nature neuroscience

Molecular taxonomy of major neuronal classes in the adult mouse forebrain

Ken Sugino[1,3], Chris M Hempel[1,3], Mark N Miller[1], Alexis M Hattox[1], Peter Shapiro[1], Caizi Wu[2], Z Josh Huang[2] & Sacha B Nelson[1]

Identifying the neuronal cell types that comprise the mammalian forebrain is a central unsolved problem in neuroscience. Global gene expression profiles offer a potentially unbiased way to assess functional relationships between neurons. Here, we carried out microarray analysis of 12 populations of neurons in the adult mouse forebrain. Five of these populations were chosen from cingulate cortex and included several subtypes of GABAergic interneurons and pyramidal neurons. The remaining seven were derived from the somatosensory cortex, hippocampus, amygdala and thalamus. Using these expression profiles, we were able to construct a taxonomic tree that reflected the expected major relationships between these populations, such as the distinction between cortical interneurons and projection neurons. The taxonomic tree indicated highly heterogeneous gene expression even within a single region. This dataset should be useful for the classification of unknown neuronal subtypes, the investigation of specifically expressed genes and the genetic manipulation of specific neuronal circuit elements.

The mammalian forebrain is a tissue of stunning complexity comprised of tens to hundreds of areas, each containing a comparable number of distinct cell types. The identification and classification of these many cell types has long been recognized as a prerequisite to understanding brain function. For example, an enduring question in the study of the neocortex is the degree to which a single canonical circuit comprised of a set of canonical cell types can be recognized across cortical areas[1]. A major impediment to answering such questions has been the difficulty of objectively defining cell types. Cell types have traditionally been defined on the basis of a wide variety of characteristics including anatomical location, intrinsic firing properties, synaptic physiology, somatodendritic and axonal morphology, and the presence or absence of particular marker genes, such as those encoding neuropeptides and calcium binding proteins[2,3]. However, these efforts have not been able to provide a consensus view on the division of neurons into functional subtypes[4].

It has been suggested that global gene expression profiling could provide a useful alternative strategy for the identification and characterization of neuronal subtypes[3,4]. Here, we implemented this approach by using DNA microarrays to reveal global expression differences among 12 distinct neuronal populations. These were selected to span major recognized categories of forebrain neurons, including excitatory projection neurons and inhibitory interneurons. They were also chosen to permit comparisons within and between neocortical regions, between major subdivisions of the telencephalon including the neocortex, hippocampus and amygdala, and between the diencephalon and telencephalon. The results reveal a hitherto unappreciated degree of molecular heterogeneity among forebrain neurons and permit the construction of an objective neuronal taxonomy on the basis of gene expression.

RESULTS

Twelve neuronal subpopulations in the mouse forebrain

We obtained 11 populations of fluorescently labeled neurons from four transgenic mouse lines in which green or yellow fluorescent protein (GFP and YFP, respectively) was expressed in distinct subsets of neurons. These lines were (i) YFPH: YFP expressed under control of the *Thy1* promoter[5], (ii) GIN: GFP under a 2.8 kilobase (kb) *Gad1* promoter[6], (iii) G42: GFP under a *Gad1* bacterial artificial chromosome (BAC) clone[7] and (iv) G30: GFP under a 5.5 kb *Gad2* promoter[8]. A 12th population was labeled in cingulate cortex by the injection of a retrograde tracer into the medial dorsal nucleus of the thalamus. We refer to each population using an abbreviated mouse strain name (or the prefix CT6 for the retrogradely labeled corticothalamic neurons in layer 6) followed by the brain region from which the population was taken (**Fig. 1** and **Table 1**).

In the cingulate cortex (CG), the five populations represented three major classes of interneurons[2,9] (parvalbumin-positive fast-spiking basket cells, somatostatin-positive regular-spiking bipolar and multipolar cells, and cholecystokinin-positive irregular-spiking bipolar and multipolar cells) and two major classes of projection neurons (thick-tufted layer 5 nonadapting pyramidal neurons and layer 6 adapting corticothalamic neurons; **Fig. 1**). For clarity and consistency, all genes will hereafter be referred to by their NCBI gene symbol (for example, *Pvalb*, *Sst* and *Cck* encode the peptides parvalbumin, somatostatin and

[1]Department of Biology and National Center for Behavioral Genomics, Brandeis University, MS 008, 415 South Street, Waltham, Massachusetts 02454-9110, USA. [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to S.B.N. (nelson@brandeis.edu).

Received 6 October; accepted 22 November; published online 20 December 2005; doi:10.1038/nn1618

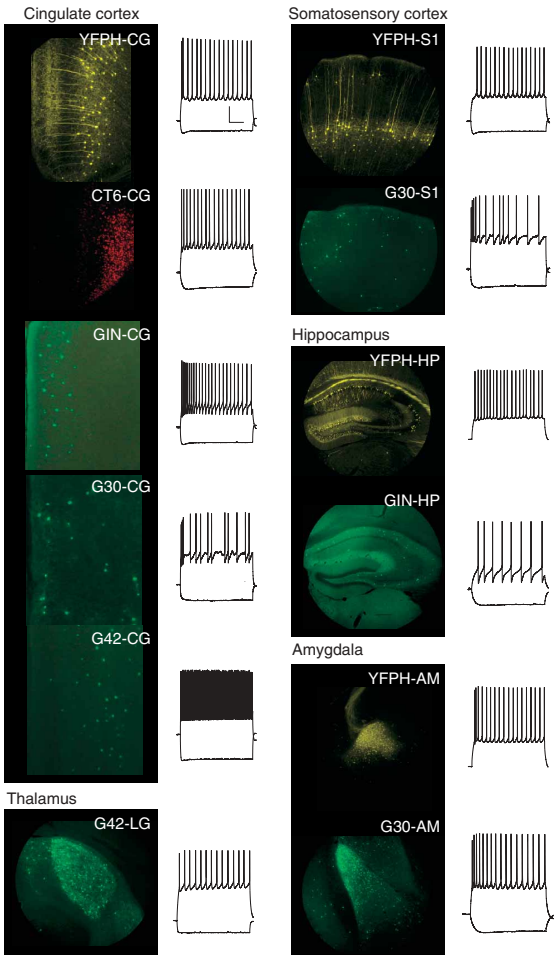

Figure 1 Twelve neuronal populations. Micrographs and representative current-clamp recordings of the neuronal populations studied. See main text for the naming convention. Scale bar shown with YFPH-CG current-clamp trace is 20 mV, 200 ms and applies to all traces.

cholecystokinin, respectively). These five populations were almost entirely nonoverlapping on the basis of marker expression (**Supplementary Note** online), laminar distributions and distinct firing properties (**Table 1**) that were reproducible within cell populations and across mice (**Supplementary Note**). The two populations in primary somatosensory cortex (S1), YFPH-S1 and G30-S1, were very similar to their CG counterparts in laminar distribution, morphology and electrophysiological properties. The remaining five populations were from outside the neocortex and included two from the hippocampus (HP), two from the amygdala (AM) and one from the lateral geniculate nucleus of the thalamus (LG).

On the basis of shared properties, we operationally defined each of these cell populations as a cell type. Within a cell type, there was reduced variability of anatomical, physiological and biochemical characteristics relative to the entire population of neurons. This definition of the cell type was intrinsically hierarchical: each cell type belonged to a larger, functionally defined cell type (for example, interneurons and pyramidal neurons) and may have also comprised further recognizable subtypes. The mechanisms restricting transgene expression to these particular neuronal populations are unknown and presumably result from the interaction between the transgenic promoters and the genomic integration sites.

Gene expression profiles from neuronal subpopulations

mRNA obtained from 27–120 hand-sorted neurons was reverse transcribed, amplified, labeled and hybridized to Affymetrix microarrays (MOE430A) containing 22,690 probe sets representing 13,232 distinct transcripts. These included 12,556 (90.1%) known genes (including 1750 Riken Genes) and 676 (4.9%) expressed sequence tags (ESTs). In our analysis, we did not include 233 (1.7%) nonspecific probes and 467 (3.3%) nonannotated probes (that is, no symbol, LocusLink or GeneBank information in the latest (2005/03/31) Affymetrix annotation). Each neuronal population was profiled with three biological replicates from different mice. The results were highly reproducible (average correlation coefficient between replicate pairs was 0.982 ± 0.005; mean ± s.d.). Genes associated with non-neuronal brain cells, such as glia and red blood cells, were detected at much lower levels in sorted neurons than in homogenized brain tissue (**Supplementary Note**).

