

**Transcript profiling of the cone-only Nrl-knockout retina using custom cDNA-
microarrays: identification of novel targets of Nrl and of signaling pathways
involved in photoreceptor function**

by

Jindan Yu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in The University of Michigan
2004

Doctoral Committee:

Professor Anand Swaroop, Co-Chair
Assistant Professor Alan J. Hunt, Co-Chair
Associate Professor Thomas M. Glaser
Professor Alfred O. Hero
Professor Peter F. Hitchcock

To my mother and my husband

ACKNOWLEDGMENTS

Many people helped me throughout my degree. My dissertation committee, including Drs. Anand Swaroop, Alan J. Hunt, Tom M. Glaser, Alfred O. Hero and Peter F. Hitchcock, has continuously supported and advised me over the last couple years. I am especially indebted to my thesis advisor, Dr. Anand Swarop, for his guidance in both my professional development and personal growth.

I am grateful to past and present members of Swaroop lab for providing me a wonderful environment. Particularly, Rafal Farjo was involved in sequencing and annotating of 10,000 clones and printing microarray slides. James S. Friedman and Shirley He performed the chromatin immunoprecipitation assay. James also commented on the writing of this dissertation. Sean P. MacNee picked up thousands of clones and conducted PCR amplification for printing microarray slides. Mohammad I. Othman managed the microarray facility and helped in optimizing microarray hybridization protocols. Sepideh Zarepari was involved in testing the effects of starting amounts of total RNA on microarray hybridization and provided helpful discussion on this dissertation. My undergraduate student, Rusha Patel, carried out some of the real-time PCR experiments. I also thank Shigeo Yoshida for teaching me molecular biology and Alan J. Mears for valuable discussions.

In addition, thanks are due to many other friends and colleagues, including Brian J. Akerley, Deborah M Eadie, Debashis Ghosh, David Hicks, Dider Job, David H. Kohn, Mirna Mustapha, Donald G. Puro, Terry Speed, Maria E. Steele and Sandra A. Staneff.

Lastly, I am grateful to my husband, Chang-Sheng, for being incredibly supportive throughout. I thank my late parents, my sister and brother for their love. I have also been very fortunate to have my mother-in-law's help in taking care of my baby daughter. I want to express my extreme gratitude to her for the love she has extended to me. I also thank Emily for being such a sweet daughter.

TABLE OF CONTENTS

DEDICATION.....	ii
ACKNOWLEDGMENTS.....	iii
LIST OF FIGURES.....	vii
LIST OF TABLES.....	ix
CHAPTER	
1. INTRODUCTION TO THE STUDY.....	1
1.1 Problem Statement.....	1
1.2 Purpose of Study.....	2
1.3 Specific Aims.....	2
1.4 Dissertation Outline.....	3
1.5 References.....	10
2. REVIEW OF THE LITERATURE ON PHOTORECEPTOR DIFFERENTIATION AND FUNCTION.....	14
2.1 Eye Anatomy and Structure.....	14
2.2 Photoreceptor Function.....	17
2.3 Retinal Development and Photoreceptor Differentiation.....	25
2.4 Nrl and the Nrl Knockout.....	26
2.5 Photoreceptor Degeneration and Cone Survival.....	30
2.6 Microarray Technology.....	30
2.7 Summary.....	33
2.8 References.....	34
3. ANNOTATION AND ANALYSIS OF 10,000 ESTS FROM DEVELOPING MOUSE EYE AND ADULT RETINA.....	39
3.1 Abstract.....	39
3.2 Introduction.....	40
3.3 Materials and Methods.....	42
3.4 Results.....	45
3.5 Discussion.....	64
3.6 References.....	70
4. GENERATION OF EYE-GENE MICROARRAYS.....	73
4.1 Abstract.....	73
4.2 Introduction.....	73
4.3 Methods.....	75

	4.4 Results.....	79
	4.5 Discussion.....	86
	4.6 References.....	88
5.	EVALUATION AND OPTIMIZATION OF PROCEDURES FOR TARGET LABELING AND HYBRIDIZATION OF CDNA MICROARRAYS.....	91
	5.1 Abstract.....	91
	5.2 Introduction.....	92
	5.3 Methods.....	93
	5.4 Results and Discussion.....	99
	5.5 References.....	107
6.	ALTERED EXPRESSION OF GENES OF THE BMP/SMAD AND WNT/CALCIUM SIGNALING PATHWAYS IN THE CONE-ONLY NRL-KNOCKOUT MOUSE RETINA, REVEALED BY GENE PROFILING USING CUSTOM CDNA MICROARRAYS.....	110
	6.1 Summary.....	110
	6.2 Introduction.....	111
	6.3 Materials and Methods.....	114
	6.4 Results.....	118
	6.5 Discussion.....	133
	6.6 References.....	138
7.	GENE EXPRESSION PROFILES OF RETINA DURING POSTNATAL DEVELOPMENT AND COMPARISON TO THE NRL^{-/-} RETINA.....	144
	7.1 Summary.....	144
	7.2 Introduction.....	145
	7.3 Materials and Methods.....	147
	7.4 Results.....	150
	7.5 Discussion.....	160
	7.6 Reference.....	165
8.	CONCLUSIONS.....	170
	8.1 Concluding Remarks.....	170
	8.2 Implications.....	171
	8.3 Future Directions.....	172
	8.4 References.....	173

LIST OF FIGURES

Figures

Figure 2.1. The anatomy and structure of the adult human eye.....	15
Figure 2.2. The structure of the retina.....	16
Figure 2.3. The structure of photoreceptors.....	18
Figure 2.4. The phototransduction cascade.....	20
Figure 2.5. Retinoid cycles.	22
Figure 2.6. The retinal cellular pathways.....	24
Figure 2.7. A model for photoreceptor differentiation.	27
Figure 2.8. Opsin immunohistochemistry of the wild-type (+/+), Nrl heterozygous (+/-), and Nrl knockout (-/-) retina.	29
Figure 2.9. Schematic overview of cDNA microarray technology.	32
Figure 3.1. Schematic representation of the EST analysis and annotation processes.....	48
Figure 3.2. Screen shot of I-gene database of annotated ESTs.....	50
Figure 3.3. Histograms showing the number of genes or ESTs at each level of redundancy in the three libraries.....	53
Figure 3.4. Functional categorization of ESTs with known-gene match.....	55
Figure 3.5. qRT-PCR validation of <i>in silico</i> expression profiles.....	65
Figure 4.1. A schematic representation of the microarray strategy.....	80
Figure 4.2. Pie charts representing different clone types identified in our mouse eye libraries.....	82
Figure 4.3. A typical microarray experiment with identical target RNAs.....	84

Figure 4.4. Scatter plot analysis of microarray hybridization experiments with identical target RNAs.	85
Figure 5.1. Target labeling procedures.	96
Figure 5.2. False color overlaid images.	102
Figure 5.3. Scatter plot of self-self hybridization.	103
Figure 5.4. Scatter plots of two targets hybridization.	104
Figure 5.5. Signal-to-noise ratio (SNR) for serial hybridization.	106
Figure 6.1. Representative scatter-plots showing differential expression of two copies each of S-opsin (circled spots above 0 axis) and Rhodospin (circled spots below 0 axis).	120
Figure 6.2. Differential expression of three genes: <i>Gnb3</i> , <i>Gnb1</i> and MRA-1648.	127
Figure 6.3. Expression analysis by qRT-PCR.	130
Figure 6.4. Chromatin immunoprecipitation (ChIP).	132
Figure 6.5. A schematic representation of signaling molecules and corresponding pathways altered in the <i>Nrl</i> ^{-/-} mouse retina.	136
Figure 7.1. Gene expression profiling of the developing wild-type and <i>Nrl</i> ^{-/-} retinas.	151
Figure 7.2. Time course of genes with altered expression during the development of the wild-type or <i>Nrl</i> ^{-/-} retinas.	154
Figure 7.3. Time course of genes differentially expressed in the wild-type and <i>Nrl</i> ^{-/-} retinas.	156
Figure 7.4. ChIP analyses of selected genes.	158
Figure 7.5. qRT-PCR validation of microarray data.	159

LIST OF TABLES

Tables

Table 3.1. Summary of ESTs.	46
Table 3.2. Highly expressed genes in the M15E, M2PN and MRA libraries, excluding crystallin genes.	51
Table 3.3. Ocular gene distribution in mouse and human chromosomes.	57
Table 3.4. Candidate ocular disease genes identified in three mouse retina / eye cDNA libraries.	58
Table 3.5. <i>In silico</i> temporal expression profiles of selected ESTs.	60
Table 3.6. Retina/eye-enriched genes, excluding crystallin and phototransduction genes.	62
Table 3.7. Retina/eye-enriched unknown ESTs.	63
Table 6.1. Putative NRE and the PCR primers used for ChIP enrichment tests.	119
Table 6.2. The number of genes commonly identified by both direct and indirect comparisons as having the highest odds of differential expression in the top 40, 50 and 70 lists.	122
Table 6.3. The top 50 differentially expressed genes. Microarray and qRT-PCR columns show fold-change, where (-) and (+) indicate lower and higher expression in $Nrl^{-/-}$ retina, respectively.	124
Table 6.4. The top 24 differentially expressed unknown genes / ESTs. Microarray and qRT-PCR columns show fold-change, where (-) and (+) indicate lower and higher expression in $Nrl^{-/-}$ retina, respectively.	128

CHAPTER 1

INTRODUCTION TO THE STUDY

1.1 Problem Statement

Degeneration of photoreceptors, rods and cones, is the primary cause of blindness in retinal dystrophies (RDs), including retinitis pigmentosa (RP), age-related macular degeneration (AMD) and cone-rod dystrophies (CRD). RP has a prevalence rate of 1/3000, and AMD is the leading cause of visual impairment in the elderly population (Klein et al. 1995; Saleem and Walter 2002). Both diseases are characterized by initial rod degeneration, followed by the loss of cone function (Dryja and Li 1995; Hageman et al. 2001). In contrast, CRDs often have early onset of cone dysfunction with subsequent loss of rod-mediated vision (Szlyk et al. 1993). To date, mutations in over 130 genes have been associated with RDs (Retnet, <http://www.sph.uth.tmc.edu/Retnet>). However, little is known about their etiology, due partly to the mixed effects of rod and cone degeneration. In order to investigate the pathways involved in rod and cone survival and degeneration, it is essential to understand the molecular differences between the two types of photoreceptors during development and functional maintenance (homeostasis).

Elucidation of rod versus cone differences at the molecular level has long been hindered, as cones constitute only 3-5% of photoreceptors in most mammals. Recently, a cone-only mouse mutant deficient in Nrl (Neural retinal leucine zipper), the $Nrl^{-/-}$, has been generated (Mears et al. 2001). Nrl is a basic motif-leucine zipper transcription factor that is expressed specifically in rod photoreceptor cells and regulates rod-specific gene expression; mutations in the human NRL gene are associated with autosomal

dominant RP (Bessant et al. 1999; DeAngelis et al. 2002; Mitton et al. 2000; Swain et al. 2001). The deletion of *Nrl* in mice leads to morphological and functional switch of rods to cones; the mutant retina contains no rods and a greater than normal number of cones (Mears et al. 2001). This mouse model, together with the rod-rich wild-type, provides an ideal tool for comparative investigation of rod and cone differentiation and function.

1.2 Purpose of Study

Cellular activities are defined by their gene expression profiles, i.e., transcriptomes (the set of genes expressed within a specific cell). An analysis of transcriptional differences between wild-type and *Nrl*^{-/-} mouse retinas may point to the underlying mechanisms of rod- versus cone-mediated visual functions, some of which may be directly regulated by *Nrl*. The goals of this thesis were: (1) to identify genes whose expression is regulated by *Nrl*; (2) to reveal rod- or cone-specific molecules and pathways; and (3) to elucidate the role of *Nrl* in these pathways. The results should assist in further delineating the regulatory networks controlled by *Nrl* and facilitate a better understanding of RDs in the context of rod and/or cone function.

