Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 296 (2006) 545–560
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Expression profiling of glial genes during *Drosophila* embryogenesis

Benjamin Altenhein [a,1], Angela Becker [a,1], Christian Busold [b], Boris Beckmann [b], Jörg D. Hoheisel [b], Gerhard M. Technau [a,*]

[a] *Institute of Genetics, University of Mainz, Germany*
[b] *Division of Functional Genome Analysis, Deutsches Krebsforschungszentrum Heidelberg, Germany*

Received for publication 22 March 2006; revised 19 April 2006; accepted 21 April 2006
Available online 5 May 2006

Abstract

In the central nervous system of *Drosophila*, the induction of the glial cell fate is dependent on the transcription factor *glial cells missing* (*gcm*). Though a considerable number of other genes have been shown to be expressed in all or in subsets of glial cells, the course of glial cell differentiation and subtype specification is only poorly understood. This prompted us to design a whole genome microarray approach comparing *gcm* gain-of-function and, for the first time, *gcm* loss-of-function genetics to wildtype in time course experiments along embryogenesis. The microarray data were analyzed with special emphasis on the temporal profile of differential regulation. A comparison of both experiments enabled us to identify more than 300 potential *gcm* target genes. Validation by in situ hybridization revealed expression in glial cells, macrophages, and tendon cells (all three cell types depend on *gcm*) for 70 genes, of which more than 50 had been unknown to be under *gcm* control. Eighteen genes are exclusively expressed in glial cells, and their dependence on *gcm* was confirmed in situ. Initial considerations regarding the role of the newly discovered glial genes are discussed based on gene ontology and the temporal profile and subtype specificity of their expression. This collection of glial genes provides an important basis for the clarification of the genetic network controlling various aspects of glial development and function.

Keywords: Glial development; *gcm*; Glial genes; Microarrays; *Drosophila* embryogenesis

Introduction

The central and peripheral nervous systems (CNS and PNS) comprise two major cell types: neurons, which receive, transmit and integrate information, and glial cells that ensheath the neurons and their axons and fulfil several accessory functions. In both vertebrates and invertebrates, these two cell types are derived from multipotent neural stem cells. In *Drosophila*, neural stem cells delaminate from the neurogenic ectoderm in a stereotypic spatial and temporal pattern. Each of these stem cells can be addressed by its subectodermal position and specific gene expression (Doe, 1992; Urbach and Technau, 2003) as well as by its characteristic cell lineage (Bossing et al., 1996; Schmidt et al., 1997). According to their progeny, these cells can be divided into neuroblasts (NB) that give rise only to neurons, neuroglioblasts (NGB) generating both neurons and glial cells and glioblasts (GB) that produce only glia. The decision to acquire particular neuronal or glial cell fates needs precise regulation in order to generate a functional nervous system. In *Drosophila*, the glial cell fate is induced by the transcription factor *glial cells missing*/*glial cells deficient* (*gcm*/*glide*) (Hosoya et al., 1995; Jones et al., 1995; Vincent et al., 1996). *gcm* is transiently expressed in all embryonic glial cells, except for the mesectoderm-derived midline glia. A second *gcm* gene, *gcm2*, is also expressed in all lateral glial cells, but much weaker than *gcm* itself (Kammerer and Giangrande, 2001). Both paralogs form a gene complex with shared enhancer elements (Kammerer and Giangrande, 2001; Ragone et al., 2003; Jones et al., 2004; Jones, 2005). The key role for *gcm* in glial cell development is underlined by the fact that *gcm* mutant embryos lack nearly all lateral glial cells, whereas ectopically expressed *gcm* transforms presumptive neurons into glial cells (Hosoya et al., 1995; Jones et al., 1995; Vincent et al., 1996; Akiyama-Oda et al., 1998). Like *gcm*, *gcm2* promotes the glial fate upon ectopic expression, but *gcm2* loss-of-function has

* Corresponding author. Fax: +49 6131 392 4584.
E-mail addresses: technau@uni-mainz.de, balt@uni-mainz.de (G.M. Technau).
[1] These authors contributed equally to this work.

0012-1606/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2006.04.460

only a mild effect on glial cells (Kammerer and Giangrande, 2001).

Beside their role in early glial cell development, recent studies demonstrated a role for *gcm* as well as *gcm2* in both glial and neuronal differentiation in the optic lobes of third instar larval stages (Chotard et al., 2005). Hence, the mode of neuronal or glial downstream target gene activation by Gcm appears to be context-dependent. Furthermore, both genes are expressed in cells of the hematopoietic anlage, the prohemocytes, and induce the differentiation into plasmatocytes and finally into macrophages (Bernardoni et al., 1997; Alfonso and Jones, 2002; Bataille et al., 2005).

To shed light on the regulation of gliogenesis, the promoter region of *gcm* was subject to intense studies. Transcriptional regulation of *gcm* as the nodal point of gliogenesis in the embryo is achieved by multi-component upstream pathways (reviewed in Jones, 2005), as well as auto- and cross-regulation with *gcm2* (Miller et al., 1998; Kammerer and Giangrande, 2001; Jones et al., 2004). Because of the transient expression of *gcm*, it is necessary to activate glial-specific genes, which accomplish the differentiation and maintenance of the glial cell fate. To date, only a few downstream target genes of Gcm have been well characterized, like *reversed polarity* (*repo*), *pointed* (*pnt*), and *tramtrack* (*ttk*), which are all expressed in lateral glial cells (Klämbt, 1993; Klaes et al., 1994; Xiong et al., 1994; Halter et al., 1995; Giesen et al., 1997). The homeodomain protein Repo cooperates with Pnt to promote the glial fate, as well as with Ttk to suppress the neuronal fate (Badenhorst et al., 2002; Yuasa et al., 2003). Synergistic cooperations between the transcription factors Gcm, Repo, Pnt and Ttk seem to be an additional way to activate glial-specific genes, like the expression of *locomotion defects* (*loco*) (Granderath et al., 1999, 2000). Both *repo* and *loco* are exclusively expressed in all lateral glial cells, while *pnt* and *ttk* are also expressed in neurons and/or the Gcm-independent midline glia. Many other genes, which have been described to be expressed in lateral glial cells, are expressed in subsets of these cells only. The composition of these subsets varies in number and identity of the cells and does not necessarily reflect the classification of glial subtypes as introduced by Ito et al. (1995). This classification is based on the position and the morphology of the cells in the embryonic CNS. Whether a developmental program underlying this subtype specification exists has not been shown so far. The clarification of the processes underlying the determination and differentiation of glial subtypes and their specific functions in the developing and mature nervous system requires the identification of the genetic network acting downstream of or in cooperation with the master regulator Gcm.

Recently, two screens for genes acting downstream of *gcm* have been published, which took advantage of the microarray technique to compare gene expression in wildtype embryos versus embryos overexpressing *gcm* throughout the CNS (Egger et al., 2002; Freeman et al., 2003). Using an in silico approach and a public database search in addition to the microarray approach, Freeman et al. (2003) identified about 40 new glial genes. Nevertheless, there was only a minor overlap of candidate genes uncovered by the three approaches of Freeman et al. as well as between the two microarray screens (Egger et al., 2002; Freeman et al., 2003). This prompted us to design a new-microarray-based screen for further *gcm* target genes. We used a whole-genome microarray approach that compares ectopic expression of *gcm* to wildtype and, for the first time, *gcm* loss-of-function to the wildtype situation. To ectopically express *gcm*, we used a Gal4 driver line, which is exclusively expressed in the CNS at the same time as the endogenous *gcm*. The loss-of-function experiment was achieved by sorting staged, living *gcm* mutant embryos by means of different green fluorescent protein (GFP) carrying balancer chromosomes and an embryo sorter (Furlong et al., 2001b). To identify the dynamic changes in gene expression, we generated a time course experiment throughout embryogenesis. This antagonistic approach in combination with time course gene expression profiling, various quality controls and carefully selected filtering methods enabled us to identify about 70 novel *gcm* target genes, 18 of which exclusively expressed in glial cells.

Materials and methods

Fly strains

Mz1060-Gal4, UAS-*gcm* (Bloomington stock B-#5446), $gcm^{N7\text{-}4}$ (Bloomington stock B-#4104) (Vincent et al., 1996), *Kr*-GFP (Bloomington stock B-#5194) and *Ubi*-GFP (Bloomington stock B-#4888) balancer chromosomes (Casso et al., 1999, 2000), wildtype Oregon R.

Embryo collection

Eggs were collected for 1 h at 25°C on standard apple juice plates and shifted to either 18°C, 25°C or 29°C for further development. Embryos were dechorionated using 7.5% hypochlorite, washed in water and either fixed for antibody staining, snap-frozen in liquid nitrogen for RNA preparation or transferred to PBT (PBS, 3% Tween-20) and GFP-sorted.

Antibody staining

For each embryo collection, a small fraction was fixed as described (Patel, 1994) and stained for the glial marker protein Repo using a rabbit anti-Repo antibody (Halter et al., 1995) and alkaline-phosphatase-conjugated donkey anti-rabbit secondary antibodies. In situ hybridizations were counterstained with rabbit anti-Repo antibodies and biotin-conjugated secondary antibodies (Dianova, Hamburg, Germany). Immunolabelings were performed with rabbit anti-Repo and rat anti-Elav antibodies (7E8A10, DSHB Iowa, USA) and donkey anti-rabbit/anti-mouse secondary antibodies conjugated with Fitc/Cy5 respectively (Dianova, Hamburg, Germany). Embryos were staged according to standard morphological markers (Campos-Ortega and Hartenstein, 1997).

Sorting

$gcm^{N7\text{-}4}$ mutant flies were balanced with either *Kr*-GFP (for stages 10 to 13) or *Ubi*-GFP (for stages 14 to 16) balancer chromosomes. Eggs were collected as described above, transferred into PBT buffer according to Furlong et al. (2001b) and homozygous mutant embryos were automatically sorted with the Copas™ Select embryo sorter (Union Biometrica, Somerville, MA, USA) by their lack of GFP expression. The two different balancer chromosomes used showed different characteristics concerning the ability to distinguish between homozygous and heterozygous embryos at different stages. In Fig. 1, two examples of the Copas™ sorter plots are given. An aliquot of sorted embryos was fixed for antibody staining as described above, and the major portion was snap-frozen in liquid nitrogen.

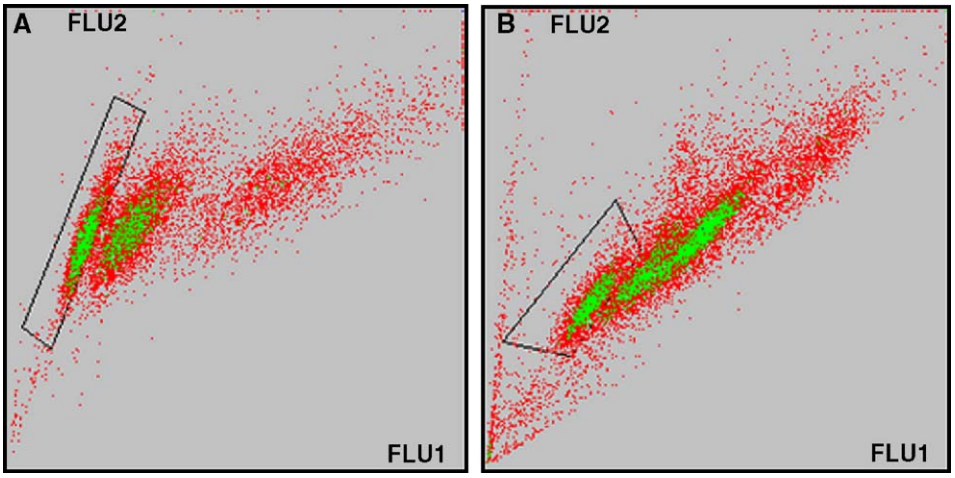

Fig. 1. Sorting of homozygous mutant embryos. (A, B) Two different GFP balancer chromosomes were used to balance the *gcm*$^{N7-4}$ mutation. *Kr*-GFP (A) was used to balance and sort stages 10–12, *Ubi*-GFP (B) was used for stages 13–16. Sorter plots of the Copas™ embryo sorter (Union Biometrica, Somerville, USA) with green fluorescent signal (FLU1) plotted against autofluorescence (FLU2) show the three populations of genotypes. The black rectangles indicate the sorted fraction which lacks GFP signal.

