Functional genomics identifies neural stem cell sub-type expression profiles and genes regulating neuroblast homeostasis

Travis D. Carney, Michael R. Miller, Kristin J. Robinson, Omer A. Bayraktar, Jessica A. Osterhout, Chris Q. Doe *

Institute of Molecular Biology, Institute of Neuroscience, Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403, USA

A R T I C L E   I N F O

Article history:
Received for publication 15 September 2011
Revised 17 October 2011
Accepted 18 October 2011
Available online 25 October 2011

Keywords:
Drosophila
Neuroblast
Self-renewal
Expression profiling
Transit-amplifying
Neural progenitors

A B S T R A C T

The Drosophila larval central brain contains about 10,000 differentiated neurons and 200 scattered neural progenitors (neuroblasts), which can be further subdivided into ~95 type I neuroblasts and eight type II neuroblasts per brain lobe. Only type II neuroblasts generate self-renewing intermediate neural progenitors (INPs), and consequently each contributes more neurons to the brain, including much of the central complex. We characterized six different mutant genotypes that lead to expansion of neuroblast numbers; some preferentially expand type II or type I neuroblasts. Transcriptional profiling of larval brains from these mutant genotypes versus wild-type allowed us to identify small clusters of transcripts enriched in type II or type I neuroblasts, and we validated these clusters by gene expression analysis. Unexpectedly, only a few genes were found to be differentially expressed between type I/II neuroblasts, suggesting that these genes play a large role in establishing the different cell types. We also identified a large group of genes predicted to be expressed in all neuroblasts but not in neurons. We performed a neuroblast-specific, RNAi-based functional screen and identified 84 genes that are required to maintain proper neuroblast numbers; all have conserved mammalian orthologs. These genes are excellent candidates for regulating neural progenitor self-renewal in Drosophila and mammals.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Drosophila neuroblasts are a powerful model system for understanding the molecular control of stem cell self-renewal versus differentiation. The majority of neuroblasts (type I neuroblasts) repeatedly divide asymmetrically with respect to size and fate to self-renew and produce a smaller daughter cell called a ganglion mother cell (GMC) that divides only once to produce two post-mitotic neurons or glia (reviewed in: Chia et al., 2008; Doe, 2008; Knoblich, 2008). Neuroblast/GMC fate differences are due in part to the asymmetric partitioning of proteins into the GMC during neuroblast cell division. These factors include the transcription factor Prospero (Pros), the Notch inhibitor Numb, and the translational repressor Brain tumor (Brat) (Betschinger et al., 2006; Broadus et al., 1998; Hirata et al., 1995; Knoblich et al., 1995; Lee et al., 2006c; Spana and Doe, 1995). Proper segregation of Pros, Numb, and Brat into the GMC require the scaffolding protein Miranda (Mira) and the WD40-domain protein Lethal giant larvae (Lgl) (Betschinger et al., 2006; Lee et al., 2006c; Ohshiro et al., 2000; Peng et al., 2000). In the GMC, Pros enters the nucleus and promotes cell cycle exit and differentiation by directly activating differentiation genes and repressing self-renewal and cell cycle regulatory genes (Choksi et al., 2006; Li and Vaessen, 2000). This process allows for a single neuroblast to generate a lineage of many differentiated neurons and glia and for a relatively small number of neuroblasts to generate the thousands of cells found in the central nervous system of the adult fly.

Recently, “type II” neuroblasts were identified in the larval brain which divide asymmetrically to produce small transit-amplifying progenitors (intermediate neural progenitors, INPs). INPs themselves undergo molecularly asymmetric cell divisions to generate 4–6 GMCs, each of which typically generates two post-mitotic neurons or glia (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008; Izergina et al., 2009; Viktorin et al., 2011). Six type II neuroblasts inhabit the dorso-medial region of the lobe and are designated DM1–6, and two occupy more lateral positions (Bayraktar et al., 2010; Izergina et al., 2009). Type II neuroblasts behave in a manner similar to mammalian neural stem cells in that they generate transit-amplifying INPs. Transit-amplifying progenitors are important in the development of the nervous system in mammals as well as in flies because they permit the rapid amplification of neuronal progeny (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008; Merkle and Alvarez-Buylla, 2006; Morrison and Kimble, 2006). Thus, while there are ~95 type I neuroblasts per larval brain lobe and only eight type II neuroblasts, these few type II lineages produce a considerable fraction – approximately a quarter – of the neurons of the adult brain (Izergina et al., 2009). Type II neuroblasts and INPs have thus become a model for the study of transit-amplifying neural progenitors, and determining how these cells are specified and maintained may shed light on the function of mammalian neural stem cells.
Type II neuroblasts are known to differ from type I neuroblasts in several ways. First, they generate INPs and thus make much larger lineages than type I neuroblasts (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008). Second, they are the major contributors of the intrinsic neurons of the adult central complex (intrinsinc neurons have projections entirely within the central complex) (Bayraktar et al., 2010; Izergina et al., 2009). Third, they are more susceptible to tumor formation (Boone and Doe, 2008; Bowman et al., 2008; Ouyang et al., 2011; Weng et al., 2010). Numerous genotypes have been identified that cause the production of ectopic larval brain neuroblasts, and several of these specifically affect type II neuroblasts. For example, mutations in brat lead to “overgrowth” of just the type II neuroblasts, and mutations in lgl affect type II much more strongly than type I neuroblasts. On the other hand, loss of Aurora-A (Aur) or neuroblast misexpression of membrane-tethered atypical protein kinase C (aPKCCAAX) leads to ectopic type I and type II neuroblasts (Bowman et al., 2008). It is likely that the lack of Pros in the new-born INP renders this daughter cell sensitive to the loss of a second growth inhibitor/differentiation factor, making it easier for this cell to revert to a type II neuroblast identity.