1.3 Specific Aims

To achieve the goals of this study, a comprehensive approach combining molecular biology, genomics, bioinformatics and statistics was undertaken. A major component of this study was microarray analysis, a high-throughput method that enables researchers to examine gene expression on a genome scale (Spellman et al. 1998). In a typical microarray experiment, microarray slide containing genes of interest is hybridized simultaneously by two cDNA targets prepared from two different RNA sources. The difference in transcript levels of each gene in each source is based on their hybridization strength (Ramsay 1998; Schena et al. 1998; Schena et al. 1995). Since eye-expressed

genes (eye-genes) were under-represented on commercially available microarrays, it was necessary to produce custom microarrays containing consolidated eye-genes. We used custom cDNA microarrays of eye-expressed genes for comparative transcriptional profiling of wild-type and *Nrl*^{-/-} retinas. The five specific aims of this study were:

- Aim 1. To generate and annotate expressed sequence tags (ESTs) from cDNA libraries constructed from eye tissues.**
- Aim 2. To produce custom microarrays (I-gene microarrays) containing cDNAs of eye-genes.**
- Aim 3. To determine optimal conditions for microarray experimentation.**
- Aim 4. To identify differential gene expression between mature wild-type and *Nrl*^{-/-} retinas and discover rod- or cone-specific pathways.**
- Aim 5. To define gene expression patterns during development and delineate gene expression changes in the developing *Nrl*^{-/-} retina.**

Specific aims 1 through 3 made it possible to efficiently examine the expression levels of thousands of eye-genes. In specific aim 4, comparison of transcripts in two mouse models revealed a set of genes differentially expressed in rods and cones. This helped achieve the goals of identifying *Nrl* target genes and rod- or cone-specific molecules and pathways. Examination of genes and pathways during development, in specific aim 5, further elucidated their roles in rod differentiation and their regulation by *Nrl*.

1.4 Dissertation Outline

This dissertation contains 8 chapters. Literature review in Chapter 2 is followed by five chapters, corresponding to the five specific aims. Chapter 8 summarizes the overall project and provides concluding remarks. A brief summary of each chapter is

given below. The structure of summary for Chapters 3-7 is as follows: the research question, the goal, the primary results, and the significance.

Chapter 2: Review of literature on photoreceptor differentiation and function.

This chapter begins by providing background information about the vertebrate eye, retina and photoreceptors. It then analyzes published reports on photoreceptor development, including *Nrl* and *Nrl*^{-/-}, and photoreceptor degeneration. Emphasis is placed on the differences between rods and cones. This chapter ends with a review of microarray technology and its application in understanding regulatory networks.

Chapter 3: Annotation and analysis of 10,000 ESTs from developing mouse eyes and adult retinas.

To investigate retinal gene expression under various conditions, it is necessary to produce a thorough catalogue of eye- and/or retina-expressed genes. When this study began, such a task was far from complete. Retinal cDNAs, especially those expressed at embryonic or early postnatal ages, were greatly under-represented in public EST databases. Only 30,000 out of 1.8 million mouse cDNA sequences (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>) were isolated from retinas. Of these, approximately half were derived from adult retinas. This under-representation of mouse retinal cDNAs had greatly hampered large-scale study of retinal gene expression. The goal of this chapter was to generate a comprehensive set of eye-expressed ESTs.

Three cDNA libraries were constructed from embryonic day 15.5 eyes, postnatal day 2 eyes and adult retinas, respectively, corresponding to 3 important stages of development. Over 10,000 clones were isolated and sequenced. Annotation and analysis of 8,633 high-quality ESTs, corresponding to 7,019 unique gene clusters, revealed 57% known genes and 43% unknown or novel ESTs. Of these, 210 were enriched in eye tissues and many showed varying transcript levels during eye development.

The annotated EST sets help define the transcriptomes of the eye, and are useful for functional genomic studies. These ESTs can be analyzed by computational approaches to identify additional candidate genes for ocular diseases (Bortoluzzi et al. 2000; Katsanis et al. 2002). In addition to this study, another two large-scale EST sequencing projects have been carried out to obtain ESTs from embryonic day 14.5 (Mu et al. 2001) and adult retinas (Wistow 2002). These kinds of studies assist in putting together the eye and/or retina transcriptomes at various developmental stages.

Chapter 4: Mouse I-gene microarrays for investigating ocular development and diseases

Microarrays have been successfully used to measure gene expression profiles in a high-throughput manner (Alizadeh and Staudt 2000; Golub et al. 1999). Such profiles may lead to the identification of transcriptional regulations (Cho et al. 2001; Mayanil et al. 2001). However, large-scale expression profiling of the retina has not been attempted, due in part to the under-representation of eye-derived ESTs in public databases and thus on commercial slides. Studies using small-scale cDNA microarrays containing less than 1000 eye-expressed ESTs have illustrated the genetic network of Crx (Livesey et al. 2000) and identified a potential target of Brn-3b (Mu et al. 2001). It is, therefore, of great interest to generate large-scale microarrays containing cDNAs of eye-genes.

In this study, slides containing over 4,000 ESTs, a subset of the total 10,000 ESTs isolated, were utilized to optimize procedures for preparing and printing cDNAs onto glass slides. The quality of these slides was then examined by self-self hybridization - microarray hybridization designed to compare the relative expression of genes in two aliquots of the same sample. The data demonstrated positive hybridization for a majority of genes on the slide, mostly with equal strength of hybridization in both aliquots.

cDNA microarrays of eye-expressed genes (called I-gene microarrays) provide a useful resource for systematic investigation of transcriptomes of the eye under various

conditions. Similar efforts have recently been attempted to generate custom microarrays of genes expressed in human retinas (Chowers et al. 2003a; Chowers et al. 2003b)

Chapter 5: Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays

In two-channel microarray experiments, differential gene expression between two mRNA sources was examined by simultaneously hybridizing their corresponding cDNAs, usually labeled by either Cy3 or Cy5 fluorescent dyes, onto a single slide (Bowtell 1999; Schena et al. 1995). The results are prone to a large pool of systematic sources of variations (Churchill 2002; Watson et al. 1998). To dissect the real expression differences between two samples from artificial changes, it is necessary to minimize all sources of variations. This study was intended to evaluate and optimize various microarray experimental methodologies, with an emphasis on the cDNA-labeling techniques. The goal was to identify a protocol capable of generating reproducible results with minimal technical variations.

Four protocols, differing primarily in labeling techniques, were applied to detected differential gene expression between wild-type and $Nrl^{-/-}$ samples during dye-swap experiments (switch of Cy3 and Cy5 for labeling of wild-type and $Nrl^{-/-}$ samples). Genisphere 3DNA-labeling protocol consistently demonstrated that (1) expression levels of a majority of genes did not change between the 2 samples and (2) differentially expressed genes showed swapped Cy3- and Cy5-signal in dye-swap experiments. Using this protocol, 3 μg of starting RNA was sufficient to produce reliable expression data. The optimization of microarray experimentation laid a foundation for meaningful study of gene expression. This optimization is the first of its kind (Dobbin et al. 2003; Manduchi et al. 2002; t Hoen et al. 2003). These protocols serve as templates for researchers who intend to use microarrays.

Chapter 6: Altered expression of genes of the Bmp/Smad and Wnt/Calcium signaling pathways in the mature cone-only $Nrl^{-/-}$ retina, revealed by gene profiling using custom cDNA microarrays

The two major types of photoreceptors, rods and cones, have different visual functions. Rod photoreceptors have high sensitivity to light and are responsible for night vision. In contrast, cone photoreceptors mediate high-acuity daylight color vision (Dowling 1988). Little is known about how a similar scheme of phototransduction leads to different functions in rods and cones (Ebrey and Koutalos 2001). There have been several hypotheses. Cones have additional pigment regeneration pathways leading to faster phototransduction kinetics (Arshavsky 2002; Hisatomi and Tokunaga 2002). Cone photoreceptor cells possess specific isoforms of phototransduction proteins that are similar to yet different from those in rods (Dhallan et al. 1992; Ebrey and Koutalos 2001). These two types of photoreceptors also differ in Ca^{2+} homeostasis (Krizaj and Copenhagen 2002; Korenbrot and Rebrik 2002; Ohyama et al. 2002; Rebrik and Korenbrot 1998) and in cellular pathways - the architecture of cells transmitting visual information from photoreceptors to the brain (Masland 2001). However, none of these studies has systematically examined rod- versus cone-mediated vision to elucidate the mechanisms underlying their differences. Taking advantage of the cone-only $Nrl^{-/-}$ mouse, the present study aimed to distinguish the genetic basis of cone function from that of rod.

Expression analysis using I-gene microarrays demonstrated induction of cone genes, repression of rod genes, disruption of Ca^{2+} homeostasis and reorganization of cytoskeletal structures in the adult $Nrl^{-/-}$ retinas. Together with quantitative real-time PCR (qRT-PCR), microarray results suggested biased utilization of Bmp/Smad and Wnt/ Ca^{2+} signaling pathways in rod- versus cone-mediated visual functions.

Genes with increased or decreased expression in the $Nrl^{-/-}$, relative to the wild-type, are candidate rod- or cone-genes, defined as genes participate mainly in either rod

or cone functions, respectively. A few of these down-regulated genes may be direct targets of Nrl. Functional analysis of these genes may help elucidate Nrl-regulated genetic network. This study is the first report that proposed potential signaling mechanisms underlying differential rod and cone functions. These findings, if confirmed, will greatly assist in the understanding of photoreceptor degeneration.

Chapter 7: Patterns of gene expression during retinal development and comparative microarray analysis of wild-type versus Nrl^{-/-} retina

Microarray analysis of mature retina indicated that Bmp/Smad and Wnt/Ca²⁺ signaling pathways are differentially used in rods and cones. Our studies also suggested a regulatory role for Nrl in modulating the Bmp/Smad pathway. The goal of studies presented in this chapter was to define gene expression changes during retinal development and their association with gene regulatory networks mediated by Nrl.

Microarray experiments were conducted using wild-type and Nrl^{-/-} retinas at 5 developmental ages: postnatal day (P) 0, 2, 6, 10 and 21. Analyses of gene expression during wild-type retinal development identified 2 groups of genes, characteristic of developing and mature retinas respectively. “Retina-early” genes, represented by those encoding protein biosynthesis and cell structural proteins, demonstrated expression levels high during P0 – P6 and low during P10 – P21. “Retina-late” genes, with temporal expression pattern opposite to retina-early genes, included enzymes, crystallins and genes of rod phototransduction cascade. The Nrl^{-/-} retina illustrated delayed expression of retina-early genes involved in protein biosynthesis and cell structure and dramatically reduced expression of retina-late genes associated with rod phototransduction. Through clustering and chromatin immunoprecipitation analysis, promoter regions of a group of

genes showing expression pattern similar to that of rhodopsin were demonstrated to bind Nrl.

This study defined expression patterns of the wild-type and Nrl^{-/-} retina at 5 developmental stages. It identified a number of genes and ESTs as candidate target genes of Nrl transcriptional regulation. Co-expression of Bmp/Smad signaling genes with those of rod phototransduction cascade indicates a pivotal role of Bmp/Smad pathway in rod development and function. These results highlight the molecular differences between rods and cones and should assist in delineating Nrl-regulated transcriptional networks.

Chapter 8: Conclusion

This chapter summarizes the key findings of this thesis. It also analyzes the implications of our work and the directions for future investigations.

1.5 References

- Alizadeh, A.A. and Staudt, L.M. 2000. Genomic-scale gene expression profiling of normal and malignant immune cells. *Curr Opin Immunol* 12: 219-225.
- Arshavsky, V. 2002. Like night and day: rods and cones have different pigment regeneration pathways. *Neuron* 36: 1-3.
- Bessant, D.A., Payne, A.M., Mitton, K.P., Wang, Q.L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A., and Bhattacharya, S.S. 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nat Genet* 21: 355-356.
- Bortoluzzi, S., d'Alessi, F., and Danieli, G.A. 2000. A novel resource for the study of genes expressed in the adult human retina. *Invest Ophthalmol Vis Sci* 41: 3305-3308.
- Bowtell, D.D. 1999. Options available--from start to finish--for obtaining expression data by microarray. *Nat Genet* 21: 25-32.
- Cho, Y.S., Kim, M.K., Cheadle, C., Neary, C., Becker, K.G., and Cho-Chung, Y.S. 2001. Antisense DNAs as multisite genomic modulators identified by DNA microarray. *Proc Natl Acad Sci U S A* 98: 9819-9823.
- Chowers, I., Gunatilaka, T.L., Farkas, R.H., Qian, J., Hackam, A.S., Duh, E., Kageyama, M., Wang, C., Vora, A., Campochiaro, P.A. et al. 2003a. Identification of novel genes preferentially expressed in the retina using a custom human retina cDNA microarray. *Invest Ophthalmol Vis Sci* 44: 3732-3741.
- Chowers, I., Liu, D., Farkas, R.H., Gunatilaka, T.L., Hackam, A.S., Bernstein, S.L., Campochiaro, P.A., Parmigiani, G., and Zack, D.J. 2003b. Gene expression variation in the adult human retina. *Hum Mol Genet* 12: 2881-2893.
- Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32 Suppl: 490-495.
- DeAngelis, M.M., Grimsby, J.L., Sandberg, M.A., Berson, E.L., and Dryja, T.P. 2002. Novel mutations in the NRL gene and associated clinical findings in patients with dominant retinitis pigmentosa. *Arch Ophthalmol* 120: 369-375.
- Dhallan, R.S., Macke, J.P., Eddy, R.L., Shows, T.B., Reed, R.R., Yau, K.W., and Nathans, J. 1992. Human rod photoreceptor cGMP-gated channel: amino acid sequence, gene structure, and functional expression. *J Neurosci* 12: 3248-3256.
- Dobbin, K., Shih, J.H., and Simon, R. 2003. Statistical design of reverse dye microarrays. *Bioinformatics* 19: 803-810.
- Dowling, J.E. 1988. The retina: an approachable part of the brain. Belknap Press.