We identified specific markers for the profiled neuronal populations (**Fig. 2**). Many of these were expressed in a manner consistent with previously published observations. For example, the ionotropic serotonin receptor *Htr3a* and the cannabinoid receptor *Cnr1* are known to be present in irregular-spiking cholecystokinin (CCK)-positive interneurons[10,11] and were enriched in G30-CG. *Tac1*, the precursor for substance P, is known to be expressed in *Pvalb*-positive basket cells[12] and was found in G42-CG. Three transcripts known to be enriched in corticospinal neurons (*Crym*, *Pcp4* and *Diap3*)[13] were also enriched in YFPH-CG. However, most of the differentially expressed genes discovered were not previously identified with neuronal function. The identification of the cell types in which these genes are specifically expressed should aid in elucidating their function.

Many genes were expressed in patterns that defined higher-order cell types. For example, we identified a number of genes known to be differentially expressed in glutamatergic neurons and GABAergic neurons, including the genes encoding the GABA transporters Slc32a1 and Slc6a1, the GABA synthesis enzymes Gad1 and Gad2, the kainate receptor Grik1 (ref. 14), the scaffolding protein Grip1 (ref. 15), the Dlx family of transcription factors[16–18], Arx[19,20] and the vesicular glutamate transporter Slc17a6 (Vglut2). In some cases, these genes were expressed in all GABAergic neurons tested; in others, they were expressed only in the telencephalic GABAergic neurons (for example, *Arx* and *Dlx1* were not expressed in the diencephalic G42-LG). To the best of our knowledge, we correctly identified all genes

Table 1 Descriptions of the 12 neuronal populations

Population designation	Brain region	Subregion or cortical layer	Neurotransmitter	Firing properties
CT6-CG	Cingulate cortex	Layer 6	Glu	A (4/5)
G30-AM	Amygdala	Lateral	Glu	Mixed
G30-CG	Cingulate cortex	Layers 1–6	GABA	IS (4/4)
G30-S1	Somatosensory cortex	Layers 1–6	GABA	IS (3/3)
G42-CG	Cingulate cortex	Layers 4–6	GABA	FS, NA (6/6)
G42-LG	Thalamus	Dorsal LGN	GABA	Mixed
GIN-CG	Cingulate cortex	Layers 2–4	GABA	A (21/22)
GIN-HP	Hippocampus	CA1-CA3	GABA	A (3/3)
YFPH-AM	Amygdala	Basolateral	Glu	D, NA (4/4)
YFPH-CG	Cingulate cortex	Layer 5	Glu	D, NA (5/5)
YFPH-HP	Hippocampus	CA1	Glu	NA (3/3)
YFPH-S1	Somatosensory cortex	Layers 5,6	Glu	NA (4/4)

See Methods for definitions of firing properties. A, adapting; IS, irregular spiking; FS, fast spiking; NA, nonadapting; D, doublet firing; Glu, glutamate.

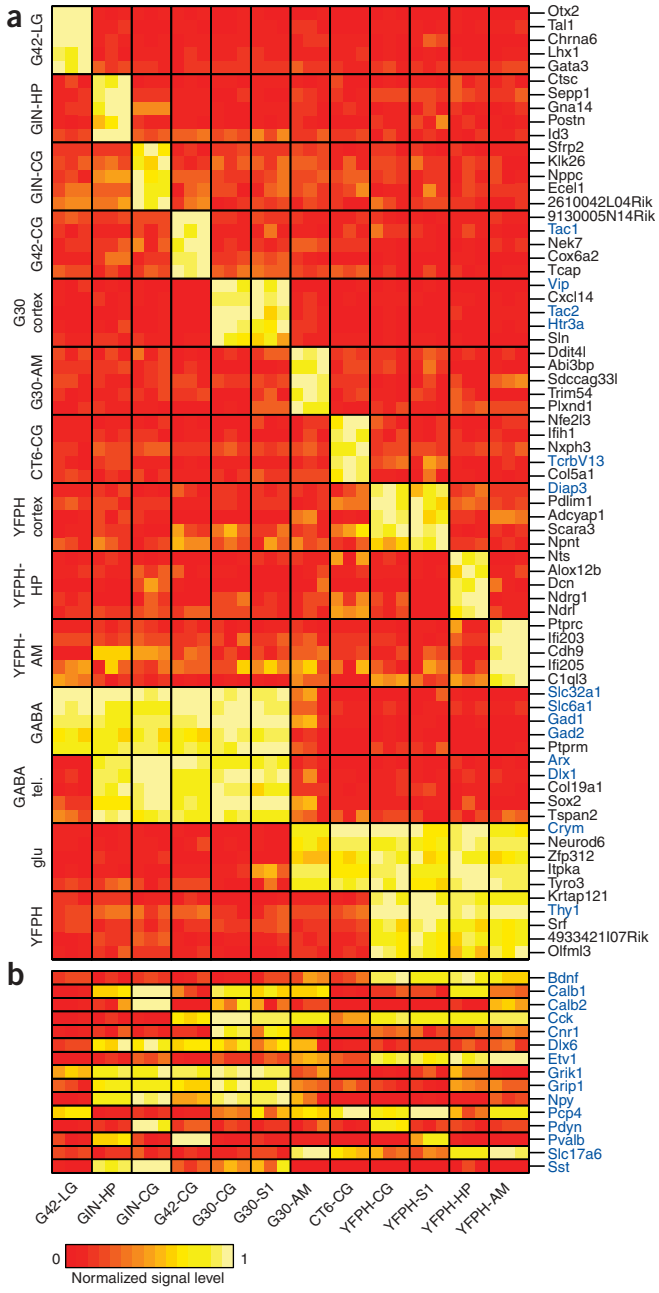

Figure 2 Selected differentially expressed genes among 12 neuronal populations. (**a**) Expression levels for selected differentially expressed genes. For each gene, the Affymetrix signal level is plotted for 36 samples (three replicates across 12 populations). Signal levels are normalized from minimum (dark red) to maximum (bright yellow) values for each gene. Genes were selected on the basis of which of 14 templates they matched best (see Methods); the five best matched genes are shown for each template. Labels on the left describe these templates, either by the name of a single population or by a description of a set of populations (G30 cortex, both neocortical populations derived from line G30; YFPH cortex, both neocortical populations derived from line YFPH; GABA, GABAergic neurons; glu, glutamatergic neurons; GABA tel., GABAergic telencephalic neurons; YFPH, all populations derived from line YFPH). Genes whose expression pattern is consistent with previously published observations are in blue. (**b**) Expression patterns of transcripts previously found to be present in at least one of the populations studied, but that did not meet the template match criteria in **a**.

that were previously shown to be universally expressed in GABAergic or glutamatergic neurons and that were represented on the microarray; this demonstrated the sensitivity and accuracy of our expression analysis. Our results also identified numerous genes not previously known to be differentially expressed between these two major neuronal classes in the adult. Among these were genes encoding transcription factors such as Sox2 and Neurod6 and signaling molecules such as Ptprm, Itpka and Tyro3.

To assess whether our ability to detect rare transcripts was limited by the small number of cells used, we plotted the number of transcripts scored as present (by Affymetrix MAS 5.0 software) against the number of sorted cells used. We found that the 'present' count increased with the number of cells, from 1 (27% present, $n = 2$) to 5 (31% present, $n = 2$) to 30 or more (46% present, $n = 42$). However, further increases in the number of cells—beyond 30 and even up to 500—did not increase the number of transcripts scored as present (**Supplementary Note**). This indicated that 30 cells were sufficient to reliably measure the population transcriptome. It should be noted that this general relationship was calculated using pooled expression data from all 12 profiled cell types. It is possible that the precise relationship varied slightly from cell type to cell type so that some cell types required more cells to reach maximum present count than did others.

We tested the accuracy of our microarray expression data using four independent methods: quantitative real-time polymerase chain reaction (*q*PCR), *in situ* hybridization (ISH), immunocytochemistry (ICC) and electrophysiological recording. Relative transcript levels measured by *q*PCR were strongly correlated with those measured by microarray (**Supplementary Note**; correlation coefficient = 0.93, $n = 19$). However, on average, the microarray underestimated fold change (defined as the ratio of signal levels) by a factor of 4.5. ISH confirmed that three of the transcripts enriched in YFPH-CG, namely *Igfbp4* (**Supplementary Note**), *Ckmt1* and *Nefh*, were elevated in layer 5 of the cingulate cortex. ICC confirmed that the products of three known marker genes (*Sst*, *Pvalb* and *Vip*) and *Ank1* were expressed in the populations predicted by the microarray (**Supplementary Note**). Finally, the microarray correctly predicted the relative magnitude of two ionic currents in populations G30-CG and G42-CG (**Fig. 3**): the apamin-sensitive intermediate calcium-dependent potassium current (I_{SK}) to which the potassium channel subunit Kcnn2 contributes[21], and the rapidly activating hyperpolarization-activated cation current (I_h) to which the cationic channel subunit Hcn1 contributes[22].