In spite of the marked differences between type I and type II neuroblasts in proliferative potential and susceptibility to tumor formation, only two molecular differences are known: type II neuroblasts lack the transcription factors Pros and Aense (Ase), while both are present in all type I neuroblasts (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008). It is currently unknown how many other genes are regulated differentially between type I and type II neuroblasts, and which of them regulates each type. In order to discover such transcriptional differences, relatively pure populations of each neuroblast sub-type must be isolated from which to extract RNA. Complicating these efforts, the Drosophila central nervous system contains only a small number of neuroblasts which are dispersed throughout a complex population of thousands of neurons and glia, making it difficult to physically separate neuroblast sub-types from each other and from other cell types. Thus, comparing the transcriptional outputs of neuroblast sub-types is technically challenging due to the difficulty of isolating cell type-specific RNA.

In order to enrich for each type of neuroblast, here we make use of published and unpublished mutants in which type I and type II neuroblasts exhibit differential overproliferation phenotypes. We perform microarray-based whole-genome transcriptional profiling to compare each of these different mutant brains to wild-type; thus we are able to probe the transcriptional differences not only between each mutant and wild-type, but also between type I and type II neuroblasts. We identify only a small number of genes exhibiting transcriptional differences between type I and type II neuroblasts, providing a highly specific group of genes to screen for a function in establishing each type of neuroblast. We identify a large group of genes which are likely expressed in neuroblasts but not neurons, and we verify the neuroblast function of a subset of these genes using an RNAi-based targeted loss of function screen. Using this approach we identify 84 genes required to maintain neuroblast numbers in larval brains, all of which have conserved mammalian orthologs.

Materials and methods

Fly stocks

All fly stocks used in this study have been previously described except for lgl1, lgl2, in which the lgl locus has been spontaneously lost (Jason Boone, unpublished data) from the lgl2 chromosome (Jaekel and Klein, 2006). Other fly stocks used were: lgl1, lgl2, lgl1; lgl2; P[acman]act5C (Rolls et al., 2003); ntrA-B7 (Lee et al., 2006a); lgl1; P[acman]act5C and UAS-aPKC-CAS (Lee et al., 2006b); brat11 (Lee et al., 2006c); wor-gal4 (Albertson et al., 2004); R9D11-Gal4 and R19H09-Gal4 (Bayraktar et al., 2010); UAS-mCD8::GFP and UAS-Dicer2 (Bloomington Stock Center).

Microarray analysis

Mutant larvae were dissected at 144 h after larval hatching (ALH); wild-type heterozygous larvae that were at the same developmental stage were dissected at 96 h ALH as wild-type controls. The only exceptions were the aPKC-CAS experiments, in which the experimental larvae were raised at 30°C and exhibited enlarged brains packed with ectopic neuroblasts, and genetically identical control larvae were raised at room temperature, where the ectopic neuroblast phenotype is much weaker. Total RNA was extracted from larval brain lobes using TRIzol extraction methods according to the manufacturer’s instructions (Invitrogen, Carlsbad, CA, USA). First strand synthesis and amplification of Cy3- and Cy5-labeled RNA were accomplished using the Agilent Low Input Quick Amp Labeling Kit (Agilent, Santa Clara, CA, USA). Four experimental replicates were performed for each mutant genotype: two standard replicates (Cy5-labeled mutant RNA and Cy3-labeled wild-type RNA) and two dye-swapped replicates. Exceptions were the aPKC-CAS experiments, in which one standard and one swapped replicates were used for clustering, and the lgl experiments, in which three standard and one swapped replicates were used (Fig. S1). Hybridization was performed as previously described (Miller et al., 2009), except for each replicate, 825 ng of both mutant and wild-type RNA were mixed and hybridized to Agilent microarrays. The slides were scanned using an Axon GenePix 4000B, and GenePix software was used for feature extraction. Microarray data will be made publicly available through the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) upon publication.

