
TGgccacagtCAGCTGgatCAGATGttataagcttcttcttctcctctgtcctctcagcctcctaattgatcacggetacctgtg
gagctgcccctcaatcttatttttagccctctcaaggattagattgattgagctgtagggcacttaagtattgtattgaaatctggtttatc
tattggtggtgtcttcttctcagaaggtatcggggttaggatgaaa

Zfp238EEL2 (69 bp)

ttaaaggagcaCAGCTGgaggtagcatttcCATCTGaacatgatgtcagagCAGCTGacagcctgcat

HEN1 promo (210 bp)

GagtctctctgcctgctacctctagtgtgggaggtgagggggtCATCTGtgccctgccctggtcttgcctcagccccccccttc
cccagctctctatttctgctatgaaatgcatggagaaCAGATGtcctctccccactccccccccctcaggccagcccc
ccccagctccaaggggtccagatcc

B. Sequences of non-functional EEL and ECR enhancers

mGadd45g mini (2)

tcccCACGTGgccttctggcagagccgagccctgccctcattgcatcagaaagcgggcgctggccaatgggagtgagcctc
ccgcCAGCTGgc

mNeuroD4 ECR1 (3)

tctgagccccagttctcactttctatagctctgattatgcctctggaccCATTGtttgaagtagtggtcttaattcaaaaaggCA
AGTGgtctccaaggaagaggagaagggggaCATCTGgggacagtaccaacagcgtctgtgcttccagcc

mNeuroD4 ECR2+3 (4)

tctctgctcaatggctctagtgctcagggccagtgccagCAACTGagtgctggtctggtctcttttccactcagCAGATGgtt
ctaacaggttacctcagatctttacaaaaagcaatagagaaagggagggctgtggactggccctctggccaggaaatgtggggg
aggagaatgaatCAAATGaccttaattcttttctacggcagtcactccctcttttctgattgcacagcctttctgccttctgc
tctctcagaattcaacagccaacccaCAACTGtcaaggggtttgac

mDl1 ECR1 mini1 (3)

cagacagaactCAGATGCACTTGttgtggctttaaattctgattattatgcaacctgagctgcctatccaccattgcaattattcag
attgggctggggggggggggCATGTGgaaaaaatcaactaccaccagac

mDl1 ECR2 mini1 (2)

gggtgctagatctttgaagcctCATTGttctataataaacaggtttttaaagtgggatctaacctgectttctcacctcagccttga
gtattataCACATGgcttttggtaactcttctgattgtctgtgagttgg

mDl1 ECR2 mini2 (3)

taggcaggctgggtcatcagaaaggagctgtaattCAGTTGcCAGATGgccaacaCAGATGattctgccagtaactg
ctagattctgttagcagtgcttctctggcatgca

mDl1 ECR2 mini4 (3)

gtgacgggtgtgaaaggttatgtgtgggggtgaggtggggggagctttcaaaCATGTGccggacattgtgcagaggccgccc
gtgcgggaggggagctcttctcgcattgtgcagagagCAGGTGctgtctgattaccataCAGCTGagcgcacaaagagc
cactgattcagcctgcacaataaca

mElav13 promo mini (2)

gcactcgttcCACCTGcccgccgctgggagccgggagggagccggagaccaccgtctcagccaatgagcgacggcgctg
gCAGCTGcctctat

mElav13 enh mini (3)

aaggaaaagggtcagaaatacaaggttggtctgtacCACGTGgtggccctgactgtgtgcttagatgccactcaaagccctatt
gtcaccagatccaggCATCTGtactgtgactggagccagaggccCAGCTGgaggccaaggggtctgagga

mHEN1 ECR1 mini (4)

tggcaatgatgatgcaagaaccaagccgcCACCTGatccctgccccccccctcctcagcccttctccctccagctcagccC
ATCTGctggcgcctgggCACCTGtctggcctgccCAGGTGggagttggcacctccagcctgggggtggggagagga
gcgtgagctt

Key:

CAGATG-type (and reverse complements for each motif)
 CAGCTG-type
 CAAATG-type
 Highlighted are Pou site, degenerate consensus gmttkgmw

C. Types of E-boxes found in all ECR enhancers versus those in functional ECRs

Types of E-boxes found in all ECR fragments:

CAGGTG: 5 (17.2)	CAAGTG: 2 (6.9)	CATGTG: 3 (10.3)	CACGTG: 2 (6.9)
CAGATG: 7 (24.1)	CAAATG: 3 (10.3)	CATATG: 0	
CAGTTG: 3 (10.3)	CAATTG: 0		
CAGCTG: 4 (13.8)			total: 29

Types of E-boxes found in functional (Ngn/NeuroD-responsive) ECR fragments:

In Ebf3 ECR1 (4), Ebf3 ECR2 (5), HEN1 promo (2), and Dll1 ECR1 mini1 (2): total 13

CAGATG: 7/13 = 53.8%	14/42 = 33.3%
CAGCTG: 3/13 = 23.1%	7/42 = 16.7%
CAAATG: 3/13 = 23.1%	6/42 = 14.3%

D. Weight matrices (experimental and computational)

>Ngn2 experimental consensus

```
A 3 1 0 7 2 6 0 0 0 0
C 2 4 7 0 0 1 0 0 0 6
G 1 2 0 0 5 0 0 7 3 1
T 1 0 0 0 0 0 7 0 4 0
```

>NeuroD1 experimental consensus

```
A 1 1 0 5 0 4 0 0 0 0
C 2 3 5 0 0 1 0 0 0 5
G 2 0 0 0 5 0 0 5 2 0
T 0 1 0 0 0 0 5 0 3 0
```

>Ngn/NeuroD position frequency matrix-CompareProspector

```
A 13 27 0 57 3 36 0 0
C 10 5 57 0 0 21 0 0
G 27 12 0 0 48 0 0 57
T 7 13 0 0 6 0 57 0
```

>POU position frequency matrix

```
A 4 23 3 0 0 0 6 25
C 3 11 0 0 0 0 28 0
G 27 0 0 0 9 34 0 0
T 0 0 31 34 25 0 0 9
```

E. Artificial enhancer primer sequences

5' GA E box	ctagcCAGATGtcgactgtacgacatgcCAGATGtcgaatgtaggacttgcCAGATGtca
3' GA E box	gatctgaCATCTGgcaagtcctacattcgaCATCTGgcatgtctacagtgcgaCATCTGg
5' GC E box	ctagcCAGCTGtcgactgtacgacatgcCAGCTGtcgaatgtaggacttgcCAGCTGtca
3' GC E box	gatctgaCAGCTGgcaagtcctacattcgaCAGCTGgcatgtctacagtgcgaCAGCTGg
5' AA E box	ctagcCAAATGtcgactgtacgacatgcCAAATGtcgaatgtaggacttgcCAAATGtca
3' AA E box	gatctgaCATTTGgcaagtcctacattcgaCATTTGgcatgtctacagtgcgaCATTTGg
5' AT E box	ctagcCAATTGtcgactgtacgacatgcCAATTGtcgaatgtaggacttgcCAATTGtca
3' AT E box	gatctgaCAATTGgcaagtcctacattcgaCAATTGgcatgtctacagtgcgaCAATTGg
5' TA E box	ctagcCATATGtcgactgtacgacatgcCATATGtcgaatgtaggacttgcCATATGtca
3' TA E box	gatctgaCATATGgcaagtcctacattcgaCATATGgcatgtctacagtgcgaCATATGg
5' CG E box	ctagcCACGTGtcgactgtacgacatgcCACGTGtcgaatgtaggacttgcCACGTGtca
3' CG E box	gatctgaCACGTGgcaagtcctacattcgaCACGTGgcatgtctacagtgcgaCACGTGg
5' GG E box	ctagcCAGGTGtcgactgtacgacatgcCAGGTGtcgaatgtaggacttgcCAGGTGtca
3' GG E box	gatctgaCACCTGgcaagtcctacattcgaCACCTGgcatgtctacagtgcgaCACCTGg
5' GT E box	ctagcCAGTTGtcgactgtacgacatgcCAGTTGtcgaatgtaggacttgcCAGTTGtca
3' GT E box	gatctgaCAACTGgcaagtcctacattcgaCAACTGgcatgtctacagtgcgaCAACTGg
5' AG E box	ctagcCAAGTGtcgactgtacgacatgcCAAGTGtcgaatgtaggacttgcCAAGTGtca
3' AG E box	gatctgaCACTTGgcaagtcctacattcgaCACTTGgcatgtctacagtgcgaCACTTGg
5' TG E box	ctagcCATGTGtcgactgtacgacatgcCATGTGtcgaatgtaggacttgcCATGTGtca
3' TG E box	gatctgaCACATGgcaagtcctacattcgaCACATGgcatgtctacagtgcgaCACATGg

Seo_Supplemental Figure S3. Ngn and NeuroD target genes that regulate cell migration/morphology

Gene name (Gene symbol)	Xenopus Unigene	Protein structure/roles	References
Calponin2 (Cnn2)	Xl.554	Thin, filament-associated protein that regulates migration of endothelial cells in zebrafish	Tang, J., G. Hu, J. Hanai, G. Yadlapalli, Y. Lin, B. Zhang, J. Galloway, N. Bahary, S. Sinha, B. Thisse, C. Thisse, J.P. Jin, L.I. Zon, and V.P. Sukhatme. 2006. A critical role for calponin 2 in vascular development. <i>J Biol Chem</i> 281 : 6664-72.
Frizzled homolog 7 (Fzd7)	Xl.663	Planar cell polarity (PCP) pathway; regulates convergent extension during gastrulation	Djiane, A., J. Riou, M. Umbhauer, J. Boucaut, and D. Shi. 2000. Role of frizzled 7 in the regulation of convergent extension movements during gastrulation in <i>Xenopus laevis</i> . <i>Development</i> 127 : 3091-100. Sumanas, S. and S.C. Ekker. 2001. <i>Xenopus</i> frizzled-7 morphant displays defects in dorsoventral patterning and convergent extension movements during gastrulation. <i>Genesis</i> 30 : 119-22.
Prickle 1 (Prickle)	Xl.7556/ Xl.12111	Planar cell polarity (PCP) pathway; regulates posterior migration of motor neurons in zebrafish	Carreira-Barbosa, F., M.L. Concha, M. Takeuchi, N. Ueno, S.W. Wilson, and M. Tada. 2003. Prickle 1 regulates cell movements during gastrulation and neuronal migration in zebrafish. <i>Development</i> 130 : 4037-46.
MGC68463 similar to growth arrest-specific 6 (Gas-6)	Xl.4807	Regulates migration of gonadotropin-releasing hormone (GnRH) neurons by controlling actin cytoskeleton reorganization and promoting lamellipodial extension, membrane ruffling, and chemotaxis	Allen, M.P., D.A. Linseman, H. Udo, M. Xu, J.B. Schaack, B. Varnum, E.R. Kandel, K.A. Heidenreich, and M.E. Wierman. 2002. Novel mechanism for gonadotropin-releasing hormone neuronal migration involving Gas6/Ark signaling to p38 mitogen-activated protein kinase. <i>Mol Cell Biol</i> 22 : 599-613.
Similar to LIM and senescent cell antigen-like domains 1 (Lims1, PINCH1)	Xl.8188	Regulates assembly of the actin cytoskeleton	Li, S., R. Bordoy, F. Stanchi, M. Moser, A. Braun, O. Kudlacek, U.M. Wewer, P.D. Yurchenco, and R. Fassler. 2005. PINCH1 regulates cell-matrix and cell-cell adhesions, cell polarity and cell survival during the peri-implantation stage. <i>J Cell Sci</i> 118 : 2913-21.
Transcribed locus similar to Myosin 10	Xl.24147	Mediates interactions between the cytoskeleton and plasma membrane and stimulates formation of filopodia	Sousa, A.D. and R.E. Cheney. 2005. Myosin-X: a molecular motor at the cell's fingertips. <i>Trends Cell Biol</i> 15 : 533-9. Zhang, H., J.S. Berg, Z. Li, Y. Wang, P. Lang, A.D. Sousa, A. Bhaskar, R.E. Cheney, and S. Stromblad. 2004. Myosin-X provides a motor-based link between integrins and the cytoskeleton. <i>Nat Cell Biol</i> 6 : 523-31
LOC494829 similar to Drebin-1 (Dbn1)	Xl.49084	F-actin binding protein that localizes to dendrites and is involved in spinogenesis and synaptogenesis in dendrites	Aoki, C., Y. Sekino, K. Hanamura, S. Fujisawa, V. Mahadomrongkul, Y. Ren, and T. Shiraio. 2005. Drebrin A is a postsynaptic protein that localizes in vivo to the submembranous surface of dendritic sites forming excitatory synapses. <i>J Comp Neurol</i> 483 : 383-402. Mizui, T., H. Takahashi, Y. Sekino, and T. Shiraio. 2005. Overexpression of drebrin A in immature neurons induces the accumulation of F-actin and PSD-95 into dendritic filopodia, and the formation of large abnormal protrusions. <i>Mol Cell Neurosci</i> 30 : 630-8.
Polo-like kinase 3 (Plk3; Plx3 in <i>Xenopus</i>)	Xl.16835	Regulates the actin cytoskeleton	Duncan, P.I., N. Pollet, C. Niehrs, and E.A. Nigg. 2001. Cloning and characterization of Plx2 and Plx3, two additional Polo-like kinases from <i>Xenopus laevis</i> . <i>Exp Cell Res</i> 270 : 78-87. Liu, X. and R.L. Erikson. 2002. Activation of Cdc2/cyclin B and inhibition of centrosome amplification in cells depleted of Plk1 by siRNA. <i>Proc Natl Acad Sci U S A</i> 99 : 8672-6.
Similar to angiomin-2 (Amot2)	Xl.53293/Xl.76708	Like mammalian Angiomin, which localizes to the leading edge of migrating endothelial cells and promotes cell migration and tube formation	Bratt, A., O. Birot, I. Sinha, N. Veitonmaki, K. Aase, M. Ernkvist, and L. Holmgren. 2005. Angiomin regulates endothelial cell-cell junctions and cell motility. <i>J Biol Chem</i> 280 : 34859-69. Trojanovsky, B., T. Levchenko, G. Mansson, O. Matvienko, and L. Holmgren. 2001. Angiomin: an angiostatin binding protein that regulates endothelial cell migration and tube formation. <i>J Cell Biol</i> 152 : 1247-54.

Seo_Supplemental Figure S4. Expression Profiles of Ngn/NeuroD target genes

A. Expression profiles of Ngn/NeuroD-responsive target genes from microarray

	# of genes with expression in tissue	Percentage of genes with expression in tissue	Nhlh1	Neurod4	Pou3f1	Pou3f2	Ebf2	Ebf3	Hes5	Dll1	Elavl3	Sat1	Gadd45g	RIKEN cDNA 9330182L06
adipose tissue	1	8%	0	0	0	0	0	0	0	0	0	1486	0	0
blood	0	0%	0	0	0	0	22	0	0	0	0	88	33	0
bone	1	8%	0	0	0	0	0	56	0	0	0	56	0	0
bone marrow	0	0%	0	0	0	0	0	13	0	0	6	34	13	6
brain	7	58%	22	16	9	9	29	45	74	3	110	38	35	97
connective tissue	3	25%	0	0	0	0	148	49	0	0	0	49	297	0
dorsal root ganglion	6	50%	0	371	0	0	371	297	74	222	0	74	0	74
embryonic tissue	2	17%	6	0	0	0	42	12	0	12	0	54	12	42
endocrine	3	25%	0	117	0	0	58	58	0	0	0	88	29	176
epididymis	0	0%	0	0	0	0	0	0	0	0	0	0	0	0
extraembryonic tissue	1	8%	0	0	0	0	0	13	0	13	0	26	173	0
eye	7	58%	10	183	10	0	16	145	91	48	113	48	5	32
female genital	1	8%	0	0	0	0	24	24	0	0	0	74	49	0
gastrointestinal tract	1	8%	0	16	0	0	32	0	0	0	8	82	74	0
head and neck	8	67%	15	0	0	0	176	107	53	30	30	46	46	53
heart	2	17%	0	0	0	0	36	90	0	0	0	0	36	0
inner ear	4	33%	0	0	0	25	355	50	0	25	0	0	25	0
limb	0	0%	0	0	0	0	0	33	0	0	0	33	0	0
liver	1	8%	0	0	0	0	26	17	0	0	0	62	35	0
lung	3	25%	0	0	0	0	0	0	0	38	0	241	135	9
lymph node	2	17%	0	0	0	0	0	0	0	63	0	255	0	0
mammary gland	1	8%	0	0	0	0	12	28	0	12	0	338	22	9
muscle	1	8%	0	0	0	0	0	0	0	0	0	224	0	0
pancreas	4	33%	9	0	0	0	45	72	0	27	0	18	82	18
prostate	1	8%	0	0	0	0	0	0	0	0	0	355	0	0
skin	2	17%	0	0	11	0	11	22	0	0	0	11	0	44
spinal cord	7	58%	0	74	0	0	187	187	74	37	299	0	0	112
spleen	1	8%	0	0	0	0	0	20	10	0	0	10	40	0
sympathetic ganglion	2	17%	0	0	0	0	358	0	0	0	0	0	0	448
testis	1	8%	0	0	0	0	16	16	0	0	0	162	0	0
thymus	1	8%	0	0	0	0	37	0	29	0	0	29	0	0
urinary	3	25%	0	0	0	0	0	28	0	21	7	141	56	0
vascular	2	17%	0	0	0	0	160	80	0	0	0	0	0	0
vesicular gland	0	0%	0	0	0	0	0	0	0	0	0	0	0	0
AVERAGE (in transcripts per million)			2	23	1	1	64	43	12	16	17	121	35	33

B. Expression profiles of non-responsive target genes from microarray

	# of genes with expression in tissue	Percentage of genes with expression in tissue	Twist1	Btd6	Zfx1b	Slc43a2	Irx3	b4galt3	Plk3	Prickle
adipose tissue	0	0%	0	0	0	0	0	0	0	0
blood	2	25%	0	0	66	231	0	121	0	0
bone	4	50%	0	28	113	142	0	56	0	113
bone marrow	1	13%	0	6	224	81	0	61	6	6
brain	5	63%	0	29	110	84	0	129	35	107
connective tissue	0	0%	0	0	0	0	0	0	0	0
dorsal root ganglion	2	25%	0	0	445	0	0	0	0	222
embryonic tissue	2	25%	12	0	79	6	0	36	0	134
endocrine	3	38%	58	0	117	0	29	29	0	29
epididymis	0	0%	0	0	0	0	0	0	0	0
extraembryonic tissue	1	13%	0	13	26	106	13	93	0	0
eye	3	38%	0	16	156	21	32	37	10	75
female genital	5	63%	0	0	174	24	24	174	74	74
gastrointestinal tract	1	13%	0	8	32	65	0	41	16	32
head and neck	1	13%	15	15	76	46	7	69	0	38
heart	3	38%	0	36	72	54	18	90	0	18
inner ear	3	38%	25	0	152	50	0	0	0	126
limb	4	50%	67	67	0	0	67	100	0	0
liver	2	25%	0	80	17	8	0	17	35	0
lung	7	88%	19	38	106	115	77	115	28	125
lymph node	1	13%	0	0	191	0	0	0	0	0
mammary gland	3	38%	25	16	70	25	125	125	3	32
muscle	2	25%	0	112	37	0	37	37	0	0
pancreas	1	13%	0	0	118	9	0	9	0	18
prostate	1	13%	0	0	129	32	0	0	0	32
skin	7	88%	0	33	166	110	22	88	33	221
spinal cord	2	25%	0	0	112	37	0	224	0	0
spleen	1	13%	0	0	120	60	0	30	10	40
sympathetic ganglion	3	38%	0	89	269	0	0	0	89	0
testis	0	0%	0	8	32	16	0	16	8	40
thymus	1	13%	0	0	82	67	0	149	0	44
urinary	4	50%	0	28	198	99	28	78	21	35
vascular	2	25%	0	0	160	80	0	320	0	0
vesicular gland	1	13%	0	0	0	0	0	457	0	0
AVERAGE (in transcripts per million)			7	18	107	46	14	79	11	46

