Stochastic yet biased expression of multiple *Dscam* splice variants by individual cells

Guilherme Neves¹, Jacob Zucker¹, Mark Daly¹ & Andrew Chess^{1,2}

The *Drosophila melanogaster* gene *Dscam* is essential for axon guidance and has 38,016 possible alternative splice forms. This diversity can potentially be used to distinguish cells. We analyzed the *Dscam* mRNA isoforms expressed by different cell types and individual cells. The choice of splice variants expressed is regulated both spatially and temporally. Different subtypes of photoreceptors express broad yet distinctive spectra of *Dscam* isoforms. Single-cell RT-PCR documented that individual cells express several different *Dscam* isoforms and allowed an estimation of the diversity that is present. For example, we estimate that each R3/R4 photoreceptor cell expresses 14–50 distinct mRNAs chosen from the spectrum of thousands of splice variants distinctive of its cell type. Thus, the *Dscam* repertoire of each cell is different from those of its neighbors, providing a potential mechanism for generating unique cell identity in the nervous system and elsewhere.

The *D. melanogaster* homolog of the Down syndrome cell adhesion molecule (DSCAM) is a cell-surface protein containing ten immunoglobulin domains and six fibronectin domains in its extracellular surface. *Dscam* is necessary for correct axon targeting in the Bolwig's organ in the embryo¹ and in the mushroom bodies² and the antennal lobe³ of the adult fruit fly. The gene encoding *Dscam* is capable of generating 38,016 different isoforms through alternative splicing, most of which contain a different assortment of immunoglobulin domains and therefore probably have differential adhesive properties.

The discovery of 38,016 possible alternative splice forms of *Dscam*¹ raises a number of interesting questions. In particular, it is not known whether *Dscam* expression can provide a signature of a distinct subpopulation of cells or whether an individual cell expresses only one isoform of *Dscam*. The large number of possible splice variants makes it technically difficult to use standard methods of analyzing expression, such as *in situ* hybridization and immunohistochemistry.

To circumvent these technical difficulties, we used an approach that combined a customized oligonucleotide microarray, isolation of distinct populations of cells and sensitive RT-PCR that allowed examination of *Dscam* expression by single cells. Here we show that *Dscam* alternative splicing is differentially regulated both during development and in different tissues. A given cell type expresses a broad, yet distinctive, spectrum of splice variants. Individual cells express a unique repertoire of *Dscam* splice variants. We estimate that individual R3/R4 photoreceptors express 14–50 distinct mRNA molecules chosen from a subset of thousands of *Dscam* isoforms characteristic of the population. This would ensure that the *Dscam* repertoire of each cell is different from those of its neighbors.

RESULTS

A microarray for analyzing *Dscam* alternative splicing

The genomic structure of *Dscam*¹ is shown in Figure 1a. There are four alternatively spliced exons. In this study we focused on three alternatively spliced exons, which may create many possible isoforms (Fig. 1a): the cluster in exon 4 has 12 alternative exons, the cluster in exon 6 has 48 alternative exons and the cluster in exon 9 has 33 alternative exons. Each of these highly variable exons can form a portion of an immunoglobulin domain in the extracellular region of the protein. We designed a microarray containing probes for all 93 alternative exons in these three clusters. We amplified cloned DNAs (whose identities were confirmed by sequencing) by PCR to create a reference library that equally represents all the alternative exons present in the microarray. To test the specificity of the different probes in the microarray, we used combinations of RNAs in controlled amounts (Fig. 1b). RNA samples were *in vitro* transcribed from cloned cDNAs containing different combinations of exons 4, 6 and 9. We carried out a quantitative analysis of four independent experiments in which eight different samples were combined in different amounts spanning two orders of magnitude. Throughout this study, we expressed sample:reference ratios as a percentage of use of alternative exons at each cluster, which provided an internal control allowing comparisons between different arrays (details on printing, hybridization and analysis of microarrays are available in **Supplementary Methods** online). The observed linear relationship (Fig. 1c) shows that the array provides an accurate estimate of the relative levels of alternative exons present in a given RNA sample. As expected due to the similar size of the PCR amplification products, we identified no significant PCR-introduced bias (up to 35 cycles of amplification) among the eight different RNAs analyzed, which represented 22 distinct alternative exons.

¹Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. ²Massachusetts Institute of Technology, Department of Biology, Cambridge, Massachusetts 02139, USA. Correspondence should be addressed to A.C. (chess@wi.mit.edu).