- Dryja, T.P. and Li, T. 1995. Molecular genetics of retinitis pigmentosa. *Hum Mol Genet* 4 Spec No: 1739-1743.
- Ebrey, T. and Koutalos, Y. 2001. Vertebrate photoreceptors. *Prog Retin Eye Res* 20: 49-94.
- Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A. et al. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531-537.
- Hageman, G.S., Luthert, P.J., Victor Chong, N.H., Johnson, L.V., Anderson, D.H., and Mullins, R.F. 2001. An integrated hypothesis that considers drusen as biomarkers of immune-mediated processes at the RPE-Bruch's membrane interface in aging and age-related macular degeneration. *Prog Retin Eye Res* 20: 705-732.
- Hisatomi, O. and Tokunaga, F. 2002. Molecular evolution of proteins involved in vertebrate phototransduction. *Comp Biochem Physiol B Biochem Mol Biol* 133: 509-522.
- Katsanis, N., Worley, K.C., Gonzalez, G., Ansley, S.J., and Lupski, J.R. 2002. A computational/functional genomics approach for the enrichment of the retinal transcriptome and the identification of positional candidate retinopathy genes. *Proc Natl Acad Sci U S A* 99: 14326-14331.
- Klein, R., Wang, Q., Klein, B.E., Moss, S.E., and Meuer, S.M. 1995. The relationship of age-related maculopathy, cataract, and glaucoma to visual acuity. *Invest Ophthalmol Vis Sci* 36: 182-191.
- Korenbrod, J.I. and Rebrik, T.I. 2002. Tuning outer segment Ca²⁺ homeostasis to phototransduction in rods and cones. *Adv Exp Med Biol* 514: 179-203.
- Krizaj, D. and Copenhagen, D.R. 2002. Calcium regulation in photoreceptors. *Front Biosci* 7: d2023-2044.
- Livesey, F.J., Furukawa, T., Steffen, M.A., Church, G.M., and Cepko, C.L. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene *Crx*. *Curr Biol* 10: 301-310.
- Manduchi, E., Scearce, L.M., Brestelli, J.E., Grant, G.R., Kaestner, K.H., and Stoeckert, C.J., Jr. 2002. Comparison of different labeling methods for two-channel high-density microarray experiments. *Physiol Genomics* 10: 169-179.
- Masland, R.H. 2001. Neuronal diversity in the retina. *Curr Opin Neurobiol* 11: 431-436.
- Mayanil, C.S., George, D., Freilich, L., Miljan, E.J., Mania-Farnell, B., McLone, D.G., and Bremer, E.G. 2001. Microarray analysis detects novel Pax3 downstream target genes. *J Biol Chem* 276: 49299-49309.

- Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., and Swaroop, A. 2001. Nrl is required for rod photoreceptor development. *Nat Genet* 29: 447-452.
- Mitton, K.P., Swain, P.K., Chen, S., Xu, S., Zack, D.J., and Swaroop, A. 2000. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. *J Biol Chem* 275: 29794-29799.
- Mu, X., Zhao, S., Pershad, R., Hsieh, T.F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. et al. 2001. Gene expression in the developing mouse retina by EST sequencing and microarray analysis. *Nucleic Acids Res* 29: 4983-4993.
- Ohyama, T., Picones, A., and Korenbrot, J.I. 2002. Voltage-dependence of ion permeation in cyclic GMP-gated ion channels is optimized for cell function in rod and cone photoreceptors. *J Gen Physiol* 119: 341-354.
- Ramsay, G. 1998. DNA chips: state-of-the art. *Nat Biotechnol* 16: 40-44.
- Rebrik, T.I. and Korenbrot, J.I. 1998. In intact cone photoreceptors, a Ca²⁺-dependent, diffusible factor modulates the cGMP-gated ion channels differently than in rods. *J Gen Physiol* 112: 537-548.
- Saleem, R.A. and Walter, M.A. 2002. The complexities of ocular genetics. *Clin Genet* 61: 79-88.
- Schena, M., Heller, R.A., Thériault, T.P., Konrad, K., Lachenmeier, E., and Davis, R.W. 1998. Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol* 16: 301-306.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467-470.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* 9: 3273-3297.
- Swain, P.K., Hicks, D., Mears, A.J., Apel, I.J., Smith, J.E., John, S.K., Hendrickson, A., Milam, A.H., and Swaroop, A. 2001. Multiple phosphorylated isoforms of NRL are expressed in rod photoreceptors. *J Biol Chem* 276: 36824-36830.
- Szlyk, J.P., Fishman, G.A., Alexander, K.R., Peachey, N.S., and Derlacki, D.J. 1993. Clinical subtypes of cone-rod dystrophy. *Arch Ophthalmol* 111: 781-788.

t Hoen, P.A., de Kort, F., van Ommen, G.J., and den Dunnen, J.T. 2003. Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res* 31: e20.

Watson, A., Mazumder, A., Stewart, M., and Balasubramanian, S. 1998. Technology for microarray analysis of gene expression. *Curr Opin Biotechnol* 9: 609-614.

Wistow, G. 2002. A project for ocular bioinformatics: NEIBank. *Mol Vis* 8: 161-163.

CHAPTER 2

REVIEW OF THE LITERATURE ON PHOTORECEPTOR DIFFERENTIATION AND FUNCTION

2.1 Eye Anatomy and Structure

The **eye**, an organ for sight, is part of the central nervous system (Wright 1997). The wall of the eyeball has three layers of tissues (Figure 2.1). The outermost layer is a tough, fibrous tissue with very few blood vessels. It is comprised of two continuous parts: the “cornea”, a transparent tissue, covers the front one-third of the eyeball; and the “sclera” surrounds the back two-thirds of the eyeball. The middle vascular layer includes, from the front to the back of the eyeball, the iris, the ciliary body and the choroid. The vascular layer nourishes the retina, which forms the innermost layer of the eyeball.

The **retina** is comprised of three layers of nerve cell bodies, which include the ganglion cell layer, inner nuclear layer and outer nuclear layer, separated by two layers of synapses, termed inner and outer plexiform layers respectively (Figure 2.2) (Dowling 1987; Wright 1997). In the vertebrate retina, there are six major types of neurons, which include ganglion cells, amacrine cells, bipolar cells, horizontal cells, and rod and cone photoreceptor cells (Livesey and Cepko 2001; Sterling 1990). This network of neurons senses light, converts images into electric signals, and transmits these signals to the brain. The ganglion cell layer contains the cell bodies of ganglion cells. At the inner plexiform layer, the dendrites of these cells form synaptic contacts with the axons of bipolar cells, which are located in the inner nuclear layer. Visual information transmitted from the

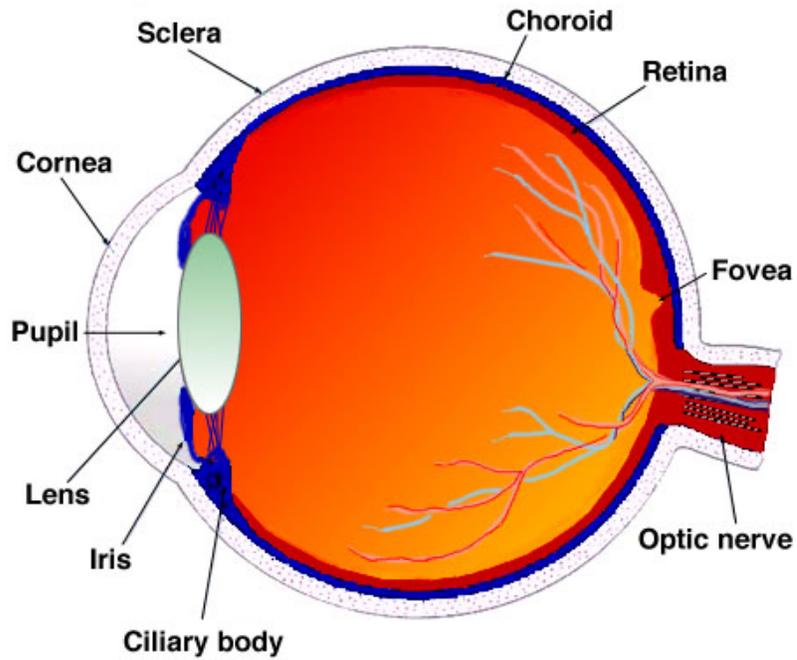


Figure 2.1. The anatomy and structure of the adult human eye.

The three layers of eyeball are shown in different colors: white for the cornea and the sclera, blue for the choroid, and red for the retina. This figure was adapted from Webvision (<http://webvision.med.utah.edu/anatomy.html>).

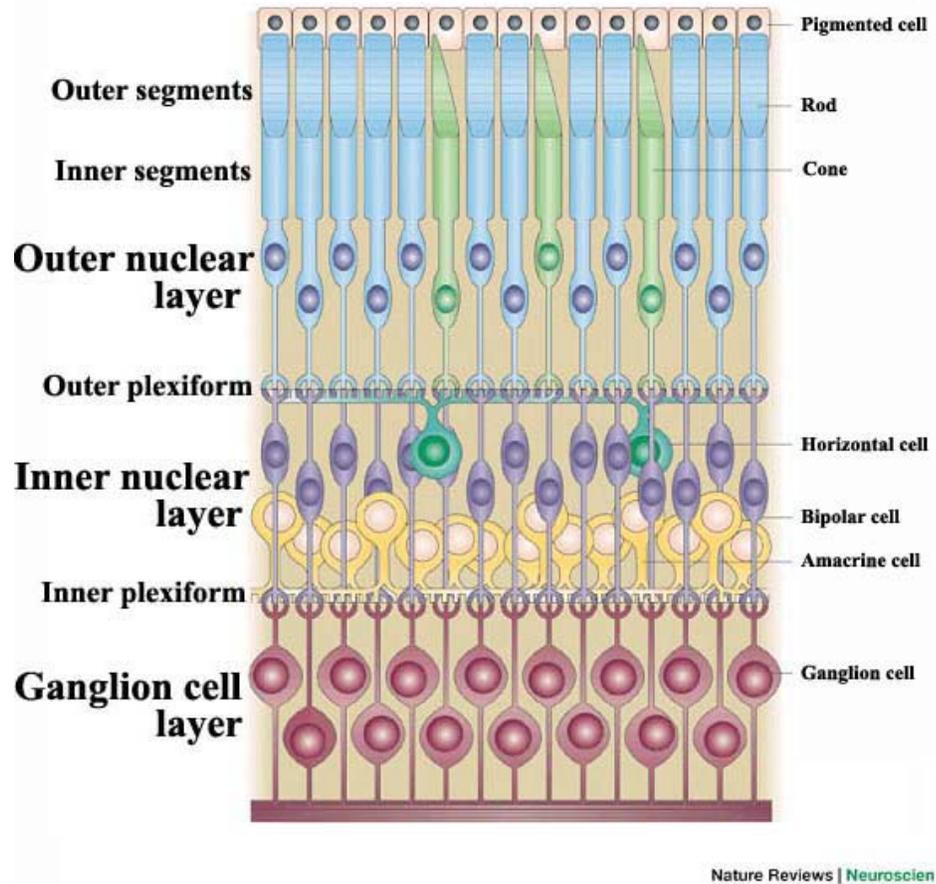


Figure 2.2. The structure of the retina.