C. Expression profiles for PAP-predicted NeuroD-responsive target genes

	# genes enriched in location	Percentage of genes with expression in tissue	DCC	DCX	EPHA3	KCND3	NOTCH1	PLXNA2	SYNPR	SEMA6D	UNC5C
adipose tissue	0	0%	0	0	0	0	0	0	0	0	0
blood	1	11%	0	0	0	0	275	0	0	0	0
bone	0	0%	0	0	0	0	28	0	0	0	0
bone marrow	0	0%	0	0	0	0	13	13	0	0	0
brain	7	78%	162	61	55	38	38	74	822	12	207
connective tissue	0	0%	0	0	0	0	0	0	0	0	0
dorsal root ganglion	5	56%	371	371	0	0	148	0	0	520	743
embryonic tissue	1	11%	97	6	12	6	6	24	0	12	42
endocrine	4	44%	0	88	0	0	176	0	58	58	0
epididymis	1	11%	0	0	0	0	304	0	0	0	0
extraembryonic tissue	1	11%	0	0	0	0	0	53	0	0	0
eye	7	78%	254	167	102	59	189	124	264	27	54
female genital	1	11%	24	0	0	0	273	24	0	24	0
gastrointestinal tract	0	0%	49	0	0	8	32	32	0	8	8
head and neck	5	56%	92	92	237	7	69	168	23	30	84
heart	2	22%	0	0	0	0	54	72	0	72	54
inner ear	2	22%	25	0	25	0	126	50	25	25	25
limb	2	22%	0	0	67	0	167	33	0	0	33
liver	1	11%	0	8	0	0	26	8	0	71	26
lung	2	22%	0	19	48	0	67	9	28	38	19
lymph node	1	11%	0	0	63	0	0	0	0	0	0
mammary gland	0	0%	0	0	3	3	74	3	0	6	16
muscle	2	22%	0	0	0	0	37	112	0	37	0
pancreas	6	67%	0	0	218	27	100	145	9	118	118
prostate	1	11%	32	0	64	0	32	0	0	32	32
skin	1	11%	0	0	11	0	11	66	11	0	0
spinal cord	6	67%	487	374	149	37	0	299	374	0	0
spleen	3	33%	0	0	50	160	120	10	10	20	30
sympathetic ganglion	3	33%	89	269	0	0	179	0	0	0	89
testis	0	0%	8	0	0	8	24	8	0	16	40
thymus	2	22%	0	0	14	0	261	14	67	14	0
urinary	1	11%	0	7	7	14	28	7	14	7	35
vascular	2	22%	0	0	0	80	80	0	0	0	400
vesicular gland	0	0%	0	0	0	0	0	0	0	0	0
AVERAGE (in transcripts per million)			50	43	33	13	86	40	50	34	60

D. Expression profiles of PAP-predicted non-responsive target genes

	# genes enriched in tissue	Percentage of genes with expression in tissue	DAB1	EDA	GAS7	ICA1	NRXN1	ODZ1	PHF17	SLIT2	SLIT3
adipose tissue	0	0%	0	0	0	0	0	0	0	0	0
blood	1	11%	0	0	44	11	0	0	44	0	0
bone	1	11%	0	0	0	0	0	0	113	28	28
bone marrow	3	33%	0	13	122	27	238	0	170	115	0
brain	6	67%	155	12	81	120	419	81	38	35	45
connective tissue	1	11%	0	0	0	49	0	0	49	49	0
dorsal root ganglion	3	33%	0	0	74	0	0	74	74	594	0
embryonic tissue	5	56%	24	30	0	42	6	18	231	60	109
endocrine	2	22%	0	0	0	88	58	0	147	29	29
epididymis	2	22%	0	0	0	0	0	0	304	609	0
extraembryonic tissue	2	22%	0	0	0	0	0	0	360	0	160
eye	3	33%	145	5	59	16	54	10	86	194	5
female genital	3	33%	49	24	74	0	0	0	224	24	24
gastrointestinal tract	3	33%	82	0	8	74	0	41	41	24	24
head and neck	6	67%	38	23	7	30	76	30	122	207	46
heart	1	11%	0	0	0	54	0	0	90	0	0
inner ear	0	0%	0	0	25	0	25	0	0	101	25
limb	3	33%	0	335	134	0	0	33	0	33	0
liver	0	0%	0	26	0	8	0	0	44	0	8
lung	2	22%	0	48	28	19	9	0	48	154	19
lymph node	1	11%	0	0	0	0	0	0	127	0	0
mammary gland	1	11%	0	12	3	51	0	0	77	77	41
muscle	2	22%	0	0	0	0	0	0	149	373	0
pancreas	5	56%	0	0	0	72	72	0	164	1093	109
prostate	3	33%	0	129	0	0	0	32	0	0	225
skin	4	44%	0	22	55	77	0	0	121	66	155
spinal cord	5	56%	37	0	112	37	712	0	74	112	112
spleen	2	22%	40	0	70	0	10	10	90	0	0
sympathetic ganglion	4	44%	0	89	0	0	448	0	0	448	89
testis	2	22%	16	8	64	97	32	0	97	56	40
thymus	2	22%	0	7	7	22	14	37	112	0	37
urinary	2	22%	35	7	7	28	0	0	85	21	56
vascular	5	56%	0	160	160	0	0	80	480	0	80
vesicular gland	0	0%	0	0	0	0	0	0	0	0	0
		0%									
AVERAGE (in transcripts per million)			18	28	33	27	64	13	111	132	43

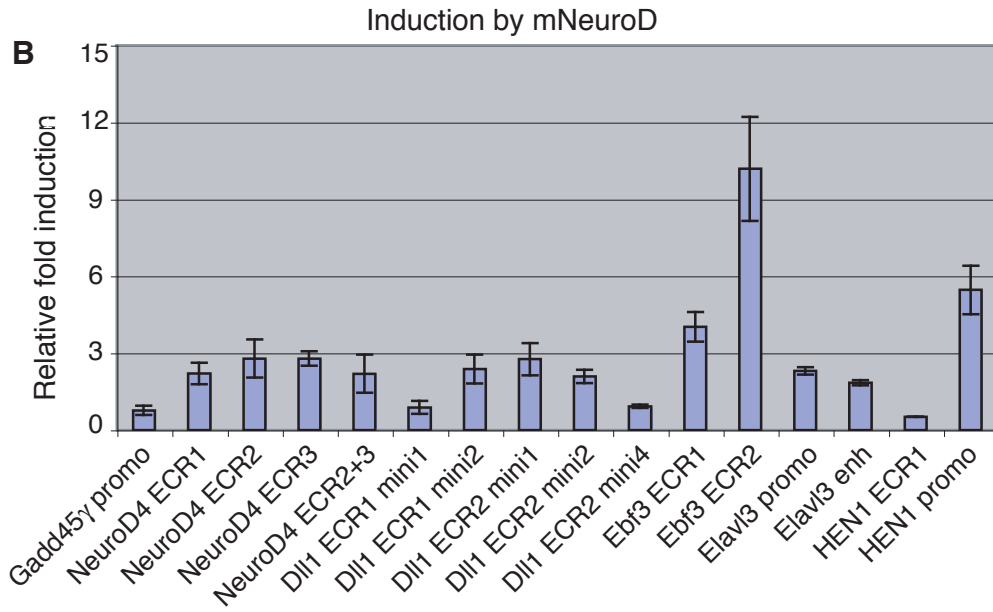
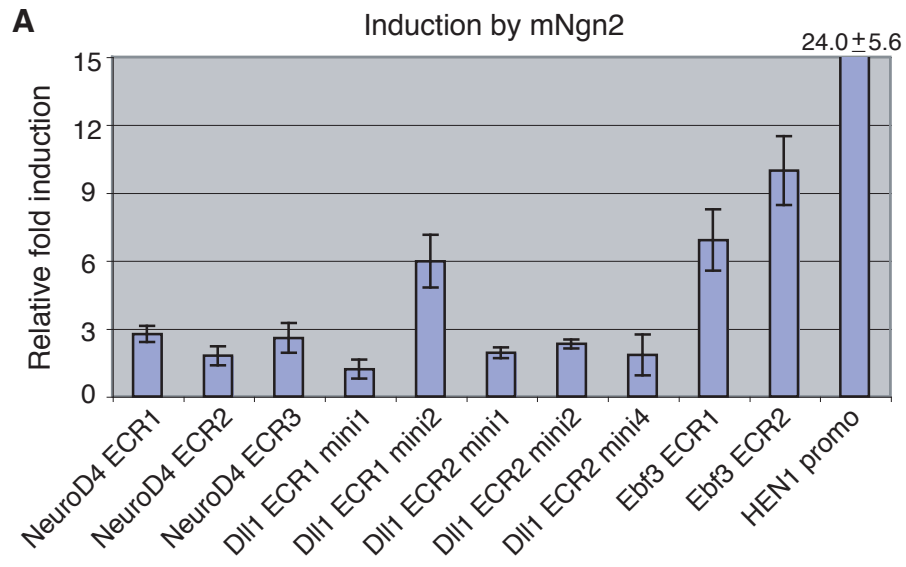
E. Summary of expression profiles for 21 Ngn/NeuroD responsive target genes

	12 experimental targets: # of genes with expression in tissue		9 PAP derived/confirmed targets: # genes enriched in location	Sum genes enriched in location		SUM TOTAL: Percentage of genes enriched in location
adipose tissue	1	adipose tissue	0	1	adipose tissue	5%
blood	0	blood	1	1	blood	5%
bone	1	bone	0	1	bone	5%
bone marrow	0	bone marrow	0	0	bone marrow	0%
brain	7	brain	7	14	brain	67%
connective tissue	3	connective tissue	0	3	connective tissue	14%
dorsal root ganglion	6	dorsal root ganglion	5	11	dorsal root ganglion	52%
embryonic tissue	2	embryonic tissue	1	3	embryonic tissue	14%
endocrine	3	endocrine	4	7	endocrine	33%
epididymis	0	epididymis	1	1	epididymis	5%
extraembryonic tissue	1	extraembryonic tissue	1	2	extraembryonic tissue	10%
eye	7	eye	7	14	eye	67%
female genital	1	female genital	1	2	female genital	10%
gastrointestinal tract	1	gastrointestinal tract	0	1	gastrointestinal tract	5%
head and neck	8	head and neck	5	13	head and neck	62%
heart	2	heart	2	4	heart	19%
inner ear	4	inner ear	2	6	inner ear	29%
limb	0	limb	2	2	limb	10%
liver	1	liver	1	2	liver	10%
lung	3	lung	2	5	lung	24%
lymph node	2	lymph node	1	3	lymph node	14%
mammary gland	1	mammary gland	0	1	mammary gland	5%
muscle	1	muscle	2	3	muscle	14%
pancreas	4	pancreas	6	10	pancreas	48%
prostate	1	prostate	1	2	prostate	10%
skin	2	skin	1	3	skin	14%
spinal cord	7	spinal cord	6	13	spinal cord	62%
spleen	1	spleen	3	4	spleen	19%
sympathetic ganglion	2	sympathetic ganglion	3	5	sympathetic ganglion	24%
testis	1	testis	0	1	testis	5%
thymus	1	thymus	2	3	thymus	14%
urinary	3	urinary	1	4	urinary	19%
vascular	2	vascular	2	4	vascular	19%
vesicular gland	0	vesicular gland	0	0	vesicular gland	0%

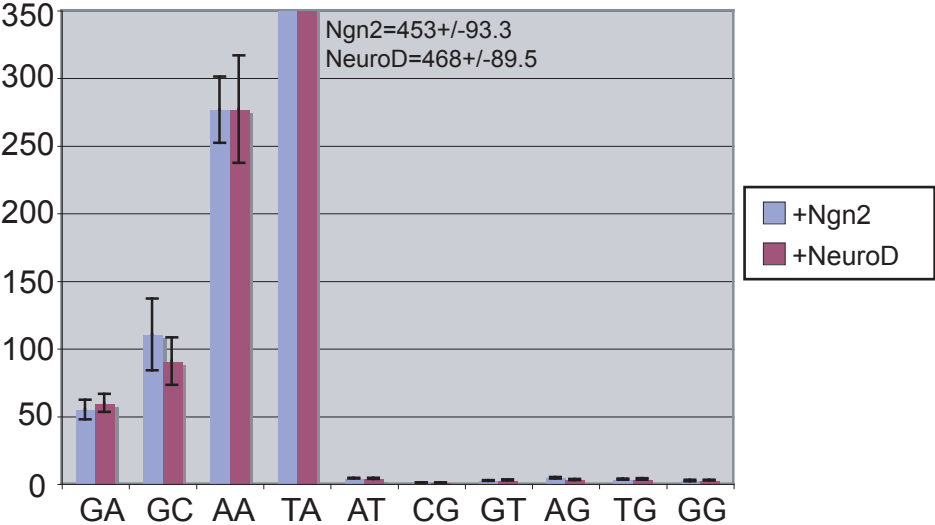
F. Summary of expression profiles for 17 non-responsive target genes

	8 experimental targets: # of genes with expression in tissue		9 PAP derived/confirmed targets: # genes enriched in location	Sum genes enriched in location		SUM TOTAL: Percentage of genes enriched in location
adipose tissue	0	adipose tissue	0	0	adipose tissue	0%
blood	2	blood	1	3	blood	17%
bone	4	bone	1	5	bone	28%
bone marrow	1	bone marrow	3	4	bone marrow	22%
brain	5	brain	6	11	brain	61%
connective tissue	0	connective tissue	1	1	connective tissue	6%
dorsal root ganglion	2	dorsal root ganglion	3	5	dorsal root ganglion	28%
embryonic tissue	2	embryonic tissue	5	7	embryonic tissue	39%
endocrine	3	endocrine	2	5	endocrine	28%
epididymis	0	epididymis	2	2	epididymis	11%
extraembryonic tissue	1	extraembryonic tissue	2	3	extraembryonic tissue	17%
eye	3	eye	3	6	eye	33%
female genital	5	female genital	3	8	female genital	44%
gastrointestinal tract	1	gastrointestinal tract	3	4	gastrointestinal tract	22%
head and neck	1	head and neck	6	7	head and neck	39%
heart	3	heart	1	4	heart	22%
inner ear	3	inner ear	0	3	inner ear	17%
limb	4	limb	3	7	limb	39%
liver	2	liver	0	2	liver	11%
lung	7	lung	2	9	lung	50%
lymph node	1	lymph node	1	2	lymph node	11%
mammary gland	3	mammary gland	1	4	mammary gland	22%
muscle	2	muscle	2	4	muscle	22%
pancreas	1	pancreas	5	6	pancreas	33%
prostate	1	prostate	3	4	prostate	22%
skin	7	skin	4	11	skin	61%
spinal cord	2	spinal cord	5	7	spinal cord	39%
spleen	1	spleen	2	3	spleen	17%
sympathetic ganglion	3	sympathetic ganglion	4	7	sympathetic ganglion	39%
testis	0	testis	2	2	testis	11%
thymus	1	thymus	2	3	thymus	17%
urinary	4	urinary	2	6	urinary	33%
vascular	2	vascular	5	7	vascular	39%
vesicular gland	1	vesicular gland	0	1	vesicular gland	6%

