Factorial microarray analysis of zebrafish retinal development

Yuk Fai Leung[†‡§¶], Ping Ma[§‖††], Brian A. Link[‡‡], and John E. Dowling[†¶]

[†]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138; [‡]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907; [‖]Department of Statistics and [††]Institute for Genomic Biology, University of Illinois, Champaign, IL 61820; and [‡‡]Department of Cell Biology, Neurobiology, and Anatomy, Medical College of Wisconsin, Milwaukee, WI 53226

Contributed by John E. Dowling, June 22, 2008 (sent for review April 17, 2008)

In a zebrafish recessive mutant *young* (*yng*), retinal cells are specified to distinct cell classes, but they fail to morphologically differentiate. A null mutation in a *brahma-related gene 1* (*brg1*) is responsible for this phenotype. To identify retina-specific Brg1-regulated genes that control cellular differentiation, we conducted a factorial microarray analysis. Gene expression profiles were compared from wild-type and *yng* retinas and stage-matched whole embryos at 36 and 52 hours postfertilization (hpf). From our analysis, three categories of genes were identified: (*i*) Brg1-regulated retinal differentiation genes (731 probesets), (*ii*) retina-specific genes independent of Brg1 regulation (3,038 probesets), and (*iii*) Brg1-regulated genes outside the retina (107 probesets). Biological significance was confirmed by further analysis of components of the Cdk5 signaling pathway and Irx transcription factor family, representing genes identified in category 1. This study highlights the utility of factorial microarray analysis to efficiently identify relevant regulatory pathways influenced by both specific gene products and normal developmental events.

Brg1-regulated genes | neuronal differentiation | cellular differentiation | signal transduction pathways | transcription factors

Microarray analysis is a common and important tool for phenotype analysis. Most often, gene expression differences between a limited number of biological conditions are compared pairwise or to a standard reference. Alternatively, gene expression values are used as markers for disease classification. Numerous statistical tools have been developed over the years to address these needs (1). However, it is difficult to study combinatorial effects on gene expression when multiple biological perturbations are introduced. This frequently arises in analysis of pleiotropic mutants, in which multiple tissue-specific phenotypes and developmental delays exist. Factorial analysis is a statistical tool for analyzing the effects of several independent variables on a dependent variable (2, 3), a solution for this situation. Here we describe the use of a factorial design to identify specific molecular controls for retinal development using the zebrafish *young (yng)* mutant.

The retina consists of six classes of neurons and one major glial cell that originate from the same retinal progenitor cell population. The cell classes are specified around the time of cell cycle withdrawal by both intrinsic and extrinsic signals (4, 5). After cell-type fate commitment, cells undergo terminal differentiation, elaborate neuronal processes, and begin synaptogenesis. The final product of retinal development is a complex neural tissue that conveys visual information to the rest of the brain. Cell proliferation and neuronal specification have been extensively studied (5, 6). However, regulation of terminal differentiation remains poorly understood.

In the zebrafish *yng* mutant, retinal cells fail to terminally differentiate, although each major cell class is specified (7). The causative mutation disrupts the *brg1*/*smarca4* gene, which codes for the ATPase of the SWI/SNF chromatin remodeling complex (8). These complexes are under tight spatial and temporal control, and can create local alterations in chromatin structure and permit gene transcription at specific developmental stages. Although MAP kinase (MAPK) activity is affected by loss of Brg1 (9), the ultimate targets of Brg1-mediated cell differentiation pathways remain elusive. In addition to retinal defects, *yng* embryos show dysgenesis of other organs including the ears and heart. Although the mutation in *brg1* is the underlying genetic cause for these developmental problems, the altered downstream target profiles are likely different for each organ. Therefore, tissue specificity must be accounted for when identifying Brg1-regulated genes.

To identify retinal genes controlled by Brg1 and understand better the dynamic mechanisms of terminal differentiation of retinal cells, we compared gene expression levels in wild-type (WT) and *yng* retinas at two key developmental stages: 36 and 52 hours postfertilization (hpf). These times mark the initiation of cellular morphogenesis and synaptogenesis, respectively. By comparing retinal tissue to the whole embryo, we also identified genes specifically expressed in the retina but not differentially expressed between WT and *yng*. Finally, we identified genes that are regulated by *brg1* outside the retina, providing insights to Brg1 function in other tissues.

Results

Factorial Analysis. ***Overview of factorial design.*** Retinas at 36 and 52 hpf from WT (WR36 and WR52) and *yng* (YR36 and YR52) larvae were isolated as previously described (10). Stage-matched whole embryos (WA36, WA52, YA36, YA52) were also collected. At 36 hpf, initial GCs differentiation is underway (11) and MAPK activity is first observed in the ventral part of the anterior retina (12). Synaptogenesis and overt retinal lamination is yet to occur [supporting information (SI) Fig. S1]. By 50–52 hpf, several cell types are differentiating and lamination is visible (11, 13) (Fig. 3 and Fig. S1). Widespread MAPK activity is also observed (12). Thus, gene expression changes at 36 and 52 hpf identify transcripts regulated by the Brg1-MAPK retinal differentiation pathway. Inclusion of whole embryos in the study provides information about retinal specificity of the gene expression.

We analyzed the effects of three factors on gene expression: (*i*) mutation (*M*), (*ii*) tissue (*R*), and (*iii*) time (*T*). Each of these factors has two levels: mutant vs. WT (*M*); retina versus whole body (*R*); and 36 versus 52 hpf (*T*), making a total of 8 conditions. To handle these complicated relationships, we used a $2 \times 2 \times 2$ factorial design (Fig. 1) (2, 3). The analysis delineates the effects of each factor and their combined effects on gene expression. The analytical strategy is to first fit a full three-way analysis of variance (ANOVA) model

Author contributions: Y.F.L., P.M., B.A.L., and J.E.D. designed research; Y.F.L. performed research; Y.F.L., P.M., B.A.L., and J.E.D. analyzed data; and Y.F.L., P.M., B.A.L., and J.E.D. wrote the paper.