The retina contains a network of neurons, including rod and cone photoreceptors, bipolar cells, horizontal cells, amacrine cells and ganglion cells. The cell bodies of these neurons are organized into 3 nuclear layers, separated by 2 plexiform layers. This figure was adapted from (Livesey and Cepko 2001).

bipolar cells is processed by ganglion cells and output to the brain through the optic nerve. In addition to bipolar cells, the inner nuclear layer also contains cell bodies of inter-neurons: amacrine cells and horizontal cells. The inter-neurons form a network of synapses with photoreceptor cells, bipolar cells and ganglion cells and aid in the integrating and processing of visual signals (Masland 1996; Wright 1997). The outer nuclear layer contains the cell bodies of photoreceptors. On the interior side, photoreceptors have axons transmitting electric signals to bipolar cells. Protruding from the exterior side of the photoreceptor cells is enlarged outer segments (Figure 2.3). These outer segments contain, in their flat membrane discs, many photopigment molecules. New membrane discs are continuously formed at the base; and old discs shed off from the top of photoreceptor outer segments. The inner segment is connected to the outer segment by the connecting cilium and is packed of organelles for the metabolism of the cell.

2.2 Photoreceptor Function

Vertebrate retina contains two types of **photoreceptors**: rods and cones (Sterling 1990). Rod photoreceptors have high light sensitivity: rods are capable of detecting a single photon. However, rods may saturate under steady daylight or bright light (Rodieck 1998). In contrast, cones are less sensitive but are adapted to operate under bright light to provide high acuity vision. Many vertebrate retinas contain several subtypes of cones, which differ in their peak spectral sensitivities. The simultaneous activation of these cone subtypes by photons generates color vision. The spectral sensitivities of photoreceptors are determined by the visual pigments (called opsin) available in their outer segments. Each photoreceptor subtype generally possesses only one type of opsins. Once synthesized in the inner segment, these opsin molecules are transported to the outer segment. With its membrane discs and enfoldings, the outer segment is capable of

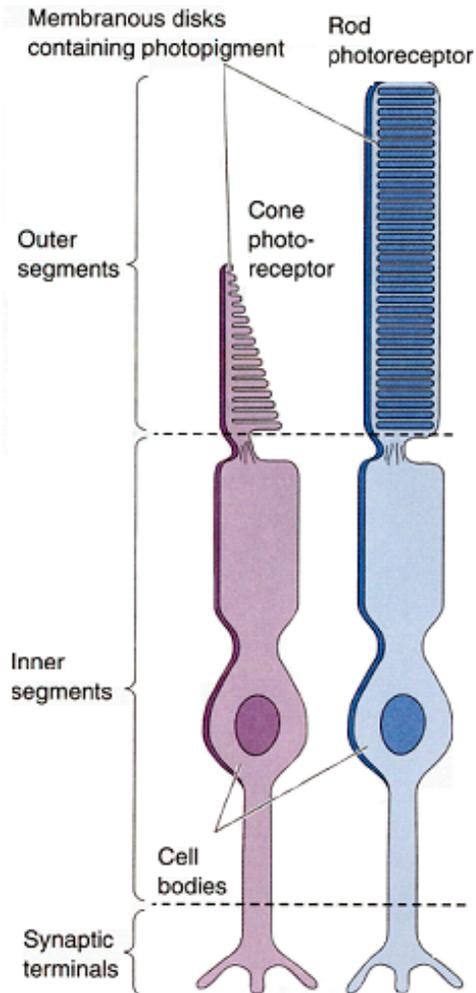


Figure 2.3. The structure of photoreceptors.

Rods and cones have differently shaped outer segments. The outer segments are rich in photopigment molecules, located in their membrane discs. These discs are consistently updated by synthetic activity of inner segments. The synaptic terminals of photoreceptors are connected to bipolar cells.

holding a large number of opsins (Figure 2.3). Each opsin contains a protein fraction, which is distinct for different photoreceptor subtypes, and an 11-cis retinal-based chromophore portion. Photoactivation, the isomerization of the chromophore by a photon, is the first step of a biochemical cascade, termed phototransduction, that leads to visual perception.

Phototransduction is the basic scheme for the detection of light stimulus in all photoreceptors (Polans et al. 1996). The phototransduction cascade has been studied primarily in rod photoreceptor cells (Figure 2.4) (Dowling 1987; Ebrey and Koutalos 2001; Molday 1998). Light enters through the pupil, passes several layers of eye tissues and finally reaches the retina at the back of the eye. It then penetrates almost all layers of the retina and is absorbed by photoreceptor outer segments. In the membrane discs of these outer segments, the energy from absorbed photons changes the shape of chromophores of opsins, and hence activates these molecules. In rods, the activated rhodopsin temporarily binds to a G protein, called transducin, which has GDP associated with its G_{α} subunit. This binding results in the replacement of GDP by GTP and the separation of GTP-bound G_{α} (activated G_{α}) from the $G_{\beta\gamma}$ subunit. The activated G_{α} molecule then binds to a cyclic GMP (cGMP)-hydrolyzing enzyme, called cGMP phosphodiesterase (PDE). The interaction between G_{α} and PDE displaces the PDE_{γ} subunit and thus exposes the catalytic site on either PDE_{α} or PDE_{β} subunit. The activated PDE acts on cGMP molecules and converts them into GMP molecules, leading to the closure of the cGMP-gated cation channels. In the dark, the influx of cations in the photoreceptor outer segments, through the cGMP-gated channels, approximately balances the outflow of cations in the inner segments. This balance maintains the transmembrane potential of photoreceptor cells. In the light, the closure of cGMP-gated channels generates a net outward current and causes hyperpolarization of the photoreceptor cell membrane. This marks the end of phototransduction cascade and the start of visual signal transmission between neurons. For most of these phototransduction proteins, rods

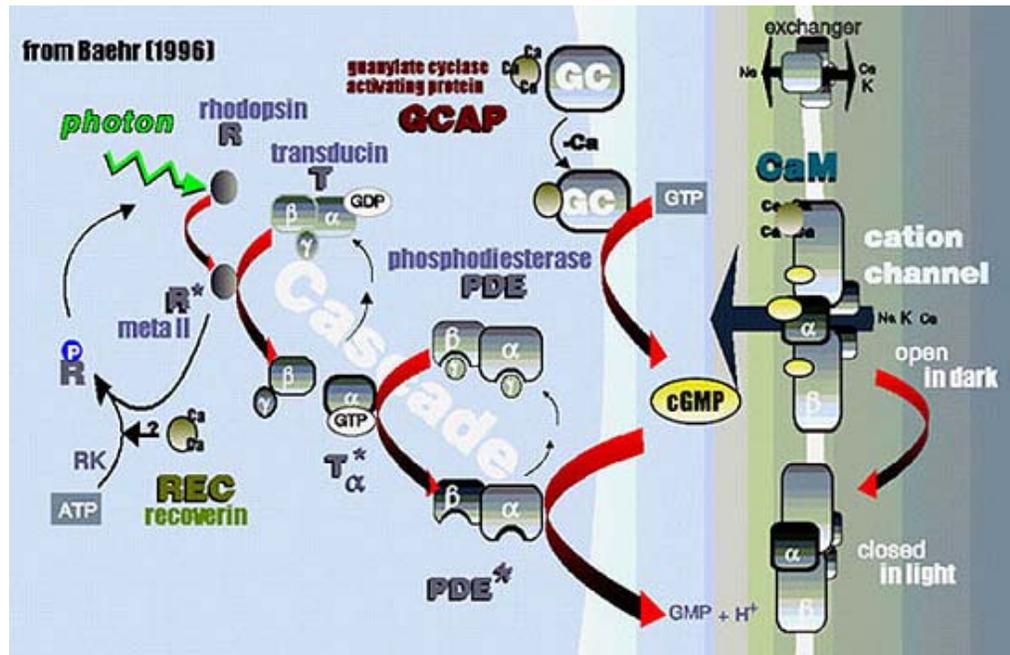


Figure 2.4. The phototransduction cascade.

The phototransduction cascade involves a chain of biochemical reactions, and operates a cGMP-gated cation channel. This figure was adapted from (Polans et al. 1996).

and cones have different isoforms. The specific molecular characteristics of these isoforms may influence their physiological responses and are perhaps associated with differential rod and cone function (Tachibanaki et al. 2001). Characterization of the regulation and function of these isoforms may facilitate our understanding of rod and cone function.

Visual cycle, a process that regenerates visual pigments, greatly affects phototransduction dynamics. A visual cycle includes two separate processes: the recovery of opsin molecules from the photoisomerized form and the regeneration of 11-cis retinal, termed the retinoid cycle. The activated rhodopsin is recovered through phosphorylation by the rhodopsin kinase under the regulation of recoverin (Figure 2.4) (Chen 2002; Molday 1998). Arrestin recognizes phosphorylated rhodopsin and changes the rhodopsin molecule to its original non-photoisomerized form. Studies of visual pigment phosphorylation demonstrated that, in the outer segments of cones, this reaction is over 20 times faster than that of rods (Tachibanaki et al. 2001). This faster decay of active opsin explains, in part, the smaller amplitude, lower light sensitivity and shorter photo-response in cones than in rods. However, further investigations of underlying molecular mechanisms are yet to be carried out.

In the **retinoid cycle**, the 11-cis retinal is regenerated from all-trans-retinal through a chain of enzymatic activities. A traditional cycling pathway involves the participation of the RPE cells located outside of the photoreceptor layer (McBee et al. 2001; Saari et al. 2000). Both rods and cones are able to use this RPE-mediated pathway for pigment regeneration. Mata et al. recently reported an additional pathway requiring the function of Muller cells (Figure 2.5) (Arshavsky 2002; Mata et al. 2002). The authors indicated that this pathway is significantly faster (over 20 folds) than the RPE-mediated pathway and is exclusively used by cones. Their findings suggest a possible mechanism for the cones to remain functional, when the rods saturate, under bright illumination: cones remain responsive by always keeping some visual pigments unbleached through

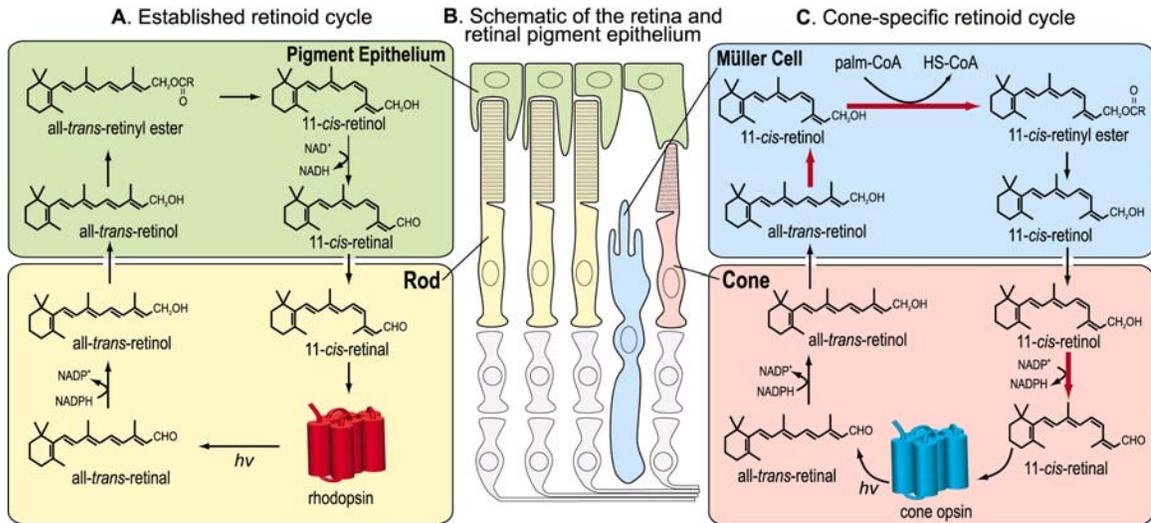


Figure 2.5. Retinoid cycles.

(A) An established retinoid cycle involves the RPE, and is used by both rods and cones. (B) The schematic of the retina and RPE. (C) A novel retinoid cycle, for cone photoreceptors only, involves Müller cells. This figure was adapted from (Arshavsky 2002).

rapid regeneration. It is to be understood, however, how the nature has evolved in such a way that cones, but not rods, are able to use this effective pigment regeneration pathway.

