Supporting Text

SAGE Libraries and SAGE Tag Extraction. Hypothalamus SAGE libraries were made according to SAGE protocol, Version 1.0d (1) with some modifications (1-4). For the neocortex, mRNA from female mice was obtained and data from three independent libraries (2 right, 1 left) were combined and normalized. All tags representing linkers were removed. SAGE tag extraction was performed by using either the program SAGE2000 (www.sagenet.org) or specifically developed software (5) from DNA sequence output files (Applied Biosystems). Whereas tags from the hypothalamus averaged 14 nt in length, improvements to SAGE technology allowed 21-nt tags (LongSAGE, ref. 6) to be obtained from the neocortex from which short tags (14 nt) were subsequently extracted.

Statistical Analysis of SAGE Data. A Fisher's exact test was used to identify differentially expressed genes between the hypothalamic and neocortical libraries. Given the counts for a certain tag and the sum of all tags present in the libraries, Fisher's exact test computed the *P* value that these counts could have been observed by chance if the tag was equally represented in the libraries under comparison. The Benjamini and Hochberg correction (7) was used to control the false-discovery rate associated with a large-scale multiple-testing environment. Tags with a corrected *P* value <0.10 were considered to be significant and were differentially expressed between libraries. More detail is provided in *Appendix*.

GO Database Annotations. To link tag identity with putative gene function, all identified genes were annotated by using the Mouse Genome Informatics (www.informatics.jax.org) and GO (www.geneontology.org) databases. These annotations were separately considered under three main classifications: cellular component, biological process and molecular function. Each of these main classifications contain various descriptors (or GO database terms) of gene localization or gene function. For each GO database term, all associated genes from hypothalamus and neocortex were separately classified, and a combined abundance value was computed for the GO database term in question (after correction for false-discovery rate, ref. 7). Only GO-selected terms with significant differences between hypothalamus and neocortex (χ^2 test, $P < 0.05$) are displayed (8).

SAGE Tag Mapping to Genes and Chromosomal Positions. A comparative chromosomal view of expression levels in hypothalamus and neocortex was generated by assigning chromosomal positions for tags corresponding to known genes after matching to the RefSeq database at the National Center for Biotechnology Information (February 2003 release; Mm3 at http://genome.ucsc.edu). Where a tag matches several genes, a unique identity was assigned using the ranking priority by the RefSeq database: NM (curated mRNA) > XM (not-curated, mRNA). In cases of multiple matches within the same category, no identity was assigned. Tag counts were normalized to 100,000 tags per library. To correlate tag abundances with gene-poor and gene-rich chromosomal regions, *in silico* reconstruction of Giemsa bands in the mouse genome was carried out using the

methodology previously applied for the human genome (9). To detect differential expression between the two tissues, a series of data points was generated (at $\log_2$ scale) to depict the ratio of expression levels between the two tissues. By using a 10-Mb window, all tags present for the hypothalamus (or neocortex) were summed to yield a data point at the midpoint of the 10-Mb window. Serial data points for each library were generated along each chromosome by moving the window forward by one tag position at a time. To compute for the significance of these ratios, an identical process was performed on a randomized gene order 10,000 times.

1. Margulies, E. H., Kardia, S. L. R. & Innis, J. W. (2001) *Nucleic Acids Res.* **29,** e60 (1-8).

2. Gunnersen, J. M., Augustine, C., Spirkoska, V., Kim, M., Brown, M. & Tan, S. S. (2002) *Mol. Cell. Neurosci.* **19,** 560-573.

3. Powell, J. (1998) *Nucleic Acids Res.* **26,** 3445-3446.

4. Kenzelmann, M. & Mühlemann, K. (1999) *Nucleic Acids Res.* **27,** 917-918.

5. Beissbarth, T., Hyde, L., Smyth, G. K., Job, C., Boon, W. M., Tan, S. S., Scott, H. S. & Speed, T. P. (2004) *Bioinformatics* **20,** Suppl. 1, I31-I39.

6. Saha, S., Sparks, A. B., Rago, C., Akmaev, V., Wang, C. J., Vogelstein, B., Kinzler, K. W. & Velculescu, V. E. (2002) *Nat. Biotechnol.* **20,** 508-512.

7. Benjamini, Y. & Hochberg, Y. (1995) *J. R. Stat. Soc. B* **57,** 289-300.

8. Beissbarth, T. & Speed, T.P. (2004) *Bioinformatics* **20,** 1464-1465.

9. Niimura, Y & Gojobori, T. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 797-802.