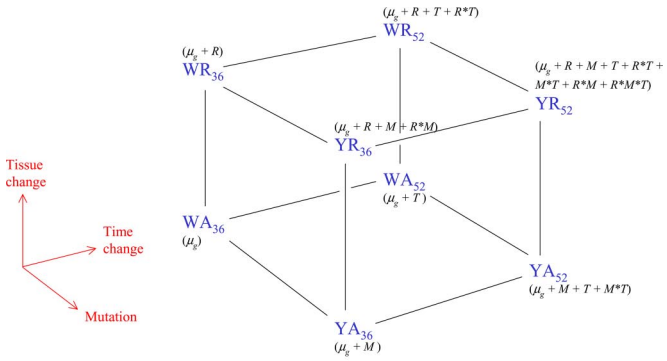
The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE8874).

[§]Y.F.L. and P.M. contributed equally to this work.

[¶]To whom correspondence may addressed. E-mail: yfleung@purdue.edu or dowling@mcb.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806038105/DCSupplemental.



Null hypothesis to test	Interpretations when the null hypothesis is rejected (corresponding fold change)
1. 1-way (main factor) models	
1.H_0: M = 0	There is a non-retinal specific significant differential expression at 36hpf. (*yng* vs WT at the same time point)
2. 2-way (first order) models (A model is regarded as significant if either contrast 2.1 or 2.2 is significant)	
*1. H_0: M + M*R = 0*	There is a retinal-specific significant differential expression at 36hpf. (YR36/WR36)
*2. H_0: M + M*R + T*M = 0*	There is a retinal-specific significant differential expression at 52hpf. (YR52/WR52)
3. 3-way (second order) models (A model is regarded as significant if either contrast 3.1 or 3.2 is significant)	
*1. H_0: M + M*R = 0*	There is a retinal-specific significant differential expression at 36hpf (YR36/WR36)
*2. H_0: M + M*R + T*M + T*M*R = 0*	There is a retinal-specific significant differential expression at 52hpf. (YR52/WR52)

Fig. 1. Factorial microarray experimental design and specific contrasts for inferring Brg1-regulated retinal terminal differentiation genes. See *Factorial Analysis* for details.

(Eq. **1**; see details below) with all possible main effects (T, M, R) and their interactions ($T*M$, $M*R$, $R*M$, and $T*M*R$) to the gene expression data for each probeset. The coefficients of the models are obtained through a maximum likelihood estimation method. Then a backward elimination strategy is used to remove insignificant interactions to get the most parsimonious ANOVA model (one-, two- or three-way) for each probeset. Finally, a group of contrasts (specific linear combinations of coefficients; see *SI Appendix* for all of the contrasts used), and their corresponding fold changes, are used to infer whether that probeset is significantly associated with a particular biological process.

(Step 1) Model fitting. A full three-way ANOVA model (Eq. **1**) with all of the coefficients ($T, M, R, T*M, M*R, R*M$, and $T*M*R$) was fitted for each probeset by using a maximum likelihood method.

$$y_g = \mu_g + R + M + T + M*R + T*M + T*R + T*M*R + \varepsilon \qquad [1]$$

y_g is the expression value of the probeset g in a particular replicate as measured by the microarray; μ_g is the basal expression value of the probeset in WA36, also known as the intercept term (abbreviated as I in *SI Appendix* and Table S8, Table S10, and Table S12); and ε is the error term.

Each model consists of a family of 8 equations to represent gene expression in the eight different conditions. For example, the three-way model in Eq. **1** consists of the 8 equations as shown in the brackets in Fig. 1. Also, there were three replicates for each condition/equation. As a result, each probeset was represented by an ANOVA model with 24 equations. Second, a backward elimination strategy was used to find the minimal model that could best explain the expression values of a particular probeset.

If a probeset had an insignificant three-way interaction coefficient ($T*M*R$) [false discovery rate (FDR) q-value >0.05, (14)], a simpler two-way model (Eq. **2**), without the insignificant $T*M*R$ coefficient, was refitted for the probeset.

$$y_g = \mu_g + R + M + T + M*R + T*M + T*R + \varepsilon \qquad [2]$$

If the two-way interaction coefficients ($M*R$, $T*M$, or $T*R$) were again insignificant after fitting the two-way model, a one-way model (Eq. **3**) with only the main effects coefficients (T, M, R) was refitted for this probeset.

$$y_g = \mu_g + R + M + T + \varepsilon \qquad [3]$$

Ultimately, the most parsimonious model with the fewest higher order interacting terms (in this case the minimal eight equations) was identified to explain the gene expression for each probeset. The identification of a probeset in a higher model suggests a qualitatively specific retinal change, because the presence of the other factors with the R factor (i.e., $M*R$, $T*R$, or $T*M*R$) has an interaction effect on gene expression not present when the factors are present individually. In other words, these ANOVA models have qualitatively categorized the probesets by relative retinal specificity. The three-way models have the most retina-specific change, followed by the two-way models and then the one-way models. This categorization, which is not possible by pairwise comparisons alone, will help prioritize functional characterization of the significant genes.

(Step 2) Identification of significant genes (i.e., significance inference). After the best fit model for each probeset was obtained, specific contrasts, a combination of the coefficients from the best fit model for finding the effects of factors and their combined effects on gene expression levels (2, 15) were tested to infer three groups of significant genes. A probeset was classified as differentially expressed if the contrast of interest had an FDR q-value <0.05 (14) and the corresponding fold change of that contrast was >2, which is equivalent to testing significance at a level of ≈0.05 with known variance. For the same comparison, for example, YA52 versus WA52, different probesets would have different contrasts, depending on their best fit ANOVA model (i.e., one-, two-, or three-way; Fig. 1 and *SI Appendix*).