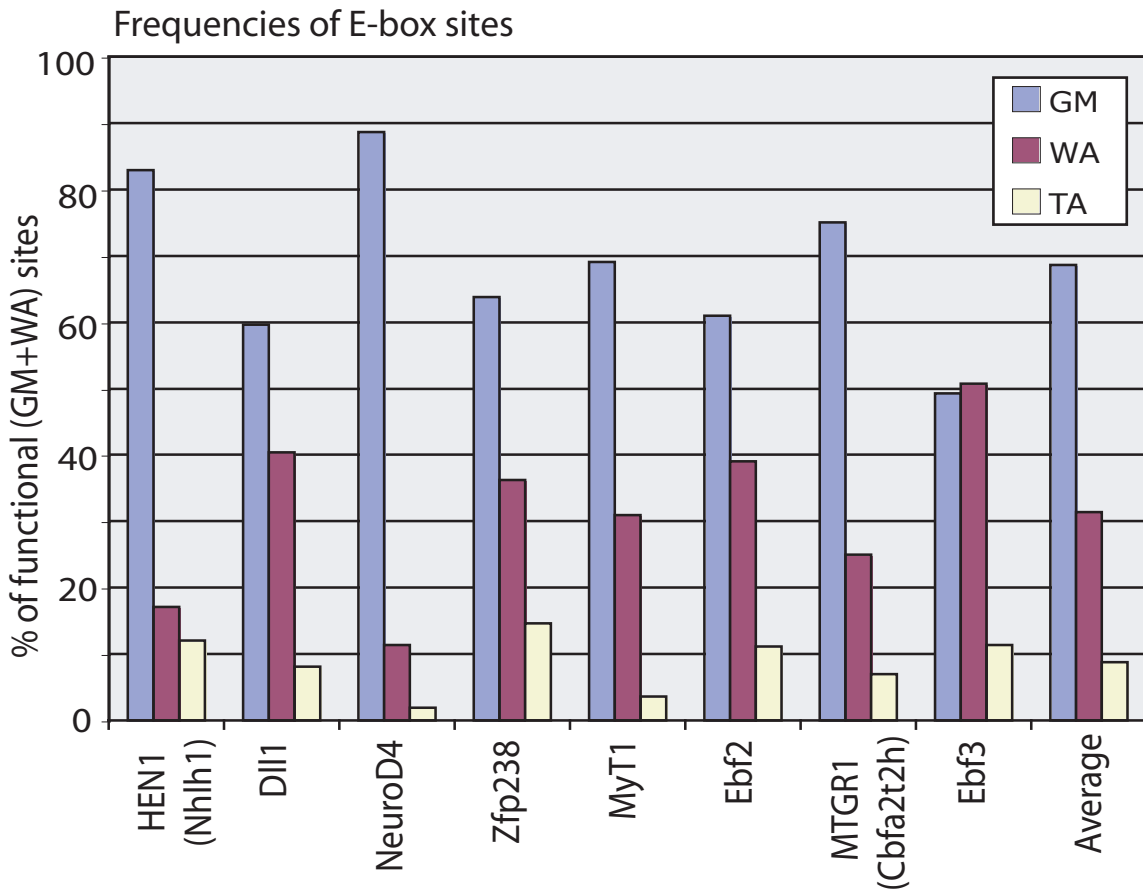
Seo_Supplemental Fig. S5



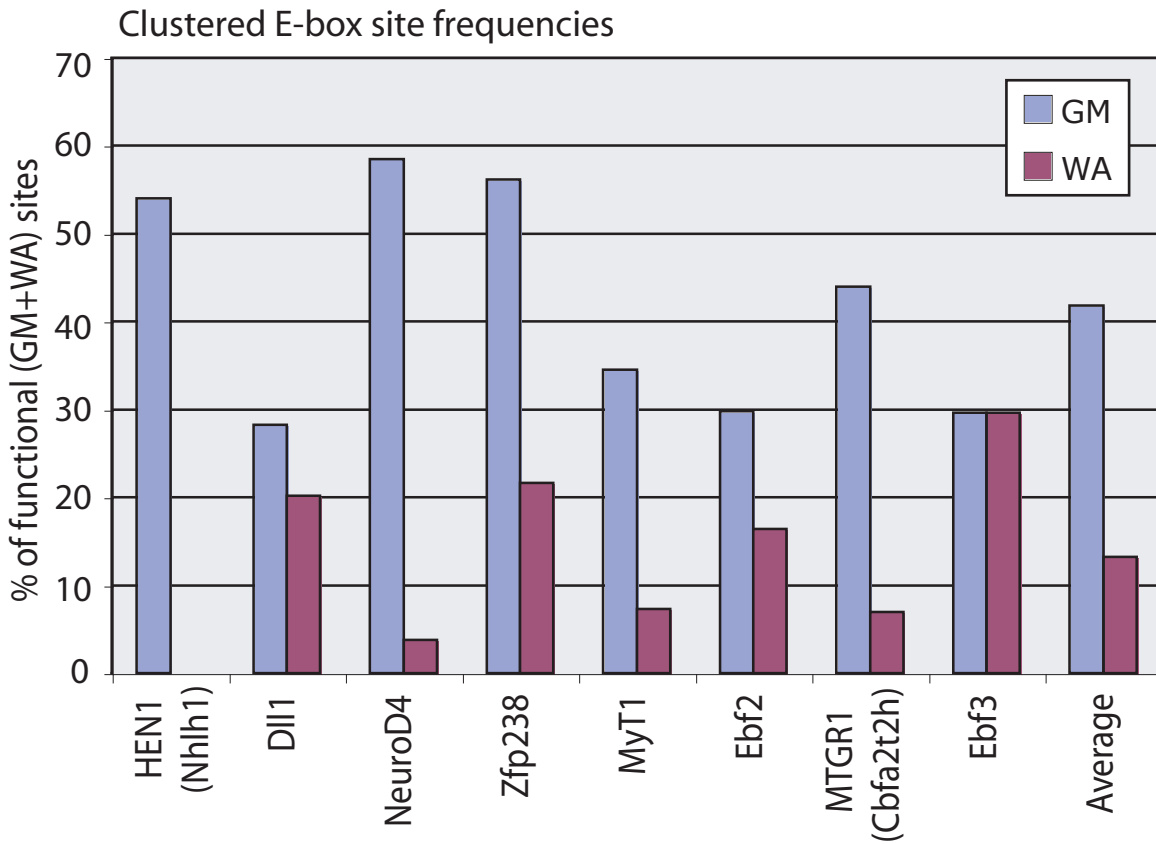
Seo_Supplemental Fig. S6.



A



B



Seo_Supplemental Figure S7.

(C) Quantitation of E-box frequencies. Raw numerical data and calculations of frequency of conserved E-boxes seen in pairwise comparisons of each locus between species (with mouse as the base organism).

DII1							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	18	11	1	12	6	3	1
opossum	13	6	7	10	6	4	2
dog	11	6	5	7	4	3	2
chick	7	2	0	3	0	1	2
frog	6	3	1	5	2	1	1
fugu	0	0	0	0	0	0	0
zebrafish	4	0	1	3	2	0	0
Totals	59	28	15	40	20	12	8
Total functional (GM+WA) E-box=		99					

DII1 Summary of E-box frequency	
% GM= (as %GM+WA)	60
% WA= (as %GM+WA)	40
% GM clustered= (as %GM+WA)	28
% WA clustered= (as %GM+WA)	20
% GM in WA cluster= (as % GM)	25
% WA in GM cluster= (as % WA)	30
(% TA) (as %GM+WA)	08

Ebf2							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	69	35	10	41	18	13	10
opossum	29	13	3	22	9	5	8
dog	62	30	15	42	17	13	11
chick	21	11	3	12	4	3	4
frog	4	2	1	2	2	0	1
fugu	1	0	0	0	0	0	0
zebrafish	0	0	0	0	0	0	0
Totals	186	91	32	119	50	34	34
Total functional (GM+WA) E-box=		305					

Ebf2 Summary of E-box frequency	
% GM= (as %GM+WA)	61
% WA= (as %GM+WA)	39
% GM clustered= (as %GM+WA)	30
% WA clustered= (as %GM+WA)	16
% GM in WA cluster= (as % GM)	17
% WA in GM cluster= (as % WA)	29
(% TA) (as %GM+WA)	11

Ebf3							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	64	38	23	68	44	23	11
opossum (1)	34	16	3	34	22	1	10
dog	55	31	33	59	33	10	11
chick	39	27	18	46	30	15	9
frog	35	22	9	34	20	9	8
fugu	11	7	1	11	2	4	7
zebrafish	17	12	2	10	2	4	3
Totals	255	153	89	262	153	66	59
Total functional (GM+WA) E-box=		517					

Ebf3 Summary of E-box frequency	
% GM= (as %GM+WA)	49
% WA= (as %GM+WA)	51
% GM clustered= (as %GM+WA)	30
% WA clustered= (as %GM+WA)	30
% GM in WA cluster= (as % GM)	35
% WA in GM cluster= (as % WA)	25
(% TA) (as %GM+WA)	11

MTGR1/Cbfa2t2h							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	29	16	2	13	5	4	3
opossum	8	6	0	1	0	1	1
dog	33	20	2	10	2	4	2
chick	4	2	0	1	0	0	1
frog	0	0	0	0	0	0	0
fugu	0	0	0	0	0	0	0
zebrafish	1	0	0	0	0	0	0
Totals	75	44	4	25	7	9	7
Total functional (GM+WA) E-box=		100					

MTGR1/Cbfa2t2h Summary of E-box frequency	
% GM= (as %GM+WA)	75
% WA= (as %GM+WA)	25
% GM clustered= (as %GM+WA)	44
% WA clustered= (as %GM+WA)	07
% GM in WA cluster= (as % GM)	05
% WA in GM cluster= (as % WA)	36
(% TA) (as %GM+WA)	07

MyT1							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	18	7	3	10	4	3	1
opossum	4	2	0	1	0	0	0
dog	13	7	0	5	0	2	1
chick	3	3	0	1	0	0	0
frog	0	0	0	0	0	0	0
fugu	0	0	0	0	0	0	0
zebrafish	0	0	0	0	0	0	0
Sums	38	19	3	17	4	5	2
Total functional (GM+WA) E-box=		55					

MyT1 Summary of E-box frequency	
% GM= (as %GM+WA)	69
% WA= (as %GM+WA)	31
% GM clustered= (as %GM+WA)	35
% WA clustered= (as %GM+WA)	07
% GM in WA cluster= (as % GM)	08
% WA in GM cluster= (as % WA)	29
(% TA) (as %GM+WA)	04

NeuroD4							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	21	14	2	4	2	3	1
opossum	8	3	0	0	0	0	0
dog	18	14	0	2	0	1	0
chick (2)	did not align with mouse						
frog (2)	did not align with mouse						
fugu (2)	did not align with mouse						
zebrafish (2)	did not align with mouse						
Totals	47	31	2	6	2	4	1
Total functional (GM+WA) E-box=		53					

NeuroD4 Summary of E-box frequency	
% GM= (as %GM+WA)	89
% WA= (as %GM+WA)	11
% GM clustered= (as %GM+WA)	58
% WA clustered= (as %GM+WA)	04
% GM in WA cluster= (as % GM)	04
% WA in GM cluster= (as % WA)	67
(% TA) (as %GM+WA)	02

HEN1/Nhlh1							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human (2)	10	8	0	3	0	0	2
opossum	9	7	0	2	0	0	2
dog	13	7	0	2	0	0	1
chick	2	0	0	0	0	0	0
frog	0	0	0	0	0	0	0
fugu	0	0	0	0	0	0	0
zebrafish	0	0	0	0	0	0	0
Totals	34	22	0	7	0	0	5
Total functional (GM+WA) E-box=		41					

HEN1/Nhlh1 Summary of E-box frequency	
% GM= (as %GM+WA)	83
% WA= (as %GM+WA)	17
% GM clustered= (as %GM+WA)	54
% WA clustered= (as %GM+WA)	00
% GM in WA cluster= (as % GM)	00
% WA in GM cluster= (as % WA)	00
(% TA) (as %GM+WA)	12

Zfp238							
mouse base	GM-type total # sites	GM-type # of clustered sites	GM sites in WA cluster	WA-type total # sites	WA-type # clustered sites	WA sites in GM cluster	TA-type total # sites
human	32	27	13	21	17	8	6
opossum	21	18	4	12	6	6	5
dog	30	26	5	15	4	8	5
chick	11	10	0	8	4	4	3
frog	5	5	5	2	2	2	2
fugu	5	5	5	2	2	2	2
zebrafish	5	5	5	2	2	2	2
Totals	109	96	37	62	37	32	25
Total functional (GM+WA) E-box=		171					

Zfp238 Summary of E-box frequency	
% GM= (as %GM+WA)	64
% WA= (as %GM+WA)	36
% GM clustered= (as %GM+WA)	56
% WA clustered= (as %GM+WA)	22
% GM in WA cluster= (as % GM)	34
% WA in GM cluster= (as % WA)	52
(% TA) (as %GM+WA)	15

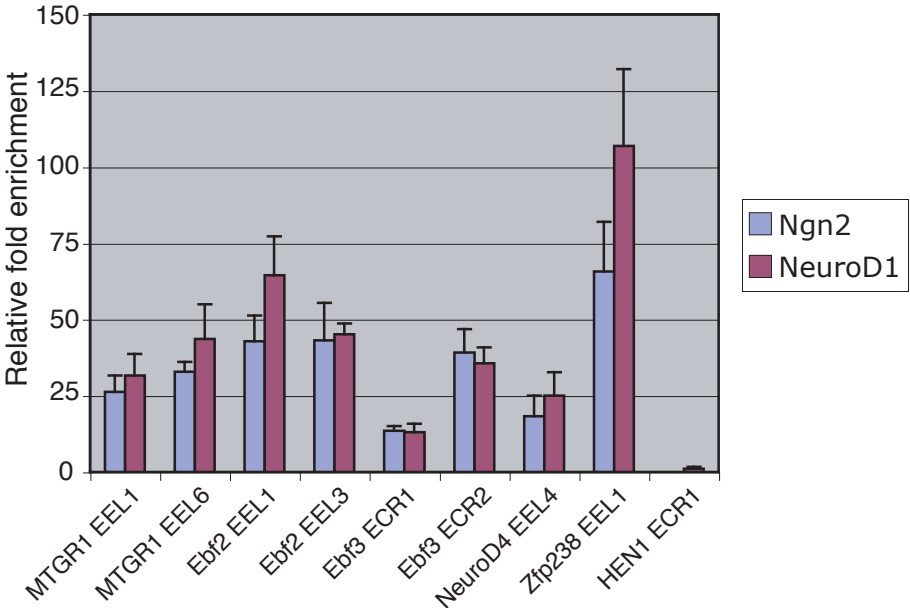
(1) For Ebf3 locus mouse/ opossum syntenic alignment, 4 partial alignments spanned the region of the locus including 50kb 5', ORF+introns, and 50kb 3' sequences. Sites were summed across these 4 alignments

(2) For NeuroD locus, mouse locus did not have sufficient synteny with chicken, frog, fugu or zebrafish loci for rVista analysis to be performed

(3) For Nhlh1 locus human/mouse alignment, only 25 kb of 5' and 3' sequences was analyzed as ECR browser could not perform Rvista site analysis with longer alignments(4)

Zfp238 locus human/mouse was analyzed as two 5' and 3' alignment reads as Rvista site analysis could not be done on a single 100+ kb alignment

Seo_Supplemental Fig. S8.



Seo_Supplemental Figure_S9. Computational genome-wide prediction of Ngn/NeuroD target genes

A. Ngn/NeuroD target genes predicted by PAP (347 genes)

Intersection			p-value	
Ngn/NeuroD matrix			h/m	h/m/r
Gene ID	Symbol	Gene Name		
90	ACVR1	ACTIVIN A RECEPTOR, TYPE I	0	0
238	ALK	ANAPLASTIC LYMPHOMA KINASE (KI-1)	0	0
23245	ASTN2	ASTROTACTIN 2	0	0
463	ATBF1	AT-BINDING TRANSCRIPTION FACTOR 1	0	0
84913	ATOH8	ATONAL HOMOLOG 8 (DROSOPHILA)	0	0
54828	BCAS3	BREAST CARCINOMA AMPLIFIED SEQUENCE 3	0	0
64919	BCL11B	B-CELL CLL/LYMPHOMA 11B (ZINC FINGER PROTEIN)	0	0
54796	BNC2	BASONUCLIN 2	0	0
8945	BTRC	BETA-TRANSDUCIN REPEAT CONTAINING	0	0
27291	C10orf28	CHROMOSOME 10 OPEN READING FRAME 28	0	0
55799	CACNA2D3	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, ALPHA 2/DELTA 3 SUBUNIT	0	0
9139	CBFA2T2	CORE-BINDING FACTOR, RUNT DOMAIN, ALPHA SUBUNIT 2; TRANSLOCATED TO, 2	0	0
863	CBFA2T3	CORE-BINDING FACTOR, RUNT DOMAIN, ALPHA SUBUNIT 2; TRANSLOCATED TO, 3	0	0
55118	CRTAC1	CARTILAGE ACIDIC PROTEIN 1	0	0
51523	CXXC5	CXXC FINGER 5	0	0
1600	DAB1	DISABLED HOMOLOG 1 (DROSOPHILA)	0	0
1641	DCX	DOUBLECORTEX; LISSENCEPHALY, X-LINKED (DOUBLECORTIN)	0	0
1740	DLG2	DISCS, LARGE HOMOLOG 2, CHAPSYN-110 (DROSOPHILA)	0	0
54567	DLL4	DELTA-LIKE 4 (DROSOPHILA)	0	0
1780	DYNC111	DYNEIN, CYTOPLASMIC 1, INTERMEDIATE CHAIN 1	0	0
2035	EPB41	ERYTHROCYTE MEMBRANE PROTEIN BAND 4.1 (ELLIPTOCYTOSIS 1, RH-LINKED)	0	0
2048	EPHB2	EPH RECEPTOR B2	0	0
2104	ESRRG	ESTROGEN-RELATED RECEPTOR GAMMA	0	0
8326	FZD9	FRIZZLED HOMOLOG 9 (DROSOPHILA)	0	0
2719	GPC3	GLYPICAN 3	0	0
2904	GRIN2B	GLUTAMATE RECEPTOR, IONOTROPIC, N-METHYL D-ASPARTATE 2B	0	0
3459	IFNGR1	INTERFERON GAMMA RECEPTOR 1	0	0
26280	IL1RAPL2	INTERLEUKIN 1 RECEPTOR ACCESSORY PROTEIN-LIKE 2	0	0
3752	KCND3	POTASSIUM VOLTAGE-GATED CHANNEL, SHAL-RELATED SUBFAMILY, MEMBER 3	0	0
3784	KCNQ1	POTASSIUM VOLTAGE-GATED CHANNEL, KQT-LIKE SUBFAMILY, MEMBER 1	0	0
57710	KIAA1614	KIAA1614	0	0
79817	MOBK12B	MOB1, MPS ONE BINDER KINASE ACTIVATOR-LIKE 2B (YEAST)	0	0
4684	NCAM1	NEURAL CELL ADHESION MOLECULE 1	0	0
4784	NFIX	NUCLEAR FACTOR I/X (CCAAT-BINDING TRANSCRIPTION FACTOR)	0	0
4808	NHLH2	NESCIENT HELIX LOOP HELIX 2	0	0
64067	NPAS3	NEURONAL PAS DOMAIN PROTEIN 3	0	0
8828	NRP2	NEUROFILIN 2	0	0
9378	NRXN1	NEUREXIN 1	0	0
57451	ODZ2	ODZ, ODD OZ/TEN-M HOMOLOG 2 (DROSOPHILA)	0	0
26011	ODZ4	ODZ, ODD OZ/TEN-M HOMOLOG 4 (DROSOPHILA)	0	0
5087	PBX1	PRE-B-CELL LEUKEMIA TRANSCRIPTION FACTOR 1	0	0
29951	PDZRN4	PDZ DOMAIN CONTAINING RING FINGER 4	0	0
389158	PLSCR5	PHOSPHOLIPID SCRAMBLASE FAMILY, MEMBER 5	0	0
5362	PLXNA2	PLEXIN A2	0	0
5521	PPP2R2B	PROTEIN PHOSPHATASE 2 (FORMERLY 2A), REGULATORY SUBUNIT B (PR 52), BETA ISOFORM	0	0
5579	PRKCB1	PROTEIN KINASE C, BETA 1	0	0
5629	PROX1	PROSPERO-RELATED HOMEODOMAIN 1	0	0
5690	PSMB2	PROTEASOME (PROSOME, MACROPAIN) SUBUNIT, BETA TYPE, 2	0	0
11122	PTPRT	PROTEIN TYROSINE PHOSPHATASE, RECEPTOR TYPE, T	0	0
64901	RANBP17	RAN BINDING PROTEIN 17	0	0
5914	RARA	RETINOIC ACID RECEPTOR, ALPHA	0	0
9904	RBM19	RNA BINDING MOTIF PROTEIN 19	0	0
6091	ROBO1	ROUNDABOUT, AXON GUIDANCE RECEPTOR, HOMOLOG 1 (DROSOPHILA)	0	0
6095	RORA	RAR-RELATED ORPHAN RECEPTOR A	0	0
862	RUNX1T1	RUNT-RELATED TRANSCRIPTION FACTOR 1; TRANSLOCATED TO, 1 (CYCLIN D-RELATED)	0	0
23328	SASH1	SAM AND SH3 DOMAIN CONTAINING 1	0	0
8631	SCAP1	SRC FAMILY ASSOCIATED PHOSPHOPROTEIN 1	0	0
10806	SDCCAG8	SEROLOGICALLY DEFINED COLON CANCER ANTIGEN 8	0	0
10371	SEMA3A	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3A	0	0
89874	SLC25A21	OXODICARBOXYLATE CARRIER	0	0
6586	SLIT3	SLIT HOMOLOG 3 (DROSOPHILA)	0	0
23648	SSBP3	SINGLE STRANDED DNA BINDING PROTEIN 3	0	0
9705	ST18	SUPPRESSION OF TUMORIGENICITY 18 (BREAST CARCINOMA) (ZINC FINGER PROTEIN)	0	0
8224	SYN3	SYNAPSIN III	0	0
6925	TCF4	TRANSCRIPTION FACTOR 4	0	0
10867	TSPAN9	TETRASPANIN 9	0	0
26137	ZBTB20	ZINC FINGER AND BTB DOMAIN CONTAINING 20	0	0
9839	ZFH1B	ZINC FINGER HOMEODOMAIN 1B	0	0
23090	ZNF423	ZINC FINGER PROTEIN 423	0	0
57616	ZNF537	ZINC FINGER PROTEIN 537	0	0
27044		STAPHYLOCOCCAL NUCLEASE DOMAIN CONTAINING 1	0	0
50863		NEUROTRIMIN	0	0
121256		TRANSMEMBRANE PROTEIN 132D	0	0