Developmental analysis of *Dscam* alternative splicing

We used the microarray to analyze comprehensively the expression of splice variants of *Dscam* at different stages of development (in embryos; first, second and third instar larvae; and adults; **Fig. 1d**). Most alternative exons were expressed at some or all stages of development, but exon 6.11 was never detected in any of our analyses, consistent with previous studies involving sequencing of cloned cDNAs^{1,3} (G.N., J.Z. and A.C., unpublished observations). There were several significant differences in the profile of *Dscam* expression at different developmental stages, indicating that the splicing of *Dscam* is developmentally regulated, particularly with respect to exon 9. Five of the thirty-three isoforms of exon 9 were expressed at a significantly higher level in embryos than at the larval stages, with the relative expression level increasing again towards the adult (**Fig. 1e**). In contrast, several forms of exon 9 were almost undetectable in embryos but were robustly expressed at later stages of development. Alternative exons 4 and 6 showed only moderate changes in expression during development. Our results also confirm previous observations that the relative representation of exon 4.2 increases from embryos to adults⁴.

Dscam expression in different cell types

To determine whether alternative splicing is differentially regulated across different tissues, we analyzed various tissues of the third instar larvae. We observed that the repertoire of *Dscam* isoforms expressed in RNA extracted from whole brains was very different from the profile observed in RNA isolated from eye-antennal imaginal disks (**Fig. 2a**). Our analysis focused on the exon 9 cluster, as the differences observed at the other clusters were very small (data not shown).

Particular cell populations may express a small subset of all available *Dscam* isoforms, so that the particular forms expressed by the population act as a 'molecular label'³. To examine this possibility, we analyzed RNA prepared from distinct cell populations in the developing eye disk. We used the GAL4-UAS system⁵ to label specific populations of photoreceptors with green fluorescent protein (GFP). Cells were then dissociated and individual, viable, GFP-labeled cells were isolated by fluorescence activated cell sorting (FACS)⁶. Analysis of postmitotic neurons (mostly photoreceptors) isolated from dissected eye-antennal imaginal disks (using GFP driven by the pan-neuronal ELAV promoter) showed that the repertoire of *Dscam* isoforms was similar to the profile observed in isolated eye-antennal imaginal disks (as predicted by the mostly neuronal expression of *Dscam*)¹.

Next, we analyzed restricted populations of photoreceptors that project to separate regions in the optic lobe. The m δ promoter drives expression in R3/R4 photoreceptors that project to the lamina region⁷ (**Fig. 2b,c**), and the PM181 promoter labels R7 photoreceptors that project to the medulla⁸ (**Fig. 2d,e**). Both R3/R4 and R7 populations of photoreceptors expressed many different isoforms at all three exon clusters. The highest frequency exon choices at cluster 9 (exon 9.8 for R3/R4 and 9.9 for R7; **Fig. 2a**) represented only 10% and 9% of the total pool of exon 9 isoforms for the R3/R4 photoreceptors and the R7 photoreceptor populations, respectively. These results rule out a model in which a specific class of photoreceptor that projects to a particular brain region expresses one or a small number of distinct *Dscam* isoforms. The observed broad spectra notwithstanding, clustering analysis indicates that the profiles of *Dscam* isoforms expressed by the two populations are significantly different from each other and from those of the entire eye-antennal imaginal disk. Thus, each specific cell type expresses a broad yet distinctive spectrum of *Dscam* isoforms.

Figure 1 Developmental regulation of alternative splicing. **(a)** Genomic structure of the *Dscam* locus. The exons shown in red are always included in the mature mRNA. The two exon 17 variants (green) code for the single transmembrane domain of the protein. The custom microarray contained probes for all the alternative exons in the extracellular side of the protein (blue). **(b)** Representative microarray hybridization with a Cy3-labeled sample containing equal amounts of seven different RNA molecules and the Cy5-labeled reference library. Most spots are colored red, reflecting binding of the reference library. Spots corresponding to the alternative exons present in the RNA sample are colored yellow or green. The reference library provides a positive control for each array probe. **(c)** Analysis of microarray results obtained from control mixtures of *in vitro*-synthesized RNA. The straight line describing the data has a slope of 0.72 and an intercept of 3.1% ($r^2 = 0.93$). Error bars represent s.e.m. ($n = 4$). **(d)** Relative levels of differentially spliced exons of *Dscam* at several stages during development. The level of expression for each exon is represented relative to the level expected for a random distribution (1/number of alternative exons at a given cluster; shown in yellow), with red representing preferential selection and blue diminished representation. The black bar denotes the alternative exons shown in more detail in **e**. **(e)** Relative expression levels of five alternative forms of exon 9 throughout development. The RNA used in this analysis was extracted from whole organisms. Error bars represent s.e.m. ($n \geq 2$).