Three groups of significant probesets were classified: (*i*) Brg1-regulated retinal differentiation genes (Fig. 1; *SI Appendix*) include all of the probesets with one-way models that have a significant M coefficient, and two- and three-way models that have significant M-related interaction coefficients ($T*M$, $M*R$ *or* $T*M*R$). This group contains all of the probesets in which expression levels were affected by the Brg1 mutation (M). For two- and three-way models, a probeset would be inferred as differentially expressed if the corresponding contrast at either 36 or 52 hpf was statistically significant. (*ii*) Retina-specific genes independent of Brg1 regulation (*SI Appendix*) include probesets that are excluded by *i* and not affected by the Brg1 mutation (M), and that meet one of the following criteria: (*a*) one-way models with a significant T coefficient and two- and three-way models with significant T and $T*R$ interaction coefficients (this indicates a retina-specific expression with a significant temporal expression change), or (*b*) models with a significant intercept $I(\mu_g)+R$ contrast but not classified as significant by criterion (*a*) (these genes are specifically expressed in retina but the expression does not change over time). (*iii*) Brg1-regulated genes outside the retina (*SI Appendix*) are probesets with a significant M coefficient but an insignificant R coefficient. Such transcripts are affected by the Brg1 mutation but not differentially expressed in the retina.

One potential pitfall of the selection criteria for the third group of probesets is that the fold change cutoff might be too stringent for highly specific expression changes in a small tissue such as the otic

Table 1. A summary of significant probesets identified by the factorial analysis

	One-way models	Two-way models	Three-way models	Total
1. Brg1-regulated retinal differentiation probesets	100	437	194	731
2. Retinal specific probesets independent of Brg1 regulation				
a. With a temporal change	75	797	10	882
b. Without a temporal change	1,376	755	25	2,156
3. Brg1-regulated probesets outside retina*	28 (2,415)	40 (405)	30 (54)	98 (2,874)

*The numbers in parentheses are probesets selected by the same contrasts but not the fold change cutoff (see text for rationale)

vesicle, which could be masked by averaging mRNA expression levels within the whole embryo. As a result, the results for this group are selected only by the contrast q-values (Table S13, Table S14, and Table S15). The fold change cutoff can be lowered to obtain more potential candidate genes for a specific organ. For example, several genes that expressed in the otic vesicle (*sepm*, *foxi1*, *dlx4b*, *gata3*) and in the heart and vasculature (*jak1*, *igfbp1*, *jun*) as shown by the zfin *in situ* expression data (http://zfin.org) were identified by lowering the fold change cutoff to 1.4. These candidate genes might be specific Brg1 targets in those tissues.

A summary of the significant probesets selected by these criteria is shown in Table 1. Further temporal breakdown of the significant probesets is shown in Table S1, Table S2, and Table S3. The contrasts are listed in Fig. 1 and *SI Appendix*. The q-values and fold changes for the significant probesets are listed in Table S4, Table S5, and Table S6 for group 1; Table S7, Table S8, Table S9, Table S10, Table S11, and Table S12 for group 2; and Table S13, Table S14, and Table S15 for group 3. The average expression values for each condition for all of the probesets are listed in Table S16.

Overview of the Factorial Analysis Results. ***Brg1-regulated retinal differentiation genes.*** A group of 731 probesets showed significant *M*-related coefficients in the contrasts of one-, two-, or three-way ANOVA models, and was inferred as significantly regulated in the retina by Brg1 (Table 1, group 1; also see Table S1, Table S4, Table S5, and Table S6). Many of these probesets are higher two- or three-way models that contain the retinal tissue change factor in the interaction coefficients ($T*R$, $M*R$, and $T*M*R$). This suggests that these genes are specifically expressed in retina.

Among the 731 probesets analyzed, 199 were annotated genes and the remaining 532 were unannotated genes or expressed sequence tags (ESTs). Of the 199 annotated probesets, 52 were three-way models, 127 were two-way, and 20 were one-way (Table S1). These 199 known probesets were reviewed and clustered based on potential functions (Fig. 2). First, a group of genes potentially involved in processes directly related to terminal differentiation, including neurite outgrowth, cytoskeletal regulation, cellular adhesion, and synaptogenesis, were found to be suppressed in *yng* retinas. Second, because the *yng* mutation was shown to act in a non-cell autonomous manner by mosaic cell analysis (7), we looked for altered expression of extracellular signal genes but not their receptor genes. Delta-Notch and Fgf signaling pathways were identified based on this selection criterion. Third, we noted reduced expression of several transcription regulatory genes, including the *irx*, *tfap2*, and *id2* families. Finally, specific cell cycle genes, which