55084		HYPOTHETICAL PROTEIN FLJ10159	0	0
54715		ATAXIN 2-BINDING PROTEIN 1	0	0
51474		LIM DOMAIN AND ACTIN BINDING 1	0	0
54622	ARL15	ADP-RIBOSYLATION FACTOR-LIKE 15	0.0001	0
460	ASTN	ASTROTACTIN	0.0001	0
9254	CACNA2D2	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, ALPHA 2/DELTA SUBUNIT 2	0.0001	0
50515	CHST11	CARBOHYDRATE (CHONDROITIN 4) SULFOTRANSFERASE 11	0.0001	0
1630	DCC	DELETED IN COLORECTAL CARCINOMA	0.0001	0
4916	NTRK3	NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 3	0.0001	0
5076	PAX2	PAIRED BOX GENE 2	0.0001	0
10472	ZNF238	ZINC FINGER PROTEIN 238	0.0001	0
441115		NA	0.0001	0
29789		GTP-BINDING PROTEIN PTD004	0.0001	0
832	CAPZB	CAPPING PROTEIN (ACTIN FILAMENT) MUSCLE Z-LINE, BETA	0.0002	0
54927	CHCHD3	COILED-COIL-HELIX-COILED-COIL-HELIX DOMAIN CONTAINING 3	0.0002	0
9442	CRSP8	COFACTOR REQUIRED FOR SP1 TRANSCRIPTIONAL ACTIVATION, SUBUNIT 8, 34KDA	0.0002	0
85407	NKD1	NAKED CUTICLE HOMOLOG 1 (DROSOPHILA)	0.0002	0
8829	NRP1	NEUROFILIN 1	0.0002	0
9353	SLIT2	SLIT HOMOLOG 2 (DROSOPHILA)	0.0002	0
7903	ST8SIA4	ST8 ALPHA-N-ACETYL-NEURAMINIDE ALPHA-2,8-SIALYLTRANSFERASE 4	0.0002	0
90390	THRAP6	THYROID HORMONE RECEPTOR ASSOCIATED PROTEIN 6	0.0002	0
60528	ELAC2	HYPOTHETICAL PROTEIN FLJ10530	0.0004	0
2042	EPHA3	EPH RECEPTOR A3	0.0004	0
2047	EPHB1	EPH RECEPTOR B1	0.0004	0
2066	ERBB4	V-ERB-A ERYTHROBLASTIC LEUKEMIA VIRAL ONCOGENE HOMOLOG 4 (AVIAN)	0.0005	0
54901	CDKAL1	CDK5 REGULATORY SUBUNIT ASSOCIATED PROTEIN 1-LIKE 1	0.0007	0
9935	MAFB	V-MAF MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE HOMOLOG B (AVIAN)	0.0008	0
4692	NDN	NECDIN HOMOLOG (MOUSE)	0.0011	0
168667	BMPER	BMP BINDING ENDOTHELIAL REGULATOR	0.0012	0
64411	CENTD3	CENTAURIN, DELTA 3	0.0013	0
1523	CUTL1	CUT-LIKE 1, CCAAT DISPLACEMENT PROTEIN (DROSOPHILA)	0.0014	0
59269	HIVEP3	HUMAN IMMUNODEFICIENCY VIRUS TYPE I ENHANCER BINDING PROTEIN 3	0.0024	0
9420	CYP7B1	CYTOCHROME P450, FAMILY 7, SUBFAMILY B, POLYPEPTIDE 1	0	0.0001
55137	FIGN	FIDGETIN	0	0.0001
3479	IGF1	INSULIN-LIKE GROWTH FACTOR 1 (SOMATOMEDIN C)	0	0.0001
203447	NRK	NIK RELATED KINASE	0	0.0001
10891	PPARGC1A	PEROXISOME PROLIFERATIVE ACTIVATED RECEPTOR, GAMMA, COACTIVATOR 1, ALPHA	0	0.0001
26508	HEYL	HAIRY/ENHANCER-OF-SPLIT RELATED WITH YRPW MOTIF-LIKE	0.0001	0.0001
4045	LSAMP	LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN	0.0001	0.0001
4654	MYOD1	MYOGENIC DIFFERENTIATION 1	0.0001	0.0001
9423	NTN1	NETRIN 1	0.0001	0.0001
55697	VAC14	TAX1 (HUMAN T-CELL LEUKEMIA VIRUS TYPE I) BINDING PROTEIN 2	0.0001	0.0001
23414	ZFPM2	ZINC FINGER PROTEIN, MULTITYPE 2	0.0001	0.0001
79191	IRX3	IROQUOIS HOMEODOMAIN PROTEIN 3	0.0002	0.0001
132204	SYNPR	SYNAPOPORIN	0.0002	0.0001
5566	PRKACA	PROTEIN KINASE, CAMP-DEPENDENT, CATALYTIC, ALPHA	0.0003	0.0001
171425	CLYBL	CITRATE LYASE BETA LIKE	0.0004	0.0001
10194	SDCCAG33	SEROLOGICALLY DEFINED COLON CANCER ANTIGEN 33	0.0006	0.0001
1804	DPP6	DIPEPTIDYL-PEPTIDASE 6	0.0008	0.0001
28992		LRP16 PROTEIN	0.001	0.0001
8038	ADAM12	ADAM METALLOPEPTIDASE DOMAIN 12 (MELTRIN ALPHA)	0.0017	0.0001
1910	EDNRB	ENDOTHELIN RECEPTOR TYPE B	0.0028	0.0001
89796	NAV1	HYPOTHETICAL PROTEIN MGC14961	0	0.0002
56245	C21orf62	CHROMOSOME 21 OPEN READING FRAME 62	0.0001	0.0002
84303	CHCHD6	COILED-COIL-HELIX-COILED-COIL-HELIX DOMAIN CONTAINING 6	0.0001	0.0002
84662	GLIS2	GLIS FAMILY ZINC FINGER 2	0.0001	0.0002
116449		MAST CELL IMMUNORECEPTOR SIGNAL TRANSDUCER	0.0001	0.0002
3779	KCNMB1	POTASSIUM LARGE CONDUCTANCE CALCIUM-ACTIVATED CHANNEL, SUBFAMILY M, BETA MEMBER 1	0.0002	0.0002
64374	SIL1	SIL1 HOMOLOG, ENDOPLASMIC RETICULUM CHAPERONE (S. CEREVISIAE)	0.0002	0.0002
80031	SEMA6D	KIAA1479 PROTEIN	0.0003	0.0002
80731		KIAA1679 PROTEIN	0.0005	0.0002
8618	CADPS	CA2+-DEPENDENT SECRETION ACTIVATOR	0.0008	0.0002
9627	SNCAIP	SYNUCLEIN, ALPHA INTERACTING PROTEIN (SYNPILIN)	0.0008	0.0002
4010	LMX1B	LIM HOMEODOMAIN TRANSCRIPTION FACTOR 1, BETA	0.0009	0.0002
3757	KCNH2	POTASSIUM VOLTAGE-GATED CHANNEL, SUBFAMILY H (EAG-RELATED), MEMBER 2	0.001	0.0002
214	ALCAM	ACTIVATED LEUKOCYTE CELL ADHESION MOLECULE	0.0012	0.0002
4853	NOTCH2	NOTCH HOMOLOG 2 (DROSOPHILA)	0.0014	0.0002
7150	TOP1	TOPOISOMERASE (DNA) I	0.0014	0.0002
1307	COL16A1	COLLAGEN, TYPE XVI, ALPHA 1	0.0016	0.0002
6468	FBXW4	F-BOX AND WD-40 DOMAIN PROTEIN 4	0.0016	0.0002
4851	NOTCH1	NOTCH HOMOLOG 1, TRANSLLOCATION-ASSOCIATED (DROSOPHILA)	0.002	0.0002
223075		HYPOTHETICAL PROTEIN LOC223075	0.002	0.0002
64131	XYLT1	XYLOSYLTRANSFERASE I	0.0031	0.0002
4212	MEIS2	MEIS1, MYELOID ECOTROPIC VIRAL INTEGRATION SITE 1 HOMOLOG 2 (MOUSE)	0.0054	0.0002
90134	KCNH7	POTASSIUM VOLTAGE-GATED CHANNEL, SUBFAMILY H (EAG-RELATED), MEMBER 7	0	0.0003
151531	UPP2	URIDINE PHOSPHORYLASE 2	0	0.0003
1013	CDH15	CADHERIN 15, M-CADHERIN (MYOTUBULE)	0.0001	0.0003
143458	LDLRAD3	LOW DENSITY LIPOPROTEIN RECEPTOR CLASS A DOMAIN CONTAINING 3	0.0001	0.0003
26018	LRIG1	LEUCINE-RICH REPEATS AND IMMUNOGLOBULIN-LIKE DOMAINS 1	0.0001	0.0003
10451	VAV3	VAV 3 ONCOGENE	0.0002	0.0003
83989	C5orf21	CHROMOSOME 5 OPEN READING FRAME 21	0.0006	0.0003

23769	FLRT1	FIBRONECTIN LEUCINE RICH TRANSMEMBRANE PROTEIN 1	0.0008	0.0003
5954	RCN1	RETICULOCALBIN 1, EF-HAND CALCIUM BINDING DOMAIN	0.0008	0.0003
4919	ROR1	RECEPTOR TYROSINE KINASE-LIKE ORPHAN RECEPTOR 1	0.0013	0.0003
2674	GFRA1	GNDF FAMILY RECEPTOR ALPHA 1	0.0016	0.0003
339855	KY	KYPHOSCOLIOSIS PEPTIDASE	0.0017	0.0003
64328	XPO4	EXPORTIN 4	0.0019	0.0003
4880	NPPC	NATRIURETIC PEPTIDE PRECURSOR C	0.003	0.0003
4651	MYO10	MYOSIN X	0.0045	0.0003
118	ADD1	ADDUCIN 1 (ALPHA)	0	0.0004
4131	MAP1B	MICROTUBULE-ASSOCIATED PROTEIN 1B	0	0.0004
146760	RTN4RL1	RETICULON 4 RECEPTOR-LIKE 1	0	0.0004
3397	ID1	INHIBITOR OF DNA BINDING 1, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN	0.0001	0.0004
132	ADK	ADENOSINE KINASE	0.0003	0.0004
255275		SIMILAR TO MYELOID-ASSOCIATED DIFFERENTIATION MARKER	0.0003	0.0004
10783	NEK6	NIMA (NEVER IN MITOSIS GENE A)-RELATED KINASE 6	0.0004	0.0004
91522	COL23A1	COLLAGEN, TYPE XXIII, ALPHA 1	0.001	0.0004
7067	THRA	THYROID HORMONE RECEPTOR, ALPHA (ERYTHROBLASTIC LEUKEMIA VIRAL (V-ERB-A) ONCOGENE HOMOLOG, AVIAN)	0.0016	0.0004
51646	YPEL5	YIPPEE-LIKE 5 (DROSOPHILA)	0.0018	0.0004
9844	ELMO1	ENGULFMENT AND CELL MOTILITY 1	0.002	0.0004
93664	CADPS2	HYPOTHETICAL PROTEIN FLJ20761	0.0029	0.0004
4761	NEUROD2	NEUROGENIC DIFFERENTIATION 2	0.0001	0.0005
30820	KCNIP1	KV CHANNEL INTERACTING PROTEIN 1	0.0004	0.0005
5793	PTPRG	PROTEIN TYROSINE PHOSPHATASE, RECEPTOR TYPE, G	0.0005	0.0005
79960	PHF17	PHD FINGER PROTEIN 17	0.0006	0.0005
1785	DNM2	DYNAMIN 2	0.0013	0.0005
10435	CDC42EP2	CDC42 EFFECTOR PROTEIN (RHO GTPASE BINDING) 2	0.0017	0.0005
64788		HYPOTHETICAL PROTEIN FLJ12681	0.0021	0.0005
4978	OPCML	OPIOID BINDING PROTEIN/CELL ADHESION MOLECULE-LIKE	0.0023	0.0005
1996	ELAVL4	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 4 (HU ANTIGEN D)	0.0001	0.0006
2892	GRIA3	GLUTAMATE RECEPTOR, IONOTROPIC, AMPA 3	0.0014	0.0006
6641	SNTB1	SYNTROPHIN, BETA 1 (DYSTROPHIN-ASSOCIATED PROTEIN A1, 59KDA, BASIC COMPONENT 1)	0.0015	0.0006
116987	CENTG2	CENTAURIN, GAMMA 2	0.0024	0.0006
9242	MSC	MUSCULIN (ACTIVATED B-CELL FACTOR-1)	0.0028	0.0006
4807	NHLH1	NESCIENT HELIX LOOP HELIX 1	0.0029	0.0006
182	JAG1	JAGGED 1 (ALAGILLE SYNDROME)	0.0036	0.0006
6870	TACR3	TACHYKININ RECEPTOR 3	0.0054	0.0006
9782	MATR3	MATRIN 3	0.0076	0.0006
10000	AKT3	V-AKT MURINE THYMOMA VIRAL ONCOGENE HOMOLOG 3 (PROTEIN KINASE B, GAMMA)	0.0001	0.0007
26235	FBXL4	F-BOX AND LEUCINE-RICH REPEAT PROTEIN 4	0.0001	0.0007
84529	C15orf41	CHROMOSOME 15 OPEN READING FRAME 41	0.0011	0.0007
445	ASS	ARGININOSUCCINATE SYNTHETASE	0.0013	0.0007
388336		FLJ45455 PROTEIN	0.0031	0.0007
8633	UNC5C	UNC-5 HOMOLOG C (C. ELEGANS)	0.0032	0.0007
9482	STX8	SYNTAXIN 8	0.005	0.0007
57419	SLC24A3	SOLUTE CARRIER FAMILY 24 (SODIUM/POTASSIUM/CALCIUM EXCHANGER), MEMBER 3	0.0002	0.0008
2649	NR6A1	NUCLEAR RECEPTOR SUBFAMILY 6, GROUP A, MEMBER 1	0.0004	0.0008
148979	GLIS1	GLIS FAMILY ZINC FINGER 1	0.0005	0.0008
55504	TNFRSF19	TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 19	0.0005	0.0008
1496	CTNNA2	CATENIN (CADHERIN-ASSOCIATED PROTEIN), ALPHA 2	0.0008	0.0008
4255	MGMT	O-6-METHYLGUANINE-DNA METHYLTRANSFERASE	0.0016	0.0008
7482	WNT2B	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 2B	0.002	0.0008
348980	HCN1	HYPERPOLARIZATION ACTIVATED CYCLIC NUCLEOTIDE-GATED POTASSIUM CHANNEL 1	0.0021	0.0008
9543	PUNC	PUTATIVE NEURONAL CELL ADHESION MOLECULE	0.0023	0.0008
10620	ARID3B	AT RICH INTERACTIVE DOMAIN 3B (BRIGHT- LIKE)	0.0059	0.0008
10973	ASCC3	ACTIVATING SIGNAL COINTEGRATOR 1 COMPLEX SUBUNIT 3	0.0002	0.0009
10404		PLASMA GLUTAMATE CARBOXYPEPTIDASE	0.0012	0.0009
57529	RGAG1	RETROTRANSPOSON GAG DOMAIN CONTAINING 1	0.0015	0.0009
90161	HS6ST2	HEPARAN SULFATE 6-O-SULFOTRANSFERASE 2	0.0017	0.0009
3399	ID3	INHIBITOR OF DNA BINDING 3, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN	0.0033	0.0009
3605	IL17	INTERLEUKIN 17 (CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED SERINE ESTERASE 8)	0.0035	0.0009
28514	DLL1	DELTA-LIKE 1 (DROSOPHILA)	0.0048	0.0009
9779	TBC1D5	TBC1 DOMAIN FAMILY, MEMBER 5	0	0.001
57178	RAI17	RETINOIC ACID INDUCED 17	0.0002	0.001
160335	TMTC2	TRANSMEMBRANE AND TETRATRICOPEPTIDE REPEAT CONTAINING 2	0.0092	0.001
2900	GRIK4	GLUTAMATE RECEPTOR, IONOTROPIC, KAINATE 4	0	0.0011
23335	WDR7	WD REPEAT DOMAIN 7	0.0005	0.0011
9450	LY86	LYMPHOCYTE ANTIGEN 86	0.001	0.0011
23037	PDZD2	PDZ DOMAIN CONTAINING 2	0.0052	0.0011
4088	SMAD3	SMAD, MOTHERS AGAINST DPP HOMOLOG 3 (DROSOPHILA)	0.0002	0.0012
139263		NA	0.0018	0.0012
1879	EBF	EARLY B-CELL FACTOR	0.0022	0.0012
60412	EXOC4	EXOCYST COMPLEX COMPONENT 4	0.0045	0.0012
1896	EDA	ECTODYSPLASIN A	0.0005	0.0013
131377	KBTBD5	KELCH REPEAT AND BTB (POZ) DOMAIN CONTAINING 5	0.0012	0.0013
55161	TMEM33	TRANSMEMBRANE PROTEIN 33	0.0025	0.0013
79877	DCAKD	DEPHOSPHO-COA KINASE DOMAIN CONTAINING	0.0038	0.0013
4781	NFIB	NUCLEAR FACTOR I/B	0	0.0014
57144	PAK7	P21(CDKN1A)-ACTIVATED KINASE 7	0.0004	0.0014
80326	WNT10A	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 10A	0.0014	0.0014
80301	PLEKHQ1	PLECKSTRIN HOMOLOGY DOMAIN CONTAINING, FAMILY Q MEMBER 1	0.0022	0.0014
153443	SRFBP1	SERUM RESPONSE FACTOR BINDING PROTEIN 1	0.0052	0.0014