Because cell types were, for the most part, assayed in different lines of mice, it is possible that differences in genetic background contributed to the observed differences in gene expression. To test this directly, we profiled a single anatomically defined cell type (CT6-CG) in four different transgenic lines (YFPH, $n = 3$; GIN, $n = 1$; G42, $n = 1$; and G30, $n = 1$). The average correlation coefficient between these samples (mean of the pairwise comparisons, 15 pairs) was not significantly different from the average correlation coefficient between all replicate samples (mean ± s.d.: 0.979 ± 0.006 versus 0.982 ± 0.005; $P = 0.093$, two-tailed *t*-test). In comparison, the average correlation coefficient between CT6-CG and the other populations was 0.946 ± 0.012 (198 pairs) and was significantly different from the correlation coefficients within the CT6-CG samples ($P = 1.6 \times 10^{-14}$, two-tailed *t*-test). This demonstrated that for at least one population (CT6-CG) and among the strains used for these experiments (see Methods), the influence of strain background on expression profile was small relative to the contribution of cell type.

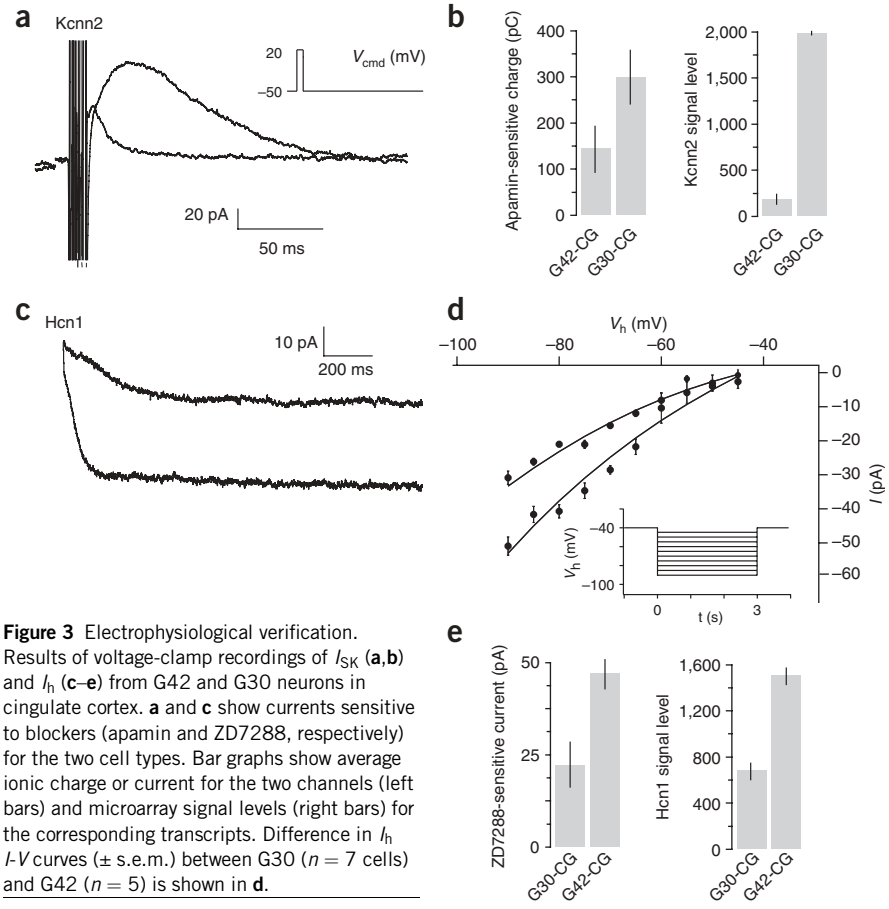

Figure 3 Electrophysiological verification. Results of voltage-clamp recordings of I_{SK} (**a**,**b**) and I_h (**c**–**e**) from G42 and G30 neurons in cingulate cortex. **a** and **c** show currents sensitive to blockers (apamin and ZD7288, respectively) for the two cell types. Bar graphs show average ionic charge or current for the two channels (left bars) and microarray signal levels (right bars) for the corresponding transcripts. Difference in I_h *I-V* curves (± s.e.m.) between G30 ($n = 7$ cells) and G42 ($n = 5$) is shown in **d**.

Heterogeneous gene expression across neuronal cell types

Although cell types have been traditionally defined in part by their differential expression of a handful of marker genes, the genome-wide extent of differential gene expression is unknown. To estimate this, we performed an analysis of variance (ANOVA) for each gene across samples, using both the original expression data and shuffled datasets in which the data were grouped into 12 triplicates at random (**Fig. 4a**). Randomization yielded distributions similar to those expected by chance. Because we tested roughly 10,000 genes, *P*-values of 10^{-5} or below were never observed. In contrast, the unshuffled expression data revealed 3,082 genes with $P < 10^{-5}$ and 4,412 genes with $P < 10^{-4}$, corresponding to 23.3% and 33.3% of the genes tested. We obtained a less conservative estimate of the number of differentially expressed genes by subtracting the shuffled histogram from the unshuffled histogram. This yielded 10,315 genes, corresponding to 78% of the genes tested. In most of these cases, the magnitude of differential expression was modest, but for many it was both statistically significant and large (**Fig. 4b**). For example, 2,301 genes (17.4% of those tested) registered ANOVA *P*-values below 10^{-5} and showed a maximum fold change (expression ratio between the most- and least-enriched populations) of more than a factor of 2. This degree of heterogeneity could not be previously appreciated because most previous cell type–specific gene expression studies monitored levels of only one or a few genes. The numbers, though large, are likely to represent an underestimate of the number of differentially expressed genes in the forebrain, because we tested only a subset of the existing neuronal populations. In addition, the microarray tends to substantially underestimate fold changes (**Supplementary Note**), and not all known genes were represented on the microarray.

Classes of differentially expressed genes

To determine the functional classes of the genes distinguishing neuronal subpopulations, we performed an over-representation analysis using the Gene Ontology (GO) database[23]. GO terms are nested functional categories that summarize the known molecular functions and biological processes associated with each gene, as well as its cellular localization. We determined which of these terms were associated with greater or smaller numbers of differentially expressed genes (defined by ANOVA, $P < 10^{-5}$) than predicted by chance using a significance threshold of $P < 10^{-3}$ (see Methods). Over-represented cellular components included the axon, the extracellular matrix, the synapse and the synaptic vesicles, whereas the nucleus and ribosomes were under-represented (**Fig. 5a**). These results show that protein products of differentially expressed genes are preferentially targeted to parts of the cell directly involved in synaptic connectivity and communication. Over-represented biological processes (**Supplementary Note**) also included categories associated with cell-cell communication. Unexpectedly, carbohydrate metabolism pathways were also over-represented, perhaps reflecting different metabolic needs of different cell types. Both over-represented biological processes and molecular functions (**Fig. 5b**) contained terms associated with the actin cytoskeleton. These may contribute to differences in the dendritic and axonal morphology of different cell types. Both also contained terms associated with ion transport and ion channels that may contribute to electrophysiological differences between cell types.

Differentially expressed genes are enriched for paralogs

What evolutionary mechanisms could produce the diversity of cellular transcriptomes that we observed? It has been hypothesized that gene duplication is a powerful mechanism that produces variation by permitting independent evolution of homologous genes together

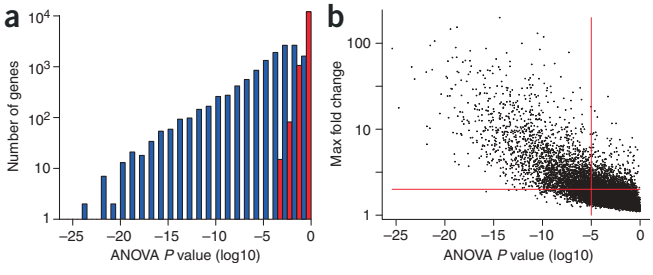

Figure 4 Heterogeneity of gene expression within neuronal subpopulations. (**a**) Frequency distribution of ANOVA *P*-values (see Methods) for shuffled (red) and unshuffled (blue) data. ANOVA was performed for each gene on the microarray across 36 samples from the 12 neuronal populations profiled. (**b**) ANOVA *P*-values plotted against maximum fold change for each gene.

with their regulatory regions[24]. If this is the case, the set of differentially expressed genes should be enriched for genes that belong to families. By categorizing genes as belonging to or not belonging to families using the SwissProt[25] database, we found that 1179 (39.7%) of 2,973 differentially expressed genes (defined by ANOVA, $P < 10^{-5}$) belonged to families. This was significantly greater ($P = 1.1 \times 10^{-11}$) than the 1,019.3 ± 22.9 (34.3%) expected by chance (**Fig. 6a**) (see Methods). The degree of enrichment was greatest for the most highly differentially expressed genes (**Fig. 6b**). This supports the idea that duplicated genes contribute to the diversification of cell types.