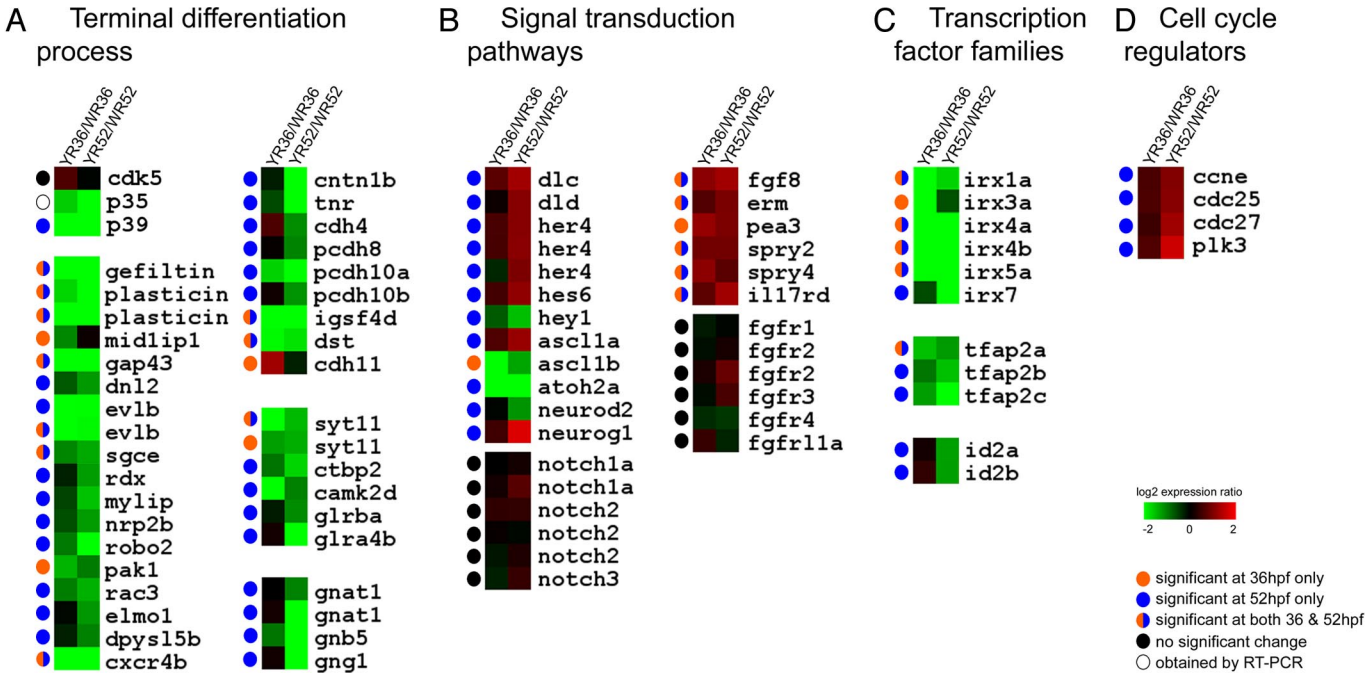


Fig. 2. Heat maps of four kinds of pathways or gene families that are controlled by the Brg1-regulated retinal differentiation program. The log2 YR36/WR36 and YR52/WR52 expression ratios of four kinds of pathways, or gene families, are shown in a heat map format. The timing of significance is shown by the colored circles on the left hand side of the heat maps (see *Factorial Analysis (Step 2)* for details on significance inference). Orange circle: probesets inferred as significant at 36 hpf only; blue circle: probesets inferred as significant at 52 hpf only; orange and blue circle: probesets inferred as significant at both 36 and 52 hpf; black circle: probesets that are not significantly changed; open circle: no probeset on the microarray, expression values obtained by RT-PCR. (*A*) Processes directly related to terminal differentiation: neurite outgrowth and cytoskeletol regulation (*Left*), adhesion molecules (*Top Right*), synaptogenesis (*Middle Right*), and phototransduction (*Bottom Right*). (*B*) Signal transduction pathways: Delta-Notch and Fgf. (*C*) Transcription factor families: *irx*, *tfap2*, and *id2*. (*D*) Specific cell cycle regulators. Several genes (*plasticin, evlb, syt11, gnat1, her4, notch1a, notch2*, and *fgfr2*) have multiple probesets on the arrays, and the respective results are shown in the heat maps.

have been implicated in progenitor cell maintenance and in preventing cell differentiation, were overexpressed in *yng* retinas.

Retinal-specific genes independent of Brg1 regulation. A group of 3,038 probesets was found specifically expressed in the retina but independent of Brg1 regulation. These had significant *R*-related coefficients but insignificant *M*-related coefficients (Table 1, group 2; Table S7, Table S8, Table S9, Table S10, Table S11, and Table S12), i.e., the expression levels between WT and *yng* retinas were not significantly different. Among them, 882 had a significant temporal change, whereas the remaining 2,156 did not. This group of probesets contain genes that are involved in processes that occur before or are independent of Brg1-regulated terminal differentiation. Several transcription factors that control both early- and late-born cell fate specifications were identified in this group (Table S18). A subset of these (*prox1*, *crx*, *vsx1*, and *foxn4*) had a time-dependent change, whereas others (*ath5*/*atoh7* and *bhlbh5*) did not. *Pax6a* and *pax6b*, genes that are responsible for specifying the whole eye field earlier in development were also identified. Further, these probesets were all from the two- or three-way models, suggesting their functions are specific to the retina.

Brg1-regulated genes outside the retina. Genes regulated by Brg1 outside the retina were also identified by the analysis. All models that contained significant *M*-related coefficients were selected by the relevant contrasts (*SI Appendix*). The probesets identified in group 1 (retina-specific regulation) were eliminated; thus, remaining probesets were regulated by Brg1 outside the retina. These criteria identified 107 probesets (Table 1, group 3; Table S13, Table S14, and Table S15). Although this process selects those probesets outside the retina, it does not mean that those specifically altered in the retina have no functional relevance in development elsewhere. A probeset deemed critical for Brg1-regulated retinal differentiation can have a different functional role in other tissues. For example, *brg1* was significantly altered in both the retina and whole embryo. This was expected because in *yng*, the mutated *brg1* would presumably undergo nonsense-mediated decay and its expression level would be lower compared to that of WT.

Functional Validations of Brg1-Regulated Retinal Differentiation Genes. ***Probesets directly related to terminal differentiation process.*** Several probesets listed as components of the axonal guidance pathway (hsa04360) in the Kyoto Encyclopedia of Genes and Genomes, including *pak1*, *rac3*, *robo2*, *cxcr4b*, and *dpysl5b*, were found to be significantly suppressed in the *yng* retinas (Fig. 2A). Although the expression of cdk5, a key player in the neurite outgrowth and differentiation (16), was not altered in *yng* retinas, its activators *p39*/*cdk5r2* were significantly suppressed. Three neurite pathfinding signals, *gap43*, *rdx*, and *nrp2b*, were also suppressed in *yng* retinas as well as probesets that control various cytoskeletal dynamics including *mid1ip1*, *dnl2*, *elmo1*, *evlb*, *gelfiltin*, *plasticin*, *mylip*, and *sgce*. They were identified in higher (two- and three-way) models (Table S5 and Table S6), suggesting that the deregulation was retina specific.