Cluster analysis

We selected genes for cluster analysis if, in at least one of the mutant genotypes, their average transcript levels over all biological replicate experiments deviated from the wild-type RNA sample by greater than two-fold. This criterion resulted in the selection of 2781 genes. To investigate the reproducibility of replicate experiments of the same mutant genotype and the overall transcriptional similarities between experiments, we first performed cluster analysis without averaging individual replicate experiments. This analysis allowed us to determine whether, for each experiment, the clustering analysis grouped replicates of the same genotype together (Fig. S1). To identify groups of genes with similar transcript patterns over the different mutant genotypes, we averaged replicates and performed cluster analysis on these averages in order to avoid artificial gene clustering relationships due to technical noise between experimental replicates. After performing cluster analysis, we found that a large group of genes with increased expression in mutant brains clustered with high correlation. Group A was defined by a tree branch that represented a large decrease in correlation, from >0.6 to ~0.38. We similarly found that a long branch in the genes with reduced expression in mutant brains caused a decrease in correlation that passed the 0.6 cutoff (from >0.7 to ~0.54); thus we considered all the genes clustering below this branch to comprise group C. The remainder of the genes exhibited variable expression patterns and did not cluster with high correlation, and was defined as group B. The results of the clustering analysis can be viewed in PDF format in Supplementary File 1; in addition, results may be viewed in Treeview software (Eisen et al., 1998) using the files compressed in Supplementary File 2.

RNAi screen

Knock-downs were performed on group A genes with annotated human orthologs and transgenic RNAi stocks available from the
Vienna Drosophila RNAi Center (VDRC). We used these lines to knock down genes cell type-specifically by crossing RNAi line males to wor-Gal4 UAS-Dicer2 virgins. The progeny of these crosses were raised at 30 °C and scored for lethality. Each knock-down was performed at least twice to judge the consistency of the phenotype. For those RNAi constructs that caused lethality, we performed crosses again and dissected brains from wandering third instar larvae. We took confocal stacks of these brains (see below) and determined the number of central brain neuroblasts per brain lobe using antibodies against the neuroblast-specific proteins Deadpan (Dpn) and Mira. Optic lobe neuroblasts were excluded from these counts based on their small size, tight clustering, and stereotyped lateral position in the brain lobe.

Fixation, antibody staining, and confocal microscopy

The following antibodies were used for immunohistochemical staining of larval brains: guinea pig anti-Dpn, 1:2000 (J. Skeath); rat anti-Dpn, 1:1–1:50 (Doe lab); rat anti-Elav, 1:50 (Developmental Studies Hybridoma Bank 7E8A10); rabbit anti-Ase, 1:2000 (Brand et al., 1993); rabbit anti-Optix, 1:500 (Kenyon et al., 2005); rabbit anti-Rx, 1:2000 (Davis et al., 2003); chicken anti-GFP, 1:1000 (Aves Laboratories, Tigard, OR, USA); mouse anti-Pros, 1:1000 (MR1A, Doe lab). Brains were dissected in Schneider’s medium (Sigma Aldrich, St. Louis, MO, USA), fixed in PBST (phosphate-buffered saline + 0.1% Triton-X100; Sigma Aldrich) with 4% formaldehyde for 20 min, rinsed for 30 min in PBST, blocked for 30 min using PBSBT (PBST + 1% bovine serum albumin) or PBST + 5% normal goat serum. Brains were incubated in primary antibody overnight at 4 °C with rocking, and then rinsed in PBST + block for 1 h. Brains were incubated with secondary antibodies (1:500; Molecular Probes, Eugene, OR, USA) or Jackson Immunoresearch, West Grove, PA, USA) for 2 h at room temperature with rocking, and then rinsed for 1 h with PBST and stored in Vectashield (Vector Laboratories, Inc., Burlingame, CA, USA) until microscopy could be performed. Microscopy images were taken using a Bio-Rad Radiance or Zeiss700 confocal microscope.

Results

Transcriptional profiling of larval brains containing ectopic type I or type II neuroblasts

We analyzed six different genotypes that generate ectopic neuroblasts in the third instar larval brain (Table 1). The brat and lgl single mutants produce primarily ectopic type II neuroblasts, whereas aurA mutation or misexpression of membrane-tethered aPKC proteins Deadpan (Dpn) and Mira.