200734	SPRED2	SPROUTY-RELATED, EVH1 DOMAIN CONTAINING 2	0.0074	0.0014
5727	PTCH	PATCHED HOMOLOG (DROSOPHILA)	0.0096	0.0014
5453	POU3F1	POU DOMAIN, CLASS 3, TRANSCRIPTION FACTOR 1	0	0.0015
9555	H2AFY	H2A HISTONE FAMILY, MEMBER Y	0.0016	0.0015
286514	MAGEB18	MELANOMA ANTIGEN FAMILY B, 18	0.0045	0.0016
1588	CYP19A1	CYTOCHROME P450, FAMILY 19, SUBFAMILY A, POLYPEPTIDE 1	0.0087	0.0016
1131	CHRM3	CHOLINERGIC RECEPTOR, MUSCARINIC 3	0.002	0.0017
64211	LHX5	LIM HOMEOBOX 5	0.0026	0.0017
5142	PDE4B	PHOSPHODIESTERASE 4B, CAMP-SPECIFIC (PHOSPHODIESTERASE E4 DUNCE HOMOLOG, DROSOPHILA)	0.0026	0.0017
9110	MTMR4	MYOTUBULARIN RELATED PROTEIN 4	0.0044	0.0017
7483	WNT9A	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 9A	0.0048	0.0017
3617	IMPG1	INTERPHOTORECEPTOR MATRIX PROTEOGLYCAN 1	0.0089	0.0017
5139	PDE3A	PHOSPHODIESTERASE 3A, CGMP-INHIBITED	0.0071	0.0018
93986	FOXP2	TRINUCLEOTIDE REPEAT CONTAINING 10	0.0093	0.0018
3164	NR4A1	NUCLEAR RECEPTOR SUBFAMILY 4, GROUP A, MEMBER 1	0.0088	0.0019
11248	NXPH3	NEUREXOPHILIN 3	0.0045	0.002
161357	MAMDC1	MAM DOMAIN CONTAINING 1	0.0089	0.002
5596	MAPK4	MITOGEN-ACTIVATED PROTEIN KINASE 4	0.0009	0.0021
23157	SEPT6	SEPTIN 6	0.003	0.0021
117154	DACH2	DACHSHUND HOMOLOG 2 (DROSOPHILA)	0.0041	0.0021
70	ACTC	ACTIN, ALPHA, CARDIAC MUSCLE	0.008	0.0021
92369	SPSB4	SPLA/RYANODINE RECEPTOR DOMAIN AND SOCS BOX CONTAINING 4	0.0096	0.0021
4209	MEF2D	MADS BOX TRANSCRIPTION ENHANCER FACTOR 2, POLYPEPTIDE D (MYOCYTE ENHANCER FACTOR 2D)	0.0018	0.0022
340784	HMX3	HOMEOBOX (H6 FAMILY) 3	0.0036	0.0023
54843	SYTL2	SYNAPTOTAGMIN-LIKE 2	0.0037	0.0023
54541	DDIT4	DNA-DAMAGE-INDUCIBLE TRANSCRIPT 4	0.0071	0.0023
23439	ATP1B4	ATPASE, (NA+)/K+ TRANSPORTING, BETA 4 POLYPEPTIDE	0.0074	0.0023
27324	TNRC9	TRINUCLEOTIDE REPEAT CONTAINING 9	0.0001	0.0024
63951	DMRTA1	DMRT-LIKE FAMILY A1	0.0026	0.0024
30010	NXPH1	NEUREXOPHILIN 1	0.0029	0.0024
3382	ICA1	ISLET CELL AUTOANTIGEN 1, 69KDA	0.0043	0.0024
3788	KCNS2	POTASSIUM VOLTAGE-GATED CHANNEL, DELAYED-RECTIFIER, SUBFAMILY S, MEMBER 2	0	0.0025
1620	DBC1	DELETED IN BLADDER CANCER 1	0.0026	0.0026
818	CAMK2G	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE (CAM KINASE) II GAMMA	0.0032	0.0026
25834	MGAT4C	MANNOSYL (ALPHA-1,3-)-GLYCOPROTEIN BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE, ISOZYME C (PUTATIVE)	0.0006	0.0028
9577	BRE	BRAIN AND REPRODUCTIVE ORGAN-EXPRESSED (TNFRSF1A MODULATOR)	0.0076	0.0028
91752	C2orf10	CHROMOSOME 2 OPEN READING FRAME 10	0.0081	0.0028
26207	PITPNC1	PHOSPHATIDYLINOSITOL TRANSFER PROTEIN, CYTOPLASMIC 1	0.0054	0.003
57591	MKL1	MEGAKARYOBLASTIC LEUKEMIA (TRANSLOCATION) 1	0.001	0.0031
845	CASQ2	CALSEQUESTRIIN 2 (CARDIAC MUSCLE)	0.0031	0.0032
10266	RAMP2	RECEPTOR (CALCITONIN) ACTIVITY MODIFYING PROTEIN 2	0.0063	0.0033
91768	CABLES1	CDK5 AND ABL ENZYME SUBSTRATE 1	0.0018	0.0034
4897	NRCAM	NEURONAL CELL ADHESION MOLECULE	0.0049	0.0034
23468	CBX5	CHROMOBOX HOMOLOG 5 (HP1 ALPHA HOMOLOG, DROSOPHILA)	0.0058	0.0034
90423	ATP6V1E2	ATPASE, H+ TRANSPORTING, LYSOSOMAL 31KDA, V1 SUBUNIT E2	0.0078	0.0035
7881	KCNAB1	POTASSIUM VOLTAGE-GATED CHANNEL, SHAKER-RELATED SUBFAMILY, BETA MEMBER 1	0.0098	0.0035
131368	ZPLD1	ZONA PELLUCIDA-LIKE DOMAIN CONTAINING 1	0.0085	0.0036
7074	TIAM1	T-CELL LYMPHOMA INVASION AND METASTASIS 1	0.0033	0.0037
4163	MCC	MUTATED IN COLORECTAL CANCERS	0.0093	0.0037
144165	PRICKLE1	PRICKLE-LIKE 1 (DROSOPHILA)	0.0027	0.0038
25925	ZNF521	ZINC FINGER PROTEIN 521	0	0.0039
8525	DGKZ	DIACYLGLYCEROL KINASE, ZETA 104KDA	0.0071	0.0039
23645	PPP1R15A	PROTEIN PHOSPHATASE 1, REGULATORY (INHIBITOR) SUBUNIT 15A	0.0084	0.004
10178	ODZ1	ODZ, ODD OZ/TEN-M HOMOLOG 1(DROSOPHILA)	0.0003	0.0041
84929	FIBCD1	FIBRINOGEN C DOMAIN CONTAINING 1	0	0.0042
10495	COVA1	CYTOSOLIC OVARIAN CARCINOMA ANTIGEN 1	0.0001	0.0042
90627	STARD13	START DOMAIN CONTAINING 13	0.0012	0.0042
318	NUDT2	NUDIX (NUCLEOSIDE DIPHOSPHATE LINKED MOIETY X)-TYPE MOTIF 2	0.0092	0.0042
105	ADARB2	ADENOSINE DEAMINASE, RNA-SPECIFIC, B2 (RED2 HOMOLOG RAT)	0.0041	0.0043
10659	CUGBP2	CUG TRIPLET REPEAT, RNA BINDING PROTEIN 2	0.0022	0.0044
10010	TANK	TRAF FAMILY MEMBER-ASSOCIATED NFKB ACTIVATOR	0.0023	0.0044
4593	MUSK	MUSCLE, SKELETAL, RECEPTOR TYROSINE KINASE	0.0013	0.0045
1295	COL8A1	COLLAGEN, TYPE VIII, ALPHA 1	0.0022	0.0045
6195	RPS6KA1	RIBOSOMAL PROTEIN S6 KINASE, 90KDA, POLYPEPTIDE 1	0.0056	0.0046
8840	WISP1	WNT1 INDUCIBLE SIGNALING PATHWAY PROTEIN 1	0.0097	0.0047
56920	SEMA3G	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3G	0.0005	0.0048
2596	GAP43	GROWTH ASSOCIATED PROTEIN 43	0.0008	0.0048
8522	GAS7	GROWTH ARREST-SPECIFIC 7	0.0064	0.0048
488	ATP2A2	ATPASE, CA++ TRANSPORTING, CARDIAC MUSCLE, SLOW TWITCH 2	0.0065	0.0048
2138	EYA1	EYES ABSENT HOMOLOG 1 (DROSOPHILA)	0.0004	0.0049
7158	TP53BP1	TUMOR PROTEIN P53 BINDING PROTEIN, 1	0.0014	0.0049
4771	NF2	NEUROFIBROMIN 2 (BILATERAL ACOUSTIC NEUROMA)	0.0025	0.005
54549	SDK2	SIDEKICK HOMOLOG 2 (CHICKEN)	0.0051	0.0055
83700	JAM3	JUNCTIONAL ADHESION MOLECULE 3	0.0043	0.0057
23179	RGL1	RAL GUANINE NUCLEOTIDE DISSOCIATION STIMULATOR-LIKE 1	0.0029	0.0058
1394	CRHR1	CORTICOTROPIN RELEASING HORMONE RECEPTOR 1	0.009	0.0059
140469	MYO3B	MYOSIN IIIB	0.0093	0.006
9501	RPH3AL	RABPHILIN 3A-LIKE (WITHOUT C2 DOMAINS)	0.0097	0.006
4035	LRP1	LOW DENSITY LIPOPROTEIN-RELATED PROTEIN 1 (ALPHA-2-MACROGLOBULIN RECEPTOR)	0.006	0.0062
51360	MBTPS2	MEMBRANE-BOUND TRANSCRIPTION FACTOR PEPTIDASE, SITE 2	0.0042	0.0063
23176	SEPT8	SEPTIN 8	0.0064	0.0065

79885	HDAC11	HISTONE DEACETYLASE 11	0.0064	0.0066
6885	MAP3K7	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 7	0.0012	0.0067
22986	SORCS3	SORTILIN-RELATED VPS10 DOMAIN CONTAINING RECEPTOR 3	0	0.0069
117531	TMC1	TRANSMEMBRANE CHANNEL-LIKE 1	0.0025	0.0069
54149	C21orf91	CHROMOSOME 21 OPEN READING FRAME 91	0.0017	0.0072
2534	FYN	FYN ONCOGENE RELATED TO SRC, FGR, YES	0.0057	0.0076
58158	NEUROD4	NEUROGENIC DIFFERENTIATION 4	0.0073	0.0076
153090	DAB2IP	NGAP-LIKE PROTEIN	0.0058	0.0077
257194	NEGR1	NEURONAL GROWTH REGULATOR 1	0.0069	0.0077
7090	TLE3	TRANSDUCIN-LIKE ENHANCER OF SPLIT 3 (E(SP1) HOMOLOG, DROSOPHILA)	0.0034	0.0078
27249	C2orf25	CHROMOSOME 2 OPEN READING FRAME 25	0.0041	0.0078
5530	PPP3CA	PROTEIN PHOSPHATASE 3 (FORMERLY 2B), CATALYTIC SUBUNIT, ALPHA ISOFORM (CALCINEURIN A ALPHA)	0.0012	0.0079
29	ABR	ACTIVE BCR-RELATED GENE	0.0056	0.0079
1268	CNR1	CANNABINOID RECEPTOR 1 (BRAIN)	0.0026	0.0081
5188	PET112L	PET112-LIKE (YEAST)	0.0008	0.0082
57658	CALCOCO1	CALCIUM BINDING AND COILED-COIL DOMAIN 1	0.0022	0.0082
411	ARSB	ARYLSULFATASE B	0.0039	0.0084
64359	NXN	NUCLEOREDOXIN	0.0074	0.0084
8671	SLC4A4	SOLUTE CARRIER FAMILY 4, SODIUM BICARBONATE COTRANSPORTER, MEMBER 4	0.0007	0.0085
25959	ANKRD25	ANKYRIN REPEAT DOMAIN 25	0.0061	0.0086
3732	CD82	CD82 ANTIGEN	0.0001	0.0092
113278	C20orf54	CHROMOSOME 20 OPEN READING FRAME 54	0.0057	0.0092
54584	GNB1L	GUANINE NUCLEOTIDE BINDING PROTEIN (G PROTEIN), BETA POLYPEPTIDE 1-LIKE	0.0087	0.0093
26259	FBXW8	F-BOX AND WD-40 DOMAIN PROTEIN 8	0.0042	0.0095
79139	DERL1	DER1-LIKE DOMAIN FAMILY, MEMBER 1	0.0066	0.0095
1263	PLK3	POLO-LIKE KINASE 3 (DROSOPHILA)	0.0031	0.0096
655	BMP7	BONE MORPHOGENETIC PROTEIN 7 (OSTEOGENIC PROTEIN 1)	0.0062	0.0096
4217	MAP3K5	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 5	0.0082	0.0098
6857	SYT1	SYNAPTOTAGMIN I	0.0012	0.0099
596	BCL2	B-CELL CLL/LYMPHOMA 2	0.0059	0.0099
23767	FLRT3	FIBRONECTIN LEUCINE RICH TRANSMEMBRANE PROTEIN 3	0.0083	0.0099

Ngn/NeuroD target genes identified by microarray and in PAP computational analysis

Human/Mouse		
Ngn/NeuroD matrix		
Gene ID	Symbol	P-value
9139	CBFA2T2	0
28514	DLL1	0.0048
79191	IRX3	0.0002
58158	NEUROD4	0.0073
4807	NHLH1	0.0029
1263	PLK3	0.0031
5453	POU3F1	0
5454	POU3F2	0.004
144165	PRICKLE1	0.0027
9839	ZFH1B	0
10472	ZNF238	0.0001

Human/Mouse/Rat		
Ngn/NeuroD matrix		
Gene ID	Symbol	P-value
9139	CBFA2T2	0
28514	DLL1	0.0009
1995	ELAVL3	0.005
79191	IRX3	0.0001
4661	MYT1	0.0035
58158	NEUROD4	0.0076
4807	NHLH1	0.0006
1263	PLK3	0.0096
5453	POU3F1	0.0015
144165	PRICKLE1	0.0038
9839	ZFH1B	0
10472	ZNF238	0

Intersection	
Ngn/NeuroD matrix	
Gene ID	Symbol
9139	CBFA2T2
28514	DLL1
79191	IRX3
58158	NEUROD4
4807	NHLH1
1263	PLK3
5453	POU3F1
144165	PRICKLE1
9839	ZFH1B
10472	ZNF238

B. M01034 E-box target genes predicted by PAP (365 genes)

Gene ID	Gene Symbol	Gene name
29	ABR	ACTIVE BCR-RELATED GENE
43	ACHE	ACETYLCHOLINESTERASE (YT BLOOD GROUP)
97	ACYP1	ACYLPHOSPHATASE 1, ERYTHROCYTE (COMMON) TYPE
271	AMPD2	ADENOSINE MONOPHOSPHATE DEAMINASE 2 (ISOFORM L)
310	ANXA7	ANNEXIN A7
364	AQP7	AQUAPORIN 7
377	ARF3	ADP-RIBOSYLATION FACTOR 3
439	ASNA1	ARSA ARSENITE TRANSPORTER, ATP-BINDING, HOMOLOG 1 (BACTERIAL)
445	ASS	ARGININOSUCCINATE SYNTHETASE
468	ATF4	ACTIVATING TRANSCRIPTION FACTOR 4 (TAX-RESPONSIVE ENHANCER ELEMENT B67)
498	ATP5A1	ATP SYNTHASE, H+ TRANSPORTING, MITOCHONDRIAL F1 COMPLEX, ALPHA SUBUNIT 1, CARDIAC MUSCLE
498	ATP5A1	ATP SYNTHASE, H+ TRANSPORTING, MITOCHONDRIAL F1 COMPLEX, ALPHA SUBUNIT 1, CARDIAC MUSCLE
516	ATP5G1	ATP SYNTHASE, H+ TRANSPORTING, MITOCHONDRIAL F0 COMPLEX, SUBUNIT C1 (SUBUNIT 9)
527	ATP6V0C	ATPASE, H+ TRANSPORTING, LYSOSOMAL 16KDA, V0 SUBUNIT C
533	ATP6V0B	ATPASE, H+ TRANSPORTING, LYSOSOMAL 21KDA, V0 SUBUNIT B
558	AXL	AXL RECEPTOR TYROSINE KINASE
576	BAI2	BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 2
598	BCL2L1	BCL2-LIKE 1
745	C11ORF9	CHROMOSOME 11 OPEN READING FRAME 9
747	C11ORF11	CHROMOSOME 11 OPEN READING FRAME 11
784	CACNB3	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, BETA 3 SUBUNIT
844	CASQ1	CALSEQUESTRIN 1 (FAST-TWITCH, SKELETAL MUSCLE)
863	CBFA2T3	CORE-BINDING FACTOR, RUNT DOMAIN, ALPHA SUBUNIT 2; TRANSLOCATED TO, 3
896	CCND3	CYCLIN D3
1019	CDK4	CYCLIN-DEPENDENT KINASE 4
1152	CKB	CREATINE KINASE, BRAIN