Figure 2 A large number of different *Dscam* isoforms is expressed by homogeneous cell populations. **(a)** Relative expression of exon 9 isoforms in RNA samples isolated from distinct tissues of wandering third instar larvae. Whole brains and eye-antennal imaginal disks were dissected. Different photoreceptor populations were isolated by FACS using GFP driven by specific promoters (ELAV for photoreceptors, *mδ* for R3/R4 cells and PM181 for R7 cells). Using a threshold for detection of 50% of the expected representation, we detected in the R7 photoreceptor population 25 different exon 9 isoforms, 41 exon 6 isoforms and 11 exon 4 isoforms. We therefore estimate that a R7 photoreceptor could choose between 11,275 of the 19,008 possible combinations of these three exons. These values are underestimations because single-cell RT-PCR can detect expression of even rarely chosen alternative exons (**Fig. 5d**), and the splicing of exon 17 is not taken into consideration. **(b)** Three-dimensional reconstruction of R3/R4 photoreceptor cells in the eye disk of third instar larvae. **(b,d)** Whole ommatidia were stained red using antibody to ELAV; two adjacent photoreceptors (R3 and R4) were labeled per ommatidium. **(c)** Three-dimensional reconstruction of the axon projections of R3/R4 photoreceptors in the optic lobe of a wandering third instar larvae. **(d)** R7 cells in the eye disk of third instar larvae. One cell per ommatidium is visible. **(e)** Axon projections of R7 photoreceptors visualized with membrane-bound GFP. **(c,e)** Medial is to the bottom, anterior to the right; arrows indicate the position of the lamina and brackets the position of the medulla. Genotype for **b** and **c** was *mδ-GAL4/+; UAS-mCD8GFP/+* and for **d** and **e** was *PM181-GAL4; UAS-mCD8GFP/+*. Scale brackets: **b**, 7.6 μ m; **c,e**, 14.7 μ m; **d**, 8.9 μ m.

Dscam expression in fruit fly cell lines

Although R3/R4 and R7 cells are relatively homogeneous populations, the possibility of subtle differences among individual neurons in a population is inherent in any *in vivo* analysis. We therefore examined *Dscam* expression in a more uniform population of cells: S2 cell lines and sub-clones thereof, which express a conspicuous profile of *Dscam* isoforms (**Figs. 3** and **4a**). Use of different isoforms of exons 4 and 6 was roughly equal, whereas only 5 of the 33 isoforms for exon 9 were used. This observation was confirmed using restriction endonuclease digestion analysis of 78 cloned cDNAs from S2 cell lines. Over 97% of the analyzed cDNAs contained one of these five isoforms of exon 9. In contrast, we observed an almost random distribution of isoforms of exon 6 (**Fig. 3**).

S2 cells express a variety of hemocyte markers⁹ and can be used as an *in vitro* model for phagocytosis^{10,11}. Two other cell lines that have been suggested to share similar properties with hemocytes (Mbn-2 (ref. 12) and Kc^{9,13}) have a profile of *Dscam* isoforms similar to that of S2 cells (data not shown). We isolated hemocytes by FACS from third instar larvae that express GFP under the control of the promoter region of the hemocyte marker hemolectin (*hml*)¹⁴. Consistent with a potential hemocyte origin of these cell lines, the spectrum of *Dscam* isoforms of primary hemocytes was overall similar to and had the same restriction of exon 9 use as the cell lines. The same five exon 9 isoforms were over-represented in whole embryos, larval eye-antennal disks and larval hemocytes. The extent of expression of these isoforms varied greatly

between the different cell populations, and it is unclear whether the use of these isoforms is functionally relevant or whether they represent a 'default' splicing option that can be overwritten in more differentiated cells.

To test whether the alternative splicing choices at different exons were independent, we determined whether the choice of one particular alternative exon biased the choice of exons 4 or 6. We amplified cDNA generated from S2 cell lines with one primer that bound on the constitutive exon 3 and one primer that bound in one of the alternative isoforms of exon 9. No specific exon 4 and 6 isoform specifically associated with either exon 9.9 or exon 9.31 (**Supplementary Fig. 1** online). These results suggest that alternative splicing occurs independently at exons 4, 6 and 9.

To determine whether the restricted *Dscam* repertoire of S2 cells and primary hemocytes was the result of the uniformity of the cell populations, we analyzed two S2 cell lines maintained independently

Figure 3 Independent confirmation of microarray results using restriction endonuclease analyses of cloned cDNAs. **(a)** Restriction endonuclease analyses of the distribution of exon 9 isoforms in 73 *Dscam* cDNA clones generated from S2 cell lines. The results are compared with microarray analyses of the same population. **(b)** Restriction endonuclease analyses of the distribution of exon 6 isoforms in the same set of 73 *Dscam* cDNA clones as in **a**. Capital letters denote groups of alternative exons that cannot be further dissected using this combination of restriction endonucleases. Results in **a** and **b** are expressed as a percentage of the total exons detected at each cluster. **(c)** Representative example of the restriction endonuclease digestion analysis (clone 1, isoform 9.13; clone 2, isoform 9.30). The validity of this approach was confirmed by sequencing. **(d)** A similar analysis was carried out for alternative exon 6. Clone 1 (same clones as in **c**) contained alternative forms 4.6 and 6.25, and clone 2 contained isoforms 4.4 and 6.48.

in separate labs (nonclonal populations). Furthermore, we generated several subclonal populations. All independent clonal cell lines showed the same narrow spectrum of *Dscam* isoforms, including those lines generated by two successive rounds of subcloning and subclonal lines analyzed soon after subcloning (Fig. 4). These results suggest that individual cells express a large number of different alternative exons, but the most direct way to determine the number of different isoforms expressed by an individual cell is to analyze the RNA expressed by isolated cells.