Table 1

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Synonym</th>
<th>Type I/II neuroblast phenotype</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>lgl124</td>
<td>lgl</td>
<td>Ectopic type II</td>
<td>Bowman et al. (2008)</td>
</tr>
<tr>
<td>brat11</td>
<td>brat</td>
<td>Ectopic type II</td>
<td>Bowman et al. (2008)</td>
</tr>
<tr>
<td>wor-gal4 UAS-aPKC&lt;sup&gt;CAAX&lt;/sup&gt;</td>
<td>aPKC&lt;sup&gt;CAAX&lt;/sup&gt;</td>
<td>Ectopic type I (some II)</td>
<td>Bowman et al. (2008)</td>
</tr>
<tr>
<td>aur&lt;sup&gt;PTES&lt;/sup&gt;</td>
<td>aur</td>
<td>Ectopic type I (some II)</td>
<td>Bowman et al. (2008)</td>
</tr>
<tr>
<td>lgl&lt;sup&gt;−&lt;/sup&gt; lgl&lt;sup&gt;−/−&lt;/sup&gt;</td>
<td>lgl</td>
<td>Ectopic type I (some II)</td>
<td>This work</td>
</tr>
<tr>
<td>lgl&lt;sup&gt;−&lt;/sup&gt;-pins&lt;sup&gt;−/−&lt;/sup&gt;</td>
<td>lgl pins</td>
<td>Ectopic type I (some II)</td>
<td>This work</td>
</tr>
</tbody>
</table>

Identification of genes transcribed preferentially in type I or type II neuroblasts

Identification of genes transcribed preferentially in type I or type II neuroblasts

To identify genes expressed differentially between type I and type II neuroblasts, we sequenced the transcriptome of type I and type II neuroblasts, and we found that the brat single mutant generates primarily ectopic type II Neuroblasts whereas aurA<sup>−</sup> brains contain more type I Ase<sup>+</sup> neuroblasts, although there is also an increase in type II neuroblasts (Fig. 1E). Moreover, we found that the lgl lgl<sup>−</sup> double mutant is strongly enriched for type II neuroblasts, and lgl pins brains contain both neuroblast types with an enrichment of type I neuroblasts. We noted that in lgl pins brains, distinct regions of type I and type II neuroblast overproliferation are discernible based on the lack of Ase and Pros in ectopic cells derived from type II neuroblasts (Fig. 1E and inset). We conclude that the six genotypes used here exhibit a range of type I/type II differential overproliferation phenotypes, with brat, lgl, and lgl pins brains representing enriched pools of type II neuroblasts, and with aur, aPKC<sup>CAAX</sup>, and lgl pins being more enriched for type I neuroblasts.

Next we used transcriptional profiling of the larval brain lobes from each of these six genotypes to identify (a) genes differentially regulated in type I vs. type II neuroblasts, that may function in establishing the striking differences between these two types of progenitors, and (b) genes expressed in all neuroblasts, that may function to regulate self-renewal or asymmetric cell division. For each experiment, we isolated RNA from mutant and wild-type brain lobes, amplified and fluorescently labeled each RNA sample, and hybridized them directly against each other to microarrays representing the entire complement of protein-coding Drosophila genes with at least two-fold redundancy (Fig. 1F). We used cluster analysis to group genes according to transcriptional pattern similarities in the different experiment parameters. Genes exhibiting no change between wild-type and mutant were not included in the cluster analysis (see Materials and methods). Biological replicates and dye-swap experiments cluster much more closely to one another than to replicates for any other mutant (Fig. S1); this demonstrates that our data are highly reproducible and that each mutant exhibits a distinct transcriptional profile.

We sorted genes into three groups based on their transcript pattern in the six genotypes (Fig. 2A). Group A contains 1045 genes with elevated expression in the mutant genotypes (Fig. 2B); these genes are likely to be expressed in neuroblasts and are good candidates for genes expressed in neurons or glia but not in neuroblasts. Group B contains 467 genes that have variable expression between mutants. Group C contains 1269 genes with decreased transcript levels in each mutant (Fig. 2A and B); thus the genes in this large group are good candidates for genes expressed in neurons or glia but not in neuroblasts. Interestingly, we did not see clustering of the genotypes that generate ectopic type II neuroblasts (lgl, lgl<sup>−</sup> lgl<sup>−</sup>, and brat) – note that the dendrogram at the top of Fig. 2A shows that each of these genotypes has a more closely related genotype that generates ectopic type I neuroblasts – suggesting type I and type II neuroblasts are much more similar transcriptionally than different.
sub-cluster exhibits such a pattern. This suggests that the other nine genes in the cluster may also be specifically expressed in type I neuroblasts, like pros and ase, and that these are potentially the only genes that exhibit this unique pattern.