1158	CKM	CREATINE KINASE, MUSCLE
1185	CLCN6	CHLORIDE CHANNEL 6
1269	CNR2	CANNABINOID RECEPTOR 2 (MACROPHAGE)
1271	CNTFR	CILIARY NEUROTROPHIC FACTOR RECEPTOR
1627	DBN1	DREBRIN 1
1654	DDX3X	DEAD (ASP-GLU-ALA-ASP) BOX POLYPEPTIDE 3, X-LINKED
1788	DNMT3A	DNA (CYTOSINE-5-)-METHYLTRANSFERASE 3 ALPHA
1841	DTYMK	DEOXYTHYMIDYLATE KINASE (THYMIDYLATE KINASE)
1870	E2F2	E2F TRANSCRIPTION FACTOR 2
1938	EEF2	EUKARYOTIC TRANSLATION ELONGATION FACTOR 2
1952	CELSR2	CADHERIN, EGF LAG SEVEN-PASS G-TYPE RECEPTOR 2 (FLAMINGO HOMOLOG, DROSOPHILA)
1969	EPHA2	EPH RECEPTOR A2
1994	ELAVL1	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 1 (HU ANTIGEN R)
1995	ELAVL3	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 3 (HU ANTIGEN C)
2006	ELN	ELASTIN (SUPRAVALVULAR AORTIC STENOSIS, WILLIAMS-BEUREN SYNDROME)
2048	EPHB2	EPH RECEPTOR B2
2077	ERF	ETS2 REPRESSOR FACTOR
2100	ESR2	ESTROGEN RECEPTOR 2 (ER BETA)
2194	FASN	FATTY ACID SYNTHASE
2251	FGF6	FIBROBLAST GROWTH FACTOR 6
2256	FGF11	FIBROBLAST GROWTH FACTOR 11
2318	FLNC	FILAMIN C, GAMMA (ACTIN BINDING PROTEIN 280)
2624	GATA2	GATA BINDING PROTEIN 2
2672	GF11	GROWTH FACTOR INDEPENDENT 1
2695	GIP	GASTRIC INHIBITORY POLYPEPTIDE
2701	GJA4	GAP JUNCTION PROTEIN, ALPHA 4, 37KDA (CONNEXIN 37)
2705	GJB1	GAP JUNCTION PROTEIN, BETA 1, 32KDA (CONNEXIN 32, CHARCOT-MARIE-TOOTH NEUROPATHY, X-LINKED)
2710	GK	GLYCEROL KINASE
2719	GPC3	GLYPICAN 3
2775	GNAO1	GUANINE NUCLEOTIDE BINDING PROTEIN (G PROTEIN), ALPHA ACTIVATING ACTIVITY POLYPEPTIDE O
2778	GNAS	GNAS COMPLEX LOCUS
2799	GNS	GLUCOSAMINE (N-ACETYL)-6-SULFATASE (SANFILIPPO DISEASE IIID)
2867	FFAR2	FREE FATTY ACID RECEPTOR 2
2876	GPX1	GLUTATHIONE PEROXIDASE 1
2886	GRB7	GROWTH FACTOR RECEPTOR-BOUND PROTEIN 7
2957	GTF2A1	GENERAL TRANSCRIPTION FACTOR IIA, 1, 19/37KDA
2965	GTF2H1	GENERAL TRANSCRIPTION FACTOR IIH, POLYPEPTIDE 1, 62KDA
3073	HEXA	HEXOSAMINIDASE A (ALPHA POLYPEPTIDE)
3131	HLF	HEPATIC LEUKEMIA FACTOR
3151	HMGN2	HIGH-MOBILITY GROUP NUCLEOSOMAL BINDING DOMAIN 2
3162	HMOX1	HEME OXYGENASE (DECYCLING) 1
3164	NR4A1	NUCLEAR RECEPTOR SUBFAMILY 4, GROUP A, MEMBER 1
3200	HOXA3	HOMEODOMAIN A3
3203	HOXA6	HOMEODOMAIN A6
3204	HOXA7	HOMEODOMAIN A7
3214	HOXB4	HOMEODOMAIN B4
3217	HOXB7	HOMEODOMAIN B7
3297	HSF1	HEAT SHOCK TRANSCRIPTION FACTOR 1
3340	NDST1	N-DEACETYLASE/N-SULFOTRANSFERASE (HEPARAN GLUCOSAMINYL) 1
3397	ID1	INHIBITOR OF DNA BINDING 1, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN
3609	ILF3	INTERLEUKIN ENHANCER BINDING FACTOR 3, 90KDA
3651	IPF1	INSULIN PROMOTER FACTOR 1, HOMEODOMAIN TRANSCRIPTION FACTOR
3757	KCNH2	POTASSIUM VOLTAGE-GATED CHANNEL, SUBFAMILY H (EAG-RELATED), MEMBER 2
3761	KCNJ4	POTASSIUM INWARDLY-RECTIFYING CHANNEL, SUBFAMILY J, MEMBER 4
3763	KCNJ6	POTASSIUM INWARDLY-RECTIFYING CHANNEL, SUBFAMILY J, MEMBER 6
3768	KCNJ12	POTASSIUM INWARDLY-RECTIFYING CHANNEL, SUBFAMILY J, MEMBER 12
3770	KCNJ14	POTASSIUM INWARDLY-RECTIFYING CHANNEL, SUBFAMILY J, MEMBER 14
3777	KCNK3	POTASSIUM CHANNEL, SUBFAMILY K, MEMBER 3
3784	KCNQ1	POTASSIUM VOLTAGE-GATED CHANNEL, KQT-LIKE SUBFAMILY, MEMBER 1
3916	LAMP1	LYSOSOMAL-ASSOCIATED MEMBRANE PROTEIN 1
4010	LMX1B	LIM HOMEODOMAIN TRANSCRIPTION FACTOR 1, BETA
4046	LSP1	LYMPHOCYTE-SPECIFIC PROTEIN 1
4149	MAX	MYC ASSOCIATED FACTOR X
4152	MBD1	METHYL-CPG BINDING DOMAIN PROTEIN 1
4287	ATXN3	ATAXIN 3
4524	MTHFR	5,10-METHYLENETETRAHYDROFOLATE REDUCTASE (NADPH)
4591	TRIM37	TRIPARTITE MOTIF-CONTAINING 37
4628	MYH10	MYOSIN, HEAVY POLYPEPTIDE 10, NON-MUSCLE
4650	MYO9B	MYOSIN IXB
4670	HNRPM	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN M
4691	NCL	NUCLEOLIN
4760	NEUROD1	NEUROGENIC DIFFERENTIATION 1
4761	NEUROD2	NEUROGENIC DIFFERENTIATION 2
4784	NFIX	NUCLEAR FACTOR I/X (CCAAT-BINDING TRANSCRIPTION FACTOR)
4821	NKX2-2	NK2 TRANSCRIPTION FACTOR RELATED, LOCUS 2 (DROSOPHILA)
4831	NME2	NON-METASTATIC CELLS 2, PROTEIN (NM23B) EXPRESSED IN
4884	NPTX1	NEURONAL PENTRAXIN I
4904	YBX1	Y BOX BINDING PROTEIN 1
4985	OPRD1	OPIOID RECEPTOR, DELTA 1
5045	FURIN	FURIN (PAIRED BASIC AMINO ACID CLEAVING ENZYME)
5076	PAX2	PAIRED BOX GENE 2

5091	PC	PYRUVATE CARBOXYLASE
5141	PDE4A	PHOSPHODIESTERASE 4A, CAMP-SPECIFIC (PHOSPHODIESTERASE E2 DUNCE HOMOLOG, DROSOPHILA)
5293	PIK3CD	PHOSPHOINOSITIDE-3-KINASE, CATALYTIC, DELTA POLYPEPTIDE
5339	PLEC1	PLECTIN 1, INTERMEDIATE FILAMENT BINDING PROTEIN 500KDA
5360	PLTP	PHOSPHOLIPID TRANSFER PROTEIN
5414	SEPT4	SEPTIN 4
5579	PRKCB1	PROTEIN KINASE C, BETA 1
5585	PKN1	PROTEIN KINASE N1
5605	MAP2K2	MITOGEN-ACTIVATED PROTEIN KINASE KINASE 2
5620	PRM2	PROTAMINE 2
5697	PYY	PEPTIDE YY
5727	PTCH	PATCHED HOMOLOG (DROSOPHILA)
5866	RAB3IL1	RAB3A INTERACTING PROTEIN (RABIN3)-LIKE 1
5883	RAD9A	RAD9 HOMOLOG A (S. POMBE)
5913	RAPSN	RECEPTOR-ASSOCIATED PROTEIN OF THE SYNAPSE, 43KD
5914	RARA	RETINOIC ACID RECEPTOR, ALPHA
5916	RARG	RETINOIC ACID RECEPTOR, GAMMA
5989	RFX1	REGULATORY FACTOR X, 1 (INFLUENCES HLA CLASS II EXPRESSION)
6097	RORC	RAR-RELATED ORPHAN RECEPTOR C
6117	RPA1	REPLICATION PROTEIN A1, 70KDA
6208	RPS14	RIBOSOMAL PROTEIN S14
6256	RXRA	RETINOID X RECEPTOR, ALPHA
6405	SEMA3F	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3F
6449	SGTA	SMALL GLUTAMINE-RICH TETRATRIPEPTIDE REPEAT (TPR)-CONTAINING, ALPHA
6455	SH3GL1	SH3-DOMAIN GRB2-LIKE 1
6525	SMTN	SMOOTHELIN
6530	SLC6A2	SOLUTE CARRIER FAMILY 6 (NEUROTRANSMITTER TRANSPORTER, NORADRENALIN), MEMBER 2
6536	SLC6A9	SOLUTE CARRIER FAMILY 6 (NEUROTRANSMITTER TRANSPORTER, GLYCINE), MEMBER 9
6612	SUMO3	SMT3 SUPPRESSOR OF MIF TWO 3 HOMOLOG 3 (YEAST)
6776	STAT5A	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5A
6809	STX3A	SYNTAXIN 3A
6812	STXBP1	SYNTAXIN BINDING PROTEIN 1
6932	TCF7	TRANSCRIPTION FACTOR 7 (T-CELL SPECIFIC, HMG-BOX)
6945	MLX	MAX-LIKE PROTEIN X
7008	TEF	THYROTROPHIC EMBRYONIC FACTOR
7064	THOP1	THIMET OLIGOPEPTIDASE 1
7142	TNP2	TRANSITION PROTEIN 2 (DURING HISTONE TO PROTAMINE REPLACEMENT)
7161	TP73	TUMOR PROTEIN P73
7186	TRAF2	TNF RECEPTOR-ASSOCIATED FACTOR 2
7205	TRIP6	THYROID HORMONE RECEPTOR INTERACTOR 6
7320	UBE2B	UBIQUITIN-CONJUGATING ENZYME E2B (RAD6 HOMOLOG)
7343	UBTF	UPSTREAM BINDING TRANSCRIPTION FACTOR, RNA POLYMERASE I
7477	WNT7B	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 7B
7483	WNT9A	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 9A
7535	ZAP70	ZETA-CHAIN (TCR) ASSOCIATED PROTEIN KINASE 70KDA
7866	IFRD2	INTERFERON-RELATED DEVELOPMENTAL REGULATOR 2
8398	PLA2G6	PHOSPHOLIPASE A2, GROUP VI (CYTOSOLIC, CALCIUM-INDEPENDENT)
8431	NROB2	NUCLEAR RECEPTOR SUBFAMILY 0, GROUP B, MEMBER 2
8508	NIPSNAP1	NIPSNAP HOMOLOG 1 (C. ELEGANS)
8662	EIF3S9	EUKARYOTIC TRANSLATION INITIATION FACTOR 3, SUBUNIT 9 ETA, 116KDA
8763	CD164	CD164 ANTIGEN, SIALOMUCIN
8775	NAPA	N-ETHYLMALEIMIDE-SENSITIVE FACTOR ATTACHMENT PROTEIN, ALPHA
8878	SQSTM1	SEQUESTOSOME 1
8913	CACNA1G	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, ALPHA 1G SUBUNIT
9066	SYT7	SYNAPTOTAGMIN VII
9093	DNAJA3	DNAJ (HSP40) HOMOLOG, SUBFAMILY A, MEMBER 3
9099	USP2	UBIQUITIN SPECIFIC PEPTIDASE 2
9110	MTMR4	MYOTUBULARIN RELATED PROTEIN 4
9114	ATP6V0D1	ATPASE, H+ TRANSPORTING, LYSOSOMAL 38KDA, V0 SUBUNIT D1
9254	CACNA2D2	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, ALPHA 2/DELTA SUBUNIT 2
9258	MFHAS1	MALIGNANT FIBROUS HISTIOCYTOMA AMPLIFIED SEQUENCE 1
9342	SNAP29	SYNAPTOSOMAL-ASSOCIATED PROTEIN, 29KDA
9377	COX5A	CYTOCHROME C OXIDASE SUBUNIT VA
9572	NR1D1	NUCLEAR RECEPTOR SUBFAMILY 1, GROUP D, MEMBER 1
9655	SOC5	SUPPRESSOR OF CYTOKINE SIGNALING 5
9885	OSBPL2	OXYSTEROL BINDING PROTEIN-LIKE 2
9967	THRAP3	THYROID HORMONE RECEPTOR ASSOCIATED PROTEIN 3
10039	PARP3	POLY (ADP-RIBOSE) POLYMERASE FAMILY, MEMBER 3
10136	ELA3B	ELASTASE 3A, PANCREATIC
10136	ELA3A	ELASTASE 3A, PANCREATIC
10159	ATP6AP2	ATPASE, H+ TRANSPORTING, LYSOSOMAL ACCESSORY PROTEIN 2
10289	EIF1B	EUKARYOTIC TRANSLATION INITIATION FACTOR 1B
10330	TMEM4	TRANSMEMBRANE PROTEIN 4
10411	RAPGEF3	RAP GUANINE NUCLEOTIDE EXCHANGE FACTOR (GEF) 3
10435	CDC42EP2	CDC42 EFFECTOR PROTEIN (RHO GTPASE BINDING) 2
10482	NXF1	NUCLEAR RNA EXPORT FACTOR 1
10501	SEMA6B	SEMA DOMAIN, TRANSMEMBRANE DOMAIN (TM), AND CYTOPLASMIC DOMAIN, (SEMAPHORIN) 6B
10564	ARFGEF2	ADP-RIBOSYLATION FACTOR GUANINE NUCLEOTIDE-EXCHANGE FACTOR 2 (BREFELDIN A-INHIBITED)
10629	TAF6L	TAF6-LIKE RNA POLYMERASE II, P300/CBP-ASSOCIATED FACTOR (PCAF)-ASSOCIATED FACTOR, 65KDA
10672	GNA13	GUANINE NUCLEOTIDE BINDING PROTEIN (G PROTEIN), ALPHA 13
10814	CPLX2	COMPLEXIN 2

10867	TSPAN9	TETRASPANIN 9
10900	RPIP8	RAP2 INTERACTING PROTEIN 8
11060	WWP2	WW DOMAIN CONTAINING E3 UBIQUITIN PROTEIN LIGASE 2
11078	TRIOBP	TRIO AND F-ACTIN BINDING PROTEIN
11152	WDR45	WD REPEAT DOMAIN 45
11240	PADI2	PEPTIDYL ARGININE DEIMINASE, TYPE II
11248	NXPH3	NEUREXOPHILIN 3
11330	CTRC	CHYMOTRYPSIN C (CALDECRIN)
11337	GABARAP	GABA(A) RECEPTOR-ASSOCIATED PROTEIN
22839	DLGAP4	DISCS, LARGE (DROSOPHILA) HOMOLOG-ASSOCIATED PROTEIN 4
22845	TMEM15	TRANSMEMBRANE PROTEIN 15
22888	UBOX5	U-BOX DOMAIN CONTAINING 5
23038	WDTC1	WD AND TETRATRICOPEPTIDE REPEATS 1
23054	NCOA6	NUCLEAR RECEPTOR COACTIVATOR 6
23116	KIAA0423	KIAA0423
23152	CIC	CAPICUA HOMOLOG (DROSOPHILA)
23187	PHLDB1	PLECKSTRIN HOMOLOGY-LIKE DOMAIN, FAMILY B, MEMBER 1
23199	KIAA0182	KIAA0182
23246	BOP1	BLOCK OF PROLIFERATION 1
23729	CARKL	CARBOHYDRATE KINASE-LIKE
23764	MAFF	V-MAF MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE HOMOLOG F (AVIAN)
25764	SERF2	SMALL EDRK-RICH FACTOR 2
25764	HYPK	SMALL EDRK-RICH FACTOR 2
25796	PGLS	6-PHOSPHOGLUCONOLACTONASE
25822	DNAJB5	DNAJ (HSP40) HOMOLOG, SUBFAMILY B, MEMBER 5
25841	ABTB2	ANKYRIN REPEAT AND BTB (POZ) DOMAIN CONTAINING 2
25911	RP11-529I10.4	DELETED IN A MOUSE MODEL OF PRIMARY CILIARY DYSKINESIA
25915	C3ORF60	CHROMOSOME 3 OPEN READING FRAME 60
26574	AATF	APOPTOSIS ANTAGONIZING TRANSCRIPTION FACTOR
27113	BBC3	BCL2 BINDING COMPONENT 3
27166	PX19	PX19-LIKE PROTEIN
27343	POLL	POLYMERASE (DNA DIRECTED), LAMBDA
28987	PSMD8BP1	PSMD8 BINDING PROTEIN 1
29078	C6ORF66	CHROMOSOME 6 OPEN READING FRAME 66
29110	TBK1	TANK-BINDING KINASE 1
29117	BRD7	BROMODOMAIN CONTAINING 7
29886	SNX8	SORTING NEXIN 8
29890	RBM15B	RNA BINDING MOTIF PROTEIN 15B
29904	EEF2K	EUKARYOTIC ELONGATION FACTOR-2 KINASE
29911	HOOK2	HOOK HOMOLOG 2 (DROSOPHILA)
29919	C18ORF8	CHROMOSOME 18 OPEN READING FRAME 8
29920	PYCR2	PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY, MEMBER 2
29942	PURG	PURINE-RICH ELEMENT BINDING PROTEIN G
30000	TNPO2	TRANSPORTIN 2 (IMPORTIN 3, KARYOPHERIN BETA 2B)
50813	COPS7A	COP9 CONSTITUTIVE PHOTOMORPHOGENIC HOMOLOG SUBUNIT 7A (ARABIDOPSIS)
50863	HNT	NEUROTRIMIN
51005	AMDHD2	AMIDOHYDROLASE DOMAIN CONTAINING 2
51035	LOC51035	UNKNOWN PROTEIN LOC51035
51043	ZBTB7B	ZINC FINGER AND BTB DOMAIN CONTAINING 7B
51063	FAM26B	FAMILY WITH SEQUENCE SIMILARITY 26, MEMBER B
51087	YBX2	Y BOX BINDING PROTEIN 2
51107	APH1A	ANTERIOR PHARYNX DEFECTIVE 1 HOMOLOG A (C. ELEGANS)
51117	COQ4	COENZYME Q4 HOMOLOG (YEAST)
51150	SDF4	STROMAL CELL DERIVED FACTOR 4
51172	NAGPA	N-ACETYLGLUCOSAMINE-1-PHOSPHODIESTER ALPHA-N-ACETYLGLUCOSAMINIDASE
51303	FKBP11	FK506 BINDING PROTEIN 11, 19 KDA
51330	TNFRSF12A	TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 12A
51343	FZR1	HOMOLOG OF YEAST CDH1/HCT1
51566	ARMCX3	ARMADILLO REPEAT CONTAINING, X-LINKED 3
51588	PIAS4	PROTEIN INHIBITOR OF ACTIVATED STAT, 4
51593	ARS2	ARS2 PROTEIN
51621	KLF13	KRUPPEL-LIKE FACTOR 13
51643	TMBIM4	TRANSMEMBRANE BAX INHIBITOR MOTIF CONTAINING 4
54361	WNT4	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 4
54443	ANLN	ANILLIN, ACTIN BINDING PROTEIN (SCRAPS HOMOLOG, DROSOPHILA)
54567	DLL4	DELTA-LIKE 4 (DROSOPHILA)
54808	DYM	DYMECLIN
54813	BTBD5	BTB (POZ) DOMAIN CONTAINING 5
54815	GATAD2A	GATA ZINC FINGER DOMAIN CONTAINING 2A
55152	DALRD3	DALR ANTICODON BINDING DOMAIN CONTAINING 3
55163	PNPO	PYRIDOXINE 5'-PHOSPHATE OXIDASE
55182	C1ORF164	HYPOTHETICAL PROTEIN LOC51261
55260	FLJ10922	HYPOTHETICAL PROTEIN FLJ10922
55352	C17ORF79	CHROMOSOME 17 OPEN READING FRAME 79
55644	OSGEP	O-SIALOGLYCOPROTEIN ENDOPEPTIDASE
55715	DOK4	DOCKING PROTEIN 4
55787	CXORF15	CHROMOSOME X OPEN READING FRAME 15
55809	TRERF1	TRANSCRIPTIONAL REGULATING FACTOR 1
55852	TEX2	TESTIS EXPRESSED SEQUENCE 2
56180	MOSPD1	MOTILE SPERM DOMAIN CONTAINING 1
57101	TMEM16B	TRANSMEMBRANE PROTEIN 16B