Differential expression of the members of a gene family could give rise to distinct physiological properties among different cell types. To identify the most differentially expressed gene families, we ranked families by the *P*-values for the over-representation of the differentially expressed genes within a family (see Methods). Over-represented families (**Table 2**) included kinases, phosphatases and small GTPases, which have key roles in intercellular and intracellular signaling. Notably, within each of these families, paralogs were expressed in different patterns across the different cell types (**Supplementary Note**), supporting the idea that paralogs contribute to functional differences between cell types.

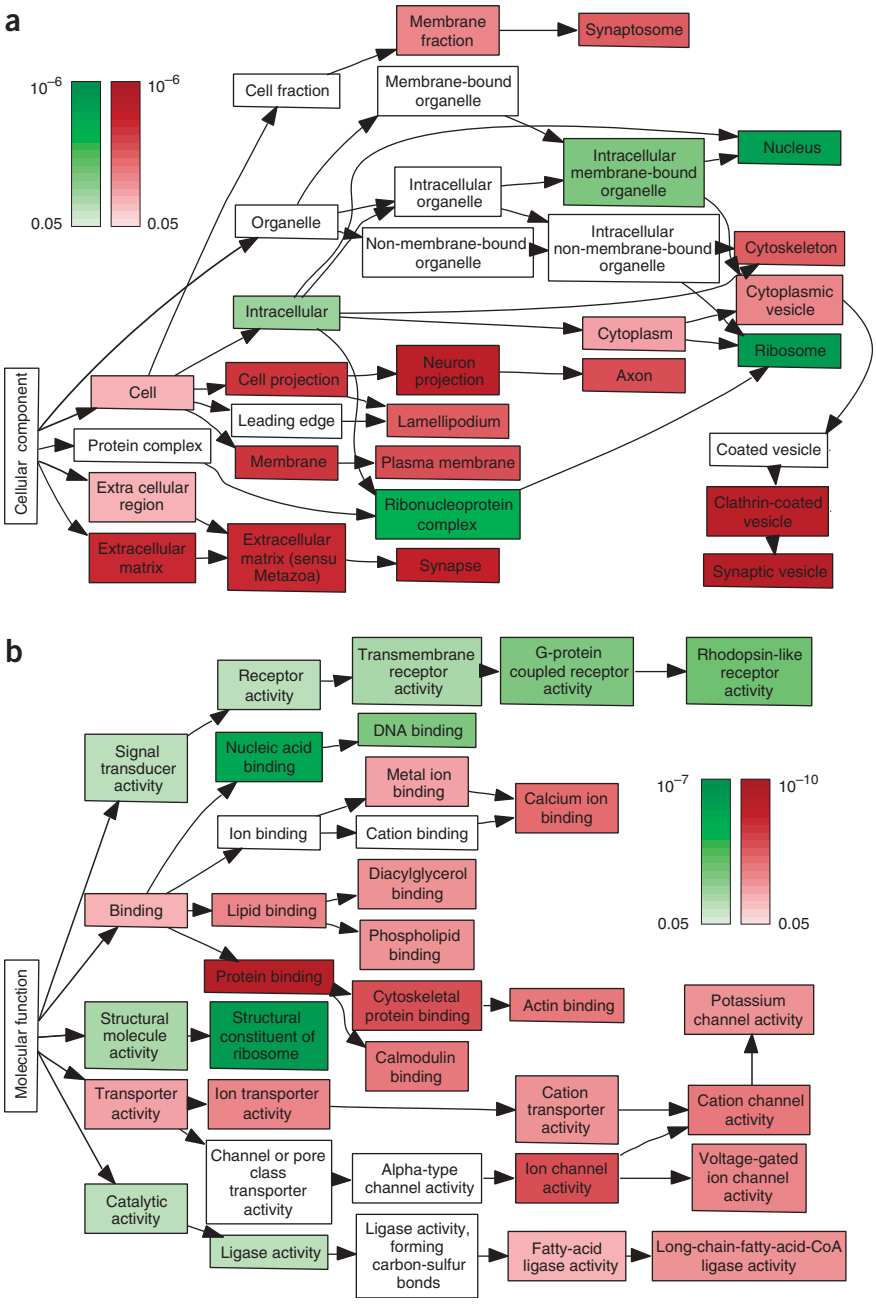

Figure 5 Gene ontology map. (**a**,**b**) Maps showing over-represented (red) and under-represented (green) GO terms with associated *P*-values. Color scales are proportional to *P*-value from over-representation analysis. White, $P > 0.05$; dark red, $P < 10^{-6}$ for **a**, $P < 10^{-10}$ for **b**; dark green, $P < 10^{-6}$ for **a**, $P < 10^{-7}$ for **b**. Owing to space considerations, only major GO terms are shown: 'cellular component' (**a**) and 'molecular function' (**b**). For 'biological process', see **Supplementary Note**.

A molecular taxonomy defined by expression distance

Heterogeneous gene expression profiles provide a potentially unbiased method for classifying neurons according to their similarity. To quantify and represent the relationships among gene expression profiles with a single parameter, we used euclidean distance because it is conceptually and computationally simple. This metric was applied pairwise to all 36 samples (see Methods). The euclidean distance correlated well with other single-parameter measures of similarity, such as the correlation coefficient and the numbers of differentially expressed genes determined by a *t*-test (data not shown). The results of this analysis are presented in the form of a distance matrix (**Fig. 6c**).

To reveal the hierarchical relationships among neuronal populations, we clustered samples according to their relative distances to render a dendrogram (**Fig. 6d**). The primary branch point on this taxonomic tree reflects the basic division between GABAergic interneurons and glutamatergic pyramidal neurons. Each of the populations grouped as GABAergic have been confirmed to be glutamic acid decarboxylase (GAD) or GABA immunopositive[6–8] (GABA ICC; data not shown), whereas those grouped as glutamatergic included the four YFPH populations and CT6-CG, all of which are spiny pyramidal neurons. Notably, G30-AM, though expressing GFP under a *Gad2* promoter, turned out to have features consistent with a glutamatergic phenotype: spiny morphology (3 of 3 cells) and expression of *Slc17a7* (*Vglut1*; single cell PCR, 16 of 18 cells), *Slc17a6* (*Vglut2*; microarray data) and glutamate ICC (48 of 121 cells). However, GABA ICC (1 of 221) and *Gad1* single cell PCR (2 of 18 cells) demonstrated that a minority of G30-AM neurons were GABAergic, showing that this is a mixed population.

The existence of a primary branch point between GABAergic and glutamatergic neurons suggests that the distinction between these phenotypes may involve the specific

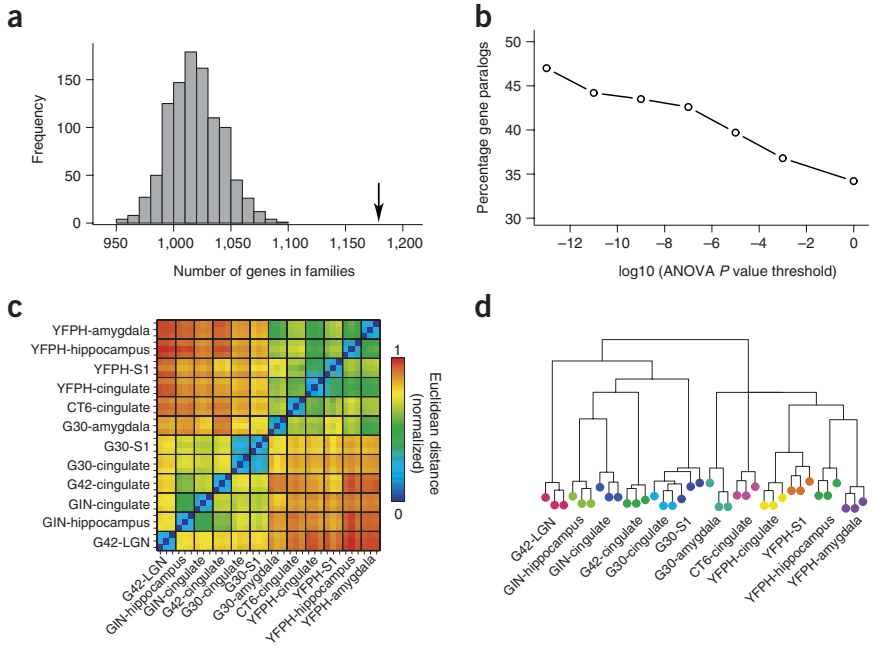

Figure 6 Differential expression of gene paralogs. (**a**) Histogram of the number of genes paralogs in repeated samples (1,000 times) of randomly selected genes represented on the microarray. Arrow indicates the number of differentially expressed genes (1,179; ANOVA, $P < 10^{-5}$) that belong to gene paralogs. (**b**) The percentage of paralogs in differentially expressed genes when the ANOVA *P*-value threshold is changed from 10^{-13} to 1. (**c**) Expression distance matrix showing normalized distances between all $36 \times 36 = 1{,}296$ pairs of samples. Color scale ranges from blue (distance of 0) to red (distance of 1, the largest distance observed). (**d**) Dendrogram based on hierarchical clustering of euclidean distances between individual microarray dataset.