Individual cells express multiple *Dscam* splice variants

The results of a single-cell RT-PCR experiment in which the conversion of mRNA to cDNA is so inefficient that only one mRNA molecule

can be amplified could be wrongfully interpreted to reflect the expression of only one *Dscam* isoform per cell. One way to address this problem is to split the reverse transcription reaction from an individual cell into multiple tubes, and then carry out independent PCR reactions to score only those cells for which sufficient cDNA for meaningful analysis was generated¹⁵. We placed individual cells in separate tubes by FACS. We carried out a reverse transcription reaction in each tube, separated its contents into 8–10 tubes and then carried out two rounds of PCR amplification independently using partially nested primers that bind on constitutive exons 3 and 7 (Fig. 5a). This method allowed us to analyze the regulation of alternative splicing at the *Dscam* locus by individual cells.

Preliminary analysis using individual hemocytes confirmed the importance of carrying out multiple independent PCR reactions from the same cell. Figure 5 outlines the methods we used to analyze single-cell RT-PCR, including one example in which we split one reverse transcription reaction into eight independent PCR amplifications, of which two yielded a product (Fig. 5b). Microarray analyses of these amplification products indicated that the PCR reactions were sensitive enough to preferentially amplify a single cDNA molecule. Different molecules were amplified in multiple independent PCR reactions from the same cell, showing that each cell expressed at least two different *Dscam* isoforms (Fig. 5d). Owing to the large number of amplification cycles used in single-cell RT-PCR, it is difficult to quantify the relative amounts of different isoforms present in a cell. Therefore, we consider detection of alternative exons above a certain threshold instead. Figure 5c shows an example cell in which all eight independent PCR amplifications were successful. When all the PCR reactions were combined, it was possible to identify at least 14 different isoforms (Fig. 5d). We used similar methods to analyze alternative splicing of exon 9. The same findings apply to differentiated neurons isolated from the brain of third instar larvae (genotype ELAV-GFP; Fig. 5e).

Figure 6 shows a more comprehensive application of these methods to analyze expression of exon 9 isoforms in single cells that belong to specific photoreceptor populations (for example, seven informative R3/R4 photoreceptors and seven informative R7 photoreceptors (out of 16 sorted cells for each type); Fig. 6a,b). As expected, there was a preference for expression of isoforms over-represented at the population level. Notably, each photoreceptor expressed a different profile of isoforms. The single-cell RT-PCR results allowed us to estimate the number of mRNA transcripts present in each cell. To accomplish this, a computer-based simulation randomly distributed different numbers of cDNA copies among ten wells and tallied the numbers of positive wells after including an empirically determined single-copy PCR failure rate¹⁶. From these simulated distributions of experimental outcomes, we determined the 99% confidence interval for the actual number of cDNAs present in each single cell after reverse transcription, based on the observed data (i.e., the fraction of successful PCRs for a given cell type). The estimated range was 7–10 cDNA copies per cell for the R3/R4 photoreceptor cells and 14–25 cDNA copies per cell for the R7 photoreceptor cells. Estimation of the number of *Dscam* mRNAs present in a given cell must also take into account the efficiency of the conversion of mRNA to cDNA. This efficiency is estimated to be on the order of 20–50% (ref. 17). Thus, for the R3/R4 cell type, the range of 7–10 cDNA copies per cell would convert to a range of 14–50 mRNAs per cell.

The single-cell RT-PCR data also allowed us to address the issue of how many different alternative splice forms are present in a given cell. For all 21 individual neurons analyzed for exon 9, the number of different isoforms detected for any given neuron increased with increasing numbers of successful independent PCR amplifications

Figure 4 *Dscam* expression in hemocytes and several clonally derived S2 cell lines. (a) Relative levels of *Dscam* alternative exon use in FACS-sorted hemocytes (based on hml-GFP fluorescence; average of three independent FACS-sorted populations) were similar to those of several clonally derived S2 cell lines. (b,c) Differential interference contrast micrograph of the clonal populations b and c, just before RNA extraction and *Dscam* analysis. We estimate, based on the number of cells, that less than ten cell divisions have taken place since the initial cloning.

(Fig. 6c). This suggests that each independent PCR-amplified cDNA contains different alternative exons (as shown in Fig. 5d and further confirmed by restriction endonuclease digestion of individual PCR products (data not shown)). Taken together, these observations are consistent with a model in which the *Dscam* mRNA profile of a given cell arises from a series of stochastic alternative splicing events for each *Dscam* transcript. The probability of selecting each individual alternative exon is a function of the splicing factors expressed by each cell type.