To test whether other genes in the small pros/ase cluster are also expressed in type II neuroblasts but not type I neuroblasts, we obtained an antibody to a candidate from this cluster, Retinal homeobox (Rx), a homeodomain-containing transcription factor (Davis et al., 2003; Eggert et al., 1998). We found that Rx is completely absent from type II neuroblasts, similar to Pros and Ase; Rx is detected in several type I neuroblasts as well as in a subset of differentiated type II progeny (Fig. 3B and C). Consistent with this expression pattern, we found that brat mutants, which overproduce type II neuroblasts, show a loss of Rx staining (Fig. 3D). In contrast, lgl pins mutants, which have ectopic type I neuroblasts, show territories of strong Rx expression which is confined to Pros + (likely type I-originating) cells (Fig. 3E). The fact that only a small patch of lgl pins mutant brain tissue is Rx + is probably because Rx is normally expressed in a subset of type I neuroblasts. We conclude that Rx, like Pros and Ase, is expressed in type I but not type II neuroblasts. Thus, most or all of the 11 genes in the pros/ase sub-cluster may be expressed in type I but not type II neuroblasts.

We next wanted to find genes expressed in type II neuroblasts but not type I neuroblasts, as there are currently no known markers specifically expressed in type II neuroblasts. We reasoned that transcripts expressed in type II neuroblasts should be enriched in genotypes that overproduce type II neuroblasts: brat, lgl and lgl lgd. We found one small cluster enriched in two of the three mutants (brat and lgl lgd) (Fig. 4A). This cluster contains just 10 genes, seven encoding transcription factors. To verify the expression pattern of this gene cluster, we examined the expression of one gene product, Optix. Optix is a conserved homeodomain-containing transcription factor required for eye development (Kenyon et al., 2005; Seimiya and Gehring, 2000; Toy et al., 1998). Consistent with our microarray data, we found that most of the Optix expression in the brain is indeed restricted to type II lineages; four of the six dorso-medial type II neuroblasts (DM1, 2, 3, and 6) express Optix, as do most of the INPs, GMCs, and neurons in these lineages (Fig. 4B and C). In addition, recent work has shown that another gene in this cluster, pointedP1, is also preferentially expressed in type II neuroblasts (Sijun Zhu and Y.
exception of the type II-negative pros and ase genes) are represented in group A, including "worniu" (wor), "deadpan", and Cycline (Table 2). Conversely, neuronal and glial genes are excluded from group A and found in group C [e.g., "elav", "gial cells missing" (gcm), and reversed polarity (repo)] (Table 2). In addition, there is a good correlation between the number of neuroblasts in each mutant brain and the level of enrichment shown for group A genes (Figs. 1B and 2B). Thus, the genes in group A are likely to be expressed in both type I and type II neuroblasts, but not in neurons or glia. Conversely, genes in group C are likely to be expressed in differentiated neurons or glia, but not in neuroblasts. We conclude that ectopic neuroblasts are similar transcriptionally to wild-type neuroblasts, and thus the mutant genotypes represent an enriched source of neuroblast-expressed mRNA.

We next determined the gene ontology (GO) terms that represent the neuroblast-enriched group A genes and the neuron/glia-enriched group C genes. We found that group A genes are strongly enriched for several GO terms, including cell cycle and ribosome biosynthesis — processes expected in neuroblasts that must repeatedly divide and grow (Fig. 5A and B). For example, a small sub-cluster in group A exhibits a very significant enrichment for genes involved in DNA replication (p<10^-15); this process is not significantly represented in any other cluster (p>0.01 for all other group A genes combined; Fig. 5A). Conversely, we found group C to be significantly enriched for the GO terms morphogenesis, signal transduction, and differentiation — all expected for post-mitotic neurons and glia (Fig. 5B). In addition, neuropeptide signaling and cell morphogenesis are both significantly enriched (p<10^-30 and p<10^-21, respectively) in distinct sub-clusters (Fig. 5C). We conclude that group A is enriched for genes that are expressed in neuroblasts but not differentiated neurons and glia, and group C is primarily composed of genes expressed in post-mitotic neurons and glia.

**Functional analysis of genes predicted to be expressed in neuroblasts but not neurons**

To determine if group A genes are required for neuroblast survival, proliferation, or self-renewal, we performed RNAi knock-down experiments. We selected genes for which transgenic RNAi stocks were available from the Vienna Drosophila RNAi Center, and we further restricted our analysis to those genes with human orthologs in order to enhance the relevance of our study to issues of human stem cell function (Fig. 6). This resulted in our analyzing 691 RNAi lines representing 595 genes. We reasoned that loss of function of genes with critical functions in neuroblasts would cause defective central nervous system development and eventual lethality, as seen in other genotypes which affect neuroblast function. Thus we screened for lethality in the progeny of males from each RNAi line crossed to "wor-Gal4 UAS-Dicer2 lines at 30 °C [wor-Gal4 drives expression in neuroblasts (Albertson et al., 2004), while Dicer2 improves RNAi efficacy (Dietzl et al., 2007)]. Of the 691 RNAi lines tested, 195 (28%) cause lethality or semi-lethality (Fig. 6A). We found that of the genes for which we tested multiple RNAi lines, 84% exhibit the same lethality phenotype for both lines. Few RNAi lines cause embryonic lethality at 30 °C, and in these cases larval stages were obtained by setting up crosses at 18 °C and shifting larvae to 30 °C after embryogenesis. The lack of a lethal phenotype in 72% of the lines may be due to either inefficient RNAi knock-down of gene expression or the non-essential function of the gene in larval neuroblasts. Hence we restricted our subsequent analysis to the 28% of lines with a lethal or semi-lethal phenotype.