expression of a large set of genes, not just a small number as has been suggested[26]. Consistent with this notion, *t*-tests revealed that 995 genes (7.5%) with $P < 10^{-5}$ and 1,603 genes (12.1%) with $P < 10^{-4}$ were differentially expressed between glutamatergic and GABAergic cell types (**Supplementary Note**). Samples could be correctly clustered as excitatory or inhibitory 80% of the time, on the basis of the amounts of expression in a random sample of only 5% of genes on the microarray. The fraction correctly clustered increased to 97% if 10% of genes were used. GO analysis of the genes differentially expressed between GABAergic and glutamatergic populations revealed an over-representation of many of the same categories that distinguish all cell types from each other, but also revealed an over-representation of other categories, including small GTPases and actin cytoskeleton–related proteins such as the Arp2/3 complex (**Supplementary Note**).

Different cortical populations derived from the same mouse line (G30 and YFPH, in CG and S1) showed the smallest distance (G30: 0.29; YFPH: 0.37) (see Methods). The relative similarity of populations from the same line also extended to comparisons between the neocortex and the hippocampus: YFPH-HP was relatively similar to the other four YFPH populations (average distance = 0.55) and GIN-HP was relatively similar to GIN-CG (0.52). This supported the idea that homologous cell types are present not only within the neocortex, but also between the neocortex and the hippocampus. This similarity of cell types isolated from the same line broke down when we compared cortical and noncortical populations. In the G30 line, cell types labeled in different brain structures were not closely related (0.79 between G30-CG and G30-AM). This was also true for the G42 line. G42-LG, the most distinct population of all, was as distant from G42-CG (0.79) as it was from GIN-HP and GIN-CG (both distances 0.76). G42-LG neurons were GABA positive (ICC, data not shown), but unlike G42-CG neurons, they were not parvalbumin positive (**Supplementary Note**) or fast-spiking (**Fig. 1**; $n = 3$). The distinct expression profile of this population may have been due to its diencephalic origin, in contrast to the telencephalic origin of the other populations.

The observation of large differences between cell types isolated from a single line was complementary to the observation of small differences between profiles of the same cell type in multiple lines (see above). Together, these results indicate that cell type, rather than genetic background, is the major determinant of these expression profiles.

DISCUSSION

Here we used microarray analysis to identify cell type–specific patterns of gene expression in the adult mammalian forebrain. The results revealed a high degree of heterogeneity of gene expression across cell types within a single region, comparable to or exceeding that previously observed across distinct tissues[27,28] or brain regions[29].

It is widely agreed that traditional classification criteria (such as axon morphology, expression of chemical markers, spiking properties and so on), though useful, have been inadequate to construct a coherent neuronal taxonomy[4,30]. Several researchers have proposed that cell type–specific gene expression profiling could help break the current impasse and lead to significant progress[4,31,32]. The gene expression profiling of 12 neuronal subtypes carried out here allowed us to construct a taxonomic tree (**Fig. 6d**). This taxonomy can readily be expanded by applying the same approach to other neuronal populations labeled genetically[33] or anatomically[13]. The strength of this approach is that it relies on a single metric, gene expression distance, that integrates over a large range of neuronal functions and can be used to compare any two neuronal populations on a uniform set of terms. For example, it can be used to quantify the similarity of cell types between related regions (such as the hippocampus and the neocortex) or across species (such as mouse and human).

Neurons isolated from two different cortical regions in a single transgenic line showed similar laminar distribution, morphology and firing properties. The existence of such functionally homologous cell types across cortical regions is predicted by the idea of a 'canonical microcircuit' reiterated throughout the neocortex[1]. However, it is known that different cortical regions make different sets of connections and process different types of information. It is also known that regionalization is directed by molecular gradients during development[34–36]. The degree to which this regionalization causes functionally homologous cell types to express different genes was unknown. We observed that the expression distance between G30 interneurons and YFPH pyramidal neurons isolated from two cortical regions were almost within the range of the expression distances between replicate experiments. These results indicate that the effects of regionalization on gene expression is small relative to the difference in expression between different cell types. The distance between YFPH pyramidal neurons was slightly larger than that between G30

Table 2 Differentially expressed gene families

Family	Selected	Total	*P*-value
Ser/Thr protein kinase family, CaMK subfamily	8	10	7.50×10^{-4}
LDH/MDH superfamily	5	5	1.50×10^{-3}
Small GTPase superfamily	40	103	3.30×10^{-3}
MAL family	4	4	5.60×10^{-3}
Stathmin family	4	4	5.60×10^{-3}
ADIPOR family	5	6	6.70×10^{-3}
ATP-dependent AMP-binding enzyme family	5	6	6.70×10^{-3}
Cyclic nucleotide phosphodiesterase family	8	13	8.20×10^{-3}
G protein γ family	5	7	1.70×10^{-2}
Small GTPase superfamily, Rab family	19	47	1.80×10^{-2}
Tyr protein kinase family, ephrin receptor subfamily	7	12	1.90×10^{-2}
Tubulin family	7	12	1.90×10^{-2}
CCN family	4	5	2.10×10^{-2}
CUT homeobox family	4	5	2.10×10^{-2}
PPP phosphatase family, PP-1 subfamily	4	5	2.10×10^{-2}
ATPase α/β chains family	3	3	2.10×10^{-2}
LDH/MDH superfamily, LDH family	3	3	2.10×10^{-2}
NDRG family	3	3	2.10×10^{-2}
PKI family	3	3	2.10×10^{-2}
RAMP family	3	3	2.10×10^{-2}
SH3BGR family	3	3	2.10×10^{-2}
bZIP family, PAR subfamily	3	3	2.10×10^{-2}
Cytochrome *b*5 family, MAPR subfamily	3	3	2.10×10^{-2}
Cytochrome *c* oxidase VIIa family	3	3	2.10×10^{-2}
Eukaryotic diacylglycerol kinase family	3	3	2.10×10^{-2}
Kinesin light chain family	3	3	2.10×10^{-2}
Opioid neuropeptide precursor family	3	3	2.10×10^{-2}
Phosphofructokinase family, two domains subfamily	3	3	2.10×10^{-2}
Phosphoglycerate mutase family, BPG-dependent PGAM subfamily	3	3	2.10×10^{-2}
Protein-tyrosine phosphatase family, receptor class 2B subfamily	3	3	2.10×10^{-2}
Vinculin/α-catenin family	3	3	2.10×10^{-2}

Selected: number of differentially expressed family members. Total: number of family members represented on the microarray. Gene families containing more differentially expressed genes than expected by chance. Selection criteria described in Methods.

interneurons. The greater similarity of expression in interneurons from the two regions may reflect the fact that interneurons migrate tangentially across cortical regions during development, whereas pyramidal neurons migrate vertically within a region[37].

The division between glutamatergic projection neurons and GABAergic interneurons is well established as one of the fundamental distinctions in the cerebral cortex. These cell types differ markedly in their morphology, electrophysiological properties, transmitter phenotype and embryological origin. The present study represents the first comprehensive picture of the widespread differences in gene expression that underlie this basic division. We also observed marked differences in expression between different populations of cortical interneurons. This finding is probably due to the fact that the populations examined included one each of the three major neurochemically defined interneuronal populations identified in previous studies of interneurons in rat neocortex[2,38,39].

A major challenge for post-genomic neurobiology is to map gene expression patterns with cellular resolution in the brain. One approach is high-throughput *in situ* hybridization[40]. For each individual gene, a complete expression map can be obtained. The approach taken here is complementary to this and other gene-by-gene methods, because it reveals patterns of coexpression within a given subset of neurons. It is these groups of coexpressed genes, organized into pathways and complexes, that probably confer cell type specific–phenotypes such as morphology, firing pattern, connectivity and synaptic transmission. In addition, identifying groups of coexpressed genes, particularly if done over a wider range of CNS cell types, will aid the search for *cis*-regulatory elements[41–43] and transcriptional modules[44,45] that underlie the elaboration of distinct neuronal cell types and the maintenance of their phenotypic identity.