Regulatory molecules that guide neurite outgrowth include cell adhesion molecules. In *yng* retinas, several adhesion molecule genes including *cntn1b*, *tnr*, *cdh4*, *pcdh8*, *pcdh10a*, *pcdh10b*, *igsf4d*, and *dst* were suppressed, whereas *cdh11* was overexpressed. Several genes involved in synaptogenesis were suppressed, including *syt11*, *ctbp2*, *camk2d*, *glrba*, and *glra4b*. Finally, genes involved in phototransduction (*gnat1*, *gnb5*, and *gng1*) were also diminished.

Cdk5 and its regulators. As noted above, *cdk5* was not altered in expression, but its activator *p39* was significantly suppressed at 52 hpf. Its isoform, *p35*/*cdk5r1*, not on the Affymetrix array, was also suppressed in *yng* retinas by more than >2-fold as shown by RT-PCR assays (data not shown). We chose these regulators of cdk5 to assess the functional significance of loss of cdk5 activity on retinal cell differentiation. Knockdown of *p35* and *p39* by antisense morpholinos (MO) affected the differentiation of the entire retina, but especially outer plexiform layer (OPL) development (Fig. 3).

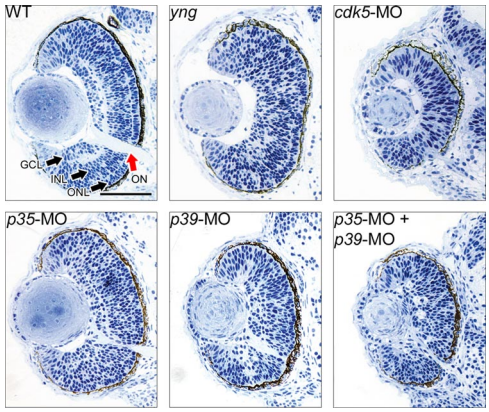


Fig. 3. Characterization of Cdk5, p35, and p39's functions in zebrafish retinal development. Retinal histology of antisense morpholino knockdown experiments at 52 hpf. *Top* (left to right): Uninjected control, *yng* mutant, *cdk5*-SMO (0.56 ng). *Bottom* (left to right): *p35*-MO (8 ng), *p39*-MO (6 ng), *p35*-MO and *p39*-MO (5.1 ng each). The black arrows indicate the location of nuclear layers GCL, INL, and ONL, while the red arrow indicates the optic nerve (ON).

Morphant embryos displayed some differentiation in the ganglion cell layer (GCL) and inner plexiform layer (IPL), but the optic nerve was reduced in size. The knockdown effect is specific to the retina, as the gross morphology of these morphants was normal (Fig. S2). The phenotype of the double-knockdown of *p35* and *p39* using a lower dose of each MO gave rise to a more severe phenotype that included disruption of the GCL and optic nerve. These data suggest an additive effect of the two activators on cdk5 activity and retinal cell differentiation. This was supported by knockdown of *cdk5*, which completely eliminated retinal lamination and formation of the optic nerve. Eye size was also significantly reduced, but gross morphology was relatively normal. The spatial expression patterns of *p35* and *p39* and their catalytic partner *cdk5* highly overlap the expression of *brg1* and *erk1* (a MAPK) in WT embryos at 52 hpf (data not shown). These results suggest that *p35* and *p39* are components of a Brg1-Erk1 differentiation pathway.

Delta-Notch and Fgf signal transduction pathways. Delta-Notch is a juxtacrine signal transduction pathway in which a cell with Notch receptors receives a Delta signal from neighbors. This activates the transcription of *hes* and *her* transcriptional repressors, which inhibit activation of proneural genes controlling differentiation (17). Overexpression of *delta* and activation of the Notch pathway thus prohibits cell differentiation. The analysis shows the probesets of several key components of the Notch pathway were deregulated in a fashion consistent with this view (Fig. 2*B*). For example, *dlc* and *dld* were significantly overexpressed, whereas the expression of *notch1a*, *notch2*, and *notch3* (the Notch receptors on the arrays) was not significantly changed. These observations are consistent with the non-cell autonomous behavior of the *yng* mutation. In addition, expression of several *hes* transcription repressors was affected, including overexpression of *her4* and *hes6*, and underexpression of *hey1*. The expression of several downstream proneural genes, including *ascl1b*, *atoh2a*, and *neuord2*, was significantly suppressed, whereas the expression of *acsl1a* and *neurog1* was significantly overexpressed. These results suggest the Notch pathway is activated in *yng* at 52 hpf and contributes to the block of retinal differentiation.

In the classic Fgf signal transduction pathway, extracellular Fgf activates a receptor tyrosine kinase (RTK), which progressively activates protein kinases including RAS, RAF, and MAPK. Ultimately, Fgf activates the transcription of several genes including *erm*, *pea3*, *sprouty(spry)* and *sef* (18). Our analysis shows probesets of several key components of this pathway were deregulated in a direction predicted by the genetic mosaic experiments (7) (Fig. 2*B*).