Ca²⁺ homeostasis differs in rods and cones. The amount of free Ca²⁺ inside photoreceptor outer segments has a profound effect on phototransduction (Pepe 2001). Cytoplasmic Ca²⁺ concentration is controlled by the rate of Ca²⁺ influx and outflow and the Ca²⁺-binding proteins (Polans et al. 1996). These proteins act as a buffer system for cytoplasmic free Ca²⁺ and consequently intervene in phototransduction. Ca²⁺ regulates the synthesis of cGMP molecules by controlling the catalytic activity of guanylate cyclase (GC) through guanylate cyclase activating proteins (Gcap) (Duda et al. 1996; Nakatani et al. 2002). On the other hand, the increase of cGMP opens cation channels and leads to the inflow of Ca²⁺. The interaction between Ca²⁺ and cGMP, through a feedback loop, is essential for keeping the transmembrane potential of photoreceptors at a stable and responsive level. There are other biochemical reactions in phototransduction that are regulated by Ca²⁺: Ca²⁺-mediated feedback pathways underlie the light adaptation of photoreceptors (Kawamura 1999); ATP-dependent deactivation and recoverin-mediated phosphorylation of rhodopsin are controlled by Ca²⁺ (Senin et al. 2002); and the breakdown of cGMP molecules is Ca²⁺-dependent. In cones, Ca²⁺ affects the cGMP affinity of the cGMP-gated cation channels by Ca-calmodulin complex. The quantitative differences in the molecular pathways regulating Ca²⁺ homeostasis in rods and cones, for example Ca²⁺-binding proteins, are proposed to explain some differences in their function (Korenbrodt and Rebrik 2002). More work needs to be done to further illustrate these molecular processes that distinguish rods and cones.

The **processing of visual information** in vertebrate retinas is often featured by two parallel processes: the rod pathway and the cone pathway (Figure 2.6) (Masland 2001; Sterling 1990). Hyperpolarization of photoreceptor membranes leads to a decrease in the rate of glutamate release at the synaptic terminals. This signal is then transmitted, from either rods or cones, to the ganglion cells and then to the brain, through the release

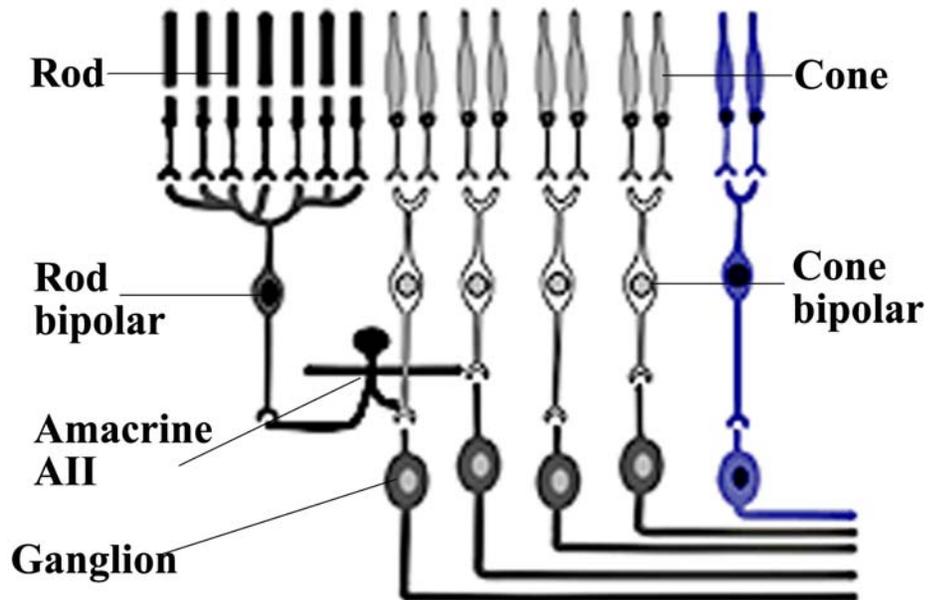


Figure 2.6. The retinal cellular pathways.

Visual signals acquired by rod photoreceptors are transmitted, through amacrine AII cells, to the dendrites of cone bipolar cells, which subsequently output rod signals to ganglion cells. This figure was adapted from (Masland 2001).

and detection of neurotransmitters at synaptic connections. Each cone is presynaptic to several cone bipolar cells, of which there are many subtypes. On the other hand, there is only one type of rod bipolar cells and each is postsynaptic to tens of rod photoreceptors. Cone bipolar cells make synaptic contacts with the processes of ganglion cells. By contrast, rod bipolar cells are unable to directly transmit rod signals to ganglion cells: there is no synaptic connection from rod bipolar cells to ganglion cells. Instead, rod bipolar cells make their synaptic output first onto the dendrites of amacrine AII cells, which subsequently spread visual signals to cone bipolar cells (Masland 1996, 2001). Such differences in rod and cone cellular pathways are of great importance in understanding the contrasts between rod and cone function.

2.3 Retinal Development and Photoreceptor Differentiation

Retinal development, particularly the determination of retinal cell fate, is controlled by intrinsic cues and extrinsic stimuli (Cepko 1996, 1999). The “competence model” suggests that, during development, retinal progenitor cells (RPCs) may undergo a series of competence states, determined by intrinsic properties - genes expressed within the cell. Each competence state, in response to various environmental cues, is able to give rise to several particular cell types (Cepko 1999). In the mouse retina, it is clear that a set of RPCs assumes a cone cell-fate first (embryonic day 10 – 18), whereas the majority of rod-fated cells are not generated until upon birth (Cepko 1996; Young 1985a, b). Since the sequential movement of RPCs from one state to the next is proposed to be irreversible, this competence model, therefore, suggests the inability of RPCs to choose, at late embryonic or early postnatal stages, a cone cell-fate (Cepko 1999). This is partially supported by the observation of retinal two-cell clones, in lineage analyses, containing rods and several other types of retinal neurons, not including cones (Turner and Cepko 1987).

Photoreceptor differentiation, defined as the development of photoreceptor progenitors to opsin-expressing cells, is regulated by many transcription factors (Cepko et al. 1998). These include potential negative regulators, such as pax6 (Hitchcock et al. 1996), chx10 (Liu et al. 1994), math5 (Brown et al. 2001) and Notch/HES signaling (Dorsky et al. 1995). There are several transcriptional factors that affect photoreceptor populations, including the cone-rod homeobox gene (Crx) (Furukawa et al. 1999), thyroid hormone receptor $\beta 2$ (Tr $\beta 2$) (Ng et al. 2001), photoreceptor-specific nuclear receptor (Nr2e3) (Haider et al. 2001), and Nrl (Mears et al. 2001). Mears et al. recently proposed a “knockout-based model” summarizing the regulation of these transcriptional factors on photoreceptor differentiation in mouse (Figure 2.7). In this model, RPCs differentiate first into a group of post-mitotic cells (PMC) committed to become cones and then to PMCs designated for either S-cones or rods. The proportional population of rods and different types of cones are controlled by the relative expression of Tr $\beta 2$, an unknown factor X, Crx, Nrl and Nr2e3. This model suggests (1) the existence of progenitor cells that are capable of giving rise to either rods or cones, but not other types of retinal neurons and (2) the reversibility of the fates of some retinal progenitors. In support of this, a recent report demonstrated that late retinal progenitors are able to generate retinal ganglion cells, a type of early-born neurons (James et al. 2003). A comprehensive study of photoreceptor differentiation may explain the controversy between the knockout-based model and the competence model.

2.4 Nrl and the Nrl Knockout

Nrl, a basic leucine zipper protein of the Maf subfamily, is expressed specifically in rod photoreceptor cells of the mammalian retina (Swain et al. 2001). Nrl regulates

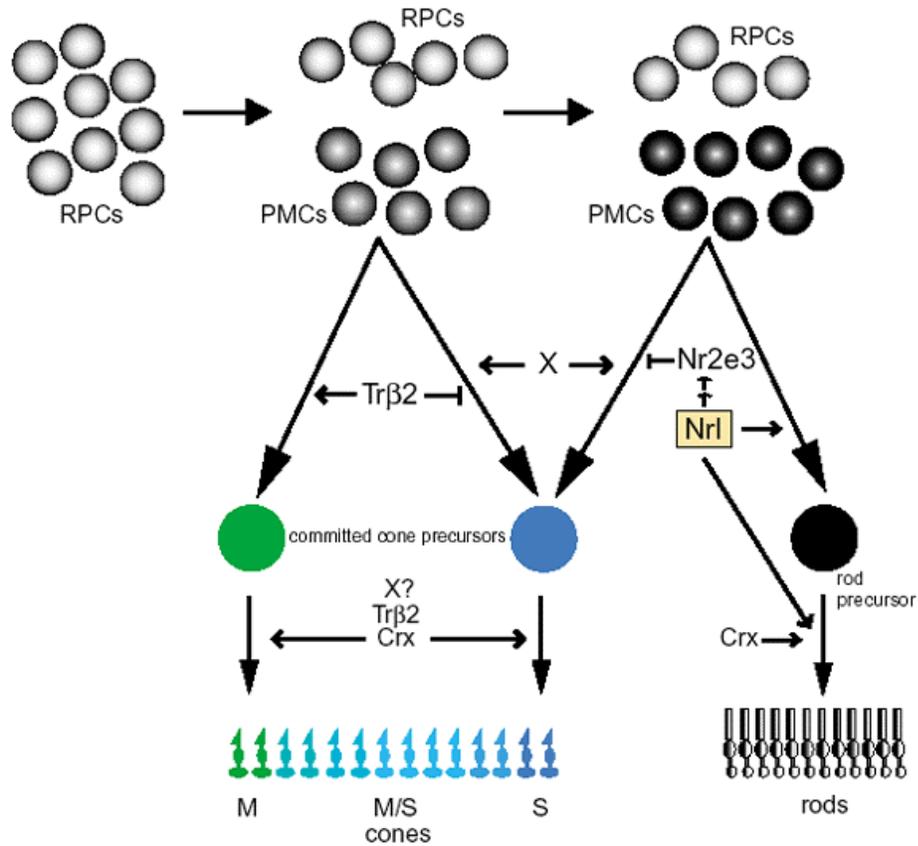


Figure 2.7. A model for photoreceptor differentiation.

In this model, some retinal progenitor cells (RPCs) become groups of post-mitotic cells (PMCs) during development. These PMCs are able to differentiate, in response to environmental stimulus - such as transcriptional factors - into rods or various types of cones. Here, Nrl functions as a molecular switch between rods and S-cones, perhaps through the regulation of Nr2e3. This figure was adapted from (Mears et al. 2001).

rod-specific gene expression, including rhodopsin (Rho) and cGMP-phosphodiesterase beta-subunit (Pdeb) (Lerner et al. 2001; Rehemtulla et al. 1996). *Nrl* and *Crx* interact with each other and synergistically transactivate the rhodopsin promoter, by binding to the adjacent NRE and Ret-4 sites respectively (Mitton et al. 2000). The importance of *Nrl* is further highlighted by the association of *NRL* mutations with autosomal dominant retinitis pigmentosa (Bessant et al. 1999; DeAngelis et al. 2002; Martinez-Gimeno et al. 2001).

The ***Nrl*-knockout** (*Nrl*^{-/-}) retina contains no rods and has an increased number of S-cones (Figure 2.8) (Mears et al. 2001). Ultrastructural examination of this retina revealed a collapse of the sub-retinal space and a shortening of the outer segments, indicating primary or secondary effects of *Nrl* on cytoskeleton. Light microscopy showed whorls and rosettes in the outer nuclear layer of the 5-week-old *Nrl*^{-/-} retina. Preliminary analyses demonstrated substantially reduced expression of genes encoding proteins of rod phototransduction cascade. The absence of the nuclear receptor *Nr2e3* transcript in the *Nrl*^{-/-} retina is of considerable interest, as in humans mutations in *NR2E3* result in enhanced S-cone syndrome and in the rd7 mouse a gain of S-cones (Haider et al. 2001). It has been hypothesized that *Nrl* directly or indirectly regulates the expression of *Nr2e3* (Mears et al. 2001). However, the molecular details of photoreceptor development and function, as controlled by *Nrl*, remain to be elucidated.

Studies of the cellular pathways for visual information identified changes of neural circuitry in the *Nrl*^{-/-} retina (Ed Pugh and A. Swaroop, unpublished data). Besides their normal connection with cone bipolar cells, cone photoreceptors establish synaptic contacts with the rod bipolar cells. These findings suggest that, in the absence of rods, cones are able to utilize rod cellular pathways for their function. The *Nrl*^{-/-} mouse model, therefore, appears ideal for the study of the molecular differences in rods and cones and their associated signaling pathways.