57192	MCOLN1	MUCOLIPIN 1
57410	SCYL1	SCY1-LIKE 1 (S. CEREVISIAE)
57524	CASKIN1	CASK INTERACTING PROTEIN 1
57597	BAHCC1	BAH DOMAIN AND COILED-COIL CONTAINING 1
57622	LRFN1	LEUCINE RICH REPEAT AND FIBRONECTIN TYPE III DOMAIN CONTAINING 1
59269	HIVEP3	HUMAN IMMUNODEFICIENCY VIRUS TYPE I ENHANCER BINDING PROTEIN 3
60625	DHX35	DEAH (ASP-GLU-ALA-HIS) BOX POLYPEPTIDE 35
63976	PRDM16	PR DOMAIN CONTAINING 16
64072	CDH23	CADHERIN-LIKE 23
64093	SMOC1	SPARC RELATED MODULAR CALCIUM BINDING 1
64211	LHX5	LIM HOMEBOX 5
64284	RAB17	RAB17, MEMBER RAS ONCOGENE FAMILY
64773	FAM113A	FAMILY WITH SEQUENCE SIMILARITY 113, MEMBER A
64926	FLJ21438	HYPOTHETICAL PROTEIN FLJ21438
64979	MRPL36	MITOCHONDRIAL RIBOSOMAL PROTEIN L36
65082	VPS33A	VACUOLAR PROTEIN SORTING 33A (YEAST)
65260	C1ORF163	CHROMOSOME 1 OPEN READING FRAME 163
78995	C17ORF53	CHROMOSOME 17 OPEN READING FRAME 53
78999	LRFN4	LEUCINE RICH REPEAT AND FIBRONECTIN TYPE III DOMAIN CONTAINING 4
79002	MGC2803	HYPOTHETICAL PROTEIN MGC2803
79064	MGC3196	HYPOTHETICAL PROTEIN MGC3196
79094	CHAC1	CHAC, CATION TRANSPORT REGULATOR-LIKE 1 (E. COLI)
79102	RNF26	RING FINGER PROTEIN 26
79624	C6ORF211	CHROMOSOME 6 OPEN READING FRAME 211
79630	C1ORF54	CHROMOSOME 1 OPEN READING FRAME 54
79794	C12ORF49	CHROMOSOME 12 OPEN READING FRAME 49
79816	TLE6	TRANSDUCIN-LIKE ENHANCER OF SPLIT 6 (E(SP1) HOMOLOG, DROSOPHILA)
79827	ASAM	ADIPOCYTE-SPECIFIC ADHESION MOLECULE
79829	NAT11	HYPOTHETICAL PROTEIN FLJ13848
79837	PIP5K2C	PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE, TYPE II, GAMMA
79924	ADM2	ADRENOMEDULLIN 2
79960	PHF17	PHD FINGER PROTEIN 17
80149	ZC3H12A	ZINC FINGER CCCH-TYPE CONTAINING 12A
80305	RP3-402G11.12	HYPOTHETICAL PROTEIN PP2447
80326	WNT10A	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 10A
80342	TRAF3IP3	TRAF3 INTERACTING PROTEIN 3
80757	TMEM121	TRANSMEMBRANE PROTEIN 121
81543	LRRC3	LEUCINE RICH REPEAT CONTAINING 3
81693	AMN	AMNIONLESS HOMOLOG (MOUSE)
81794	ADAMTS10	ADAM METALLOPEPTIDASE WITH THROMBOSPONDIN TYPE 1 MOTIF, 10
81848	SPRY4	SPROUTY HOMOLOG 4 (DROSOPHILA)
81890	QTRT1	QUEUINE TRNA-RIBOSYLTRANSFERASE 1 (TRNA-GUANINE TRANSGLYCOSYLASE)
83451	ABHD11	ABHYDROLASE DOMAIN CONTAINING 11
83475	HLRC1	HEAT-LIKE (PBS LYASE) REPEAT CONTAINING 1
83605	CCM2	CEREBRAL CAVERNOUS MALFORMATION 2
83855	KLF16	KRUPPEL-LIKE FACTOR 16
84152	PPP1R1B	PROTEIN PHOSPHATASE 1, REGULATORY (INHIBITOR) SUBUNIT 1B (DOPAMINE AND CAMP REGULATED PHOSPHOPROTEIN, DARPP-32)
84249	PSD2	PLECKSTRIN AND SEC7 DOMAIN CONTAINING 2
84254	CAMKK1	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE KINASE 1, ALPHA
84445	LZTS2	LEUCINE ZIPPER, PUTATIVE TUMOR SUPPRESSOR 2
84752	MGC4655	HYPOTHETICAL PROTEIN MGC4655
84900	TMEM118	TRANSMEMBRANE PROTEIN 118
85352	KIAA1644	KIAA1644 PROTEIN
85457	KIAA1737	KIAA1737
89796	NAV1	HYPOTHETICAL PROTEIN MGC14961
90529	C1ORF201	CHROMOSOME 1 OPEN READING FRAME 201
90701	SEC11L3	SEC11-LIKE 3 (S. CEREVISIAE)
92140	MTDH	METADHERIN
92979	MARCH9	MEMBRANE-ASSOCIATED RING FINGER (C3HC4) 9
93611	FBXO44	F-BOX PROTEIN 44
114907	FBXO32	F-BOX PROTEIN 32
115106	CCDC5	COILED-COIL DOMAIN CONTAINING 5 (SPINDLE ASSOCIATED)
115992	RNF166	RING FINGER PROTEIN 166
116534	MRGPRE	MAS-RELATED GPR, MEMBER E
124044	MGC26885	HYPOTHETICAL PROTEIN MGC26885
124842	TMEM132E	TRANSMEMBRANE PROTEIN 132E
126282	TNFAIP8L1	TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 8-LIKE 1
126526	FLJ36888	HYPOTHETICAL PROTEIN FLJ36888
126792	B3GALT6	UDP-GAL:BETAGAL BETA 1,3-GALACTOSYLTRANSFERASE POLYPEPTIDE 6
129303	FLJ90024	FASTING-INDUCIBLE INTEGRAL MEMBRANE PROTEIN TM6P1
129685	TBN	TAUBE NUSS HOMOLOG (MOUSE)
133522	PPARGC1B	PEROXISOME PROLIFERATIVE ACTIVATED RECEPTOR, GAMMA, COACTIVATOR 1, BETA
138151	BTBD14A	BTB (POZ) DOMAIN CONTAINING 14A
140606	SELM	SELENOPROTEIN M
140628	GATA5	GATA BINDING PROTEIN 5
145258	GSC	GOOSECOID
146433	MGC34647	HYPOTHETICAL PROTEIN MGC34647
146556	MGC45438	HYPOTHETICAL PROTEIN MGC45438
146956	EME1	ESSENTIAL MEIOTIC ENDONUCLEASE 1 HOMOLOG 1 (S. POMBE)
147700	KLC3	KINESIN LIGHT CHAIN 3
151648	SGOL1	SHUGOSHIN-LIKE 1 (S. POMBE)

158135	TLL11	TUBULIN TYROSINE LIGASE-LIKE FAMILY, MEMBER 11
161882	ZFPM1	ZINC FINGER PROTEIN, MULTITYPE 1
168850	LOC168850	HYPOTHETICAL PROTEIN LOC168850
197135	LOC197135	HYPOTHETICAL LOC197135
200185	KRTCAP2	KERATINOCYTE ASSOCIATED PROTEIN 2
201232	SLC16A13	SOLUTE CARRIER FAMILY 16 (MONOCARBOXYLIC ACID TRANSPORTERS), MEMBER 13
206358	SLC36A1	LYSOSOMAL AMINO ACID TRANSPORTER 1
219738	C10ORF35	CHROMOSOME 10 OPEN READING FRAME 35
220972	MARCH8	MEMBRANE-ASSOCIATED RING FINGER (C3HC4) 8
283455	KSR2	KINASE SUPPRESSOR OF RAS 2
284069	MGC34829	SIMILAR TO HYPOTHETICAL GENE SUPPORTED BY AL050367; AK022946
284358	FLJ36070	HYPOTHETICAL PROTEIN FLJ36070
284723	SLC25A34	SOLUTE CARRIER FAMILY 25, MEMBER 34
339287	LOC339287	HYPOTHETICAL PROTEIN LOC339287
348110	C15ORF38	HYPOTHETICAL PROTEIN FLJ35955
388630	LOC388630	SIMILAR TO C05G5.5
390880	LOC390880	SIMILAR TO NUCLEAR FACTOR, INTERLEUKIN 3 REGULATED

C. Myc-Max M00118 target genes predicted by PAP (127 genes)

Gene ID	Gene symbol	Gene name
95	ACY1	AMINOACYLASE 1
151	ADRA2B	ADRENERGIC, ALPHA-2B-, RECEPTOR
271	AMPD2	ADENOSINE MONOPHOSPHATE DEAMINASE 2 (ISOFORM L)
403	ARL3	ADP-RIBOSYLATION FACTOR-LIKE 3
595	CCND1	CYCLIN D1
605	BCL7A	B-CELL CLL/LYMPHOMA 7A
666	BOK	BCL2-RELATED OVARIAN KILLER
672	BRCA1	BREAST CANCER 1, EARLY ONSET
745	C11ORF9	CHROMOSOME 11 OPEN READING FRAME 9
771	CA12	CARBONIC ANHYDRASE XII
875	CBS	CYSTATHIONINE-BETA-SYNTASE
914	CD2	CD2 ANTIGEN (P50), SHEEP RED BLOOD CELL RECEPTOR
1152	CKB	CREATINE KINASE, BRAIN
1489	CTF1	CARDIOTROPHIN 1
1678	TIMM8A	TRANSLOCASE OF INNER MITOCHONDRIAL MEMBRANE 8 HOMOLOG A (YEAST)
1757	SARDH	SARCOSINE DEHYDROGENASE
1995	ELAVL3	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 3 (HU ANTIGEN C)
2135	EXTL2	EXOSTOSES (MULTIPLE)-LIKE 2
2251	FGF6	FIBROBLAST GROWTH FACTOR 6
2354	FOSB	FBJ MURINE OSTEOSARCOMA VIRAL ONCOGENE HOMOLOG B
2672	GF11	GROWTH FACTOR INDEPENDENT 1
2695	GIP	GASTRIC INHIBITORY POLYPEPTIDE
2780	GNAT2	GUANINE NUCLEOTIDE BINDING PROTEIN (G PROTEIN), ALPHA TRANSDUCING ACTIVITY POLYPEPTIDE 2
2957	GTF2A1	GENERAL TRANSCRIPTION FACTOR IIA, 1, 19/37KDA
3036	HAS1	HYALURONAN SYNTHASE 1
3151	HMG2	HIGH-MOBILITY GROUP NUCLEOSOMAL BINDING DOMAIN 2
3231	HOXD1	HOMEODOMAIN D1
3232	HOXD3	HOMEODOMAIN D3
3344	HTLF	HUMAN T-CELL LEUKEMIA VIRUS ENHANCER FACTOR
3643	INSR	INSULIN RECEPTOR
3892	KRTHB1	KERATIN, HAIR, BASIC, 1
3892	KRTHB6	KERATIN, HAIR, BASIC, 1
3925	STMN1	STATHMIN 1/ONCOPROTEIN 18
4084	MXD1	MAX DIMERIZATION PROTEIN 1
4144	MAT2A	METHIONINE ADENOSYLTRANSFERASE II, ALPHA
4261	CIITA	CLASS II, MAJOR HISTOCOMPATIBILITY COMPLEX, TRANSACTIVATOR
4297	MLL	MYELOID/LYMPHOID OR MIXED-LINEAGE LEUKEMIA (TRITHORAX HOMOLOG, DROSOPHILA)
4831	NME2	NON-METASTATIC CELLS 2, PROTEIN (NM23B) EXPRESSED IN
4893	KRAS	V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG
4893	NRAS	V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG
4893	HRAS	V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG
4953	ODC1	ORNITHINE DECARBOXYLASE 1
4985	OPRD1	OPIOID RECEPTOR, DELTA 1
5464	PPA1	PYROPHOSPHATASE (INORGANIC) 1
5536	PPP5C	PROTEIN PHOSPHATASE 5, CATALYTIC SUBUNIT
5727	PTCH	PATCHED HOMOLOG (DROSOPHILA)
5757	PTMA	PROTHYMOSIN, ALPHA (GENE SEQUENCE 28)
6449	SGTA	SMALL GLUTAMINE-RICH TETRATRICOPEPTIDE REPEAT (TPR)-CONTAINING, ALPHA
6512	SLC1A7	SOLUTE CARRIER FAMILY 1 (GLUTAMATE TRANSPORTER), MEMBER 7
6949	TCOF1	TREACHER COLLINS-FRANCESCHETTI SYNDROME 1
7023	TFAP4	TRANSCRIPTION FACTOR AP-4 (ACTIVATING ENHANCER BINDING PROTEIN 4)
7064	THOP1	THIMET OLIGOPEPTIDASE 1
7277	TUBA1	TUBULIN, ALPHA 1 (TESTIS SPECIFIC)
7291	TWIST1	TWIST HOMOLOG 1 (ACROCEPHALOSYNDACTYLY 3; SAETHRE-CHOTZEN SYNDROME) (DROSOPHILA)
7417	VDAC2	VOLTAGE-DEPENDENT ANION CHANNEL 2
8324	FZD7	FRIZZLED HOMOLOG 7 (DROSOPHILA)
8786	RGS11	REGULATOR OF G-PROTEIN SIGNALLING 11
9274	BCL7C	B-CELL CLL/LYMPHOMA 7C
9638	FEZ1	FASCICULATION AND ELONGATION PROTEIN ZETA 1 (ZYGIN I)
9639	ARHGEF10	RHO GUANINE NUCLEOTIDE EXCHANGE FACTOR (GEF) 10
9958	USP15	UBIQUITIN SPECIFIC PEPTIDASE 15
9960	USP3	UBIQUITIN SPECIFIC PEPTIDASE 3
10171	RCL1	RNA TERMINAL PHOSPHATE CYCLASE-LIKE 1
10205	EVA1	EPITHELIAL V-LIKE ANTIGEN 1
10376	K-ALPHA-1	TUBULIN, ALPHA, UBIQUITOUS
10376	TUBA6	TUBULIN, ALPHA, UBIQUITOUS
10379	RNF31	INTERFERON-STIMULATED TRANSCRIPTION FACTOR 3, GAMMA 48KDA
10379	ISGF3G	INTERFERON-STIMULATED TRANSCRIPTION FACTOR 3, GAMMA 48KDA
10411	RAPGEF3	RAP GUANINE NUCLEOTIDE EXCHANGE FACTOR (GEF) 3
10420	TESK2	TESTIS-SPECIFIC KINASE 2
10492	SYNCRIP	SYNAPTOTAGMIN BINDING, CYTOPLASMIC RNA INTERACTING PROTEIN
10501	SEMA6B	SEMA DOMAIN, TRANSMEMBRANE DOMAIN (TM), AND CYTOPLASMIC DOMAIN, (SEMAPHORIN) 6B
10528	NOL5A	NUCLEOLAR PROTEIN 5A (56KDA WITH KKE/D REPEAT)
10642	IGF2BP1	INSULIN-LIKE GROWTH FACTOR 2 MRNA BINDING PROTEIN 1
10849	CD3EAP	CD3E ANTIGEN, EPSILON POLYPEPTIDE ASSOCIATED PROTEIN
11140	CDC37	CDC37 CELL DIVISION CYCLE 37 HOMOLOG (S. CEREVISIAE)
22919	MAPRE1	MICROTUBULE-ASSOCIATED PROTEIN, RP/EB FAMILY, MEMBER 1

22986	SORCS3	SORTILIN-RELATED VPS10 DOMAIN CONTAINING RECEPTOR 3
23082	PPRC1	PEROXISOME PROLIFERATIVE ACTIVATED RECEPTOR, GAMMA, COACTIVATOR-RELATED 1
23367	LARP1	LA RIBONUCLEOPROTEIN DOMAIN FAMILY, MEMBER 1
23415	KCNH4	POTASSIUM VOLTAGE-GATED CHANNEL, SUBFAMILY H (EAG-RELATED), MEMBER 4
23544	SEZ6L	SEIZURE RELATED GENE 6 (MOUSE)-LIKE
26135	SERBP1	SERPINE1 MRNA BINDING PROTEIN 1
27043	PELP1	PROLINE, GLUTAMIC ACID AND LEUCINE RICH PROTEIN 1
27161	EIF2C2	EUKARYOTIC TRANSLATION INITIATION FACTOR 2C, 2
27293	SMPDL3B	SPHINGOMYELIN PHOSPHODIESTERASE, ACID-LIKE 3B
29920	PYCR2	PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY, MEMBER 2
51303	FKBP11	FK506 BINDING PROTEIN 11, 19 KDA
51333	LOC51333	MESENCHYMAL STEM CELL PROTEIN DSC43
51531	C9ORF156	CHROMOSOME 9 OPEN READING FRAME 156
51621	KLF13	KRUPPEL-LIKE FACTOR 13
51747	CROP	ACID-INDUCIBLE PHOSPHOPROTEIN
55234	SMU1	SMU-1 SUPPRESSOR OF MEC-8 AND UNC-52 HOMOLOG (C. ELEGANS)
55558	PLXNA3	PLEXIN A3
55617	TASP1	TASPASE, THREONINE ASPARTASE, 1
55760	DHX32	HYPOTHETICAL PROTEIN FLJ10694
55920	RCC2	REGULATOR OF CHROMOSOME CONDENSATION 2
56895	AGPAT4	1-ACYLGLYCEROL-3-PHOSPHATE O-ACYLTRANSFERASE 4 (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, DELTA)
57597	BAHCC1	BAH DOMAIN AND COILED-COIL CONTAINING 1
57645	POGK	POGO TRANSPOSABLE ELEMENT WITH KRAB DOMAIN
58478	MASA	E-1 ENZYME
64377	CHST8	CARBOHYDRATE (N-ACETYL GALACTOSAMINE 4-O) SULFOTRANSFERASE 8
64714	PDIA2	PROTEIN DISULFIDE ISOMERASE
64919	BCL11B	B-CELL CLL/LYMPHOMA 11B (ZINC FINGER PROTEIN)
79012	CAMKV	CAM KINASE-LIKE VESICLE-ASSOCIATED
79077	XTP3TPA	XTP3-TRANSACTIVATED PROTEIN A
79663	HSPBAP1	HSPB (HEAT SHOCK 27KDA) ASSOCIATED PROTEIN 1
80323	CCDC68	COILED-COIL DOMAIN CONTAINING 68
80324	PUS1	PSEUDOURIDYLATE SYNTHASE 1
81034	SLC25A32	SOLUTE CARRIER FAMILY 25, MEMBER 32
81579	PLA2G12A	PHOSPHOLIPASE A2, GROUP XIIA
81890	QTRT1	QUEUINE TRNA-RIBOSYLTRANSFERASE 1 (TRNA-GUANINE TRANSGLYCOSYLASE)
83451	ABHD11	ABHYDROLASE DOMAIN CONTAINING 11
83855	KLF16	KRUPPEL-LIKE FACTOR 16
84258	SYT3	SYNAPTOTAGMIN III
84315	MON1A	MON1 HOMOLOG A (YEAST)
84872	ZC3H10	ZINC FINGER CCCH-TYPE CONTAINING 10
84925	DIRC2	DISRUPTED IN RENAL CARCINOMA 2
85366	MYLK2	MYOSIN LIGHT CHAIN KINASE 2, SKELETAL MUSCLE
90701	SEC11L3	SEC11-LIKE 3 (S. CEREVISIAE)
90861	C16ORF34	HYPOTHETICAL PROTEIN SIMILAR TO MOUSE HN1 (HEMATOLOGICAL AND NEUROLOGICAL EXPRESSED SEQUENCE 1)
114907	FBXO32	F-BOX PROTEIN 32
116844	LRG1	LEUCINE-RICH ALPHA-2-GLYCOPROTEIN 1
118980	SFXN2	SIDEROFLEXIN 2
127733	UBXD3	UBX DOMAIN CONTAINING 3
134510	UBLCP1	UBIQUITIN-LIKE DOMAIN CONTAINING CTD PHOSPHATASE 1
145741	NLF1	NUCLEAR LOCALIZED FACTOR 1
168620	BHLHB8	BASIC HELIX-LOOP-HELIX DOMAIN CONTAINING, CLASS B, 8
222967	LOC222967	HYPOTHETICAL PROTEIN LOC222967
342918	LOC342918	HYPOTHETICAL LOC342918
388135	LOC388135	SIMILAR TO RIKEN CDNA 6030419C18 GENE
391013	RP3-340N1.3	SIMILAR TO GROUP IIC SECRETORY PHOSPHOLIPASE A2 PRECURSOR (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE GIIC) (GIIC SPLA2) (PLA2-8) (14 KDA PHOSPHOLIPASE A2)