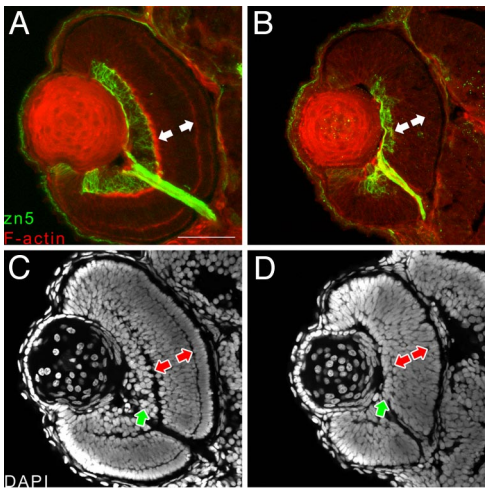


Fig. 4. *Irx7*-MO knockdown experiment. Immunohistochemistry of uninjected control (*Left*) and irx7-MO (10 ng)-injected eyes (*Right*) at 52 hpf. The phalloidin F-actin stain (red) preferentially bound to the INL and ONL (white arrows) in the WT retina, but was severely reduced or absent, respectively, in *irx7* morphants. The differentiation of RGC, as shown by the zn5 antibody (green), was compromised but not completely eliminated in the *irx7* morphants. The DAPI nuclei staining of the control and *irx7* morphants also shows the same adverse effects on the plexiform layers (red arrows) and GCL (green arrows).

Thus, *fgf8* was significantly overexpressed in *yng* retinas, while *fgfr1*, *fgfr2*, *fgfr3*, *fgfr4*, and *fgfrl1a* (the Fgf receptors on the arrays) were not significantly changed. Many downstream effectors of Fgf signaling were significantly overexpressed in *yng* retinas, including *erm*, *pea3*, *spry2*, *spry4* and *il17rd*(*sef*). Among them, *spry4* and *il17rd* were identified in the one-way model, indicating they were affected in whole *yng* embryos.

Transcriptional regulator families: irx, tfap2, and id2. All *irx* probesets present on the microarrays (*irx1a*, *3a*, *4a*, *4b*, *5a* and *7*), except *irx1b*, were strongly suppressed in *yng* retinas (Fig. 2*C*). Interestingly, *irx3a* and *irx7* were suppressed only at 36 or 52 hpf, respectively. *Irx1b*, on the other hand, was not expressed in either WT or *yng* retinas. *Irx* genes are transcription factors that activate proneural genes and are involved in many neuronal patterning processes (19–21).

Although many *irx* genes are expressed in the zebrafish retina, especially in the GCL (22)(http://zfin.org), their roles in retinal differentiation are largely unknown. Interestingly, *irx7* is specifically expressed in the INL but is not present in the retina before the INL forms (≈48 hpf). Because GCs differentiate by 36 hpf (13), most *irx* genes may be involved in the formation of GCs, their axonal projections, or the IPL, whereas *irx7* may play a role in the formation of outer retina. This was tested by an *irx7*-MO knockdown experiment (Fig. 4). In *irx7* morphants, the OPL was absent and the IPL was severely reduced. Interestingly, the differentiation of GCs in the *irx7* morphant appeared compromised, as the zn5-staining domain was smaller, along with the corresponding GCL. Also, the eye size of the *irx7* morphant was smaller. Cheng *et al*. have also found that knockdown of *irx1a* in zebrafish compromises retinal cell differentiation and lamination (23).

Additional transcription factors including probesets for *tfap2* (*tfap2a*, *tfap2b*, and *tfap2c*) and *id2* (*id2a* and *id2b*) were strongly suppressed in *yng* retinas (Fig. 2*C*). All were down-regulated at 52 hpf except *tfap2a*, which was suppressed at both 36 and 52 hpf. These factors generally control genes involved in the balance between proliferation and differentiation of many cell types, including those of the neural retina (24–26). Many (*tfap2a*, *tfap2b*, *tfap2c*, and *id2*) have been shown to express in specific regions of the retina (24, 27, 28).

Deregulation of cell cycle control. Several cell cycle genes were significantly overexpressed in *yng* retinas at 52 hpf compared to the WT, including *ccne*, *cdc25*, *cdc27*, and *plk3* (Fig. 2*D*). Further, these genes were all identified in the two-way models, suggesting these changes were retina specific.

In the cell cycle, Ccne/Cdk2 regulates G1-S phase transition, whereas Cdc25 regulates both G1-S and G2-M transitions (6). Cdc27 is part of the anaphase promoting complex (APC), which controls late M phase progression. Plk3 phosphorylates and activates Cdc25, thus stabilizing cell cycle progression (29). Up-regulation of these genes in *yng* retinas may retain cells in the cell cycle and/or prevent normal cell cycle withdrawal. Consistent with this view, previous work demonstrated a cell cycle withdrawal delay in *yng* (7).

Discussion

Factorial Analysis Can Effectively Identify Relevant Genes That Are Affected by Multiple Factors. Factorial microarray design is ideal for investigating the effects of several biological factors on gene expression levels with a minimal number of microarray experiments. Traditionally, microarray expression studies have focused on investigating one factor alone or several factors separately. Factorial analysis has several comparative advantages. First, it requires fewer replicates to get the same precision in effect estimation. Second, it gives proper estimation of the effects of factors, especially when they interact, because usage of contrast effectively handles the situation. Third, the ANOVA model grouping (i.e., one-, two-, and three-way models) is an ordinal categorization of significance. Higher models with increased levels of interaction coefficients suggest a more specific regulation of these particular probesets in the tissue of interest. This helps in the prioritization of subsequent gene characterization. Fourth, factorial analysis gives proper prediction for the conditions that are not included in an experiment because it conducts a systematic or comprehensive search over the experimental region.

ANOVA analysis has been used in several microarray studies (30–32) to account for the variations in gene expressions caused by different sources, and normalization was included as part of their models. Our analysis has separated the normalization procedure from significance inference, which identifies differentially expressed genes more reliably (33). In addition, contrast analysis is used in our significance inference. With false discovery rate control, they make our method more rigorous for detecting differentially expressed genes.