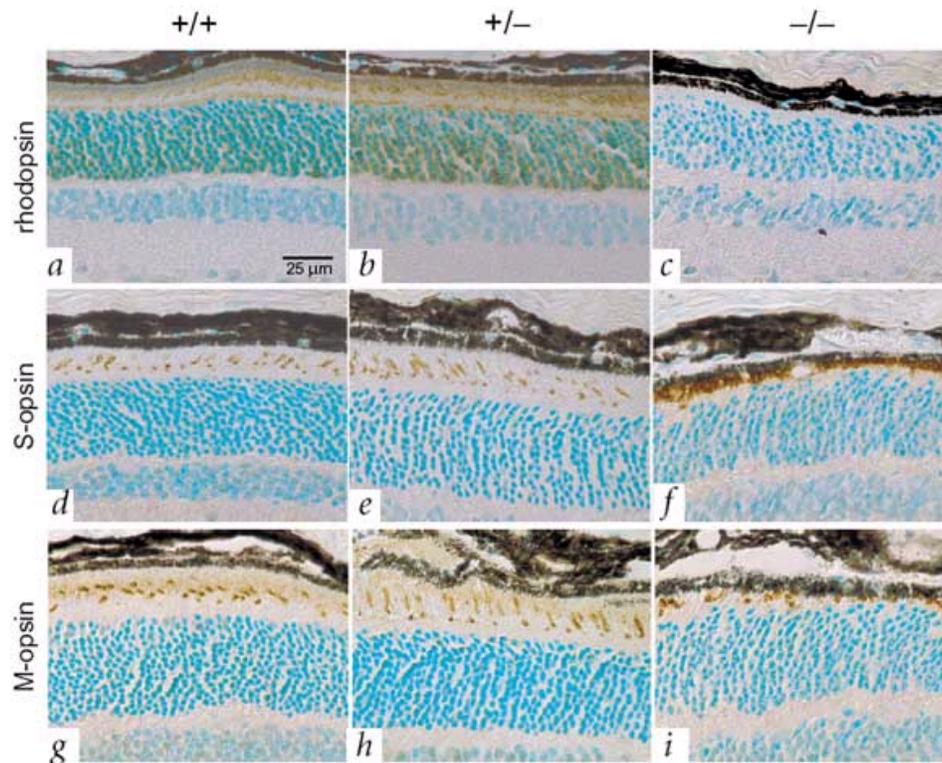


Figure 2.8. Opsin immunohistochemistry of the wild-type (+/+), Nrl heterozygous (+/-), and Nrl knockout (-/-) retina.

A comparison of immunostaining (brown staining in outer segments) of 3 genes, rhodopsin, S-opsin and M-opsin, as markers for rod, S-cone and M-cone photoreceptors respectively, are shown. This figure was adapted from (Mears et al. 2001).

2.5 Photoreceptor Degeneration and Cone Survival

Retinal dystrophies (RDs) are a group of retinal disorders that are genetically heterogeneous. Both RP and AMD primarily result in the death of photoreceptor cells, with early onset of apoptosis in rods and subsequent death of cones (Klein et al. 1995; Saleem and Walter 2002). Numerous studies have attempted to untangle the correlation between rod and cone survival using various mouse models. The rhodopsin-knockout mouse ($Rho^{-/-}$) is a model that mimics the degeneration process of typical RP (Humphries et al. 1997). In this model, the death of rods leads to that of cones, suggesting that the survival of cones is dependent on the presence of rods. It was speculated that rod photoreceptors release either survival factors at functional stages or death factors when apoptosis is engaged. In the $Rpe65^{-/-}$ mouse, the rod photoreceptors die, while the cones remain functional. These data fail to support the theory of rod death factors. In contrast, the $Nrl^{-/-}$ mouse retinas never develop rods, but still have functional cones, indicating the independence of cones on putative survival factors from rods (Mears et al. 2001). The controversy of these studies may reflect several facts. First, the degeneration pathways may vary under various conditions. Second, the current model of rod's effect on cones is perhaps overly simplistic. One approach to address this question is to examine the distinct cellular and molecular pathways important for rod or cone development and functional maintenance.

2.6 Microarray Technology

Microarray technology has enabled rapid access to gene expression information in a high-throughput manner. Microarrays have been adopted to study genome-wide gene expression (DeRisi et al. 1996), monitor changes in gene copy numbers (Pollack et al. 1999) and screen for polymorphisms (Hacia 1999). There are two major types of microarrays: oligonucleotide-based Affymetrix chips (Lipshutz et al. 1999; Lockhart et

al. 1996) and cDNA microarray slides (Brown and Botstein 1999; Schena et al. 1995). A cDNA microarray is generally a glass slide containing high-density (usually thousands) of immobilized cDNAs. These cDNAs are often clones isolated from cDNA libraries and printed onto glass slides by an arrayer or spotter (Bowtell 1999). In a typical two-color cDNA microarray experiment, total RNAs are isolated from two tissue samples and reverse-transcribed into cDNAs. These two samples of cDNAs are labeled respectively by one of two dyes (usually Cy3 or Cy5) that have different excitation spectrums and are co-hybridized onto a single microarray slide (Figure 2.9). The hybridized slide is then scanned for hybridization intensity of the Cy3- and the Cy5-channels separately. In this thesis study, customized cDNA microarrays were used to examine gene expression in mouse retinas.

The expression analysis using microarrays is limited to the genes that are available on the slide. cDNA microarrays can be either purchased or custom made. Most commercial microarray slides used IMAGE clones that were supplied by vendors, such as Genome Systems (<http://www.genomesystems.com>) and Research Genetics (<http://www.resgen.com>). However, a high percentage of sequence errors and discrepancies has been found in these clones (Halgren et al. 2001). Furthermore, eye-expressed genes or ESTs are under-represented in the NCBI databases and in the IMAGE clones. Therefore, commercially available cDNA microarray slides have limited application in profiling ocular gene expression. It is necessary to generate custom eye-gene microarrays. However, it is costly to sequence-verify IMAGE clones and consolidate eye-expressed genes. An alternative approach is to make cDNA libraries from eye tissues and isolate and print these clones onto glass slides to produce eye-gene microarrays.

Microarray data is vulnerable to many different sources of variations (Churchill 2002). Schuchhardt et al. summarizes a list of major sources of variation that may occur during a standard microarray experiment, ranging from RNA preparation to slide printing

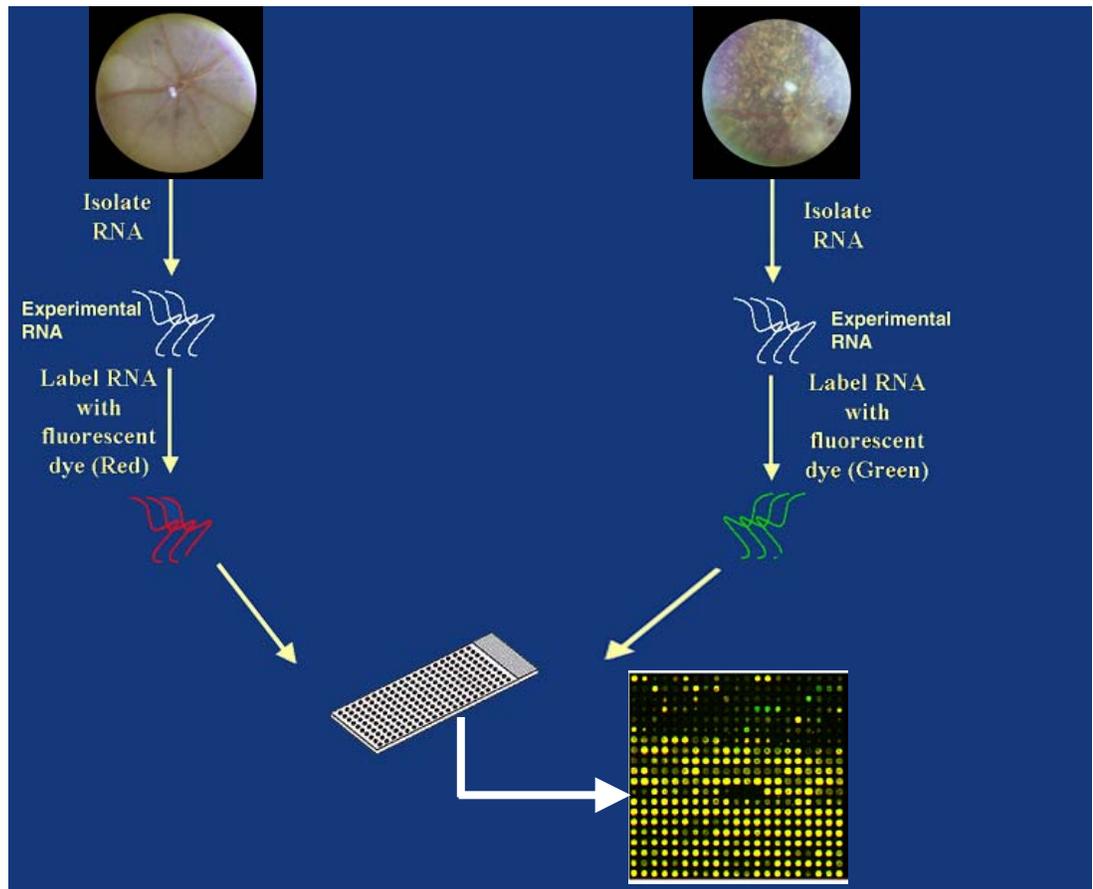


Figure 2.9. Schematic overview of cDNA microarray technology.

In a typical cDNA microarray experiment, total RNAs are isolated from two tissue samples, for instance normal and diseased retinas. These RNAs are labeled by one of two fluorescent dyes, which differ in excitation spectrum (red or green here) during reverse-transcription. Labeled probes are then hybridized to a single microarray slide to generate gene expression data.

(Schuchhardt et al. 2000). In order to provide reliable microarray data, numerous efforts have been attempted to reduce these variations. These include careful experimental design (Yang et al. 2002), sufficient replication (Pan et al. 2002; Pavlidis et al. 2003) and appropriate image and data analysis (Quackenbush 2002; Yang et al. 2002). To compare the biologically relevant differences in gene expression under various conditions, a large number of statistical models have been developed, which includes R-Limma package (Smyth et al. 2003), SAM (Tusher et al. 2001) and the ANOVA method (Kerr et al. 2000).

2.7 Summary

This review provides, not only an introduction to photoreceptor development and function, but also possible mechanisms underlying the differences between rods and cones. Several key features, including phototransduction proteins, dynamics of visual cycle, Ca^{2+} homeostasis and cellular pathways, may have causal effects. However, a comprehensive study is yet to be achieved to address this question properly. In this thesis study, microarray technology and the cone-only $\text{Nrl}^{-/-}$ retina were used for this purpose.

With respect to photoreceptor development, two models were discussed. The competence model suggests different post-mitotic progenitors, during early cell-fate determination, for cones and rods, whereas the knockout-based model implicates reversible roles of some progenitors, during late differentiation, in becoming rods or various types of cones. One role of this thesis is to determine, through transcriptome analysis of developing wild-type and $\text{Nrl}^{-/-}$ retina, the regulatory network of Nrl that distinguishes rod and cone differentiation.

2.8 References

- Arshavsky, V. 2002. Like night and day: rods and cones have different pigment regeneration pathways. *Neuron* **36**: 1-3.
- Bessant, D.A., Payne, A.M., Mitton, K.P., Wang, Q.L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A., and Bhattacharya, S.S. 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nat Genet* **21**: 355-356.
- Bowtell, D.D. 1999. Options available--from start to finish--for obtaining expression data by microarray. *Nat Genet* **21**: 25-32.
- Brown, N.L., Patel, S., Brzezinski, J., and Glaser, T. 2001. Math5 is required for retinal ganglion cell and optic nerve formation. *Development* **128**: 2497-2508.
- Brown, P.O. and Botstein, D. 1999. Exploring the new world of the genome with DNA microarrays. *Nat Genet* **21**: 33-37.
- Cepko, C.L. 1996. The patterning and onset of opsin expression in vertebrate retinæ. *Curr Opin Neurobiol* **6**: 542-546.
- Cepko, C.L. 1999. The roles of intrinsic and extrinsic cues and bHLH genes in the determination of retinal cell fates. *Curr Opin Neurobiol* **9**: 37-46.
- Cepko, C.L., Ryder, E., Austin, C., Golden, J., Fields-Berry, S., and Lin, J. 1998. Lineage analysis using retroviral vectors. *Methods* **14**: 393-406.
- Chen, C.K. 2002. Recoverin and rhodopsin kinase. *Adv Exp Med Biol* **514**: 101-107.
- Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* **32 Suppl**: 490-495.
- DeAngelis, M.M., Grimsby, J.L., Sandberg, M.A., Berson, E.L., and Dryja, T.P. 2002. Novel mutations in the NRL gene and associated clinical findings in patients with dominant retinitis pigmentosa. *Arch Ophthalmol* **120**: 369-375.
- DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* **14**: 457-460.
- Dorsky, R.I., Rapaport, D.H., and Harris, W.A. 1995. Xotch inhibits cell differentiation in the *Xenopus* retina. *Neuron* **14**: 487-496.
- Dowling, J.E. 1987. *The retina: an approachable part of the brain*. Harvard University Press, Cambridge, MA.