METHODS

Mice. Experiments were carried out on adult male mice, 57–106 d old, derived from one of four transgenic lines (designated YFPH, GIN, G42 and G30; see Results). For the retrograde tracer injection experiments, we used YFPH mice. Each line had a mixed genetic background except G42, which was pure C57BL/6J. Founders were of the following backgrounds: GIN was FVB, G30 was a mix of CBA, CD1 and C57BL/6J, and YFPH was originally a mix of CBA and C57BL/6J. Each line was outcrossed into C57BL/6J and, in some cases, also intercrossed so as to obtain homozygotes to simplify breeding. Thus, genetic background varied within and across these four transgenic colonies. We were able to reconstruct lineages of 31 of 48 mice used in our profiling experiments. The contribution of C57BL/6J to these mice ranged from 31% to 100%, with an average of 67 ± 26% (mean ± s.d.). Only heterozygous mice were used for the experiments. All experiments were conducted in accordance with the requirements of our institutional Animal Use and Care Committee.

Retrograde labeling of neurons. Red fluorescent beads (Lumafluor) were injected into the medial dorsal thalamus to retrogradely label corticothalamic cells in layer 6 of cingulate cortex in YFPH mice. Anesthesia was produced with 70 mg kg^{-1} ketamine, 3.5 mg kg^{-1} xylazine hydrochloride and 0.7 mg kg^{-1} acepromazine maleate intraperitoneally. Mice were placed in a stereotaxic device, the scalp shaved, cleaned and incised, and a small hole was drilled through the skull to allow tracer-filled micropipettes to be lowered into the brain to the appropriate target (1.2 mm caudal to bregma; 0.3 mm lateral; 2.7 mm deep). Beads were pressure ejected, the pipette removed, the hole closed with bone wax and the scalp sutured. Mice were allowed to recover for at least 2 d after surgery.

Tissue preparation. We prepared 400-μm thick fresh coronal slices as previously described[46]. After a 90 min incubation in protease solution (1 mg ml^{-1} pronase E; Sigma-Aldrich), the desired brain regions were dissected from the slices under a dissecting microscope. Dissections were guided by reference to a brain atlas[47] and, when possible, to boundaries made apparent by fluorescent protein expression.

Tissue dissociation. Microdissected tissue was placed in a 1.5 ml Eppendorf tube and triturated in artificial cerebrospinal fluid (ACSF) with a series of three Pasteur pipettes of decreasing tip diameter. The ACSF solution included 1% fetal bovine serum for this step and throughout the remainder of the procedure. The resulting cell suspension was diluted 200× with ACSF and poured over a 100 mm Petri dish with Sylgard (Dow Corning) substratum.

Cell sorting. Under visual control on a fluorescence dissecting microscope, fluorescent neurons were aspirated into a micropipette broken to a diameter of 30–50 μm; they were then transferred to a clean 35 mm Petri dish containing fresh ACSF. The neurons were then transferred to a third and then a fourth dish. Each transfer improved the purity of the sample. Finally, the cells were aspirated a last time and expelled in a small drop (1–5 μl) onto a glass-bottom dish where they could be better inspected for purity under a fluorescence compound microscope. Pure samples were immediately lysed in 50-μl XB lysis buffer (Picopure Kit, Arcturus), incubated for 30 min at 42 °C and then placed in a –20 °C freezer.

Microarray processing. All experiments were performed using Affymetrix MOE430A oligonucleotide arrays (http://www.affymetrix.com/products/

arrays/specific/mouse430.affx). Total RNA from each sample was used to prepare biotinylated target RNA, with some modifications from the manufacturer's recommendations (http://www.affymetrix.com/support/technical/manual/expression_manual.affx). Briefly, mRNA from 27–120 cells (approximately 0.25–1 ng of total RNA) was extracted using Picopure Kits (Arcturus) and amplified with two rounds of *in vitro* transcription (IVT) using T7 RNA polymerase. The amplification protocol was based on the Affymetrix GeneChip Eukaryotic Small Sample Target Labeling Assay Version 2 and typically yielded 10–60 μg of biotinylated targets. The target cDNA generated from each sample was processed according to the manufacturer's recommendation using an Affymetrix GeneChip Instrument System. Briefly, spike controls were added to 10 μg of fragmented cDNA before overnight hybridization. Arrays were then washed and stained with streptavidin-phycoerythrin before being scanned on an Affymetrix GeneChip scanner. A complete description of these procedures is available at http://mouse.bio.brandeis.edu/2005-celltype/protocol.pdf. For some of the samples, the gain and the fidelity of the amplification was assessed before and after the first round of IVT by *q*PCR with housekeeping-gene primers (*Actg*, *Ubc*, *Hprt*, *Bax*). The size distribution of labeled RNA was confirmed using an agarose gel. After scanning, array images were assessed by eye to confirm scanner alignment and the absence of significant bubbles or scratches on the chip surface. For glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β-actin, 3′/5′ ratios ranged from 2.7 to 26. BioB spike controls were found to be present on all chips, with BioC, BioD and CreX also present in increasing intensity. When scaled to a target intensity of 500 (using Affymetrix MAS 5.0 array analysis software), scaling factors for all arrays were within acceptable limits (1.1–7.2), as were the background, Q values and mean intensities. Details of quality control measures can be found at http://mouse.bio.brandeis.edu/2005-celltype/QC-RPT.xls.

Data analysis software. GeneChip data were analyzed using the software packages MAS version 5 (Affymetrix), dChip[48], R (http://www.r-project.org), the affy library in Bioconductor (http://www.bioconductor.org), Python (http://www.python.org) and IgorPro (Wavemetrics). Scanned data were first processed with MAS to convert raw image files (.DAT) to probe signal value files (.CEL). Probe signal values were normalized across samples using dChip's invariant set method. Summary values for each probe set were calculated using dChip's PM-only model.

Euclidean distance. The euclidean distances were calculated using signal levels (in log scale) of expression profiles. To reduce the noise, we restricted the calculation to 3,082 genes that varied significantly across populations (ANOVA, $P < 10^{-5}$). To facilitate comparison, distances were normalized to a scale of 0 (identity) to 1 (the largest distance observed, that between G42-LG and YFPH-HP). The distance between replicate samples was 0.25 ± 0.03 (mean ± s.d.) and ranged from 0.21 to 0.30. The distance between samples prepared from dissociated but unsorted cell suspensions from different mouse lines (G42 and GIN, $n = 3$ each) was 0.22, further supporting the fact that strain differences are not responsible for the large observed differences between cell types.

Statistics and gene selection criteria. Differentially expressed genes were selected using combinations of ANOVA, template matching and *t*-tests. The exact combination and associated stringency varied depending on the goal of the particular analysis. ANOVA *P*-values were calculated using the R statistics package for each probe set to identify those whose signal levels differed significantly between cell types. For each gene associated with multiple probe sets, a single 'best' probe set was defined as that with the smallest ANOVA *P*-value. Template matching methods were used to select probe sets specifically elevated in certain sets of cell types. Briefly, a template was constructed by assigning a value of 1 to samples corresponding to one or more cell types of interest and 0 to all others. Correlation coefficients of the template and actual signal values were then used to select specifically expressed genes. We used *t*-tests (single tail, not assuming equal variance) between enriched and impoverished samples to rank specific genes. *P*-values for GO over-representation or gene family over-representation analysis were calculated using the hypergeometric distribution.

Gene family analysis. The SwissProt database (http://us.expasy.org/sprot/) lists 1,164 gene families in *Mus musculus*, containing 5,483 genes. Of these, 4,297 are represented as known genes on the Affymetrix MOE430A microarray. (ESTs were excluded in this analysis.) To determine the representation of family members that would occur by chance, random groups of 2,973 genes were selected from the set of 12,556 known genes represented (2,973 being the number of differentially expressed known genes with ANOVA $P < 10^{-5}$). We used 1,000 trials of random sampling to calculate the mean and s.d. of the number of genes found to belong to families by chance.

***q*PCR.** *q*PCR was performed on a Rotor-Gene 3000 (Corbett Research) to confirm differential expression identified by the microarray screen. Primers were selected from within or near the Affymetrix target sequence. Primers were discarded if they did not reliably produce single products as indicated by the melting curve. To facilitate direct comparison with the microarray results, the PCR was performed on cDNA amplified once for microarray samples but not included in the labeling reaction. To estimate relative transcript abundance between two cDNA samples, the sample predicted by the microarray to contain more of the transcript was serially diluted to obtain a standard curve relating cycle threshold to relative transcript abundance. Transcript abundance of the sample predicted to contain less of the transcript was estimated using this standard curve after normalization to actin-γ which was run concurrently.