We tested each of the lethal or semi-lethal genes for a change in neuroblast number, reasoning that genes expressed in neuroblasts but not neurons may play a role in neuroblast survival, quiescence, identity, asymmetric division, or self-renewal. We performed the same RNAi experiments as above and determined the number of Dpn-" Mira-" central brain neuroblasts (optic lobe neuroblasts were not assayed). We found that nearly one half of lethal genes (86) cause a significant

---

**Identification of genes predicted to be expressed in neuroblasts but not neurons**

To determine whether the ectopic neuroblasts in the mutant brains express wild-type neuroblast genes, we tested whether genes known to be expressed primarily or exclusively in neuroblasts are found in group A; indeed all such positive control genes (with the N. Jan, personal communication). The other two dorso-medial type II lineages (DM4 and 5) exhibit some expression of Optix in a subset of neuronal progeny, but it is absent from the neuroblasts and INPs in these lineages (Fig. 4B and C and not shown). In addition, a single dorsal type I neuroblast expresses Optix (Fig. 4C). Inspection of mutant brains further confirmed the type II-biased expression of Optix, in that brat mutant brains exhibit a marked increase in Optix+ neuroblasts (Fig. 4D), and in lgl pins, the increase in Optix is almost exclusively in a Pros- (type II-originating) region of the brain (Fig. 4E). Our results indicate that our clustering relationships can be used to predict type I/type II expression bias with good accuracy. We conclude that Optix is primarily expressed in type II but not type I neuroblasts, and that Optix and the other nine genes in this cluster are excellent candidates for regulators of type II neuroblast identity.

**Fig. 2. Results of cluster analysis.** (A) Cluster analysis—categorized genes with expression changes in mutant compared to wild-type brains, divided into three groups (A, B, and C). The dendrogram at the top is labeled according to the mutant genotype; Roman numerals indicate the neuroblast subtype(s) enriched in each mutant (I = type I; II = type II). (B) Log2 expression changes (mutant/wild-type) averaged over all genes in each group.
change in central brain neuroblast numbers (Fig. 6B and C; Table S1). A majority of these changes are decreases in neuroblast number, as expected based on the predicted expression of these genes in neuroblasts. Two genes known to regulate neuroblast numbers were detected in the screen, mira and aurora borealis, thereby validating this approach. Importantly, all of these genes have clear mammalian orthologs. We conclude that our RNAi-based screening method has yielded a list of 84 new candidates for regulating neuroblast self-renewal (Table S1).

Discussion

It has previously been shown that co-clustering of genes in expression profiling data is likely to reflect physical or genetic interactions (Ge et al., 2001; Jansen et al., 2002) and participation in the same pathways (van Noort et al., 2003). Our results are consistent with these conclusions. For example, we identified a small group of 11 genes containing the only two genes known to be expressed in type I but not type II neuroblasts, and showed that a third gene has a similar pattern of expression — thus all genes in this cluster are likely to be expressed in type I but not type II neuroblasts. Furthermore, the strong enrichment of GO terms in small sub-clusters within both group A and group C (Fig. 5) indicates that genes within these clusters are likely to share similar functions or processes.

Differences between type I/type II neuroblasts are caused by a small number of genes

At the outset of this study, we expected to find a large group of genes that were differentially expressed in type II versus type I neuroblasts, because these neuroblasts have such strikingly different cell lineages. However, we were only able to identify a few gene clusters that were differentially regulated in such a type I/type II consistent manner — the 11 genes in the pros/ase cluster depleted in type II neuroblasts and the 10 genes enriched in type II neuroblasts (Figs. 3A and 4A). This suggests that the small number of genes that we identified may play a disproportionately large role in generating differences between type I and type II neuroblasts. Might pros and ase be the only genes regulating type I/type II differences? Both Ase and Pros can promote cell cycle exit (Choksi et al., 2006; Dominguez and Campuzano, 1993; Li and Vaessin, 2000; Southall and Brand, 2009; Wallace et al., 2000), which may result in the Ase⁺ Pros⁺ type I progeny taking a GMC identity and undergoing just one terminal division and the Ase⁻ Pros⁻ type II progeny taking an INP identity and continuing to proliferate. Indeed, the misexpression of either Ase or low levels of Pros in type II neuroblasts is sufficient to cause the loss of INPs and/or their premature cell cycle exit, thereby decreasing lineage size toward the size of type I neuroblasts.