- Duda, T., Goraczniak, R., Surgucheva, I., Rudnicka-Nawrot, M., Gorczyca, W.A., Palczewski, K., Sitaramayya, A., Baehr, W., and Sharma, R.K. 1996. Calcium modulation of bovine photoreceptor guanylate cyclase. *Biochemistry* **35**: 8478-8482.
- Ebrey, T. and Koutalos, Y. 2001. Vertebrate photoreceptors. *Prog Retin Eye Res* **20**: 49-94.
- Furukawa, T., Morrow, E.M., Li, T., Davis, F.C., and Cepko, C.L. 1999. Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nat Genet* **23**: 466-470.
- Hacia, J.G. 1999. Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet* **21**: 42-47.
- Haider, N.B., Naggert, J.K., and Nishina, P.M. 2001. Excess cone cell proliferation due to lack of a functional NR2E3 causes retinal dysplasia and degeneration in rd7/rd7 mice. *Hum Mol Genet* **10**: 1619-1626.
- Halgren, R.G., Fielden, M.R., Fong, C.J., and Zacharewski, T.R. 2001. Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones. *Nucleic Acids Res* **29**: 582-588.
- Hitchcock, P.F., Macdonald, R.E., VanDeRyt, J.T., and Wilson, S.W. 1996. Antibodies against Pax6 immunostain amacrine and ganglion cells and neuronal progenitors, but not rod precursors, in the normal and regenerating retina of the goldfish. *J Neurobiol* **29**: 399-413.
- Humphries, M.M., Rancourt, D., Farrar, G.J., Kenna, P., Hazel, M., Bush, R.A., Sieving, P.A., Sheils, D.M., McNally, N., Creighton, P., et al. 1997. Retinopathy induced in mice by targeted disruption of the rhodopsin gene. *Nat Genet* **15**: 216-219.
- James, J., Das, A.V., Bhattacharya, S., Chacko, D.M., Zhao, X., and Ahmad, I. 2003. In vitro generation of early-born neurons from late retinal progenitors. *J Neurosci* **23**: 8193-8203.
- Kawamura, S. 1999. Calcium-dependent regulation of rhodopsin phosphorylation. *Novartis Found Symp* **224**: 208-218; discussion 218-224.
- Kerr, M.K., Martin, M., and Churchill, G.A. 2000. Analysis of variance for gene expression microarray data. *J Comput Biol* **7**: 819-837.
- Klein, R., Wang, Q., Klein, B.E., Moss, S.E., and Meuer, S.M. 1995. The relationship of age-related maculopathy, cataract, and glaucoma to visual acuity. *Invest Ophthalmol Vis Sci* **36**: 182-191.
- Korenbrod, J.I. and Rebrik, T.I. 2002. Tuning outer segment Ca²⁺ homeostasis to phototransduction in rods and cones. *Adv Exp Med Biol* **514**: 179-203.

- Lerner, L.E., Gribanova, Y.E., Ji, M., Knox, B.E., and Farber, D.B. 2001. Nrl and Sp nuclear proteins mediate transcription of rod-specific cGMP- phosphodiesterase beta-subunit gene: involvement of multiple response elements. *J Biol Chem* **276**: 34999-35007.
- Lipshutz, R.J., Fodor, S.P., Gingeras, T.R., and Lockhart, D.J. 1999. High density synthetic oligonucleotide arrays. *Nat Genet* **21**: 20-24.
- Liu, I.S., Chen, J.D., Ploder, L., Vidgen, D., van der Kooy, D., Kalnins, V.I., and McInnes, R.R. 1994. Developmental expression of a novel murine homeobox gene (Chx10): evidence for roles in determination of the neuroretina and inner nuclear layer. *Neuron* **13**: 377-393.
- Livesey, F.J. and Cepko, C.L. 2001. Vertebrate neural cell-fate determination: lessons from the retina. *Nat Rev Neurosci* **2**: 109-118.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., et al. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* **14**: 1675-1680.
- Martinez-Gimeno, M., Maseras, M., Baiget, M., Beneito, M., Antinolo, G., Ayuso, C., and Carballo, M. 2001. Mutations P51U and G122E in retinal transcription factor NRL associated with autosomal dominant and sporadic retinitis pigmentosa. *Hum Mutat* **17**: 520.
- Masland, R.H. 1996. Processing and encoding of visual information in the retina. *Curr Opin Neurobiol* **6**: 467-474.
- Masland, R.H. 2001. Neuronal diversity in the retina. *Curr Opin Neurobiol* **11**: 431-436.
- Mata, N.L., Radu, R.A., Clemmons, R.C., and Travis, G.H. 2002. Isomerization and oxidation of vitamin a in cone-dominant retinas: a novel pathway for visual-pigment regeneration in daylight. *Neuron* **36**: 69-80.
- McBee, J.K., Van Hooser, J.P., Jang, G.F., and Palczewski, K. 2001. Isomerization of 11-cis-retinoids to all-trans-retinoids in vitro and in vivo. *J Biol Chem* **276**: 48483-48493.
- Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., and Swaroop, A. 2001. Nrl is required for rod photoreceptor development. *Nat Genet* **29**: 447-452.
- Mitton, K.P., Swain, P.K., Chen, S., Xu, S., Zack, D.J., and Swaroop, A. 2000. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. *J Biol Chem* **275**: 29794-29799.

- Molday, R.S. 1998. Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases. The Friedenwald Lecture. *Invest Ophthalmol Vis Sci* **39**: 2491-2513.
- Nakatani, K., Chen, C., Yau, K.W., and Koutalos, Y. 2002. Calcium and phototransduction. *Adv Exp Med Biol* **514**: 1-20.
- Ng, L., Hurley, J.B., Dierks, B., Srinivas, M., Salto, C., Vennstrom, B., Reh, T.A., and Forrest, D. 2001. A thyroid hormone receptor that is required for the development of green cone photoreceptors. *Nat Genet* **27**: 94-98.
- Pan, W., Lin, J., and Le, C.T. 2002. How many replicates of arrays are required to detect gene expression changes in microarray experiments? A mixture model approach. *Genome Biol* **3**.
- Pavlidis, P., Li, Q., and Noble, W.S. 2003. The effect of replication on gene expression microarray experiments. *Bioinformatics* **19**: 1620-1627.
- Pepe, I.M. 2001. Recent advances in our understanding of rhodopsin and phototransduction. *Prog Retin Eye Res* **20**: 733-759.
- Polans, A., Baehr, W., and Palczewski, K. 1996. Turned on by Ca²⁺! The physiology and pathology of Ca(2+)-binding proteins in the retina. *Trends Neurosci* **19**: 547-554.
- Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D., and Brown, P.O. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23**: 41-46.
- Quackenbush, J. 2002. Microarray data normalization and transformation. *Nat Genet* **32 Suppl**: 496-501.
- Rehmtulla, A., Warwar, R., Kumar, R., Ji, X., Zack, D.J., and Swaroop, A. 1996. The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proc Natl Acad Sci U S A* **93**: 191-195.
- Rodieck, R.W. 1998. *The first steps in seeing*. Sinauer Associates, Sunderland.
- Saari, J.C., Garwin, G.G., Haeseleer, F., Jang, G.F., and Palczewski, K. 2000. Phase partition and high-performance liquid chromatography assays of retinoid dehydrogenases. *Methods Enzymol* **316**: 359-371.
- Saleem, R.A. and Walter, M.A. 2002. The complexities of ocular genetics. *Clin Genet* **61**: 79-88.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.

- Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., and Herzog, H. 2000. Normalization strategies for cDNA microarrays. *Nucleic Acids Res* **28**: E47.
- Senin, I., Koch, K.W., Akhtar, M., and Philippov, P.P. 2002. Ca²⁺-dependent control of rhodopsin phosphorylation: recoverin and rhodopsin kinase. *Adv Exp Med Biol* **514**: 69-99.
- Smyth, G.K., Yang, Y.H., and Speed, T.P. 2003. *Statistical issues in microarray data analysis. In: Functional Genomics: Methods and Protocols*. Humana Press, Totowa, NJ.
- Sterling, P. 1990. *Retina, in The Synaptic Organization of the Brain*. Oxford University Press, New York.
- Swain, P.K., Hicks, D., Mears, A.J., Apel, I.J., Smith, J.E., John, S.K., Hendrickson, A., Milam, A.H., and Swaroop, A. 2001. Multiple phosphorylated isoforms of NRL are expressed in rod photoreceptors. *J Biol Chem* **276**: 36824-36830.
- Tachibanaki, S., Tsushima, S., and Kawamura, S. 2001. Low amplification and fast visual pigment phosphorylation as mechanisms characterizing cone photoresponses. *Proc Natl Acad Sci U S A* **98**: 14044-14049.
- Turner, D.L. and Cepko, C.L. 1987. A common progenitor for neurons and glia persists in rat retina late in development. *Nature* **328**: 131-136.
- Tusher, V.G., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**: 5116-5121.
- Wright, K.W. 1997. *Textbook of ophthalmology*. Williams & Wilkins, Baltimore.
- Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**: e15.
- Young, R.W. 1985a. Cell differentiation in the retina of the mouse. *Anat Rec* **212**: 199-205.
- Young, R.W. 1985b. Cell proliferation during postnatal development of the retina in the mouse. *Brain Res* **353**: 229-239.

CHAPTER 3
ANNOTATION AND ANALYSIS OF 10,000 ESTS FROM DEVELOPING
MOUSE EYE AND ADULT RETINA

This chapter has been published in:

J. Yu, R. Farjo, S.P. MacNee, W. Baehr, D.E. Stambolian, A. Swaroop. *Annotation and analysis of 10,000 expressed sequence tags from developing mouse eye and adult retina*. Genome Biol. 2003;4(10):R65.

3.1 Abstract

As a biomarker of cellular activities, the transcriptome of a specific tissue or cell-type during development and disease is of great interest. For this thesis, I have generated and analyzed 10,000 expressed sequence tags (ESTs) from three cDNA libraries of mouse eye tissues: embryonic day 15.5 (M15E) eye, postnatal day 2 (M2PN) eye and adult retina (MRA). Annotation of 8,633 non-mitochondrial and non-ribosomal ESTs revealed that 57% represent known genes and 43% unknown or novel ESTs; M15E had the highest percentage of novel ESTs. Of these 8,633 ESTs, 2,361 correspond to 747 unique genes and the remaining 6,272 are present only once. Genes of phototransduction cascade are preferentially identified in MRA, whereas transcripts for cell structure and regulatory proteins are highly expressed in the developing eye. Map locations of human orthologs of known genes uncovered a high density of ocular genes on chromosome 17 and identified 277 genes in the critical regions of 37 retinal disease loci. *In silico* expression profiling identified 210 genes and/or ESTs over-expressed in the eye; of these, more than 26 are known to have vital retinal function. Comparisons between cDNA libraries provided a list of temporally regulated ESTs. A few of these were validated by qRT-PCR analysis. Our studies present a large number of potentially interesting genes for

biological investigation and the annotated EST set provides a useful resource for microarray and functional genomic studies.

3.2 Introduction

Recent efforts in genomics have accomplished the daunting task of decoding the genome of several species, including human (Lander et al. 2001; Venter et al. 2001) and mouse (Waterston et al. 2002). The current estimate of transcribed genes in the mammalian genome ranges from 35,000 to 45,000, with approximately 99% of mouse genes having homologs in the human genome. Many high-throughput genomics projects have now begun to focus on the identification of cell- and tissue-specific transcriptomes since such gene expression profiles are expected to uncover fundamental insights into biological processes (Swaroop and Zack 2002). To identify genes or cellular pathways that are selectively turned on or off in response to extrinsic factors or intrinsic genetic programs, it is necessary to deduce the catalogue of mRNAs expressed in a specific cell or tissue type at various stages of development, aging, and disease.