DISCUSSION

We determined the distribution of *Dscam* isoforms in individual cells and in homogeneous cell populations. We showed that a given cell type expresses a broad, yet distinctive, spectrum of splice variants. Single-cell RT-PCR analyses showed that individual cells belonging to the same cell type express diverse complements of *Dscam* isoforms. The generation of a unique profile of cell-surface protein isoforms in each cell through alternative splicing is a potential mechanism of specifying cell identity in the nervous system. Analyses of the differences in binding properties of the different alternative exon isoforms will be important in understanding how the diversity in the DSCAM protein is used by individual cells.

Analyzing the mRNA present in an individual cell at a given time does not give the full picture of the repertoire of DSCAM proteins present in that cell at that time. Additionally, the repertoire of both mRNA and protein isoforms is certainly changing over time. Thus, analyses of the half-lives of *Dscam* mRNA and protein products will be

important to understanding the regulation and function of Dscam in *D. melanogaster*. Even a Dscam protein repertoire that is changing over time affords each cell the possibility of being distinct from otherwise identical cells in its vicinity. Being different from neighboring cells may be important during cell migration and during the growth of axons and dendrites, as well as during synapse formation and plasticity. Dscam could also contribute to the generation of diversity within each cell, giving rise to portions of a given cell that are distinct from one another. This could be accomplished by the presence of different *Dscam* transcripts in different portions of the cell, along with localized protein synthesis.

Analyses of other insect species, such as *Drosophila yakuba*⁴, *Drosophila pseudoobscura* (G.N. and A.C., unpublished observations) and *Anopheles gambiae*¹⁸ (mosquito), identify a *Dscam* ortholog with a genomic organization similar to that of *D. melanogaster* and extensive alternative splicing⁴. The fact that even a remotely related diptera, such as *A. gambiae*, shares the ability to generate a large repertoire of Dscam isoforms (with distinct immunoglobulin domains) suggests that there has been selection for this capability. Although neither the mammalian gene *DSCAM* nor the related gene *DSCAML1* have extensive alternative splicing¹⁹, other cell surface-protein families involved in neuronal development, such as neuroligins²⁰ and protocadherins²¹, have numerous differently spliced forms. Single-cell RT-PCR and microarray analyses of alternatively spliced isoforms could allow analyses of the distribution of alternatively spliced isoforms in these other genes and their possible roles in specifying neuronal identity.

Figure 5 Single cells express multiple *Dscam* isoforms. (a) Schematic representation of the strategy used for single-cell RT-PCR analysis. (b) Agarose gel electrophoresis analysis shows that two of the eight independent PCR amplifications from the cDNA synthesized from one isolated hemocyte yielded an amplification product, which was analyzed using the microarray. (c) Eight independent positive PCR reactions obtained from the same isolated hemocyte. All amplification products were purified, pooled together, labeled and hybridized to the microarray. (d) Microarray analysis of *Dscam* expression in representative individual hemocytes. We used the expected representation at each cluster as a threshold for detection in Figures 5 and 6. Detected exons are indicated by a filled box. We obtained eight independent positive PCR amplifications for cell H1 (shown in c), two for cell H2 (shown in b), five for cell H3 and five for cell H4. Left columns represent the *Dscam* repertoire of the population of hemocytes. Fewer than five different isoforms of exon 9 were observed for cells H3 and H4, which was expected from the restricted repertoire expressed by the hemocyte population. All *Dscam* isoforms detected in these cells are expected to be different at the exon 4 and 6 clusters, based on the results obtained for cells H1 and H2. (e) Microarray analysis of *Dscam* expression in representative individual postmitotic neurons isolated by FACS from brains of wandering third instar larvae. Left columns represent the *Dscam* repertoire of the population of neurons isolated using the same method (based on ELAV-GFP fluorescence; average of two independent FACS-sorted populations). Cell E1, two independent PCR amplifications; E2, four; E3, seven; E4, four.

Figure 6 Single photoreceptors from the same class express different *Dscam* repertoires. **(a)** Microarray analysis of the expression of exon 9 isoforms in seven R3/R4 photoreceptors. R3 cell 1, seven independent PCR amplifications; cell 2, four; cell 3, six; cell 4, six; cell 5, two; cell 6, five; cell 7, five. The population analysis for **a** and **b** is the same as shown in **Figure 2a**. Even though there is a preference for expression of isoforms over-represented at the population level (Pop.), each photoreceptor shows a different profile of isoforms. **(b)** Microarray analysis of the expression of exon 9 isoforms in seven R7 photoreceptors. R7 cell 1, eight independent PCR amplifications; cell 2, eight; cell 3, eight; cell 4, six; cell 5, nine; cell 6, ten; cell 7, seven. **(c)** Results from all 21 individual neurons analyzed for exon 9 show that the number of detected isoforms of *Dscam* is proportional to the number of successful PCR amplifications.