RNA preparation

Total RNA was prepared from frozen embryos using the RNeasy® kit followed by polyA$^+$-RNA extraction with the Oligotex® kit (both Qiagen, Hilden, Germany) according to manufacturer's instructions. One microgram of polyA$^+$-RNA was used for each labeling reaction. The quality of polyA$^+$-RNA preparations was checked by Northern blot and RT-PCR under standardized conditions.

Microarrays

Microarrays were used (Heidelberg FlyArray) that contain PCR products of 21,306 *Drosophila* open reading frames together with 2502 controls, spotted in two replicates on each array (~48,000 spots per array) (Hild et al., 2003).

Sample preparation and microarray hybridization

An indirect labeling method was used to reverse-transcribe 1 μg of denatured polyA$^+$-RNA with 4.5 μg random primers and 0.75 μg oligo-dT primers (both from Invitrogen, Karlsruhe, Germany). First-strand cDNA synthesis, sample preparation and microarray hybridization were performed as described in Hild et al. (2003).

Image analysis

Hybridized arrays were scanned directly after hybridization using the GeneTac™ LS IV scanner (Perkin-Elmer Life Sciences, Wellesley, USA) and the corresponding software. Fluorescent-labeled spotted controls on the arrays were used to align laser intensities and photomultiplier settings for both channels. Resulting images (16-bit gray-scale TIFF-format) for each channel were loaded into the GenePix 5.0 software (Axon Instruments, Union City, USA), and data were extracted as GenePix result files (gpr-format) for further analysis.

Replicates

For each developmental stage examined in both experiments, at least four independent replicates have been analyzed starting with independent egg collections. All replicates include balanced dye swaps, except for stage 12 in the GOF and stages 11 and 16 in the LOF, where only one dye swap was done. Only arrays with a correlation coefficient above 0.8 were used for further analysis. For every replicate, the developmental stage, the quality of the embryo collection and the accuracy of the sorting were checked by antibody staining.

Normalization, filtering and correspondence analysis

Statistical data analysis was performed with the Multi-Conditional Hybridization Intensity Processing System (M-CHiPS) (Fellenberg et al., 2002). Normalization was performed using linear regression normalization with the M-CHiPS software package. Genes were filtered out when their normalized median transcription intensities remained below a threshold of 10,000 pixels in all conditions and/or had a ratio to the control condition of less than 1.5-fold. The significance of changes was assessed by a high stringent criterion. 'Min–max separation' is calculated by taking the minimum distance between all data points of two conditions (Beissbarth et al., 2000). Genes that exhibited a min–max value of less than 0.0, 0.2 or 0.5 were discarded from analysis in three subsequent filtering steps. Correspondence analysis (CA) (Greenacre, 1984, 1993) is a clustering and projection method, which allows to plot genes and hybridization conditions in the same space. In the resulting plot, the displayed χ^2 distance is a measure of association among genes and hybridizations.

In situ hybridization

In vitro transcription and labeling of RNA with Digoxygenin was performed using the Dig-RNA labeling mix according to manufacturer's instructions (Roche Diagnostics, Mannheim, Germany). Embryos from overnight collections were dechorionated in 7.5% hypochlorite, fixed in PBS/heptane/formaldehyde (35:50:15) and incubated for 1 h in hybridization buffer (50% formamide, 5× SSC, 100 μg/ml ssDNA, 0.1% Tween-20). For fluorescent in situ hybridization, embryos were treated with 0.1% sodium borohydrid solution in 0.1% PBT (PBS, 0.1% Tween-20) for 10 min to reduce the autofluorescent background. Hybridization was performed overnight at 55°C using 10 μl of the Dig-labeled RNA probe. The following procedure is described in Tautz and Pfeifle (1989); Jiang et al. (1991); Kosman et al. (1991). For fluorescence detection of Dig-labeled RNA probes, the TSA amplification Kit with either Cy3 or Cy5 (Perkin-Elmer, Norwalk, CT) was used according to manufacturer's instructions.

Results

Microarray analysis and quality control

Prior to array hybridization, we performed several tests concerning the quality of the prepared RNA (Northern blot and RT-PCR). Furthermore, different post-hybridization quality controls were performed in order to define and reduce the

methodical background noise and, thus, to obtain high reproducibility and reliability of our microarray data. We first hybridized wildtype stage 10 mRNA labeled with Cy3 and Cy5 in three replicates to determine background noise within identical samples. In order to compare differences in gene expression between the wildtypic flies used for the ectopic expression of *gcm*, we labeled RNA samples of stage 9 embryos from Mz1060 with Cy3 and from UAS-*gcm* with Cy5 and vice versa and hybridized these samples to microarrays. All control arrays showed correlation coefficients between 0.98 and 0.99. Data were normalized and filtered with the M-CHiPS software package as described in the Materials and methods section. We applied different filter criteria for both reproducibility and ratio of normalized intensities and determined the number of genes passing both criteria. A much higher number of differentially expressed genes above 1.5-fold regulation were obtained when Gal4 and UAS flies were compared than using identical wt samples. Thus, differences in the genetic background of phenotypically wildtype fly strains lead to differences in expression levels of certain genes and increase background noise. As a consequence, embryos from both Gal4 and UAS line were collected separately for control samples in the ectopic expression experiment. RNA was extracted and then pooled in equal amounts for sample labeling and hybridization. This mixed 'wildtype control' was not used for the analysis of homozygous *gcm* mutant embryos. For this experiment, Oregon R wildtype embryos were collected and treated as the sorted embryos. Both control experiments were taken into consideration for the selection of filter criteria in M-CHiPS. Hence, filtering data with a differential regulation above 1.5-fold and a min–max separation between 0.0 and 0.5 was chosen and applied to all microarray data.

Ectopic expression of gcm and comparison with wildtype

It has been shown that ectopic expression of *gcm* in the developing nervous system can induce the glial-specific transcription factor Repo, which is generally used as a marker for all lateral glial cells (Akiyama-Oda et al., 1998). Hence, ectopic *gcm* is sufficient to induce the glial cell fate at the cost of presumptive neurons. Conversely, ectopic expression of *gcm* in epidermal cells leads to an activation of *elav*, a postmitotic neuronal marker (Akiyama-Oda et al., 1998). These findings prompted us to search for a Gal4 line which drives ectopic *gcm* expression restricted to the nervous system (and not in the ventral neuroectoderm which produces also epidermal cells) in a time window similar to endogenous *gcm* expression. The Gal4 line Mz1060 fitted these criteria. This line shows Gal4 expression in nearly 60% of the cells in the ventral nerve cord, beginning at stage 10, increasing to stage 13 and diminishing until stage 15. In Fig. 2, the Repo/Elav pattern of both wildtype embryos and embryos with ectopic expression of *gcm* (Mz1060::*gcm*) in stages 12, 14, and 16 is shown.

To analyze the effect of this ectopic activation of *gcm* on microarrays (gain-of-function, GOF), we collected embryos from eight different time points (stage 9 to stage 16). Embryos from each developmental stage were collected separately for Mz1060::*gcm* as well as for both Mz1060 and UAS *gcm* flies as 'wildtype control'. This was done in five separate experiments starting from individual fly crosses and separate egg collections for repeat experiments. Homogeneity of staging and effect of ectopic activation of *gcm* were checked by anti-Repo antibody stainings. RNA was extracted, labeled and hybridized to microarrays separately. Gene lists of differentially regulated genes were created for all eight time points (Table 1 and Supplementary Table 1).

Sorting of homozygous gcm mutant embryos and comparison with wildtype

In order to analyze differences in gene expression between *gcm* mutant embryos (loss-of-function, LOF) and wildtype, we made use of balancer chromosomes carrying the gene coding for the green fluorescent protein (GFP) (Casso et al., 1999, 2000). We balanced the $gcm^{N7\text{-}4}$ mutation (Vincent et al., 1996) with either *Kr*-GFP or *Ubi*-GFP balancer chromosomes. Homozygous mutant embryos can be detected unambiguously by their lack of GFP expression. To circumvent the need for RNA amplification, we used the Copas™ Select embryo sorter (Union Biometrica, Somerville, MA, USA) for automated sorting of living embryos as described in Furlong et al. (2001a, b). Staged egg collections of $gcm^{N7\text{-}4}$/GFP balanced parental generation were sorted in up to six independent collections per developmental stage examined, and homozygous mutant embryos were kept. Wildtype (Oregon R) embryos were staged and treated accordingly, without sorting for GFP. A fraction of each sorting was fixed and stained with an antibody against the glial marker Repo to check both staging and accuracy of the sorting. Stages 10 to 14 and stage 16 were well staged, and the sorting resulted in an accuracy of 98% homozygous mutant embryos lacking Repo staining. Unfortunately, the stage 15 egg collections turned out to be too heterogeneous with respect to their developmental stage and, thus, were excluded from further analysis. Two GFP balancer chromosomes were used for sorting of stages 10 to 12 (*Kr*-GFP) and 13 to 16 (*Ubi*-GFP), respectively, because the discrimination between the three populations of embryos (homozygous *gcm* mutant embryos without GFP, heterozygous embryos and homozygous GFP balancer carrying embryos) was easier with *Kr*-GFP in younger stages and with *Ubi*-GFP in older stages (see Fig. 1). Microarray data were generated using the same filter criteria as for the ectopic expression (GOF experiment), and gene lists for six developmental stages were generated (Table 1 and Supplementary Table 2).

Processing of microarray data

Filtering for differential expression for every single time point provided us with gene lists of up to several hundred differentially regulated genes. These lists contain log2-transformed values below the aspired threshold of 1.5-fold differential regulation (see Supplementary Tables 1 and 2). This accounts for the fact that in M-CHiPS genes pass the fold regulation filter if at least one of the measurements in all

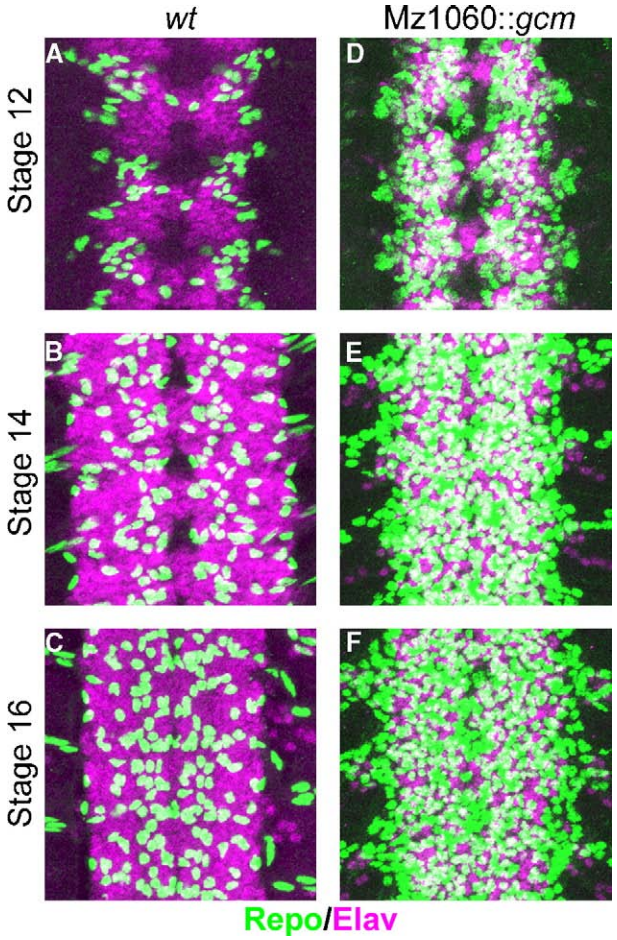

Fig. 2. Confocal images of the central nervous system of wt (A–C) and Mz1060::*gcm* (D–F) embryos stained against the glial protein Repo (green) and the postmitotic neuronal marker protein Elav (magenta). An excess of glial cells at the expense of neurons upon ectopic expression of *gcm* compared to wildtype is already detectable at stage 12 (A, D) and increases further until stage 14 (B, E). Ectopic glial cells are evenly distributed throughout the entire nervous system at stage 16 (C, F).

replicates shows differential regulation above 1.5-fold. The resulting value, though, is the median ratio of all replicates. We combined the single stages of each experiment to generate profiles for all differentially expressed genes in the GOF and LOF experiments. In Table 1, the number of differentially regulated genes for every single stage of both experiments is given. The entire (filtered) gene lists containing unique gene identifiers and normalized ratio values (log2-transformed) can be found in Supplementary Tables 1 (GOF) and 2 (LOF).