Seo_Supplemental Figure S10. Functional clustering of PAP-predicted target genes

A. Functional clusters for Ngn/NeuroD targets

GO term ID	GO term description	P-value
GO:0007399	nervous system development	4.43E-13
GO:0030154	cell differentiation	3.67E-11
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	5.04E-07
GO:0007275	development	4.80E-05
GO:0006468	protein amino acid phosphorylation	5.26E-05
GO:0008038	neuron recognition	6.60E-05
GO:0006813	potassium ion transport	1.16E-04
GO:0007219	Notch signaling pathway	1.65E-04
GO:0030326	embryonic limb morphogenesis	2.25E-04
GO:0001709	cell fate determination	3.54E-04
GO:0007155	cell adhesion	5.44E-04
GO:0045786	negative regulation of progression through cell cycle	5.94E-04
GO:0007215	glutamate signaling pathway	1.00E-03
GO:0046928	regulation of neurotransmitter secretion	1.07E-03
GO:0007411	axon guidance	1.18E-03
GO:0009953	dorsal/ventral pattern formation	2.11E-03
GO:0050793	regulation of development	2.11E-03
GO:0050772	positive regulation of axonogenesis	2.11E-03
GO:0007417	central nervous system development	3.21E-03
GO:0007605	sensory perception of sound	4.32E-03
GO:0043433	negative regulation of transcription factor activity	5.13E-03
GO:0007223	frizzled-2 signaling pathway	7.01E-03
GO:0030097	Hemopoiesis	7.01E-03
GO:0007179	transforming growth factor beta receptor signaling pathway	9.08E-03
GO:0008016	regulation of heart contraction	9.08E-03
GO:0005513	detection of calcium ion	9.34E-03

B. Functional clusters for E-box (M01034) targets

GO term ID	GO term description	P-value
GO:0007275	development	4.07E-06
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	1.08E-04
GO:0046685	response to arsenic	3.91E-04
GO:0006350	transcription	3.98E-04
GO:0030520	estrogen receptor signaling pathway	5.92E-04
GO:0007223	frizzled-2 signaling pathway	6.93E-04
GO:0015986	ATP synthesis coupled proton transport	1.10E-03
GO:0009953	Dorsal/ventral pattern formation	2.29E-03
GO:0015992	Proton transport	3.02E-03
GO:0030154	cell differentiation	3.35E-03
GO:0006366	transcription from RNA polymerase II promoter	3.45E-03
GO:0006367	transcription initiation from RNA polymerase II promoter	4.39E-03
GO:0006904	vesicle docking during exocytosis	5.90E-03
GO:0006836	neurotransmitter transport	7.20E-03
GO:0006281	DNA repair	7.69E-03
GO:0030326	embryonic limb morphogenesis	7.69E-03

C. Functional clusters for c-Myc/Max (M00118) targets

GO term ID	GO term description	P-value
GO:0007275	development	3.13E-03
GO:0000074	regulation of progression through cell cycle	4.44E-03
GO:0000083	G1/S-specific transcription in mitotic cell cycle	6.98E-03
GO:0006535	cysteine biosynthesis from serine	6.98E-03
GO:0006978	DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	6.98E-03
GO:0007019	microtubule depolymerization	6.98E-03
GO:0030887	positive regulation of dendritic cell activation	6.98E-03
GO:0035162	embryonic hemopoiesis	6.98E-03
GO:0042759	long-chain fatty acid biosynthesis	6.98E-03
GO:0051290	protein heterotetramerization	6.98E-03
GO:0046600	negative regulation of centriole replication	6.98E-03
GO:0043109	regulation of smoothened activity	6.98E-03
GO:0040015	negative regulation of body size	6.98E-03
GO:0031115	negative regulation of microtubule polymerization	6.98E-03
GO:0019343	cysteine biosynthesis via cystathione	6.98E-03
GO:0006163	Purine nucleotide metabolism	6.98E-03
GO:0008652	Amino acid biosynthesis	9.29E-03

D. Ngn/NeuroD target genes in development/neural development functional clusters

Nervous system development (45 Genes)	
ENTREZ_GENE_ID	Gene name
182	JAGGED 1 (ALAGILLE SYNDROME)
238	ANAPLASTIC LYMPHOMA KINASE (KI-1)
1131	CHOLINERGIC RECEPTOR, MUSCARINIC 3
1600	DISABLED HOMOLOG 1 (DROSOPHILA)
1630	DELETED IN COLORECTAL CARCINOMA
1641	DOUBLECORTEX; LISSENCEPHALY, X-LINKED (DOUBLECORTIN)
1910	ENDOTHELIN RECEPTOR TYPE B
2047	EPH RECEPTOR B1
2048	EPH RECEPTOR B2
2596	GROWTH ASSOCIATED PROTEIN 43
4010	LIM HOMEBOX TRANSCRIPTION FACTOR 1, BETA
4045	LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN
4684	NEURAL CELL ADHESION MOLECULE 1
4692	NECDIN HOMOLOG (MOUSE)
4761	NEUROGENIC DIFFERENTIATION 2
4807	NESCIENT HELIX LOOP HELIX 1
4808	NESCIENT HELIX LOOP HELIX 2
4853	NOTCH HOMOLOG 2 (DROSOPHILA)
4897	NEURONAL CELL ADHESION MOLECULE
4916	NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 3
5076	PAIRED BOX GENE 2
5087	PRE-B-CELL LEUKEMIA TRANSCRIPTION FACTOR 1
5453	POU DOMAIN, CLASS 3, TRANSCRIPTION FACTOR 1
6091	ROUNDABOUT, AXON GUIDANCE RECEPTOR, HOMOLOG 1 (DROSOPHILA)
6586	SLIT HOMOLOG 3 (DROSOPHILA)
7903	ST8 ALPHA-N-ACETYL-NEURAMINIDE ALPHA-2,8-SIALYLTRANSFERASE 4
8326	FRIZZLED HOMOLOG 9 (DROSOPHILA)
8522	GROWTH ARREST-SPECIFIC 7
8633	UNC-5 HOMOLOG C (C. ELEGANS)
8828	NEUROFILIN 2
8829	NEUROFILIN 1
9353	SLIT HOMOLOG 2 (DROSOPHILA)
9378	NEUREXIN 1
9423	NETRIN 1
9839	ZINC FINGER HOMEBOX 1B
9935	V-MAF MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE HOMOLOG B (AVIAN)
10178	ODZ, ODD OZ/TEN-M HOMOLOG 1(DROSOPHILA)
10371	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3A
26280	INTERLEUKIN 1 RECEPTOR ACCESSORY PROTEIN-LIKE 2
26508	HAIRY/ENHANCER-OF-SPLIT RELATED WITH YRPW MOTIF-LIKE
28514	DELTA-LIKE 1 (DROSOPHILA)
58158	NEUROGENIC DIFFERENTIATION 4
64211	LIM HOMEBOX 5
80031	KIAA1479 PROTEIN
146760	RETICULON 4 RECEPTOR-LIKE 1

Development cluster (88 Genes)	
ENTREZ_GENE_ID	Gene name
182	JAGGED 1 (ALAGILLE SYNDROME)
238	ANAPLASTIC LYMPHOMA KINASE (KI-1)
488	ATPASE, CA++ TRANSPORTING, CARDIAC MUSCLE, SLOW TWITCH 2
655	BONE MORPHOGENETIC PROTEIN 7 (OSTEOGENIC PROTEIN 1)
845	CALSEQUESTRIN 2 (CARDIAC MUSCLE)
863	CORE-BINDING FACTOR, RUNT DOMAIN, ALPHA SUBUNIT 2; TRANSLOCATED TO, 3
1131	CHOLINERGIC RECEPTOR, MUSCARINIC 3
1295	COLLAGEN, TYPE VIII, ALPHA 1
1523	CUT-LIKE 1, CCAAT DISPLACEMENT PROTEIN (DROSOPHILA)
1600	DISABLED HOMOLOG 1 (DROSOPHILA)
1630	DELETED IN COLORECTAL CARCINOMA
1641	DOUBLECORTEX; LISSENCEPHALY, X-LINKED (DOUBLECORTIN)
1879	EARLY B-CELL FACTOR
1896	ECTODYSPLASIN A
1910	ENDOTHELIN RECEPTOR TYPE B
2047	EPH RECEPTOR B1
2048	EPH RECEPTOR B2
2066	V-ERB-A ERYTHROBLASTIC LEUKEMIA VIRAL ONCOGENE HOMOLOG 4 (AVIAN)
2104	ESTROGEN-RELATED RECEPTOR GAMMA
2138	EYES ABSENT HOMOLOG 1 (DROSOPHILA)
2596	GROWTH ASSOCIATED PROTEIN 43

2719	GLYPICAN 3
3397	INHIBITOR OF DNA BINDING 1, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN
3399	INHIBITOR OF DNA BINDING 3, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN
3479	INSULIN-LIKE GROWTH FACTOR 1 (SOMATOMEDIN C)
4010	LIM HOMEBOX TRANSCRIPTION FACTOR 1, BETA
4045	LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN
4209	MADS BOX TRANSCRIPTION ENHANCER FACTOR 2, POLYPEPTIDE D (MYOCYTE ENHANCER FACTOR 2D)
4593	MUSCLE, SKELETAL, RECEPTOR TYROSINE KINASE
4654	MYOGENIC DIFFERENTIATION 1
4684	NEURAL CELL ADHESION MOLECULE 1
4692	NECDIN HOMOLOG (MOUSE)
4761	NEUROGENIC DIFFERENTIATION 2
4807	NESSCIENT HELIX LOOP HELIX 1
4808	NESSCIENT HELIX LOOP HELIX 2
4851	NOTCH HOMOLOG 1, TRANSLOCATION-ASSOCIATED (DROSOPHILA)
4853	NOTCH HOMOLOG 2 (DROSOPHILA)
4897	NEURONAL CELL ADHESION MOLECULE
4916	NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 3
5076	PAIRED BOX GENE 2
5087	PRE-B-CELL LEUKEMIA TRANSCRIPTION FACTOR 1
5362	PLEXIN A2
5453	POU DOMAIN, CLASS 3, TRANSCRIPTION FACTOR 1
5629	PROSPERO-RELATED HOMEBOX 1
5727	PATCHED HOMOLOG (DROSOPHILA)
6091	ROUNDAABOUT, AXON GUIDANCE RECEPTOR, HOMOLOG 1 (DROSOPHILA)
6468	F-BOX AND WD-40 DOMAIN PROTEIN 4
6586	SLIT HOMOLOG 3 (DROSOPHILA)
7090	TRANSDUCIN-LIKE ENHANCER OF SPLIT 3 (E(SP1) HOMOLOG, DROSOPHILA)
7482	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 2B
7483	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 9A
7903	ST8 ALPHA-N-ACETYL-NEURAMINIDE ALPHA-2,8-SIALYLTRANSFERASE 4
8038	ADAM METALLOPEPTIDASE DOMAIN 12 (MELTRIN ALPHA)
8326	FRIZZLED HOMOLOG 9 (DROSOPHILA)
8522	GROWTH ARREST-SPECIFIC 7
8633	UNC-5 HOMOLOG C (C. ELEGANS)
8828	NEUROFILIN 2
8829	NEUROFILIN 1
8840	WNT1 INDUCIBLE SIGNALING PATHWAY PROTEIN 1
9353	SLIT HOMOLOG 2 (DROSOPHILA)
9378	NEUREXIN 1
9423	NETRIN 1
9555	H2A HISTONE FAMILY, MEMBER Y
9839	ZINC FINGER HOMEBOX 1B
9935	V-MAF MUSCULOAPONEUROTIC FIBROSARCOMA ONCOGENE HOMOLOG B (AVIAN)
10178	ODZ, ODD OZ/TEN-M HOMOLOG 1(DROSOPHILA)
10194	SEROLOGICALLY DEFINED COLON CANCER ANTIGEN 33
10371	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3A
10435	CDC42 EFFECTOR PROTEIN (RHO GTPASE BINDING) 2
10495	CYTOSOLIC OVARIAN CARCINOMA ANTIGEN 1
10891	PEROXISOME PROLIFERATIVE ACTIVATED RECEPTOR, GAMMA, COACTIVATOR 1, ALPHA
26280	INTERLEUKIN 1 RECEPTOR ACCESSORY PROTEIN-LIKE 2
26508	HAIRY/ENHANCER-OF-SPLIT RELATED WITH YRPW MOTIF-LIKE
28514	DELTA-LIKE 1 (DROSOPHILA)
54567	DELTA-LIKE 4 (DROSOPHILA)
56920	SEMA DOMAIN, IMMUNOGLOBULIN DOMAIN (IG), SHORT BASIC DOMAIN, SECRETED, (SEMAPHORIN) 3G
57616	ZINC FINGER PROTEIN 537
58158	NEUROGENIC DIFFERENTIATION 4
63951	DMRT-LIKE FAMILY A1
64211	LIM HOMEBOX 5
64411	CENTAURIN, DELTA 3
79960	PHD FINGER PROTEIN 17
80031	KIAA1479 PROTEIN
80326	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 10A
89796	HYPOTHETICAL PROTEIN MGC14961
117154	DACHSHUND HOMOLOG 2 (DROSOPHILA)
146760	RETICULON 4 RECEPTOR-LIKE 1
200734	SPROUTY-RELATED, EVH1 DOMAIN CONTAINING 2

E. E-box M01034 target genes in top-ranked functional cluster

Development cluster (30 genes)	
ENTREZ_GENE_ID	Gene name
1627	DREBRIN 1
1952	CADHERIN, EGF LAG SEVEN-PASS G-TYPE RECEPTOR 2 (FLAMINGO HOMOLOG, DROSOPHILA)
1995	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 3 (HU ANTIGEN C)
2251	FIBROBLAST GROWTH FACTOR 6
3200	HOMEODOMAIN PROTEIN A3
3203	HOMEODOMAIN PROTEIN A6
3204	HOMEODOMAIN PROTEIN A7
3214	HOMEODOMAIN PROTEIN B4
3217	HOMEODOMAIN PROTEIN B7
3397	INHIBITOR OF DNA BINDING 1, DOMINANT NEGATIVE HELIX-LOOP-HELIX PROTEIN
3651	INSULIN PROMOTER FACTOR 1, HOMEODOMAIN TRANSCRIPTION FACTOR
4010	LIM HOMEODOMAIN TRANSCRIPTION FACTOR 1, BETA
4760	NEUROGENIC DIFFERENTIATION 1
4761	NEUROGENIC DIFFERENTIATION 2
4821	NK2 TRANSCRIPTION FACTOR RELATED, LOCUS 2 (DROSOPHILA)
5076	PAIRED BOX GENE 2
5620	PROTAMINE 2
7142	TRANSITION PROTEIN 2 (DURING HISTONE TO PROTAMINE REPLACEMENT)
7477	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 7B
7483	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 9A
7866	INTERFERON-RELATED DEVELOPMENTAL REGULATOR 2
10501	SEMA DOMAIN, TRANSMEMBRANE DOMAIN (TM), AND CYTOPLASMIC DOMAIN, (SEMAPHORIN) 6B
51043	ZINC FINGER AND BTB DOMAIN CONTAINING 7B
51330	TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY, MEMBER 12A
54361	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 4
54567	DELTA-LIKE 4 (DROSOPHILA)
80326	WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 10A
81693	AMNIONLESS HOMOLOG (MOUSE)
81848	SPROUTY HOMOLOG 4 (DROSOPHILA)
145258	GOOSECOID

F. Myc/Max target-genes in top-ranked functional clusters

All cell cycle cluster (14 Genes)	
ENTREZ_GENE_ID	Gene name
595	CYCLIN D1
672	BREAST CANCER 1, EARLY ONSET
2251	FIBROBLAST GROWTH FACTOR 6
2354	FBJ MURINE OSTEOSARCOMA VIRAL ONCOGENE HOMOLOG B
2672	GROWTH FACTOR INDEPENDENT 1
3925	STATHMIN 1/ONCOPROTEIN 18
4831	NON-METASTATIC CELLS 2, PROTEIN (NM23B) EXPRESSED IN
4893	V-HA-RAS HARVEY RAT SARCOMA VIRAL ONCOGENE HOMOLOG
5536	PROTEIN PHOSPHATASE 5, CATALYTIC SUBUNIT
5727	PATCHED HOMOLOG (DROSOPHILA)
5757	PROTHYMOSIN, ALPHA (GENE SEQUENCE 28)
11140	CDC37 CELL DIVISION CYCLE 37 HOMOLOG (S. CEREVISIAE)
22919	MICROTUBULE-ASSOCIATED PROTEIN, RP/EB FAMILY, MEMBER 1
55920	REGULATOR OF CHROMOSOME CONDENSATION 2

Development cluster (8 Genes)	
ENTREZ_GENE_ID	Gene name
1995	ELAV (EMBRYONIC LETHAL, ABNORMAL VISION, DROSOPHILA)-LIKE 3 (HU ANTIGEN C)
2251	FIBROBLAST GROWTH FACTOR 6
3231	HOMEODOMAIN PROTEIN D1
3232	HOMEODOMAIN PROTEIN D3
7291	TWIST HOMOLOG 1 (ACROCEPHALOSYNDACTYLY 3; SAETHRE-CHOTZEN SYNDROME) (DROSOPHILA)
8324	FRIZZLED HOMOLOG 7 (DROSOPHILA)
10501	SEMA DOMAIN, TRANSMEMBRANE DOMAIN (TM), AND CYTOPLASMIC DOMAIN, (SEMAPHORIN) 6B
57645	POGO TRANSPOSABLE ELEMENT WITH KRAB DOMAIN