Fig. 3. Differential expression of genes excluded from type II neuroblasts. (A) Position within group B of sub-cluster containing pros and ase as well as nine other genes with unknown expression patterns. Log₂ enrichment over wild-type is shown, averaged over all genes in the sub-cluster. (B) Confocal image of a wild-type brain lobe at 120 h ALH. Multiple type I neuroblasts, four type II lineages, and three type II neuroblasts are visible. (C) Enlargement of the region boxed in (B). All neuroblasts are Dpn⁺; a subset of type I neuroblasts are Rx−. (D) Confocal image of a wild-type brain lobe at 120 h ALH. INPs are also Rx−. Rx is expressed in a subset of neuronal progeny in both type I and type II lineages. (D) brut mutant brain lobe (120 h ALH) contains many Dpn⁺ neuroblasts, but these cells do not express Rx. Rx is expressed in a few of the Pros⁺ cells, all of which in this focal plane are neurons and do not express Dpn. (E) Ifg pins brain lobe (120 h ALH) in which Rx is expressed in a subset of ectopic Dpn⁺ neuroblasts. Rx expression is limited to cells expressing Pros, which in ifg pins also express Ase (Fig. 1E, inset) and are likely derived from expansion of type I neuroblasts. The Pros⁺ regions (type II-derived) are entirely Rx−. White arrow: Rx⁺ type I neuroblast; white arrowheads: Rx− type I neuroblasts; yellow arrows: Rx⁺ type II neuroblasts; mCD8::GFP driven by R9D11-Gal4 marks a subset of type II lineages, but not the type II neuroblasts themselves (Bayraktar et al., 2010). Shown outlined here with yellow dashed lines are several dorso-medial type II lineages [DM 2, 3, 4, 5, and 6 in (B); DM 2, 3, and 4 in (C)]. Scale bars: 100 μm in (B), (D), and (E); 10 μm in (C).
However, it is unclear what is required to fully transform these cells into type I neuroblasts; addressing this question will require additional molecular markers (some provided by our work here) and tracing the axon projections of the progeny of these “transformed” neuroblasts (e.g. do they now fail to make intrinsic neurons of the adult central complex?). The fact that mutants in ase and pros do not transform type I neuroblasts into type II neuroblasts (Bowman et al., 2008; Weng et al., 2010) indicates that other genes, perhaps some in the pros/ase cluster described here, are also important for specification of type I neuroblast identity.

Group A: candidate neuroblast-specific genes and neuroblast homeostasis regulators

We found that the neuroblasts in each mutant have remarkably similar expression profiles, as shown by the extensive list of similarly expressed genes in group A and by the list of genes with depleted expression in mutant brains, represented by group C. We believe that these categories provide lists of genes that are representative of those expressed in neuroblasts and neurons, respectively, based on all known neuroblast-specific genes showing up in group A and all known neuron- or glial-specific genes being excluded from group A.

Our RNAi-based screen helped to substantiate this claim, in that a substantial percentage of group A genes caused lethality when subjected to neuroblast-specific knock-down. We do not believe that off-target effects lent a significant amount of error to these

**Table 2**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cell type (in larval brain)</th>
<th>Array group</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>wor</td>
<td>Neuroblast</td>
<td>A</td>
<td>Ashraf et al. (2004)</td>
</tr>
<tr>
<td>dpn</td>
<td>Neuroblast</td>
<td>A</td>
<td>Bier et al. (1992)</td>
</tr>
<tr>
<td>cycE</td>
<td>Neuroblast</td>
<td>A</td>
<td>Caldwell and Datta (1998)</td>
</tr>
<tr>
<td>grh</td>
<td>Neuroblast</td>
<td>A</td>
<td>Uv et al. (1997)</td>
</tr>
<tr>
<td>dmyc</td>
<td>Neuroblast</td>
<td>A</td>
<td>Betschinger et al. (2006)</td>
</tr>
<tr>
<td>E(spl)/my</td>
<td>Neuroblast</td>
<td>A</td>
<td>Almeida and Bray (2005)</td>
</tr>
<tr>
<td>insc</td>
<td>Neuroblast</td>
<td>A</td>
<td>Parmentier et al. (2000)</td>
</tr>
<tr>
<td>mira</td>
<td>Neuroblast</td>
<td>A</td>
<td>Peng et al. (2000)</td>
</tr>
<tr>
<td>pros</td>
<td>Type I neuroblast, GMC neurons</td>
<td>B</td>
<td>Bello et al. (2008), Boone and Doe (2008), Bowman et al. (2008)</td>
</tr>
<tr>
<td>ase</td>
<td>Type I neuroblast, GMC neurons</td>
<td>B</td>
<td>Bowman et al. (2008)</td>
</tr>
<tr>
<td>elav</td>
<td>Neurons</td>
<td>C</td>
<td>Robinow and White (1988)</td>
</tr>
<tr>
<td>gcm</td>
<td>Glia</td>
<td>C</td>
<td>Hosoya et al. (1995)</td>
</tr>
<tr>
<td>repo</td>
<td>Glia</td>
<td>C</td>
<td>Xiong et al. (1994)</td>
</tr>
</tbody>
</table>
lethality data for two reasons: (1) a similar percentage of lethal and non-lethal RNAi lines (about 30%) had more than one non-specific target, which indicates that the observed lethality was due to specific target knock-down; and (2) a majority of genes (about 85%) caused the same lethality phenotype when targeted with multiple independent RNAi lines. Interestingly, we found that many RNAi lines caused lethality with no concomitant change in neuroblast numbers (Fig. 6). We believe this to be due to neuroblast defects which disrupt normal brain development without causing neuroblast loss, per se. For instance, neuroblast failure to make the proper number or type of progeny might be expected to cause such a phenotype. It will be interesting to investigate the specific effects these essential genes have on neuroblast function. We note that all of the putative regulators of neuroblast homeostasis identified here have mammalian orthologs; these genes are excellent candidates for regulating self-renewal of mammalian neural stem cells.