Potential glial genes were expected to be differentially upregulated upon ectopic expression of *gcm* and downregulated in the LOF background. Thus, we selected the upregulated genes of the GOF and downregulated genes of the LOF, respectively, and listed the genes according to the time point when first changes in expression appear (see Tables 2 and 3). Candidate gene selection was done by visual inspection of the differential regulation profile. The course of differential expression, the time point when first changes in expression appear, the strength of differential expression and the comparison between the two experiments (GOF and LOF) were taken into consideration to select potential candidate genes. Only few genes showed an "antagonistic" profile in both experiments (upregulated in GOF and downregulated in LOF). These were preferentially analyzed, but the percentage of positive candidate genes among those turned out to be more or less the same as for the selection of the two experiments alone. An even more important step in candidate gene selection was negative selection. We discarded all those genes from further analysis that showed differential regulation throughout either experiments without noteworthy changes or a reverted profile with respect to the predicted (and observed) profiles of *gcm* downstream targets. We also discarded genes that showed differential regulation in non-adjacent, remote stages, if differential regulation was below 2-fold. All these different filtering criteria were not necessarily applied at once but rather used in parallel. Both positive and negative filtering steps reduced the number of potential candidate genes to about 14% of all differentially regulated genes. Candidate gene selection was done in three subsequent rounds, starting with the smallest gene list obtained by the most stringent filter for reproducibility (min–max separation 0.5). Database information of the selected genes and results obtained from our in situ hybridizations was added to the lists. In the LOF experiment, the positively and negatively tested genes were scattered throughout the entire gene list (ordered according to time points of differential regulation) without any apparent preference. In the GOF, however, clusters of positively and negatively tested genes appeared. Genes within clusters that showed an accumulation of positive candidate genes were preferentially analyzed in subsequent selections. The most prominent clusters appeared for genes whose upregulation starts at stages 12 and 13 and lingers for at least one stage. Another accumulation of positive candidate genes was visible among those genes with upregulation at stage 16 only. As can be seen in Table 2, most glial-specific candidate genes of the GOF experiment were selected from these clusters.

In total, about 400 genes were selected. Literature and database searches, especially the superb and vastly expanding Berkeley *Drosophila* Genome Project's 'Patterns of gene expression in *Drosophila* embryogenesis' in situ database (http://www.fruitfly.org/cgi-bin/ex/insitu.pl), were used to get information about the 400 selected genes. More than 40 were among our microarray target genes that were already known to be expressed in glial cells. These genes are marked together with the source of information in Tables 2 and 3 and in Supplementary Tables 3 and 4. Another 39 genes were discovered by the two already published microarray screens for downstream targets of *gcm* (Egger et al., 2002; Freeman et al., 2003), which are also marked in the respective tables. For 300 of the remaining filtered genes, EST clones or cDNAs were available. These genes were analyzed by in situ hybridization for their expression in wildtype as well as in $gcm^{N7\text{-}4}/gcm^{N7\text{-}4}$ and Mz1060::*gcm* backgrounds.

Validation of candidate genes by in situ hybridization

For 300 of the selected differentially regulated genes, Digoxygenin-labeled RNA probes were synthesized and hybridized to wildtype embryos of all developmental stages. Half of these in

Table 1
Statistical analysis of the two time course microarray experiments

GOF	Stage 9	Stage 10	Stage 11	Stage 12	Stage 13	Stage 14	Stage 15	Stage 16	Total no.	Min–max-separation
No. of differentially regulated genes (total)	147	167	150	138	96	28	31	34	387	0.5
Upregulated genes (>1.5)	116	149	143	137	84	27	31	28	316	0.5
Downregulated genes (<−1.5)	31	18	7	1	12	1	0	6	71	0.5
No. of differentially regulated genes (total)	596	656	389	346	291	82	80	91	1152	0.2
Upregulated genes (>1.5)	285	447	305	331	263	74	79	74	683	0.2
Downregulated genes (<−1.5)	311	209	84	15	28	8	1	17	470	0.2
No. of differentially regulated genes (total)	877	743	425	395	388	107	107	111	1413	0.0
Upregulated genes (>1.5)	399	496	328	372	322	98	106	89	789	0.0
Downregulated genes (<−1.5)	478	247	97	23	66	9	1	22	625	0.0

LOF	Stage 10	Stage 11	Stage 12	Stage 13	Stage 14	Stage 15	Stage 16	Total no.	Min–max-separation
No. of differentially regulated genes (total)	141	135	72	94	44		141	476	0.5
Upregulated genes (>1.5)	80	31	31	50	29		124	273	0.5
Downregulated genes (<−1.5)	61	104	41	44	15		17	205	0.5
No. of differentially regulated genes (total)	414	382	344	319	111		368	1367	0.2
Upregulated genes (>1.5)	243	104	114	169	56		274	744	0.2
Downregulated genes (<−1.5)	171	278	230	150	55		94	658	0.2
No. of differentially regulated genes (total)	598	575	627	632	245		560	2086	0.0
Upregulated genes (>1.5)	353	193	180	269	161		386	1056	0.0
Downregulated genes (<−1.5)	245	382	447	363	84		174	1080	0.0

All data were filtered in M-CHiPS with a constant threshold for differential regulation of 1.5-fold and three different values for reproducibility (min–max separation 0.5, 0.2 and 0.0). Differentially regulated genes were divided into up- and downregulated genes for each stage. The total number of differentially regulated genes combines all developmental stages analyzed. Since many genes are differentially regulated at more than one stage, the total number given is smaller than the sum of all stages. Some genes appear both up- and downregulated during the time course of each experiment, especially in the LOF. Hence, total number of differentially regulated genes is smaller than the sum of up- and downregulated genes.

situ hybridizations gave no signal at any time during embryogenesis. Since we cannot rule out that in these cases the method did not work, we do not know whether these genes are expressed in the embryo and, thus, they remain uncertain candidates, still. The other approximately 140 genes showed detectable staining and were analyzed in more detail (Tables 2, 3, Figs. 3, 4; GOF and LOF, respectively). In total, 58 genes showed clear expression in the nervous system, of which 18 genes are exclusively expressed in glial cells, another 15 genes are expressed in glial cells together with single neurons, and 22 genes showed a more ubiquitous staining in the CNS and/or PNS. One gene is expressed in glial cells together with macrophages, and 7 genes are expressed in cells of the hematopoietic lineage and macrophages alone, which also depend on *gcm* expression (Bernardoni et al., 1997; Lebestky et al., 2000; Alfonso and Jones, 2002). Another 9 genes are expressed in ventral stripes representing epidermal muscle attachment sites or tendon cells, and one gene showed expression in glial cells, neurons and tendon cells. *gcm* is expressed in tendon cells from stage 12 onwards and is required for their terminal differentiation (Soustelle et al., 2004). Another 60 genes are expressed in other tissues and were considered as negative results. The aim of our microarray approach was to identify novel genes involved in glial cell differentiation and to show their dependence on *gcm*. To test whether the expression of the newly discovered glial genes is dependent on *gcm*, we performed in situ hybridizations in wildtype as well as in *gcm* mutant embryos and in embryos with ectopic *gcm*. The results of our in situ hybridizations for all tested genes are included in Supplementary Tables 3 and 4, and all genes that showed expression in the CNS as well as in tendon cells and cells of the hematopoietic lineage are given in Tables 2 and 3.

Expression pattern and profile of novel glial genes

The majority of positive candidate genes from the GOF experiment showed expression in glial cells and neurons rather than exclusive expression in glial cells (Table 2). Seven GOF candidate genes showed expression in glial cells exclusively, of which four are expressed in many if not all glial cells (CG15860, CG2893, CG3168, EG:22E5.11) and three are expressed in glial subsets only (Figs. 3, 5). These are cell body glia (CBG) (CG6218, CG6783) and peripheral glia (PG) (CG9336). The GOF resulted in only two genes specifically expressed in tendon cells (CG9796, CG1153), and no gene expressed in macrophages or cells of the hematopoietic lineage. Glial expression of some of our candidate genes was also shown by Egger et al. (2002) and Freeman et al. (2003), and some were incorporated into the BDGP in situ database in the meantime. These genes are marked in Table 2. As can be seen in Supplementary Tables 3 and 4, many genes with previously described glial expression showed differential regulation in both of our experiments. These genes are not shown here as their expression was already described elsewhere.

The LOF experiment resulted in a higher number of candidate genes with expression in certain glial subtypes and only two genes with expression in many if not all glial cells (CG2893, CG3408). Among the glial-subtype-specific expression, we found genes expressed in cell body glia (CBG) (CG6218,