METHODS

***Dscam* oligonucleotide microarrays.** Oligonucleotide probes (average size of 65.5 bp; sequences of all probes and primers available on request) designed to each of the 93 alternative exons of clusters 4, 6 and 9 of *Dscam* were synthesized by MWG and resuspended at a concentration of 1 g l^{-1} in 3 standard saline citrate with 1.5 mM betaine. In addition to the probes for *Dscam* alternative exons, we included positive controls for constitutive exons 7 and 8 and for three distinct olfactory receptors (Or22a, Or46a and Or47b). We added to our samples defined amounts of isolated DNA from cloned olfactory receptors to control hybridization conditions. We also included intergenic sequences from yeast (*Lucidea Universal Scorecard*, Amersham) as negative controls. Microarray printing, hybridization, washing and scanning were done essentially as described²². Detailed protocols for these processes and details of microarray analysis are given in **Supplementary Methods** online.

Cell sorting. We dissected eye-antennal discs from wandering third instar larvae in nonsupplemented Shields and Sang M3 medium (M3). Discs were digested for 2–4 h at room temperature in a 10 trypsin-EDTA (Sigma), 1 phosphate-buffered saline (PBS) solution. We then mechanically dissociated cells using micropipettes in M3 medium supplemented with 1% bovine serum albumin (Sigma), 10 mM MgCl₂ and 10 mM EGTA. We filtered the cell suspension through a cell strainer (35- μm nylon mesh) (Becton Dickinson) and incu-

bated it with 5 g ml^{-1} of propidium iodide. We positively selected cells based on GFP fluorescence and negatively selected them against propidium iodide accumulation using a fluorescence-activated cell sorter (MoFlo; Cytomation). We sorted cells into ice-cold SUPERase RNase inhibitor (1 U l^{-1} ; Ambion) and quickly transferred them to dry ice. We isolated RNA using TRIZOL (Invitrogen). We isolated larval brain neurons using similar methods, but the digestion conditions were 1 h at room temperature with 1 trypsin-EDTA in M3 medium. For cell sorting of hemocytes, we collected hemolymph of wandering third instar larvae by puncturing the larval surface with forceps, diluted it in PBS containing propidium iodide and filtered it through a cell strainer before sorting it using the same methods described above.

Single-cell RT-PCR. For single-cell RT-PCR analysis, we sorted GFP-positive, propidium iodide-negative cells into 96-well PCR plates containing 1 l of 5 Superscript II buffer (Invitrogen), 1 l of SUPERase RNase inhibitor (1 U l^{-1} ; Ambion), 0.25 l of Oligo dT (0.5 mg ml^{-1} ; Ambion), 0.25 l dithiothreitol (0.1 M) and 2.5 l of DEPC-treated water. Immediately after sorting, plates were sealed with aluminum foil (Axygen) and frozen in dry ice. We stored them at -80°C until use. For reverse transcription, cells were briefly spun down, heated at 65°C for 5 min and rapidly transferred to ice for 1 min. We carried out a 20- l reverse transcription reaction according to the manufacturer's instructions (Invitrogen) using Superscript III as enzyme and a specific primer that bound in exon 10 of *Dscam*. We used 2 l of the reverse transcription reaction as template for a 20- l PCR amplification. Parameters for the first-round PCR amplification were as follows: 5 min at 94°C ; 40 cycles of 30 s at 94°C , 30 s at 58°C , 1 min at 72°C ; and a final elongation step of 5 min at 72°C using *Taq* DNA polymerase (Promega) in a Robocycler thermal cycler. We used 1 l of this reaction as template for a second round of amplification (20- l volume) with partially nested primers using similar parameters but only 30 cycles. We analyzed the sizes of the amplification products ($\sim 492 \text{ bp}$ for exons 3–7 and $\sim 486 \text{ bp}$ for exons 8–10) by agarose gel electrophoresis. We detected no amplification products when we used water as template.

Preliminary single-cell RT-PCR experiments were difficult to optimize owing a combination of factors, including low levels of *Dscam* expression by hemocytes and the relative inefficiency of primers that bound on exons 3 and 7 compared with the primers that bound on exons 8 and 10. For example, of 43 individual hemocytes analyzed for alternative exons 4 and 6, we obtained at least one positive PCR reaction in 15 cells. Of these, most (10) had only one positive PCR reaction. In contrast, of 11 neurons analyzed with the same primer set, we obtained more than one independent PCR reaction for five cells. Once conditions were optimized, we studied in further detail two photoreceptor populations (R7 and R3/R4). For these analyses, we tested four independent PCRs for each of 16 individually sorted photoreceptors. At this stage, we were unable to detect any amplification product in 9 of the 16 photoreceptors of each class. These cells were probably not sorted into the reverse transcription solution or were not lysed properly and they were not analyzed further. We cannot rule out the possibility that these cells had a significantly lower level of expression of *Dscam* that prevented their analysis using single-cell RT-PCR methods. For the remaining seven cells, we carried out six additional PCR reactions (results shown in **Fig. 6**).