***In situ* hybridization.** Digoxigenin-labeled riboprobes were prepared by IVT using a nonradioactive RNA labeling kit (Roche). DNA templates for these were prepared by PCR using an upstream primer containing a 5′ T3 RNA polymerase promoter sequence for the sense control riboprobe and a downstream primer containing a 3′ T7 RNA polymerase promoter for the antisense riboprobe. RNA probes were subsequently fragmented to lengths of ~160 nucleotides. Sections from fixed brains, 14 μm thick, were permeabilized, blocked, prehybridized, hybridized, washed and stained according to previously described nonradioactive *in situ* hybridization protocols[49,50]. Processed sections were viewed under a microscope and digitally photographed.

Immunocytochemistry. Mice were perfused with 4% paraformaldehyde. Sections (40 μm) cut on a Vibratome were blocked and permeabilized in phosphate-buffered saline (PBS) with 10% normal goat serum (NGS) and 0.3% Triton X for 2 h; they were then incubated for 48 h in PBS with 5% NGS, 0.1% Triton X and primary antibody. After washing three times in PBS for 10 min, the antibody was visualized using fluorescent secondary antibody or tyramide signal amplification (TSA Kit, Molecular Probes). Antibody sources and concentrations are listed (**Supplementary Note**).

Electrophysiology. Whole-cell current and voltage-clamp recordings were obtained from fluorescently labeled neurons in coronal slices of cingulate cortex, prepared as previously described[46]. Fluorescently labeled neurons were targeted for recording under a combination of epifluorescence and differential interference contrast (DIC) illumination. To establish intrinsic electrophysiological properties, each cell was subjected to a series of 1-s long current pulses of varying amplitude. Four basic firing parameters allowed us to rigorously distinguish our 12 neuronal populations from each other. Each of these was calculated for each cell over a range of current-pulse amplitudes taken from the linear portion of the current versus spike frequency curve. The four parameters were the following. (i) Spiking regularity—that is, the number of nonmonotonic changes in the interspike interval (ISI) for successive spikes within a train, weighted by the change in the ISI and normalized to spike numbers in the train. Cells with values greater than 1.0 were considered to show 'irregular spiking' (IS). (ii) ISI decay constant. The ISI as a function of spike number within the train was fit with an exponential. Cells that fired doublets of spikes at the beginning of a train and did not adapt thereafter were considered to show 'doublet firing' (D). (iii) Maximum spike frequency. Cells with maximum frequencies >100 Hz were considered to show 'fast-spiking' (FS) cells. (iv) Adaptation. The adaptation ratio was defined as the second interspike interval divided by last interspike interval. Cells with adaptation ratios less than 0.7 were considered adapting (A); those with ratios greater than 0.7 were considered nonadapting (NA). Using the second ISI rather than the first made it possible to measure adaptation independent of spike doublets that occurred in DF cells.

Voltage-clamp studies were carried out in the following synaptic blockers: D-(−)-2-amino-5-phosphonovaleric acid (AP5; 50 μM), bicuculline (20 μM)

and 6,7-dinitroquinoxaline-2,3-dione (DNQX; 20 μM). I_h and I_{sk} were isolated pharmacologically using ZD7288 (100 μM; Tocris) and apamin (100 nM; Sigma-Aldrich), respectively. Voltage-clamp protocols were run before and after blocker application. I_h was activated with hyperpolarizing pulses, 1 s in duration, from a holding potential of –40 mV to a test potential of –100 mV. I_{sk} was activated with a 10-ms pulse, from a holding potential of –70 mV to +20 mV, and measured as a tail current at –70 mV. Subtraction of traces in the presence of the blocker from corresponding traces in the absence of the blocker yielded blocker-sensitive currents assumed to be I_h and I_{sk}.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank G. Turrigiano, P. Sengupta, D. Das and Y. Sugino for comments on the manuscript; R. Pavlyuk, Z. Zhao and Z. Meng for technical assistance; and J. Fahrenkrug (Bispebjerg University Hospital, Denmark) for a gift of VIP/PHI antibody. This work was supported by a grant from the National Eye Institute.