Table 2
GOF filtered and tested candidate genes

Gene	St. 9	St. 10	St. 11	St. 12	St. 13	St. 14	St. 15	St. 16	Expression pattern				Lit	GO-annotation
									wt	stage	*gcm+*	*gcm−*		
CG11164	+1.90	+2.17	+1.85	+2.28	+1.98	+1.81	+1.81	+1.89	CNS	9–16				unknown
CG1962	+3.29	+3.03	+3.02	+3.20	+2.14	+2.28	+2.28	+1.93	GN	11–17	++	0	*B	unknown
CG2051	+1.50	+1.52	+1.47	+1.69	+1.84	+1.50	+1.44	+1.44	CNS	11–15				histone acetyltransferase
CG16757	+1.73	+1.89	+2.04	+1.91	+1.77	+1.47	+1.35	+1.29	G	12–16				protein phosphatase 1 binding
CG4523	+1.62	+1.82	+1.77	+1.84	+1.66	+1.48	+1.44	+1.37	CNS	9–15				serine/threonine kinase
Ptp10D	+1.62	+2.04	+1.49	+1.51	+1.07	+1.21	+1.32	+1.04	G, N	10–17	++	0		tyrosine phosphatase
csw	+2.00	+2.18	+1.99	+1.56	+1.48	+1.44	+1.21	+1.25	CNS	13–16				tyrosine phopsphatase
CG9796	+2.22	+2.38	+1.49	+1.09	+1.57	−1.29	-1.02	+1.05	T	11–16				unknown
CG7433	+1.97	+2.14	+2.45	+2.78	+1.97	+2.00	+1.87	+2.07	G	13–17	0	--	*B	transaminase, amino acid metabolism
gcm	+1.03	+1.60	+1.51	+2.36	+2.54	+2.57	+2.41	+2.97						
CG1677	+1.43	+1.62	+1.59	+1.80	+1.90	+1.46	+1.51	+1.56	CNS	11–14				transcription factor
Cam	+1.44	+1.50	+1.82	+1.72	+1.94	+1.60	+1.62	+1.34	CNS	10–17	++	0		calmodulin binding, calcium ion binding
UbcD2	+1.37	+1.51	+1.51	+1.59	+1.94	+1.47	+1.56	+1.38	CNS	10–17				ubiquitin conjugating enzyme, ligase
smid	+1.41	+1.56	+1.70	+1.89	+1.43	+1.46	+1.46	+1.38	CNS	11–16			*E	ATP binding, serine type endopeptidase
CG10326	+1.59	+1.60	+1.37	+1.79	+1.47	+1.26	+1.42	+1.21	CNS	12–17				unknown
CG8171	+1.31	+1.43	+1.53	+1.79	+2.09	+1.72	+1.67	+1.89	G, N	10–16	++	0		DNA binding, replication checkpoint
CG6218	+1.00	+1.13	+1.43	+1.78	+1.63	+1.92	+2.23	+2.15	G	11–17	++	--		carbohydrate kinase
RpII140	+1.09	+1.48	+1.43	+1.58	+2.06	+1.68	+1.37	+1.53	CNS	10–17				RNA polymerase subunit
CG3227	+1.24	+1.09	+1.38	+1.64	+2.08	+1.41	+1.53	+1.20	CNS	11–15	0	0	*B	unknown
glu	+1.43	+1.36	+1.34	+1.65	+1.76	+1.54	+1.44	+1.26	CNS	11–15	++	--	*E	chromatin assembly/disassembly
Uba2	+1.15	+1.40	+1.39	+1.51	+1.89	+1.40	+1.47	+1.26	CNS	11–16				ubiquitin activating enzyme
BcDNA:LD08534	+1.17	+1.04	+1.23	+1.38	+2.09	+1.59	+1.63	+1.88	G, N	11–17			*E	dUTPase, dUTP diphosphatase
CG15141	+1.43	+1.31	+1.29	+1.46	+1.60	+1.47	+1.64	+1.51	G, N	11–16				DNA binding, ubiquitin protein ligase
Map60	+1.12	+1.08	+1.23	+1.39	+2.10	+1.35	+1.60	+1.47	G, N	11–16			*B	microtubule binding
CG4266	+1.07	+1.30	+1.27	+1.36	+1.70	+1.21	+1.18	+1.26	CNS	11–17				RNA binding
CG4936	+1.07	+1.08	+1.12	+1.25	+1.91	+1.10	+1.24	−1.06	G, N	11–17	0	0		transcription factor
CG6418	+1.08	+1.23	+1.33	+1.35	+1.60	+1.32	+1.36	+1.20	G, N	10–17	0	0		helicase, nucleic acid metabolism
pnut	+1.26	+1.31	+1.29	+1.23	+1.78	+1.27	+1.35	+1.44	CNS	11–15	++	--		structural constituent of cytoskeleton
CG1153	+1.12	−1.09	+1.07	−1.25	−1.67	−1.12	−1.42	+1.68	T	12–15				unknown
CG15860	+1.00	+1.00	+1.01	+1.00	+1.03	+1.17	+1.17	+1.77	G	14–17	++	--		calcium channel
CG2893	+1.05	+1.00	+1.05	+1.00	+1.02	+1.10	+1.17	+1.65	G	13–17	++	--	*B	cation transporter, antiporter
CG3168	+1.02	+1.00	+1.10	−1.03	−1.08	+1.02	+1.18	+2.07	G	14–17	++	--	*B	carbohydrate transporter, DNA metabolism
CG6783	+1.02	+1.02	−1.23	−1.08	−1.15	+1.11	+1.09	+2.03	G	11–17	++	--	*F	fatty acid binding, transporter
CG9336	+1.00	+1.00	−1.07	−1.06	−1.07	+1.04	+1.12	+1.75	G	11–17	++	--		unknown
EG:22E5.11	+1.04	+1.00	−1.03	+1.09	+1.10	+1.13	+1.20	+1.54	G	12–17	++	--	*F	G-protein coupled receptor

Gene identifiers are given as presented in all Supplementary data. The fold regulation for all examined developmental stages represents the log2-transformed ratio Cy5/Cy3 of median fitted intensities. Values were obtained from filtering with the M-CHiPS software package (red: upregulation above 1.5-fold; blue: downregulation below 1.5-fold). Expression pattern was confirmed by standard in situ hybridization. Expression in glial cells (G), in single neurons (N), ubiquitously in the CNS (CNS), and in tendon cells (T) was observed in stages as indicated. Gene expression in Mz1060::*gcm* (*gcm+*) and $gcm_{N7\text{-}4}/gcm_{N7\text{-}4}$ (*gcm−*) backgrounds compared to the wildtypic expression is indicated as increased expression (++), decreased/absent expression (−) or no changes in expression (0). For some of the tested candidate genes, the expression in glial cells was published during the time of data analysis. The source of information is given in the column titled 'Lit'. Expression information was either (*B) incorporated into the public in situ database (http://www.fruitfly.org/cgi-bin/ex/insitu.pl) or published by Egger et al. (2002) (*E) or by Freeman et al. (2003) (*F). GO annotations are presented according to the information given in FlyBase. (For interpretation of the references to colour in this table legend, the reader is referred to the web version of this article.)

Table 3
LOF filtered and tested candidate genes

Gene	St. 9	St. 10	St. 11	St. 12	St. 13	St. 14	St. 15	St. 16	Expression pattern				Lit	GO-annotation
									wt	stage	*gcm+*	*gcm−*		
BG:DS05899.3		−2.12	−2.61	−3.01	−3.34	−2.80		−2.05	G, H	9–17	++	--	*B	unknown
CG6218		−1.50	−1.77	−1.95	−2.41	−1.94		−1.57	G	11–17	++	--	*F	carbohydrate kinase
CG8193		−1.97	−2.94	−1.88	−3.87	−4.20		−3.96	H	10–17	0	0		monophenol monooxygenase, defense response
Caps		−1.93	−2.91	−3.37	−2.71	−1.47		−1.68	G, N	10–17				calcium/lipid binding, neurotransmitter secretion
CG4844		−2.66	−3.71	−1.91	−1.51	−1.72		+1.64	G	9–17	0	--		unknown
CG16971		−1.98	−2.29	−2.01	−1.46	−1.36		−1.38	G, H	10–17				unknown
CG15015		−1.64	−2.02	−1.22	−1.08	−1.03		−1.37	G, T	10–16				Rho interactor, signal transduction
CG8222		−1.50	−2.15	−1.03	−1.19	−1.14		+1.00	H	11–16				transmembrane receptor protein tyrosine kinase
crol		−1.95	−1.69	+1.02	−1.11	−1.01		+1.37	CNS	11–16				transcription factor
gcm		−2.04	−1.92	−1.43	−1.44	−1.45		−1.30						
Lan B1		−1.54	−1.89	−1.03	−1.53	−1.18		−1.17	H	11–16				structural molecule, cell adhesion
Nrv2		−1.90	−1.57	−1.28	−1.13	−1.04		−1.14	G, N	10–17			*F	sodium:potassium-exchange factor, ATPase
Reg-3		−1.80	−2.48	−1.38	−1.38	−1.18		+1.54	G	12–17				electron transporter, pyrimidine base metabolism
tap		−1.73	−3.22	−1.20	−1.20	−1.16		+1.11	G, N	11–16	0	0	*B	transcription factor
CG8965		−1.65	−1.15	+1.20	−1.10	−1.11		−1.16	G, N	10–16	++	--	*F	unknown
CG6783		−1.15	−2.43	−2.57	−2.71	−1.73		−1.59	G	11–17	++	--	*F	fatty acid binding, transporter
Fpps		−1.13	−2.00	−3.37	−2.53	−2.34		−1.52	CNS	11–15				transferase activity, cholesterol metabolism
scf		−1.79	−1.91	−1.85	−1.78	−1.56		−1.74	H	10–16				calmodulin binding, DNA topoisomerase
Aats-phe		−1.08	−1.89	−2.40	−2.54	−1.61		−1.17	T	12–16				phe-tRNA ligase
BcDNA:GH04753		−1.18	−1.64	−1.68	−2.23	−1.43		+1.02	H	11–14				glutathione transferase, defense response
CG11652		−1.06	−3.46	−3.47	−3.48	−3.18		−2.18	CNS	11–16				unknown
CG8066		−1.02	−1.56	−2.32	−2.98	−1.45		−1.04	T	13–16				cysteine protease inhibitor
Cys		−1.17	−1.67	−2.17	−2.28	−1.40		+1.01	T	12–16				cysteine protease inhibitor
Flo		−1.45	−2.40	−2.02	−2.18	−1.44		−1.41	CNS	11–16				structural molecule, receptor binding
CG3408		+1.14	−1.72	−1.61	−1.60	−1.17		−1.12	CNS	11–17	0	--	*B	unknown
CG7966		−1.10	−2.13	−2.40	−1.37	−1.15		−1.18	T	12–16				selenium binding
CG12139		−1.62	−2.24	+1.28	−1.06	+1.25		+1.47	CNS	11–17				LDL receptor, calcium ion binding
ttk		−1.28	−1.75	−1.03	−1.06	+1.05		+1.18						
nerfin-1		−1.21	−1.50	−1.34	−1.16	−1.01		+1.03	G, N	10–17	0	--		transcription factor
CG2852		−1.32	−1.43	−1.51	−2.10	−1.62		−1.19	H	11–16				peptidyl-prolyl cis-trans isomerase
CG5080		+1.10	−1.44	−1.78	−2.85	−1.57		−1.21	G, N	11–16	++	--		unknown
CG9796		+2.67	+1.25	−1.53	−1.70	−1.46		+1.01	T	12–16				unknown
Gdi		−1.17	−1.48	−1.83	−1.52	−1.34		−1.03	CNS	12–16				RAB GDP-dissociation inhibitor
Bc		−1.01	−1.74	−1.35	−2.63	−2.95		−2.94	H	11–17	0	0		monophenol monooxygenase, defense response
CG11910		+1.00	+1.00	−1.09	−1.54	−1.72		−2.08	G	13–17	0	--	*B	G-protein coupled receptor
CG10591		-1.08	−1.10	+1.39	−3.52	−3.27		+1.17	T	12–16				unknown
CG9196		+1.01	−1.26	+1.40	−2.11	−1.53		−1.14	T	13–16				unknown
CG9336		+1.00	−1.06	−1.32	−2.25	−1.52		−1.40	G	11–17	++	--		unknown
Argk		−1.13	−2.32	−1.34	−2.34	−1.38		+1.11	G, N	11–16			*F	arginin kinase
CG9338		−1.05	−1.04	−1.12	−1.43	−1.58		−1.23	G	11–17	++	--		unknown
CG18318		−1.04	+1.00	+1.11	−1.21	+1.05		−2.09	G	12–16	0	0		unknown
CG2893		+1.13	+1.01	−1.10	−1.41	−1.41		−1.66	G	13-17	++	--	*B	cation transport, antiporter

Gene identifiers are given as presented in all Supplementary data. The fold regulation for all examined developmental stages represents the log2-transformed ratio Cy5/Cy3 of median fitted intensities. Values were obtained from filtering with the M-CHiPS software package (red: upregulation above 1.5-fold; blue: downregulation below 1.5-fold). Expression pattern was confirmed by standard in situ hybridization. Expression in glial cells (G), in single neurons (N), ubiquitously in the CNS (CNS), in cells of the hematopoietic lineage (H), and in tendon cells (T) was observed in stages as indicated. For further legend and abbreviations, see Table 2. (For interpretation of the references to colour in this table legend, the reader is referred to the web version of this article.)