Simulations to evaluate number of cDNA copies per cell. To estimate the true number of cDNA copies present based on the number of positive wells of the ten in which the single-cell reverse transcription reactions were distributed, we designed computer simulations of the experiment to evaluate outcome probabilities. For a given number of copies, we carried out one million simulated experiments by randomly distributing the copies into ten bins. The outcome of a particular simulation was the number of wells out of ten that produced a positive PCR, with a random PCR failure rate of $\text{f}_{\text{neg}}^{\text{number of copies}}$. The false negative rate for PCR of a single copy (f_{neg}) has been estimated to be 15% (ref. 16), and we adopted this value for these calculations (varying this parameter did not alter the results substantively). Using these simulation results, we created a probability distribution for the number of cDNA copies given the number of positive PCRs by assuming a flat prior distribution over the number of cDNA copies. To obtain the final confidence limits, we calculated the joint probability of all seven replicate experiments done for each cell type (assuming a uniform expression level for each photoreceptor type).

Flies. We raised flies on standard fruit fly medium at room temperature or at 25 °C. The wild-type *D. melanogaster* strain was Oregon Red. We combined the enhancer trap line Gal4-C155 (ELAV-Gal4; ref. 23) with UAS-GFP and used it to label all postmitotic neurons for cell sorting. The HML-GFP line was originally generated by A. Goto¹⁴ (Nagoya University, Japan) and is available through the Bloomington Stock Center. We used markers specific to different R cell subtypes. PM181-Gal4 (ref. 8) is expressed in R7 photoreceptors, and mD-GAL4 (generated by P. Garrity; Massachusetts Institute of Technology) is expressed in R3/R4 photoreceptors. Both lines were independently combined with the reporter line UAS-mCD8-GFP²⁴ and used both to examine photoreceptor neurons in the third larval stage and as a marker for cell sorting. All the lines described above were a gift from P. Garrity.

Developmental time course experiments were based on timed egg laying, followed by incubation at 25 °C for the appropriate period of time. Embryos were 6–9 h old, first instar larvae 37–47 h old and second instar larvae 50–64 h old. Wandering third instar larvae were picked from the sides of bottles, and adult flies were female virgins. The RNA used in this analysis (Fig. 1d,e) was extracted from whole organisms.

Immunostaining and confocal microscopy. We dissected eye-brain complexes from wandering third instar larvae in PBS and fixed them in 4% paraformaldehyde in PBS for 1 h at room temperature. We blocked them in blocking solution (PBS plus 0.5% Triton-X100 and 10% normal goat serum) for 30 min at room temperature. We incubated samples in monoclonal antibody to rat ELAV (DSHB 7E8A10; diluted 1:50 in blocking solution) at 4 °C overnight and in Cy3-conjugated goat antibody to mouse (Jackson Immunoresearch; diluted 1:200 in blocking solution) at room temperature for 2 h. We mounted them in Vectashield (Vector). Serial confocal sections were acquired with a Carl Zeiss LSM510 confocal system, and image analysis and three-dimensional reconstructions were done using Velocity 2.0 software (Improvision).

Cell culture. We maintained S2 and Kc cells, as well as the Mbn-2 blood cell line, in Schneider's medium supplemented with 10% fetal bovine serum, 100 g ml⁻¹ of streptomycin-penicillin in an incubator at 22 °C. We carried out serial dilution as described²⁵.

Molecular biology. We isolated total RNA using TRIzol (Invitrogen) according to the manufacturer's protocols and reverse-transcribed 1 g of total RNA using Superscript III (Invitrogen) with a specific primer that binds in the constitutive exon 10. We used 2 l of the reverse transcription reaction as a template for a 25-l PCR amplification. PCR was done for 5 min at 95 °C, followed by 30–35 cycles of 30 s at 95 °C and 2.5 min at 68 °C, and a final elongation step of 5 min at 68 °C using Platinum Pfx DNA polymerase (Invitrogen) in a MJ Research thermal cycler. Primers bound at constitutive exons 3 and 10, and the size of the ~2,109-bp amplification product was confirmed by agarose gel electrophoresis. We purified the amplification products using Wizard SV PCR purification kit (Promega) and labeled them by incorporation of Cy3dUTP or Cy5dUTP (Perkin-Elmer) using the Bioprime kit (Invitrogen). Labeled DNA was purified using Cyscribe Gfx (Amersham) purification columns. We labeled the reference library in the same way as the test samples. Incorporation of labeled nucleotides was measured using a Nanodrop spectrophotometer (Nanodrop Technologies).