The major limitation of this study is that expression profiling methodology measures average gene expression levels, whereas the developing retina consists of different kinds of retinal progenitor cells at different stages of the cell cycle and differentiation. The "average" picture may not reflect the changes occurring in individual cells. Nonetheless, the use of whole retinal tissue was suitable for this study because the terminal differentiation defects in *yng* affect all retinal cells. Indeed, the initial functional characterization experiments on *cdk5*/*p35*/*p39* and *irx7* have demonstrated that an average expression profile of the whole retina can reveal insights for retinal development.

Biological Validation. So far, we have analyzed two gene groups that were identified as relevant for terminal differentiation. They include genes that control neurite outgrowth, cell migration, and cytoskeletal regulation in postmitotic neurons, as well as the Irx transcription factor family. We performed initial characterizations on *cdk5*/*p35*/*p39* and *irx7* and demonstrated their functional relevance. Our Cdk5 data, along with other published studies (34), suggest that Erk1 acts upstream of *Cdk5*/*p35*, which then functions to control aspects of terminal differentiation including neurite outgrowth and synaptogenesis.

Materials and Methods

Fish. Zebrafish [AB wild-type (WT) and *young (yng)* A50 (7)] were maintained according to standard procedures (35). All of our protocols were approved by the Harvard Standing Committee on the Use of Animals in Research and Teaching.

Egg Collection, Embryo Staging, and Collection. Embryos were collected and staged according to an optimized procedures for zebrafish retinal expression profiling (10). *Yng* embryos were staged by using their WT siblings collected at the same time.

Retinal Dissection, Sample Collection, Total RNA Extraction, and Quality Assessment. One retina was dissected as described (10), except for *yng* retinas at 36 hpf (YR36), in which case both retinas were dissected. Equal numbers of *yng* embryos were used for each YR36 replicate so genetic variability was maintained. Total RNA extraction and quality control were done as described (10).

Complementary DNA Library Preparation. Complementary DNA (cDNA) library was prepared by reverse transcribing messenger RNAs from a total RNA preparation of 2-day-old zebrafish embryos, using SuperScript II reverse transcriptase (Invitrogen) and an anchored primer 5′-dT_{20}VN-3′ (Integrated DNA Technologies).

Microarray Experiment. Gene expression levels of retinal and whole embryo samples were determined by using Affymetrix Zebrafish Genome Arrays as described (10), in a $2 \times 2 \times 2$ factorial design (Fig. 1). Three replicates were performed for each of the three conditions; the total number of microarrays used was 24.

Microarray Analysis. The probe-level data were background adjusted, normalized, and summarized by a dChip (http://www.dchip.org) algorithm using default parameters (36, 37). The fold changes of gene expression between two experimental conditions were calculated by using dChip. The estimation of contrasts and significance inference were performed by using the factDesign package of the Bioconductor project release 1.7 (http://www.bioconductor.org), and *q*-values and other statistical analyses were conducted in R statistical environment version 2.2.0 (http://www.r-project.org). Details of the model fitting and significance inference are described in the *Results* section. Heatmaps were generated using Multiexperiment Viewer (MeV) of the TM4 microarray software suite (http://www.tm4.org/). The whole dataset is available at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo/, accession number GSE8874).

Histology. Histological analyses were performed by using standard procedures (38). Five to six embryos were analyzed for each condition.

Morpholino Antisense Knockdown. Custom morpholino (MO) antisense oligomers were purchased from GeneTools or Open Biosystems. The sequences of *cdk5*, *p35*, and *p39* MOs and the dosage used in this study are listed in Table S17. The *irx7*-MO was reported in (39), and 10 ng was injected per embryo. A standard control MO was obtained from GeneTools. The MOs were microinjected to one-cell stage WT zebrafish embryos by using standard procedures (40). At least 50 embryos were injected for each analysis.

Immunohistochemistry. Immunohistochemical analysis was performed on 10-μm-thick cryosections as described (41). Antibodies and their dilutions were: mouse anti-zn5 (zn5) (1:500; University of Oregon) and Alexa Fluor 488 goat anti-mouse IgG (1:1,000; Invitrogen). Confocal images were acquired by using a Zeiss LSM510 META on an inverted microscope (Zeiss). A Z-stack was obtained for zn5 immunostaining, and the resulting images were projected and merged together. In images with DAPI and Phallodin as background stains, only one representative plane was imaged. These images were subsequently processed and merged by using Adobe Photoshop 6.0 (Adobe Systems).

ACKNOWLEDGMENTS. We thank Ellen A. Schmitt for her advice on animal staging and microdissection procedures; Jennifer Couget and ShuFen Meng for their help with the Affymetrix GeneChip experiments; Daisuke Kojima and Jeff Gross for their technical advice on microinjection and *in situ* hybridization; Sophie Caron for providing the *irx7*-MO; Jeff Gross, Leanne Godinho Misgeld, Jennifer O'Brien, Wei Pan, GuoCheng Yuan, and Donald Zack for a critical reading of the manuscript; and members from Dowling lab for helpful discussions. This work was supported by fellowships from the Croucher Foundation and Knight's Templar Foundation (to Y.F.L.), a grant from the Merck Award for Genomics Research (to Y.F.L. and J.E.D.), National Eye Institute Grant EY000811 (to J.E.D.), and National Science Foundation Grant DMS-0800631 (to P.M.).