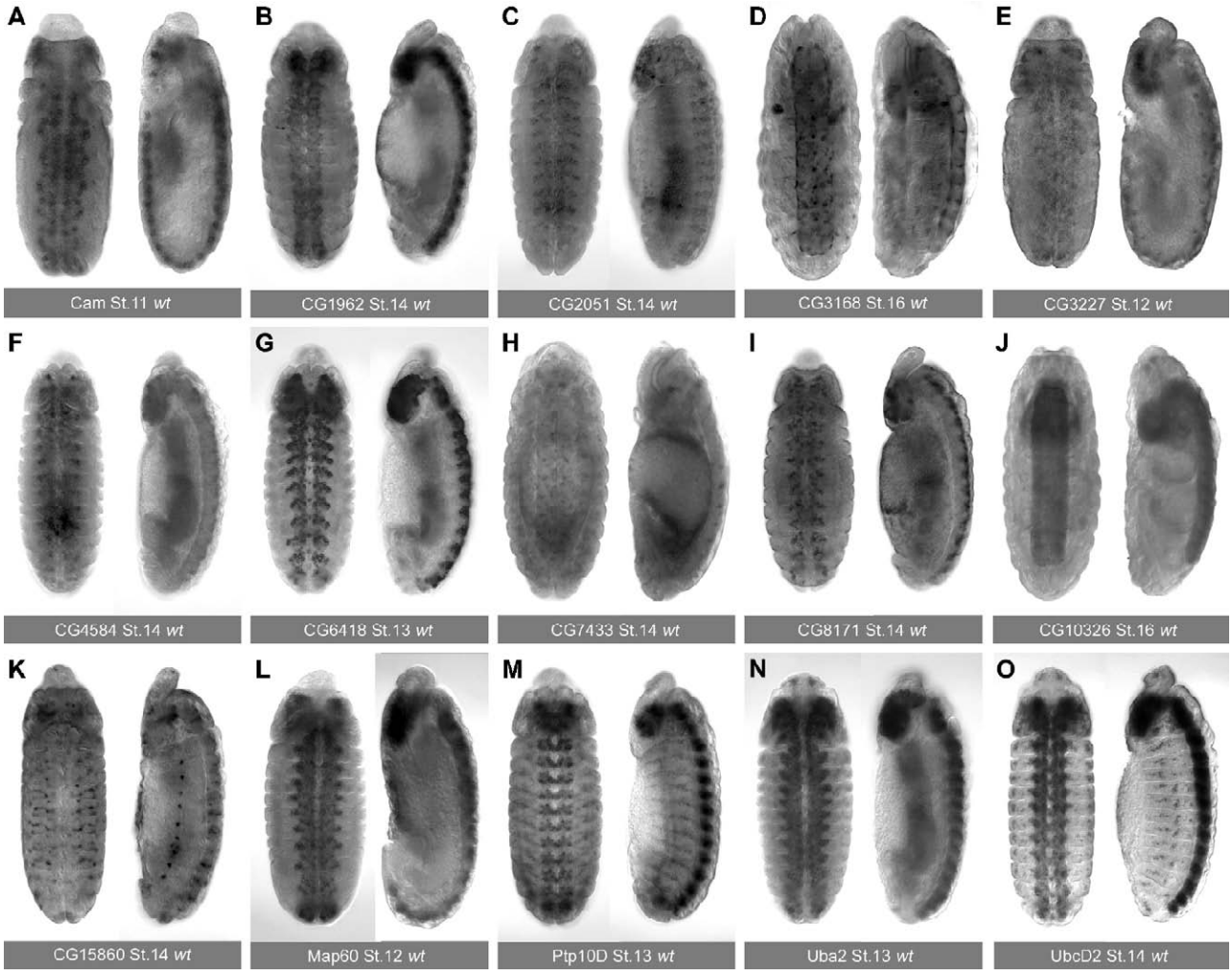

Fig. 3. In situ hybridizations of candidate genes from the GOF experiment in wildtype embryos. Genes and stages are indicated underneath the whole-mount pictures, shown in a ventral view (left) and a lateral view (right), respectively; anterior is up. If required, different focal sections were combined. Genes are sorted in alphabetic/numeric order. Examples are shown for genes expressed ubiquitously in the CNS (A, C, E, J, N, O), in single glial cells and neurons (B, F, G, I, L, M), and exclusively in glial cells (D, H, K).

CG6783), subperineural glia (SPG) (CG5080), longitudinal glia (LG) (CG11910, CG7433) and peripheral glia (PG) (CG9336, CG9338). Four of these glial-specific genes were also selected in the GOF experiment (CG2893, CG6218, CG6783, CG9336). Seven LOF candidate genes were expressed in tendon cells (Table 3, Fig. 4), and another eight genes showed expression in cells of the hematopoietic lineage including macrophages (Table 3, Fig. 4). One additional gene was selected which showed simultaneous expression in glial cells as well as in macrophages (CG16876, Table 3, Fig. 4) and one gene with expression in glial cells and tendon cells (CG15015).

Taken together, 5 genes show expression in nearly all lateral glial cells, whereas 6 genes show staining in only a subset of glial cells, like CBG or LG. Additionally, we could identify 7 genes that are expressed in more than one glial subtype, e.g. in SPG and CBG. Furthermore, 9 genes are expressed in muscle tendon cells and 8 genes show expression in cells of the hematopoietic lineage. In addition, we found one gene simultaneously expressed in glial cells and in the blood cell lineage and one gene expressed in muscle tendon cells and in glia. The temporal expression profiles for most of our candidate genes in wildtype as well as in both mutant genetic backgrounds (as revealed by in situ hybridization) resemble the observed differential regulation in both our microarray experiments at least to some extent.

In order to analyze the expression of our candidate genes in glial cells in more detail, we performed in situ hybridizations in combination with anti-Repo antibody stainings. For some of the genes with exclusive expression in glial cells, the double stainings are shown in Fig. 5. The dependence on *gcm* was tested by in situ hybridization in Mz1060::*gcm* and in *gcm* mutant backgrounds. The results are given in Tables 2 and 3. Increased expression is marked with '++', a decrease in expression is marked with '−', and no change in expression is marked with '0'. As expected, all glial-cell-specific genes lack expression in *gcm* mutant embryos. Yet, not all of these genes show an increase in expression upon ectopic activation of *gcm*. Examples are shown in Fig. 6.

The temporal expression profile of the selected candidate genes was determined by in situ hybridization in wildtype embryos. The results are given in Tables 2 and 3 (expression

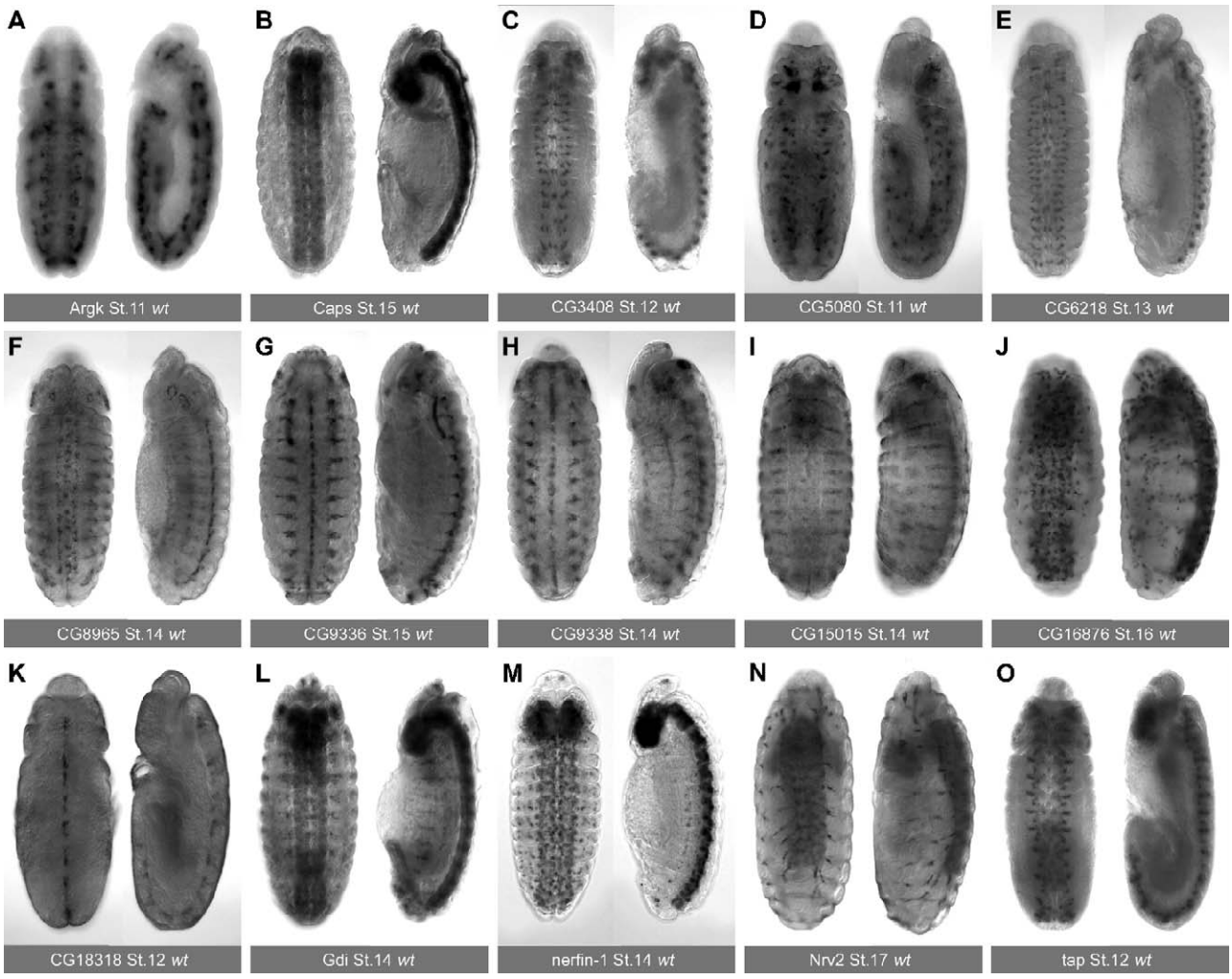

Fig. 4. In situ hybridizations of candidate genes from the LOF experiment in wildtype embryos. Genes and stages are indicated underneath the whole-mount pictures shown in a ventral view (left) and a lateral view (right), respectively; anterior is up. If required, different focal sections were combined. Genes are sorted in alphabetic/numeric order. Examples are shown for genes expressed ubiquitously in the CNS (B, L), in single glial cells and neurons (A, C, D, F, M, N, O), exclusively in glial cells (E, G, H, K), as well as in glial cells together with macrophages (J) or tendon cells (I).

pattern, stage). The first and the last stages of detectable in situ hybridization signal are indicated. For most of the selected genes, the expression in wildtype embryos is first detectable at stage 11 and continues until stage 16. For about 25% of the positively tested genes, expression was first detected at stage 12 to 13, and two genes are expressed in later stages only (Tables 2 and 3).

Discussion

Whole genome microarrays have been intensively used over the last decade. The interpretation of the huge amount of data generated by such experiments is still critical, but nonetheless crucial. Many efforts have been taken to standardize chip design, data acquisition and analysis, but recent papers demonstrate worrying discrepancies between microarray-derived data (reviewed by Tan et al. (2003); Jordan (2004). Nevertheless, many papers report on the successful use of microarrays in various biological systems, ranging from bacteria to higher organisms, and from single cells to whole organs, tissues or individuals. In any way, the DNA chip technology is highly sensitive and thus susceptible to many aspects of technical or experimental error. A straightforward experimental design including quality controls as well as the determination and reduction of technical (and genetical) background noise in a given approach and system is a prerequisite for correct and precise data interpretation.