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

ACKNOWLEDGMENTS

We thank members of the laboratories of A.C. and P. Garrity for suggestions and discussion during the course of this work, T. Volkert for help in the design and printing of microarrays at the Center for Microarray Technology at the Whitehead

Institute, B. Chazaro and J. Love for help with hybridizations and scanning, N. Watson for help with confocal microscopy at the W. M. Keck microscopy facility at the Whitehead Institute, G. Paradis for help with FACS experiments, The Bloomington Stock Center and P. Garrity laboratory for providing *D. melanogaster* stocks, several laboratories for providing the different cell lines (S2 from T. Orr-Weaver and I. Rebay, Kc from S. Bell, Mbn-2 from D. Hultmart), T. DiCesare for help with figure preparation and P. Garrity and S.L. Zipursky for critical reading of the manuscript and for sharing data before publication.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 14 October 2003; accepted 2 January 2004

Published online at <http://www.nature.com/naturegenetics/>

- Schmucker, D. *et al.* *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**, 671–684 (2000).
- Wang, J., Zugates, C.T., Liang, I.H., Lee, C.H. & Lee, T. *Drosophila* Dscam is required for divergent segregation of sister branches and suppresses ectopic bifurcation of axons. *Neuron* **33**, 559–571 (2002).
- Hummel, T. *et al.* Axonal targeting of olfactory receptor neurons in *Drosophila* is controlled by Dscam. *Neuron* **37**, 221–231 (2003).
- Celotto, A.M. & Graveley, B.R. Alternative splicing of the *Drosophila* Dscam pre-mRNA is both temporally and spatially regulated. *Genetics* **159**, 599–608 (2001).
- Brand, A.H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Jasper, H. *et al.* A genomic switch at the transition from cell proliferation to terminal differentiation in the *Drosophila* eye. *Dev. Cell* **3**, 511–521 (2002).
- Cooper, M.T. & Bray, S.J. R7 photoreceptor specification requires Notch activity. *Curr. Biol.* **10**, 1507–1510 (2000).
- Lee, C.H., Herman, T., Clandinin, T.R., Lee, R. & Zipursky, S.L. N-cadherin regulates target specificity in the *Drosophila* visual system. *Neuron* **30**, 437–450 (2001).
- Franc, N.C., Heitzler, P., Ezekowitz, R.A. & White, K. Requirement for croquemort in phagocytosis of apoptotic cells in *Drosophila*. *Science* **284**, 1991–1994 (1999).
- Ramet, M. *et al.* *Drosophila* scavenger receptor Cl is a pattern recognition receptor for bacteria. *Immunity* **15**, 1027–1038 (2001).
- Boutros, M., Agaisse, H. & Perrimon, N. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev. Cell* **3**, 711–722 (2002).
- Samakovlis, C., Asling, B., Boman, H.G., Gateff, E. & Hultmark, D. In vitro induction of cecropin genes—an immune response in a *Drosophila* blood cell line. *Biochem. Biophys. Res. Commun.* **188**, 1169–1175 (1992).
- Echalier, G. *Drosophila* cells in culture (Academic, San Diego, 1997).
- Goto, A. *et al.* A *Drosophila* haemocyte-specific protein, hemolentin, similar to human von Willebrand factor. *Biochem. J.* **359**, 99–108 (2001).
- Rhoades, K.L. *et al.* Allele-specific expression patterns of interleukin-2 and Pax-5 revealed by a sensitive single-cell RT-PCR analysis. *Curr. Biol.* **10**, 789–792 (2000).
- Mostoslavsky, R. *et al.* Kappa chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev.* **12**, 1801–1811 (1998).
- Klickstein, L.B., Neve, R.L., Golemis, E.A. & Gyuris, J. Conversion of mRNA into double-stranded cDNA. in *Current Protocols in Molecular Biology* vol. 1 (eds. Ausubel, F.M. *et al.*) (John Wiley & Sons, 1995).
- Zdobnov, E.M. *et al.* Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* **298**, 149–159 (2002).
- Barlow, G.M., Micales, B., Chen, X.N., Lyons, G.E. & Korenberg, J.R. Mammalian DSCAMs: roles in the development of the spinal cord, cortex, and cerebellum? *Biochem. Biophys. Res. Commun.* **293**, 881–891 (2002).
- Tabuchi, K. & Sudhof, T.C. Structure and evolution of neuroligin genes: insight into the mechanism of alternative splicing. *Genomics* **79**, 849–859 (2002).
- Wu, Q. *et al.* Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Res.* **11**, 389–404 (2001).
- Ren, B. *et al.* Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309 (2000).
- Lin, D.M. & Goodman, C.S. Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* **13**, 507–523 (1994).
- Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
- Ashburner, M. *Drosophila: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989).