1. Leung YF, Cavalieri D (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet* 19:649–659.
2. Montgomery DC (2005) *Design and Analysis of Experiments* (Wiley, Hoboken, NJ) 6th Ed.
3. Wu C-F, Hamada M (2000) *Experiments: Planning, Analysis, and Parameter Design Optimization* (Wiley, New York).
4. Livesey FJ, Cepko CL (2001) Vertebrate neural cell-fate determination: Lessons from the retina. *Nat Rev Neurosci* 2:109–118.
5. Agathocleous M, Harris WA (2006) Cell determination. *Retinal Development*, eds Sernagor E, Eglen S, Harris B, Wong R (Cambridge Univ Press, Cambridge), pp 75–98.
6. Ohnuma S, Harris WA (2003) Neurogenesis and the cell cycle. *Neuron* 40:199–208.
7. Link BA, Fadool JM, Malicki J, Dowling JE (2000) The zebrafish young mutation acts non-cell-autonomously to uncouple differentiation from specification for all retinal cells. *Development* 127:2177–2188.
8. Roberts CW, Orkin SH (2004) The SWI/SNF complex–chromatin and cancer. *Nat Rev Cancer* 4:133–142.
9. Gregg RG, Willer GB, Fadool JM, Dowling JE, Link BA (2003) Positional cloning of the young mutation identifies an essential role for the Brahma chromatin remodeling complex in mediating retinal cell differentiation. *Proc Natl Acad Sci USA* 100:6535–6540.
10. Leung YF, Dowling JE (2005) Gene expression profiling of zebrafish embryonic retina. *Zebrafish* 2:269–283.
11. Schmitt EA, Dowling JE (1999) Early retinal development in the zebrafish, *Danio rerio*: Light and electron microscopic analyses. *J Comp Neurol* 404:515–536.
12. Neumann CJ, Nuesslein-Volhard C (2000) Patterning of the zebrafish retina by a wave of sonic hedgehog activity. *Science* 289:2137–2139.
13. Schmitt EA, Dowling JE (1996) Comparison of topographical patterns of ganglion and photoreceptor cell differentiation in the retina of the zebrafish, *Danio rerio*. *J Comp Neurol* 371:222–234.
14. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440–9445.
15. Neter J, Kutner MH, Wasserman W, Nachtsheim CJ (1996) *Applied Linear Statistical Models* (Irwin, Chicago).
16. Dhavan R, Tsai LH (2001) A decade of CDK5. *Nat Rev Mol Cell Biol* 2:749–759.
17. Hatakeyama J, Kageyama R (2004) Retinal cell fate determination and bHLH factors. *Semin Cell Dev Biol* 15:83–89.
18. Tsang M, Dawid IB (2004) Promotion and attenuation of FGF signaling through the Ras-MAPK pathway. *Sci STKE* 2004(228):pe17.
19. Cavodeassi F, Diez Del Corral R, Campuzano S, Dominguez M (1999) Compartments and organising boundaries in the Drosophila eye: The role of the homeodomain Iroquois proteins. *Development* 126:4933–4942.
20. Kobayashi D, *et al.* (2002) Early subdivisions in the neural plate define distinct competence for inductive signals. *Development* 129:83–93.
21. Briscoe J, Pierani A, Jessell TM, Ericson J (2000) A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* 101(4):435–445.
22. Lecaudey V, Anselme I, Dildrop R, Ruther U, Schneider-Maunoury S (2005) Expression of the zebrafish Iroquois genes during early nervous system formation and patterning. *J Comp Neurol* 492:289–302.
23. Cheng CW, Yan CH, Hui CC, Strahle U, Cheng SH (2006) The homeobox gene *irx1a* is required for the propagation of the neurogenic waves in the zebrafish retina. *Mech Dev* 123(3):252–263.
24. Eckert D, Buhl S, Weber S, Jager R, Schorle H (2005) The AP-2 family of transcription factors. *Genome Biol* 6(13):246.
25. Chong SW, Nguyen TT, ChuLT, Jiang YJ, Korzh V (2005) Zebrafish id2 developmental expression pattern contains evolutionary conserved and species-specific characteristics. *Dev Dyn* 234:1055–1063.
26. Pujic Z, *et al.* (2006) Reverse genetic analysis of neurogenesis in the zebrafish retina. *Dev Biol* 293(2):330–347.
27. Luo T, *et al.* (2005) Regulatory targets for transcription factor AP2 in Xenopus embryos. *Dev Growth Differ* 47(6):403–413.
28. Li W, Cornell RA (2007) Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev Biol* 304(1):338–354.
29. Myer DL, Bahassi el M, Stambrook PJ (2005) The Plk3-Cdc25 circuit. *Oncogene* 24(2):299–305.
30. Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. *J Comput Biol* 7:819–837.
31. Jin W, *et al.* (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* 29:389–395.
32. Wolfinger RD, *et al.* (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *J Comput Biol* 8(6):625–637.
33. Yang YH, Speed T (2002) Design issues for cDNA microarray experiments. *Nat Rev Genet* 3:579–588.
34. Harada T, Morooka T, Ogawa S, Nishida E (2001) ERK induces p35, a neuron-specific activator of Cdk5, through induction of Egr1. *Nat Cell Biol* 3(5):453–459.
35. Westerfield M (2000) *The Zebrafish Book:A Guide for the Laboratory Use of Zebrafish (Danio rerio)* (University of Oregon Press, Eugene) 4th Ed.
36. Li C, Wong WH (2001) Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc Natl Acad Sci USA* 98:31–36.
37. Li C, Wong W (2001) Model-based analysis of oligonucleotide arrays: Model validation, design issues and standard error application. *Genome Biol* 2(8):RESEARCH0032.
38. Leung YF, Ma P, Dowling JE (2007) Gene expression profiling of zebrafish embryonic retinal pigment epithelium in vivo. *Invest Ophthalmol Vis Sci* 48:881–890.
39. Itoh M, Kudoh T, Dedekian M, Kim CH, Chitnis AB (2002) A role for iro1 and iro7 in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. *Development* 129:2317–2327.
40. Nusslein-Volhard C, Dahm R eds (2002) *Zebrafish: A Practical Approach* (Oxford Univ Press, Oxford UK), 1st Ed.
41. Godinho L, *et al.* (2005) Targeting of amacrine cell neurites to appropriate synaptic laminae in the developing zebrafish retina. *Development* 132:5069–5079.