Improvements in experimental design and candidate gene selection

The analysis of differential expression along the course of development and the comparison between gain- and loss-of-function genetics mark the two major novelties in our microarray screen for downstream targets of *gcm* and turned out to be extremely helpful in selecting potential candidate genes. The temporal profile of differential expression was used to either select or discard genes from further analysis. Different filtering steps enabled us to select potential target genes (about 14% of all differentially regulated genes).

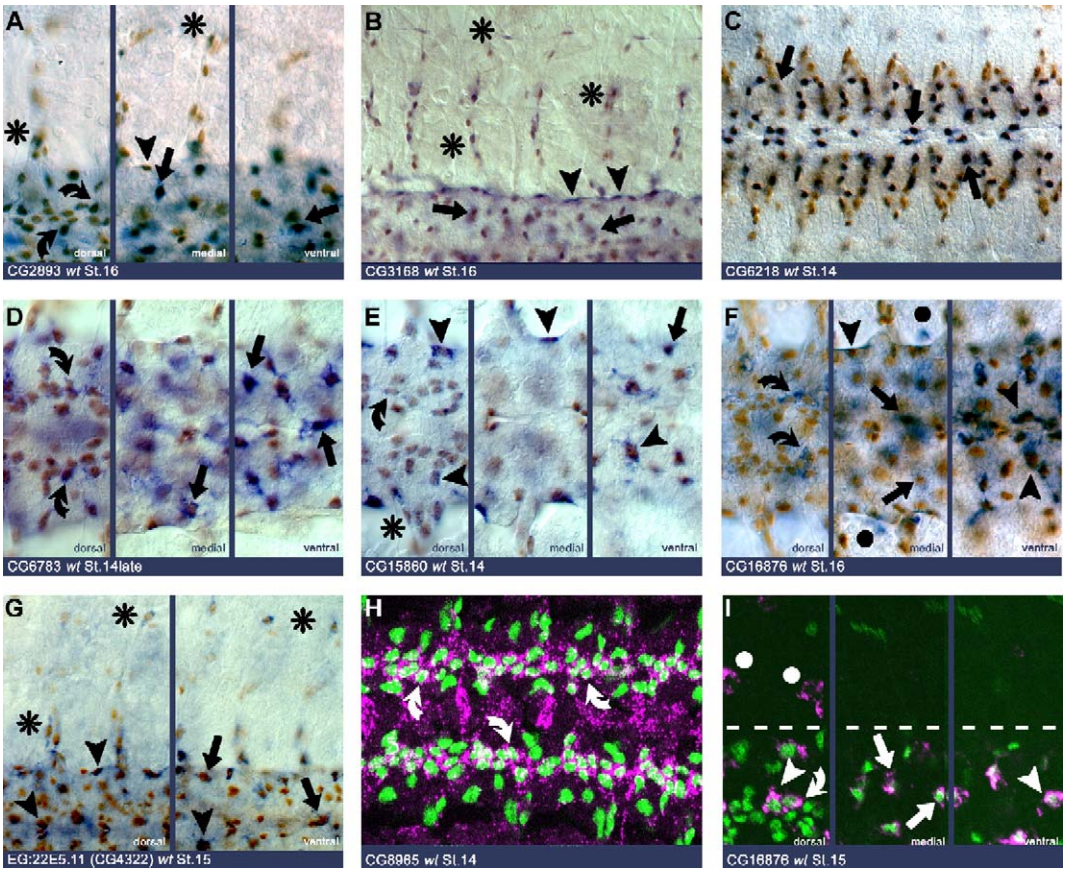

Fig. 5. In situ hybridizations against candidate genes in combination with antibody stainings against the glial protein Repo in wildtype embryos (flat preparations). Anterior to the left; in A, B, G and I, the midline is at the bottom. A–G were stained in the cytoplasm with alkaline phosphatase (in situ hybridization, blue) and nuclear DAB (anti-Repo, brown). H–I show single confocal stacks of fluorescent in situ hybridizations (magenta) counterstained with anti-Repo antibodies (green). Expression of genes can be detected in different glial subtypes such as peripheral glia (asterisk), subperineurial glia (arrow head), cell body glia (straight arrow) and longitudinal glia (bent arrow). CG16876 (BG:DS05899.3) is also expressed in macrophages (black and white dots in panels F and I).

About 400 potential candidate genes out of 2900 differentially regulated genes were selected and 300 of them were tested. After validation of these genes by in situ hybridization, 40% were found to be expressed in a *gcm*-dependent manner, 18% were rated as negative results with expression patterns outside the CNS/hemocytes/tendon cells, and 42% with no signal at all, being uncertain candidates, still. For neither of these we could see any concordance with respect to the regulation profile, the strength of differential regulation or the reproducibility. Compared to the microarray screens published by Egger et al. (2002) and Freeman et al. (2003), the percentage of positively tested *gcm* target genes is rather high. The first published oligonucleotide microarray screen for *gcm* downstream targets (Egger et al., 2002) analyzed the effect of *gcm* misexpression induced by *scabrous* (*sca*)-Gal4 in stage 11 and stage 15/16. In this approach, 400 and 1259 differentially regulated genes were discovered in stages 11 and 15/16, respectively, of which 93 genes overlapped between the two stages. Taken together, approximately 10% of all genes present on the oligonucleotide array (with ~14,000 gene-specific oligonucleotides) appeared to be differentially regulated. Though the authors could not rule out side-effect-induced differential regulation due to gross morphological abnormalities observed, they postulate an involvement of the upregulated genes in gliogenesis or neurogenesis. For none of these newly described potential *gcm* target genes the regulation by *gcm* and the expression in glial cells was shown in situ. The authors further conceded the possibility that not all glial genes were necessarily among the upregulated genes and that a *gcm* loss-of-function analysis would be needed to fill these gaps. Comparing our microarray data with the data presented by Egger et al., we find only minor overlap. Of the genes published by Egger et al., 30 genes are differentially expressed in our experiments, too. Half of these show an opposite behavior (with regard to up-/downregulation), and only three were selected by our filter criteria and tested (Table 2). These discrepancies might be due to the different microarray systems used. Inconsistent results obtained by different microarray systems have been reported (Tan et al., 2003; Jordan, 2004). However, we believe that the experimental design (specific Gal4-line, wildtype control, time course, gain- and loss-of-function) and the resulting possibilities to select potential target genes are important improvements in our microarray approach.

The second published screen for *gcm* target genes combined a cDNA microarray approach with an in silico prediction of *gcm* targets by enhancer/promoter analysis and public in situ hybridization data (Freeman et al., 2003). Ectopic expression of *gcm* was again driven by *sca*-Gal4, and only stage 12/13 was

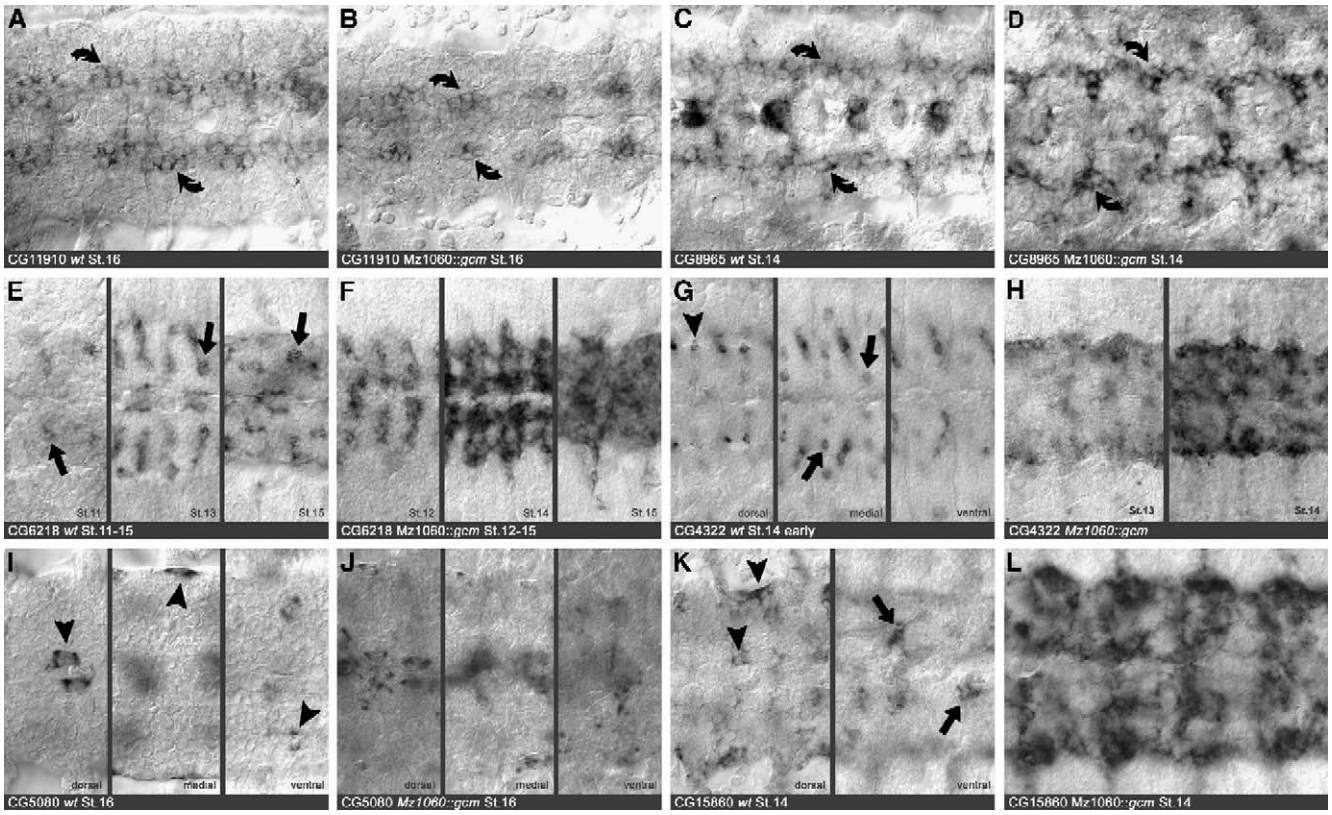

Fig. 6. In situ hybridizations of candidate genes in wildtype embryos in comparison to embryos with ectopic activation of *gcm* (Mz1060::*gcm*). Flat preparations; anterior to the left; different focal sections or developmental stages were combined, if required. Longitudinal glial-specific gene expression is not altered upon ectopic activation of *gcm* (CG11910 and CG8965, bent arrows in panels A–D), whereas all other glial subtypes can be ectopically induced in Mz1060::*gcm* (E–L). Shown are cell body glia (straight arrows in E, G, K) and subperineurial glia (arrow heads in G, I, K). Even restricted expression to single glial cells in wildtype (dorsal subperineurial glia in CG5080, I) is increased upon ectopic *gcm* (J).

compared to wildtype on microarrays containing 4386 unique *Drosophila* genes. The authors tested 153 of the upregulated genes (chosen by strength of differential regulation) of which 18 were proven to be expressed in glial cells, and their upregulation upon ectopic *gcm* was confirmed in situ. Though all three different approaches combined by Freeman and colleagues were successful and led to a description of 45 novel *gcm* target genes, only minor overlap between the three approaches was observed. The microarray approach itself resulted in the identification of 18 novel target genes representing 12% of the tested genes. The rather high rate of false positives (88%) was discussed, and the authors assumed a similar rate for the data presented by Egger and colleagues. From the 18 microarray-derived *gcm* target genes published by Freeman et al., 8 genes were among our differentially regulated genes. Another 6 candidates of those identified by the in silico screen for *gcm* binding sites and 6 of those found in the in situ database were present in our filtered gene lists. In total about 45% of all genes described by Freeman et al. appeared to be differentially regulated in our screen. Most of them (15 genes) were downregulated in the LOF or even antagonistically regulated in both experiments.