AUTHOR CONTRIBUTIONS

K.S., C.M.H. and S.B.N. designed the experiments. K.S. did the amplification, array hybridization, ICC, ISH, qPCR and all analyses except that for the electrophysiology experiments. C.M.H. did the sample sorting, ISH and qPCR. M.N.M. did the electrophysiology and the analysis for electrophysiology except for that pertaining to CT6-CG. A.M.H. did the tracer injection and electrophysiology for CT6-CG. P.S. did G30-AM fill. Z.J.H. and C.W. did Ank1 ICC. K.S., C.M.H. and S.B.N. wrote the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/natureneuroscience/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Douglas, R.J. & Martin, K.A. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* **27**, 419–451 (2004).
2. Kawaguchi, Y. & Kubota, Y. GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* **7**, 476–486 (1997).
3. Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* **5**, 793–807 (2004).
4. Mott, D.D. & Dingledine, R. Interneuron diversity series: interneuron research–challenges and strategies. *Trends Neurosci.* **26**, 484–488 (2003).
5. Feng, G. *et al.* Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000).
6. Oliva, A.A., Jr., Jiang, M., Lam, T., Smith, K.L. & Swann, J.W. Novel hippocampal interneuronal subtypes identified using transgenic mice that express green fluorescent protein in GABAergic interneurons. *J. Neurosci.* **20**, 3354–3368 (2000).
7. Chattopadhyaya, B. *et al.* Experience and activity-dependent maturation of perisomatic GABAergic innervation in primary visual cortex during a postnatal critical period. *J. Neurosci.* **24**, 9598–9611 (2004).
8. Lopez-Bendito, G. *et al.* Preferential origin and layer destination of GAD65-GFP cortical interneurons. *Cereb. Cortex* **14**, 1122–1133 (2004).
9. Xu, Q., Cobos, I., De La Cruz, E., Rubenstein, J.L. & Anderson, S.A. Origins of cortical interneuron subtypes. *J. Neurosci.* **24**, 2612–2622 (2004).
10. Ferezou, I. *et al.* 5-HT3 receptors mediate serotonergic fast synaptic excitation of neocortical vasoactive intestinal peptide/cholecystokinin interneurons. *J. Neurosci.* **22**, 7389–7397 (2002).
11. Galarreta, M., Erdelyi, F., Szabo, G. & Hestrin, S. Electrical coupling among irregular-spiking GABAergic interneurons expressing cannabinoid receptors. *J. Neurosci.* **24**, 9770–9778 (2004).
12. Vruwink, M., Schmidt, H.H., Weinberg, R.J. & Burette, A. Substance P and nitric oxide signaling in cerebral cortex: anatomical evidence for reciprocal signaling between two classes of interneurons. *J. Comp. Neurol.* **441**, 288–301 (2001).
13. Arlotta, P. *et al.* Neuronal subtype-specific genes that control corticospinal motor neuron development *in vivo*. *Neuron* **45**, 207–221 (2005).
14. Clarke, V.R. *et al.* A hippocampal GluR5 kainate receptor regulating inhibitory synaptic transmission. *Nature* **389**, 599–603 (1997).
15. Burette, A. *et al.* Differential cellular and subcellular localization of AMPA receptor-binding protein and glutamate receptor-interacting protein. *J. Neurosci.* **21**, 495–503 (2001).
16. Anderson, S.A. *et al.* Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. *Neuron* **19**, 27–37 (1997).
17. Marin, O., Anderson, S.A. & Rubenstein, J.L. Origin and molecular specification of striatal interneurons. *J. Neurosci.* **20**, 6063–6076 (2000).
18. Stuhmer, T., Anderson, S.A., Ekker, M. & Rubenstein, J.L. Ectopic expression of the Dlx genes induces glutamic acid decarboxylase and Dlx expression. *Development* **129**, 245–252 (2002).
19. Colombo, E., Galli, R., Cossu, G., Gecz, J. & Broccoli, V. Mouse orthologue of ARX, a gene mutated in several X-linked forms of mental retardation and epilepsy, is a marker of adult neural stem cells and forebrain GABAergic neurons. *Dev. Dyn.* **231**, 631–639 (2004).
20. Poirier, K. *et al.* Neuroanatomical distribution of ARX in brain and its localisation in GABAergic neurons. *Brain Res. Mol. Brain Res.* **122**, 35–46 (2004).
21. Sah, P. & Faber, E.S. Channels underlying neuronal calcium-activated potassium currents. *Prog. Neurobiol.* **66**, 345–353 (2002).
22. Robinson, R.B. & Siegelbaum, S.A. Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu. Rev. Physiol.* **65**, 453–480 (2003).
23. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
24. Ohno, S. *Evolution by Gene Duplication* (Springer-Verlag, New York, 1970).
25. Bairoch, A. *et al.* The universal protein resource (UniProt). *Nucleic Acids Res.* **33**, D154–D159 (2005).
26. Takamori, S., Rhee, J.S., Rosenmund, C. & Jahn, R. Identification of a vesicular glutamate transporter that defines a glutamatergic phenotype in neurons. *Nature* **407**, 189–194 (2000).
27. Su, A.I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. USA* **99**, 4465–4470 (2002).
28. Su, A.I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. USA* **101**, 6062–6067 (2004).
29. Evans, S.J. *et al.* DNA microarray analysis of functionally discrete human brain regions reveals divergent transcriptional profiles. *Neurobiol. Dis.* **14**, 240–250 (2003).
30. McBain, C.J. & Fisahn, A. Interneurons unbound. *Nat. Rev. Neurosci.* **2**, 11–23 (2001).
31. Diez del Corral, R. & Storey, K.G. Markers in vertebrate neurogenesis. *Nat. Rev. Neurosci.* **2**, 835–839 (2001).
32. Dougherty, J.D. & Geschwind, D.H. Progress in realizing the promise of microarrays in systems neurobiology. *Neuron* **45**, 183–185 (2005).
33. Gong, S. *et al.* A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917–925 (2003).
34. Fukuchi-Shimogori, T. & Grove, E.A. Neocortex patterning by the secreted signaling molecule FGF8. *Science* **294**, 1071–1074 (2001).
35. Lukaszewicz, A. *et al.* G1 phase regulation, area-specific cell cycle control, and cytoarchitectonics in the primate cortex. *Neuron* **47**, 353–364 (2005).
36. O'Leary, D.D. & Nakagawa, Y. Patterning centers, regulatory genes and extrinsic mechanisms controlling arealization of the neocortex. *Curr. Opin. Neurobiol.* **12**, 14–25 (2002).
37. Anderson, S., Mione, M., Yun, K. & Rubenstein, J.L. Differential origins of neocortical projection and local circuit neurons: role of Dlx genes in neocortical interneuronogenesis. *Cereb. Cortex* **9**, 646–654 (1999).
38. Gonchar, Y. & Burkhalter, A. Three distinct families of GABAergic neurons in rat visual cortex. *Cereb. Cortex* **7**, 347–358 (1997).
39. Kubota, Y. & Kawaguchi, Y. Three classes of GABAergic interneurons in neocortex and neostriatum. *Jpn. J. Physiol.* **44** (Suppl 2), S145–S148 (1994).
40. Gewin, V. A golden age of brain exploration. *PLoS Biol.* **3**, e24 (2005).
41. Bussemaker, H.J., Li, H. & Siggia, E.D. Regulatory element detection using correlation with expression. *Nat. Genet.* **27**, 167–171 (2001).
42. Roth, F.P., Hughes, J.D., Estep, P.W. & Church, G.M. Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. *Nat. Biotechnol.* **16**, 939–945 (1998).
43. Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J. & Church, G.M. Systematic determination of genetic network architecture. *Nat. Genet.* **22**, 281–285 (1999).
44. Ihmels, J., Bergmann, S. & Barkai, N. Defining transcription modules using large-scale gene expression data. *Bioinformatics* **20**, 1993–2003 (2004).
45. Ihmels, J. *et al.* Revealing modular organization in the yeast transcriptional network. *Nat. Genet.* **31**, 370–377 (2002).
46. Hempel, C.M., Hartman, K.H., Wang, X.J., Turrigiano, G.G. & Nelson, S.B. Multiple forms of short-term plasticity at excitatory synapses in rat medial prefrontal cortex. *J. Neurophysiol.* **83**, 3031–3041 (2000).
47. Paxinos, G. & Franklin, K.B.J. *The Mouse Brain in Stereotaxic Coordinates* (Academic Press, Boston, 2001).
48. Li, C. & Wong, W.H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* **98**, 31–36 (2001).
49. Esclapez, M., Tillakaratne, N.J., Tobin, A.J. & Houser, C.R. Comparative localization of mRNAs encoding two forms of glutamic acid decarboxylase with nonradioactive *in situ* hybridization methods. *J. Comp. Neurol.* **331**, 339–362 (1993).
50. Jongen-Relo, A.L. & Amaral, D.G. A double labeling technique using WGA-apoHRP-gold as a retrograde tracer and nonisotopic *in situ* hybridization histochemistry for the detection of mRNA. *J. Neurosci. Methods* **101**, 9–17 (2000).

Erratum: *DWnt4* regulates the dorsoventral specificity of retinal projections in the *Drosophila melanogaster* visual system

Makoto Sato, Daiki Umetsu, Satoshi Murakami, Tetsuo Yasugi & Tetsuya Tabata
Nat. Neurosci. **9**, 67–75 (2006)

In the print version of this article and the version initially published online, the original panel (**g**) of **Figure 5** was replaced with a duplicate of panel (**f**). The corrected figure is below. The error has been corrected in the HTML and PDF versions of the article. This correction has been appended to the PDF version.

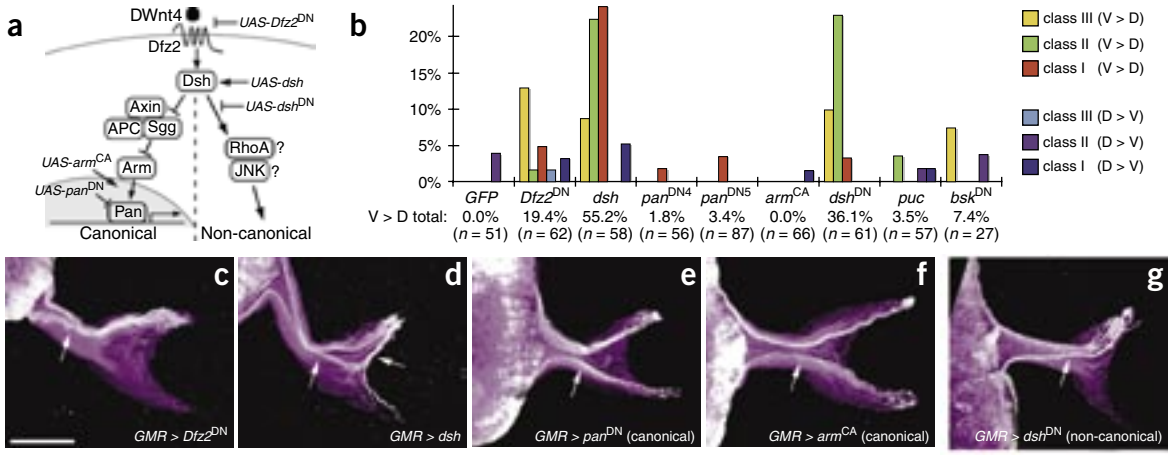

Figure 5 Autonomous requirement of noncanonical Wnt signaling in the retina. (**a**) Schematic of *D. melanogaster* Wnt signaling. (**b**) The penetrance and expressivity are compared in flies expressing various UAS transgenes under the control of *GMR-Gal4*. (**c**–**g**)UAS transgenes are expressed in the retina behind the furrow using *GMR-Gal4*. Dorsal- and ventral-most axons visualized using *omb-τlacZ* (white). R axons visualized by *UAS-GFP* (magenta). (**c**) *UAS-gpiDfz2* (*UAS-Dfz2*^{DN}) inactivates Wnt signaling. (**d**) *UAS-dsh* activates Wnt signaling. (**e**) *UAS-panΔN* (*UAS-pan*^{DN}) inactivates canonical signaling. (**f**) *UAS-Δarm* (*UAS-arm*^{CA}) activates canonical signaling. (**g**) *UAS-dshΔDEP* (*UAS-dsh*^{DN}) inactivates noncanonical signaling. Flies were raised at 29 °C for *UAS-GFP*, *UAS-pan*^{DN}, *UAS-dsh*^{DN}, *UAS-puc* and *UAS-bsk*^{DN}. Scale bars, 50 μm.

Erratum: Molecular taxonomy of major neuronal classes in the adult mouse forebrain

Ken Sugino, Chris M Hempel, Mark N Miller, Alexis M Hattox, Peter Shapiro, Caizi Wu, Z Josh Huang & Sacha B Nelson
Nat. Neurosci. **9**, 99–107 (2006)

In the print version of this article and the version initially published online, information about the dataset was missing. The complete dataset can be viewed and queried online at http://mouse.bio.brandeis.edu. The dataset has also been deposited in the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo/index.cgi with accession number GSE2882). The error has been corrected in the HTML and PDF versions of the article. This correction has been appended to the PDF version.