Group B: expression in subsets of neuroblasts or neurons?

Group B genes apparently are not expressed in all neuroblasts like the group A genes, nor in all neurons or glia like group C genes. However, group B genes are more likely to be expressed in subsets of neurons, not neuroblasts, because group B genes as a whole have an over-representation of GO terms more similar to group C than to group A (Fig. 5B). Why then are group B genes excluded from group C, the neuron cluster? One possible explanation is that different neuroblast lineages are affected in each mutant, and thus different subsets of neurons are missing in each mutant. If different neuroblast lineages express different genes (which seems likely), then each mutant would be missing a unique subset of neural differentiation genes, leading to the cluster being excluded from group C. This model raises the intriguing possibility that group B sub-clusters may represent lineage-specific genes.

It is also possible that the mutant genotypes themselves may cause unique transcriptional differences, leading to a cluster of genes in group B. For example, several small sub-clusters in group B are expressed differently only in aPKC brains (Fig. S2). These transcriptional differences are not correlated with the number of type I or type II neuroblasts. Instead, these genes appear to be differentially expressed in response to elevated aPKC. Drosophila aPKC has been best studied as a component of the apical complex in mitotic neuroblasts, and its capacity for causing ectopic self-renewal has been shown to be reliant on both its catalytic activity and its membrane localization (Atwood and Prehoda, 2009; Lee et al., 2006b). However, aPKC has been ascribed a role in neuroblast proliferation as well as in polarity (Chabu and Doe, 2008; Rolls et al., 2003), and a vertebrate homolog, PKC-ζ, was shown to possess a nuclear...
role in both proliferation of neural progenitors and neuronal cell fate specification (Sabherwal et al., 2009). These observations are consistent with a role of aPKC in causing transcriptional differences.

Conclusions

Our findings highlight the importance of expression profiling of multiple genotypes. This method allowed us to get a more reliable picture of the group A genes expressed in neuroblasts, because genes with lineage-specific or genetic background-specific changes in expression appeared to be focused into group B, where they do not interfere with the clustering of groups A and C. In addition, we identified two small sub-clusters of genes in group B that are excellent candidates for being preferentially expressed in type I or type II neuroblasts, for which there have been few examples to date. Finally, we conclude that group A genes are likely to be expressed in neuroblasts, and our functional studies have identified 84 genes that are conserved in mammals and required for regulating neuroblast numbers in Drosophila. Future phenotypic analysis in Drosophila will determine whether these genes regulate neuroblast survival, quiescence, asymmetric cell division, and/or self-renewal. Future studies on the expression and function of orthologous genes in mouse neural progenitors and human stem cells (IP or neural) will reveal whether they have conserved roles from flies to mammals.

Supplementary materials related to this article can be found online at doi:10.1016/j.ydbio.2011.10.020.

Acknowledgments

We would like to thank Sen-Lin Lai for critical reading of the manuscript, Francesca Pignoni, Uwe Walldorf, Jim Skeath, and the Developmental Studies Hybridoma Bank for antibodies, and Jason Boone for initial characterization of lgl lgd flies. We also thank the Bloomington Stock Center and the Vienna Drosophila RNAi Center for flies. This work was funded by NIH R01 HD27056 (CQD), NIH training grant T32-HD07348 (TDC, OAB), and an NSF predoctoral fellowship (MRM). CQD is an HHMI Investigator.

References


Izergina, N., Balmer, J., Bello, B., Reichert, H., 2009. Postembryonic development of tran-