Novel glial genes

Subsequent filtering and validation steps enabled us to define clusters of genes with differential glial expression at specific time points or with specific regulation profiles in both experiments. As can be seen in Tables 2 and 3, most glial-specific genes in the GOF experiment showed differential regulation in stage 16, whereas in the LOF the glial-specific genes can be found throughout all stages. We compared the profiles of differential regulation with the temporal expression profiles as revealed by in situ hybridization in wildtype embryos. Both profiles do not necessarily match with each other. The regulation profile in the GOF depends on the Gal4 driver line rather than on endogenous expression. Hence, the profiles in the LOF experiment show a better overall accordance. Still, the discrepancies between differential regulation in both experiments and the endogenous in situ expression pattern cannot be explained. Astonishingly few genes showed antagonistic regulation in both experiments. Some of these genes (e.g. CG6218, CG6783) showed glial expression exclusively. The dependence on *gcm* was tested and nicely fits to the obtained microarray data. Most of the novel glial genes, however, were differentially regulated in only one of the two experiments. In situ hybridizations were performed in *gcm* mutant backgrounds as well as in Mz1060::*gcm* embryos, and expression patterns were compared to wildtype and to the obtained microarray data. With only few exceptions, most of the novel glial genes discovered from either experiment showed expression in the respective mutant background comparable to the observed microarray profiles. Especially those genes that are also expressed in neurons show an

increase in expression upon ectopic activation of *gcm*, but no significant decrease in *gcm* mutants. Thus, the lack of glial-specific downregulation in the LOF experiment is masked by neuronal expression. Hence, these genes were only selected by upregulation in the GOF experiment. Conversely, some of the glial-specific downregulated genes in the LOF experiment are not differentially upregulated in the GOF. CG11910 for example does not show an increase in expression in situ upon ectopic *gcm*, even though it is expressed exclusively in longitudinal glial cells (Figs. 6A, B), suggesting that this gene requires factors in addition to Gcm to be activated. Hence, this gene was not previously discovered by either of the two gain-of-function microarray screens (Egger et al., 2002; Freeman et al., 2003). The same observations were made when analyzing the expression of CG8965, which is also expressed in longitudinal glial cells (Figs. 6C, D). Apart from these two, all other genes with expression restricted to glial subtypes can be activated ectopically in Mz1060::*gcm* (Figs. 6E–L). This might correlate with the origin of the respective glial subtypes, indicating that spatial, temporal or even other specific cofactors are required for glial subtype specification. Still, some of the candidate genes of the LOF experiment were not differentially regulated on microarrays of the GOF experiment, though in situ hybridizations show an increase in expression for these genes upon ectopic activation of *gcm* in the CNS (e.g. CG11652, BG:DS05899.3). We cannot explain these discrepancies. Yet, all these observations demonstrate how carefully microarray experiments need to be designed and analyzed, especially with respect to the analysis of both gain- and loss-of-function genetics rather than only one situation.

Functional categories of novel glial genes

We analyzed the composition of the filtered and tested genes with respect to their gene ontology (GO-annotation) as given in FlyBase. The GO annotation was used to categorize the putative or known function into eleven functional classes. Fig. 7 shows the distribution among these functional categories of all potential candidate genes (the filtered genes excluding those with negative in situ hybridizations). These 'functional categories' do not resemble groups of GO annotations of the same hierarchical level within the genealogic tree of GO annotations, and some genes might have annotated functions which overlap between the different categories. Hence, we do not intend to compare these categories of our filtered genes with those present in the entire genome, but rather give an overview of the functional annotations present in our filtered gene lists. For about 25% of the genes, no GO annotation was given, so their function is unknown. Signaling-related molecules and proteins involved in posttranslational protein modification as well as nucleic acid metabolism are equally distributed, representing 10% of all potential candidate genes each. Genes from these functional classes show differential regulation at various developmental stages examined in either experiments. Among the signaling-related molecules, the majority are assigned to G-protein-coupled receptor signaling or small GTPase function. Eighteen transcription factors (7%) were filtered, whose differential regulation starts in early stages (stages 9–12), which is compatible with a possible function in cell fate specification or early differentiation. Another 6% of the filtered genes encode ion transporters, antiporters, symporters or ion channels. Half of these show differential regulation at late stages only. This also appears reasonable as homeostatic control is believed to be one of the functions of glial cells. 13% of all filtered genes have putative enzymatic function and/or are involved in general metabolism. Whether this implies glial-specific or experimentally induced changes in metabolism remains unclear. Most of the filtered genes with negative in situ hybridization results were among those with putative metabolic function.

Astonishingly, neither *repo* nor *pnt* show differential regulation at any of the examined stages in both of our experiments and *ttk* appears downregulated only in the LOF at stage 11. Similar observations were described by Egger et al. (2002) and Freeman et al. (2003), where either no upregulation of these *gcm* target genes or even downregulation upon ectopic expression of *gcm* occurred. With respect to (no) differential regulation of *pnt* and *ttk*, this can be explained by the fact that for both genes alternative splice variants are expressed in neurons or other cells. For every gene on our microarray, a single exon was amplified and spotted. For *pnt* and *ttk*, the spotted exons (exons 7 and 5, respectively) are present in all splice variants. Hence, expression outside of glial cells may mask the glial-specific expression and no strong differential regulation is observed. For *repo*, there is no such obvious explanation since *repo* is expressed in all lateral glial cells exclusively. We extracted the expression values for *repo* from all our microarray hybridizations and found a strong expression of *repo* throughout the entire time course of both experiments without notable changes. Moreover, *repo* appears to be expressed on our microarrays earlier and stronger than *gcm* (already at stage 9), which does not reflect the expression of *repo* in vivo. We sequenced the spotted PCR product, and the sequence corresponds to exon 1 of the *repo* gene (as aspired). Furthermore, we hybridized Northern blots with total RNA from both wildtype and Mz1060::*gcm* embryos of different stages with a *repo* antisense probe. An increase in *repo*-mRNA upon ectopic expression of *gcm* was clearly detectable (data not shown). We conclude an unspecific cross-hybridization of another cDNA to the *repo* spot. This is obviously one of the major drawbacks of cDNA microarray approaches.

Further steps towards the understanding of glial cell development and diversity

A considerable number of genes have already been identified, which are involved in various aspects of glial cell biology in the *Drosophila* embryo: early determining factors such as *gcm*, *repo* and *pnt*, factors required for terminal glial cell differentiation and function like *loco*, EAAT1/2 or *moody*, and several genes expressed at various times in between. Most of these genes are expressed in subsets of glial cells only, and many are also expressed in other cells or tissues. The biggest collection of such genes was identified in the screen presented by Freeman et al. (2003). Our microarray experiments enlarge this collection and for the first time incorporate results from a

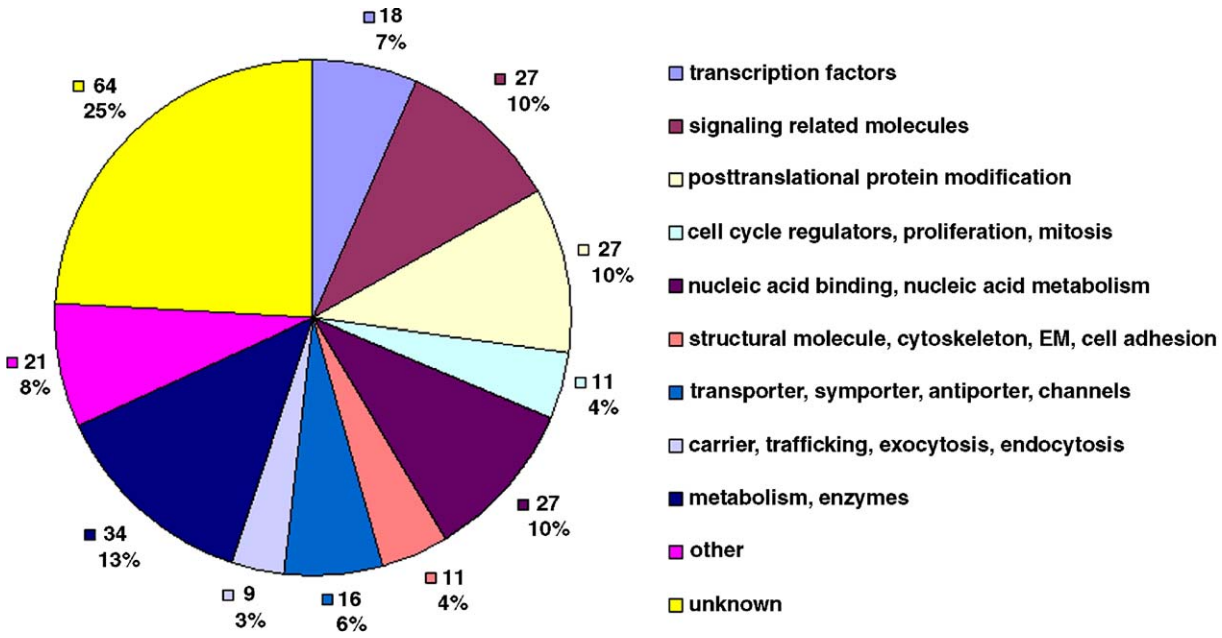

Fig. 7. Functional classification of candidate genes. All filtered (potential) glial genes (255 genes) were grouped into eleven functional categories according to the GO annotation given in FlyBase. 25% are of unknown function.

gcm loss-of-function analysis. Towards a first characterization of all these factors with respect to glial cell development, subtype specification and function, we can group the candidate genes for example according to (a) time point of expression in wildtype glial cells, (b) expression pattern in glial subtypes or (c) their annotated function. Most of the novel glial genes are expressed from stage 12/13 onwards, and the vast majority is only expressed in certain cells or subtypes. The primary determination of the glial cell fate in the embryo is achieved by Gcm, which is expressed early in development of all lateral glial cells (Hosoya et al., 1995; Jones et al., 1995; Vincent et al., 1996). Further subtype specification obviously requires spatial and/or temporal cofactors. Together, they activate a variety of other genes, ranging from signaling or transporter molecules to proteins with putative metabolic function. The restriction of gene expression to particular glial cells includes all combinations of glial cell classes. Yet, some correlations can be observed with respect to the classification of glial cells according to morphological and positional criteria (Ito et al., 1995). Cell body glia and peripheral glia for example are best confined by subtype-specific gene expressions. Whether these subtypes require the respective genes for proper development or function remains to be shown. Apparently, peripheral glial cells, which migrate over long distances from their place of origin in the CNS into the periphery, require precise subtype-specific regulation. Many of the genes expressed in more than one subtype are also expressed in subperineurial glial cells. Recent publications deal with the function of these cells in blood–brain barrier formation and demonstrate the involvement of one particular gene, *moody* (Bainton et al., 2005; Schwabe et al., 2005). This gene has already been described by Freeman et al. (CG 4322) and was also found to be differentially regulated in our screen and filtered by our criteria (EG:22E5.11). *moody* is predicted to encode a G-protein-coupled receptor molecule. The requirement of proper signaling via heterotrimeric G-protein receptors for glial cell differentiation and formation of the blood–brain barrier was already proposed earlier (Granderath et al., 1999). One of our newly described glial genes, CG11910, is also predicted to encode a G-protein-coupled receptor molecule. It is expressed from stage 14 onwards in longitudinal glia, a glial cell type that is known to enwrap the longitudinal axonal connectives. Blood–brain barrier formation requires the tight connection of subperineurial glial cells by septate junctions. Some *Drosophila* genes are known to regulate the formation of septate junctions, for example, *sinuous* or *neurexin* (Baumgartner et al., 1996; Wu et al., 2004). Both genes are differentially regulated on our microarray, too, and the latter is described to be required in blood–brain barrier formation (Baumgartner et al., 1996). This suggests that the entire lists of differentially regulated genes of both our microarray experiments (given in Supplementary Tables 1 and 2) comprise more glial-specific candidates, which may be uncovered by further filtering with respect to certain functional annotations. Yet, the collection of glial genes identified so far provides an important basis for the clarification of the genetic network controlling various aspects of glial development and function.

Acknowledgments

We are grateful to Florence Besse and Anne Ephrussi for extensive advice and assistance with the sorting and selection of embryos; to Tobias Rasse and Robert Ventzki for their introduction and help with the embryo sorter; to Joachim Urban and Gert Pflugfelder for critical comments on the manuscript; to George Dimopoulos for his help during our first experiences

with microarrays; to Kurt Fellenberg for his help in data analysis; to Vladimir Benes, Luis Teixeira, and Britta Koch for cDNA clones, ESTs and amplicons; to Angela Giangrande and the Bloomington stock center for fly strains; to Barbara Groh-Reichert, Ruth Beckervordersandforth and Irina Hein for their help with in situ hybridizations; and to Olaf Vef for his help with fly work. This work was supported by a grant from the Deutsche Forschungsgemeinschaft to B.A. and G.M.T.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2006.04.460.

References

Akiyama-Oda, Y., Hosoya, T., Hotta, Y., 1998. Alteration of cell fate by ectopic expression of *Drosophila* glial cells missing in non-neural cells. Dev. Genes Evol. 208, 578–585.

Alfonso, T.B., Jones, B.W., 2002. *gcm2* promotes glial cell differentiation and is required with glial cells missing for macrophage development in *Drosophila*. Dev. Biol. 248, 369–383.

Badenhorst, P., Finch, J.T., Travers, A.A., 2002. Tramtrack co-operates to prevent inappropriate neural development in *Drosophila*. Mech. Dev. 117, 87–101.

Bainton, R.J., Tsai, L.T., Schwabe, T., DeSalvo, M., Gaul, U., Heberlein, U., 2005. *moody* encodes two GPCRs that regulate cocaine behaviors and blood–brain barrier permeability in *Drosophila*. Cell 123, 145–156.

Bataille, L., Auge, B., Ferjoux, G., Haenlin, M., Waltzer, L., 2005. Resolving embryonic blood cell fate choice in *Drosophila*: interplay of GCM and RUNX factors. Development 132, 4635–4644.

Baumgartner, S., Littleton, J.T., Broadie, K., Bhat, M.A., Harbecke, R., Lengyel, J.A., Chiquet-Ehrismann, R., Prokop, A., Bellen, H.J., 1996. A *Drosophila* neurexin is required for septate junction and blood–nerve barrier formation and function. Cell 87, 1059–1068.

Beissbarth, T., Fellenberg, K., Brors, B., Arribas-Prat, R., Boer, J., Hauser, N.C., Scheideler, M., Hoheisel, J.D., Schutz, G., Poustka, A., Vingron, M., 2000. Processing and quality control of DNA array hybridization data. Bioinformatics 16, 1014–1022.

Bernardoni, R., Vivancos, V., Giangrande, A., 1997. *glide*/*gcm* is expressed and required in the scavenger cell lineage. Dev. Biol. 191, 118–130.

Bossing, T., Udolph, G., Doe, C.Q., Technau, G.M., 1996. The embryonic central nervous system lineages of *Drosophila melanogaster*: I. Neuroblast lineages derived from the ventral half of the neuroectoderm. Dev. Biol. 179, 41–64.

Campos-Ortega, J., Hartenstein, V., 1997. The Embryonic Development of *Drosophila melanogaster*. Springer Berlin, Heidelberg.

Casso, D., Ramirez-Weber, F.A., Kornberg, T.B., 1999. GFP-tagged balancer chromosomes for *Drosophila melanogaster*. Mech. Dev. 88, 229–232.

Casso, D., Ramirez-Weber, F., Kornberg, T.B., 2000. GFP-tagged balancer chromosomes for *Drosophila melanogaster*. Mech. Dev. 91, 451–454.

Chotard, C., Leung, W., Salecker, I., 2005. Glial cells missing and gcm2 cell autonomously regulate both glial and neuronal development in the visual system of *Drosophila*. Neuron 48, 237–251.

Doe, C.Q., 1992. Molecular markers for identified neuroblasts and ganglion mother cells in the *Drosophila* central nervous system. Development 116, 855–863.

Egger, B., Leemans, R., Loop, T., Kammermeier, L., Fan, Y., Radimerski, T., Strahm, M.C., Certa, U., Reichert, H., 2002. Gliogenesis in *Drosophila*: genome-wide analysis of downstream genes of glial cells missing in the embryonic nervous system. Development 129, 3295–3309.

Fellenberg, K., Hauser, N.C., Brors, B., Hoheisel, J.D., Vingron, M., 2002. Microarray data warehouse allowing for inclusion of experiment annotations in statistical analysis. Bioinformatics 18, 423–433.

Freeman, M.R., Delrow, J., Kim, J., Johnson, E., Doe, C.Q., 2003. Unwrapping glial biology: *gcm* target genes regulating glial development, diversification, and function. Neuron 38, 567–580.

Furlong, E.E., Andersen, E.C., Null, B., White, K.P., Scott, M.P., 2001a. Patterns of gene expression during *Drosophila* mesoderm development. Science 293, 1629–1633.

Furlong, E.E., Profitt, D., Scott, M.P., 2001b. Automated sorting of live transgenic embryos. Nat. Biotechnol. 19, 153–156.

Giesen, K., Hummel, T., Stollewerk, A., Harrison, S., Travers, A., Klämbt, C., 1997. Glial development in the *Drosophila* CNS requires concomitant activation of glial and repression of neuronal differentiation genes. Development 124, 2307–2316.

Granderath, S., Stollewerk, A., Greig, S., Goodman, C.S., O'Kane, C.J., Klämbt, C., 1999. *loco* encodes an RGS protein required for *Drosophila* glial differentiation. Development 126, 1781–1791.

Granderath, S., Bunse, I., Klämbt, C., 2000. *gcm* and pointed synergistically control glial transcription of the *Drosophila* gene loco. Mech. Dev. 91, 197–208.

Greenacre, M.J., 1984. Theory and Applications of Correspondence Analysis. Academic Press, London.

Greenacre, M.J., 1993. Correspondence Analysis in Practice. Academic Press, London.

Halter, D.A., Urban, J., Rickert, C., Ner, S.S., Ito, K., Travers, A.A., Technau, G.M., 1995. The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of *Drosophila melanogaster*. Development 121, 317–332.

Hild, M., Beckmann, B., Haas, S.A., Koch, B., Solovyev, V., Busold, C., Fellenberg, K., Boutros, M., Vingron, M., Sauer, F., Hoheisel, J.D., Paro, R., 2003. An integrated gene annotation and transcriptional profiling approach towards the full gene content of the *Drosophila* genome. Genome Biol. 5, R3.

Hosoya, T., Takizawa, K., Nitta, K., Hotta, Y., 1995. Glial cells missing: a binary switch between neuronal and glial determination in *Drosophila*. Cell 82, 1025–1036.

Ito, K., Urban, J., Technau, G., 1995. Distribution, classification, and development of *Drosophila* glial cells in the late embryonic and early larval ventral nerve cord. Roux's Arch. Dev. Biol. 204, 284–307.

Jiang, J., Kosman, D., Ip, Y.T., Levine, M., 1991. The dorsal morphogen gradient regulates the mesoderm determinant twist in early *Drosophila* embryos. Genes Dev. 5, 1881–1891.

Jones, B.W., 2005. Transcriptional control of glial cell development in *Drosophila*. Dev. Biol. 278, 265–273.

Jones, B.W., Fetter, R.D., Tear, G., Goodman, C.S., 1995. Glial cells missing: a genetic switch that controls glial versus neuronal fate. Cell 82, 1013–1023.

Jones, B.W., Abeysekera, M., Galinska, J., Jolicoeur, E.M., 2004. Transcriptional control of glial and blood cell development in *Drosophila*: *cis*-regulatory elements of glial cells missing. Dev. Biol. 266, 374–387.

Jordan, B.J., 2004. How consistent are expression chip platforms? BioEssays 26, 1236–1242.

Kammerer, M., Giangrande, A., 2001. Glide2, a second glial promoting factor in *Drosophila melanogaster*. EMBO J. 20, 4664–4673.

Klaes, A., Menne, T., Stollewerk, A., Scholz, H., Klämbt, C., 1994. The Ets transcription factors encoded by the *Drosophila* gene pointed direct glial cell differentiation in the embryonic CNS. Cell 78, 149–160.

Klämbt, C., 1993. The *Drosophila* gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. Development 117, 163–176.

Kosman, D., Ip, Y.T., Levine, M., Arora, K., 1991. Establishment of the mesoderm–neuroectoderm boundary in the *Drosophila* embryo. Science 254, 118–122.

Lebestky, T., Chang, T., Hartenstein, V., Banerjee, U., 2000. Specification of *Drosophila* hematopoietic lineage by conserved transcription factors. Science 288, 146–149.

Miller, A.A., Bernardoni, R., Giangrande, A., 1998. Positive autoregulation of the glial promoting factor *glide*/*gcm*. EMBO J. 17, 6316–6326.

Patel, N.H., 1994. Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. Methods Cell Biol. 44, 445–487.

Ragone, G., Van De Bor, V., Sorrentino, S., Kammerer, M., Galy, A., Schenck, A.,

Bernardoni, R., Miller, A.A., Roy, N., Giangrande, A., 2003. Transcriptional regulation of glial cell specification. Dev. Biol. 255, 138–150.

Schmidt, H., Rickert, C., Bossing, T., Vef, O., Urban, J., Technau, G.M., 1997. The embryonic central nervous system lineages of *Drosophila melanogaster*: II. Neuroblast lineages derived from the dorsal part of the neuroectoderm. Dev. Biol. 189, 186–204.

Schwabe, T., Bainton, R.J., Fetter, R.D., Heberlein, U., Gaul, U., 2005. GPCR signaling is required for blood–brain barrier formation in *Drosophila*. Cell 123, 133–144.

Soustelle, L., Jacques, C., Altenhein, B., Technau, G.M., Volk, T., Giangrande, A., 2004. Terminal tendon cell differentiation requires the *glide*/*gcm* complex. Development 131, 4521–4532.

Tan, P.K., Downey, T.J., Spitznagel Jr., E.L., Xu, P., Fu, D., Dimitrov, D.S., Lempicki, R.A., Raaka, B.M., Cam, M.C., 2003. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res. 31, 5676–5684.

Tautz, D., Pfeifle, C., 1989. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. Chromosoma 98, 81–85.

Urbach, R., Technau, G.M., 2003. Molecular markers for identified neuroblasts in the developing brain of *Drosophila*. Development 130, 3621–3637.

Vincent, S., Vonesch, J.L., Giangrande, A., 1996. Glide directs glial fate commitment and cell fate switch between neurones and glia. Development 122, 131–139.

Wu, V.M., Schulte, J., Hirschi, A., Tepass, U., Beitel, G.J., 2004. Sinuous is a *Drosophila* claudin required for septate junction organization and epithelial tube size control. J. Cell Biol. 164, 313–323.

Xiong, W.C., Okano, H., Patel, N.H., Blendy, J.A., Montell, C., 1994. *repo* encodes a glial-specific homeo domain protein required in the *Drosophila* nervous system. Genes Dev. 8, 981–994.

Yuasa, Y., Okabe, M., Yoshikawa, S., Tabuchi, K., Xiong, W.C., Hiromi, Y., Okano, H., 2003. *Drosophila* homeodomain protein REPO controls glial differentiation by cooperating with ETS and BTB transcription factors. Development 130, 2419–2428.