The vertebrate eye is a key component of the nervous system, the neural retina being responsible for the process of phototransduction - the signal transduction pathway by which light is converted into neural stimuli that result in perceived vision. A systematic evaluation of transcripts and their expression levels at different stages of eye or retinal development should lead to better understanding of underlying regulatory pathways of differentiation and functional maintenance. During the last decade, a number of approaches have been utilized to achieve these tasks. Serial analysis of gene expression (SAGE) provides a catalog of expressed genes of a given tissue through the sequencing of SAGE tags and quantitatively estimates transcript level based on the occurrence of corresponding tags (Cheng and Porter 2002; Jasper et al. 2002; Velculescu et al. 1995). SAGE cataloging has recently been reported for mouse and human retina

(Blackshaw et al. 2001; Sharon et al. 2002). One caveat of this technique is that tag-to-gene assignments can be ambiguous, since a specific transcript is identified by a short oligonucleotide sequence, usually a 14-20 bp SAGE tag. This is evaded in more traditional EST generation from cDNA libraries by obtaining a larger tag of 200-600 bp (Adams et al. 1991; Strachan et al. 1997). EST generation provides appreciable length of sequences of novel genes, which could be deposited into GenBank and complement public databases of ESTs (Kawamoto et al. 2000; Konno et al. 2001; VanBuren et al. 2002). Various computational approaches also rely on EST data to give validity to gene predictions, aide in the detection of functional alternatively spliced transcripts (Kan et al. 2002; Xu et al. 2002) and identify tissue-specific genes or candidate disease genes (Bortoluzzi et al. 2000; Katsanis et al. 2002; Sohocki et al. 1999). With the availability of microarray technology, slides containing a comprehensive set of ESTs that cover expressed genes of specific tissues during various developmental stages or represent specific cellular pathways, provide a powerful tool for systematic expression profiling, without repetitive sequencing.

A number of ESTs have been isolated from retina and retinal pigment epithelium (RPE) libraries (Bernstein et al. 1996; Buraczynska et al. 2002; Maubaret et al. 2002), from subtracted retina or RPE libraries (Gieser and Swaroop 1992; Sinha et al. 2000), or through computational manipulation and database mining (Bortoluzzi et al. 2000; Sohocki et al. 1999). EST generation can be especially valuable in characterizing uniquely expressed genes or identifying novel candidate genes for retinal disorders (Bernstein et al. 1995; Shimizu-Matsumoto et al. 1997; Wang et al. 1996). Recently, large-scale sequencing has been utilized to gain a clearer picture of the retinal transcriptome through the analyses of an embryonic day 14.5 retinal cDNA library (Mu et al. 2001) and a set of libraries constructed from different parts of the eye (Wistow 2002). Corresponding databases of retinal/eye ESTs, including RetinaExpress (Mu et al. 2001), RetBase (Katsanis et al. 2002) and the NEIbank database (Wistow 2002), have been

generated to serve as useful resources for eye transcriptome consolidation. In addition, microarrays containing small sets of ocular genes/ESTs have been manufactured and used for expression analysis (Farjo et al. 2002; Mu et al. 2001).

With a goal of obtain a set of ESTs with deep coverage of expressed genes of ocular tissues, we have generated, annotated, and analyzed over 10,000 ESTs derived from three cDNA libraries constructed from developing mouse eyes and adult retinas. These clones represent a significant resource for producing eye-specific cDNA microarrays (I-gene microarrays) for comprehensive gene profiling of mouse ocular development and of models of eye disease.

3.3 Materials and Methods

Library construction and cDNA sequencing

cDNA libraries were constructed from E15.5 eyes, PN2.5 eyes and adult retina, as previously described (Farjo et al. 2002). Plasmid clones were randomly selected, and colonies were inoculated into individual wells of 96-well plates containing 175 μ L LB media, covered with Breathe-Easy tape (ISC Bioexpress, Kaysville, UT, USA), and incubated at 37°C for 18-22 h. Frozen glycerol stocks were prepared by adding 77 μ L of 50% glycerol to each well, and the plate was stored at -80°C. Double-stranded cDNAs were obtained for sequencing either by miniprep (CONCERT Rapid Plasmid Miniprep System, Invitrogen, Carlsbad, CA, USA) or PCR amplification directly from frozen glycerol stocks, as described (Farjo et al. 2002). DNA sequencing from the 5' end of the cDNA insert was carried out using T7 primer with a high-throughput automated sequencer (Applied Biosystems Inc, Foster City, CA, USA) using standard protocols.

ESTs analysis and gene annotation

Raw sequences were first subjected to RepeatMasker (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>) and the repeat-masked sequence was used to query NCBI nr database with the BLAST algorithm (National Center for Biotechnology Information, Bethesda, MD, USA) (Altschul et al. 1990). Sequences matching to nr database entries with an E-value of ≤ 100 or less were classified as a positive BLAST result and RefSeq entries were preferentially selected where available since these have the greatest amount of annotation linked to them. Special consideration was given to BLAST results where our query matched a target DNA/genomic sequence over successive regions with E-values between ≤ 50 and ≤ 99 as this could represent an mRNA matching to different exons of the same gene. Further functional annotation of BLAST positive genes was performed using Perl/BioPerl scripts. In brief, accession numbers were utilized to query the Entrez nucleotide database (http://www.ncbi.nlm.nih.gov/blast/html/blastcgihelp.html#nucleotide_databases) and UniGene database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene>) for information including gene name, gene symbol, UniGene Cluster ID, LocusID, chromosome location and cDNA sources. LocusID was then utilized to query the LocusLink database (<http://www.ncbi.nlm.nih.gov/LocusLink>) for gene ontology information, and to search the human UniGene database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene>) for human homolog maps.

DNA sequences with no significant matches to NCBI nr database were further BLAST-searched against the mouse subset of dbEST (http://www.ncbi.nlm.nih.gov/blast/html/blastcgihelp.html#nucleotide_databases). Sequences matching to dbEST entries with E-value of ≤ 60 or less were considered as positive matches. The cDNA source tissues of all dbEST entries matching the sequence were obtained from the Entrez nucleotide database using Perl scripts.

To assess the redundancy of our clone sets, ESTs with known gene matches were grouped based on identical accession numbers. Unknown and novel ESTs were clustered using default parameters by the NCBI BLASTCLUST, a BLAST score-based single-linkage clustering script (<ftp://ftp.ncbi.nih.gov/blast/executables/>). Each cDNA sequence was also BLAST-searched against the collection of all sequences from the M15E, M2PN and MRA libraries using Standalone BLAST (<ftp://ftp.ncbi.nih.gov/blast/executables/>). Homology sequences with E-value of $\leq 10^{-60}$ or less were considered to be overlapping sequences belonging to one gene/EST. The quality of each sequence was examined according to objective criteria, such as sequence length and percentage of Ns (uncertain read; any of the four nucleotides) and As (adenosine) in the sequence. Integrity of each sequence was eventually determined based on these parameters, with necessary manual analysis. Basically, sequences shorter than 200 bp, with over 30% of Ns or over 50% of As were considered as low quality. Manual analysis was applied to sequences longer than 500 bp but with high-percentage of Ns or As.

Quantitative real-time PCR (qRT-PCR)

Total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) and treated with RQ1 RNase-Free DNase (Promega, Madison, WI, USA) to remove genomic DNA contamination. First-strand cDNA (+RT sample) was synthesized by reverse transcription of 2.5 μg of the total RNA using oligo-d(T) primers. Negative control sample (-RT) was obtained by incubating 2.5 μg of the total RNA from the same pool without reverse transcriptase. PCR primers were designed using the PRIMER3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). qRT-PCR was performed in an iCycler IQ real-time PCR Detection system (Bio-Rad, Hercules, CA, USA), and the thermal cycling condition was 3 minutes at 95°C, followed by 45 cycles of 95°C for 30 seconds, 57°C for 30 seconds and 72°C for 30 seconds. SYBR

Green (Molecular Probes, Eugene, OR, USA) was added into each reaction for the detection of fluorescence during amplification. PCR reactions from both +*RT* and -*RT* samples were performed in triplicate, and control reactions of *Hprt* were performed on each template to normalize the amount of cDNA present in each reaction. All reaction products were verified with melt curve analysis and agarose gel electrophoresis.

3.4 Results

Characterization of three mouse cDNA libraries

Three cDNA libraries, namely the M15E, M2PN and MRA libraries, were constructed using total RNAs from mouse embryonic day 15.5 (E15.5) eyes, postnatal day 2.5 (PN2.5) eyes and adult retinas, respectively (Farjo et al. 2002). No RNA or library amplification or normalization was applied to any of these libraries. A total of 11,057 cDNA clones (4,992 from M15E, 4,128 from M2PN and 1,937 from MRA) were randomly isolated and sequenced to generate ESTs of approximately 500 bp from the 5' end. All ESTs are downloadable from our website (<http://www.umich.edu/~igene>). We did not obtain high quality sequences for 1,419 clones: of these, 848 were from the M15E library (Table 3.1). Further evaluation showed that 577 out of these 848 clones were from the first 24 plates we sequenced from the M15E library, whereas only 134 were from the later 30 plates. Low quality sequences, perhaps caused by sequencing cross-talk, were referred as questionable (group VI, 345 clones), while others, including short sequences, and those with a high-percentage of nucleotide As or Ns and vector sequences, were classified as uninformative (group VII, 1,074 clones). All ESTs of group VI and VII were excluded from further analysis.

A total of 9,638 high quality ESTs (GenBank accession numbers CB839918-CB850489) were compared to the NCBI nr database and the mouse dbEST

Table 3.1. Summary of ESTs.

cDNA category	Library					
	M15E		M2PN		MRA	
	# ESTs	(% ESTs)	# ESTs	(% ESTs)	# ESTs	(% ESTs)
High quality ESTs						
I. Known ESTs	1956	(47.2)	2047	(55.3)	970	(54.1)
II. Unknown ESTs	463	(11.2)	557	(15.0)	299	(16.7)
III. Novel ESTs	1200	(29.0)	792	(21.4)	349	(19.5)
IV. Mitochondrial DNA	248	(6.0)	202	(5.5)	132	(7.4)
V. Ribosomal RNA, protein	277	(6.7)	103	(2.8)	43	(2.4)
Subtotal	4144	(100)	3701	(100)	1793	(100)
VI. Questionable sequences ¹	137		143		65	
VII. Uninformative sequences ¹	711		284		79	
Total	4992		4128		1937	

¹ Sequences in these categories were not deposited in the GenBank database and were excluded from further analysis.

(http://www.ncbi.nlm.nih.gov/blast/html/blastcgihelp.html#nucleotide_databases) for homology identification. Based on the analyses, they were categorized into five groups (Table 3.1). Those with significant homology in NCBI nr databases and representing well-characterized cDNAs were defined as known ESTs (Group I, 4,973 clones). There are 1,956, 2,047 and 970 known ESTs in the M15E, M2PN and MRA libraries, respectively. However, Group I did not include genes of mitochondrial origin or those coding for ribosomal proteins, which were separately classified into group IV and V, respectively. Group II (unknown ESTs, 1,319 clones) refers to ESTs with no match in the NCBI nr database, but with matches in the mouse dbEST. ESTs that have no homology to sequences in both databases were considered as novel ESTs (Group III, 2,341 clones). Approximately 50% of ESTs in each library correspond to known ESTs and 15% to unknown ESTs (Group II). The M15E library has the lowest number in these two categories, reflecting the fact that E15.5 time-point is not as well studied. The M15E library also contains the largest fraction of novel ESTs (29%), relative to 21.4% in M2PN and 19.5% in MRA, demonstrating the under-representation of E15.5 eye ESTs in public databases.

Gene annotation

To gain a better understanding of our EST sets, we performed a series of analyses (Figure 3.1). ESTs were masked to eliminate repeat sequences and examined for high quality. A number of sequence similarity searches were executed to compare every EST to those in public or in our local databases. For ESTs with known-gene matches in public databases, functional annotation was retrieved from NCBI UniGene (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene>) and LocusLink (<http://www.ncbi.nlm.nih.gov/LocusLink/>). Of the 4,973 known ESTs, 65% show

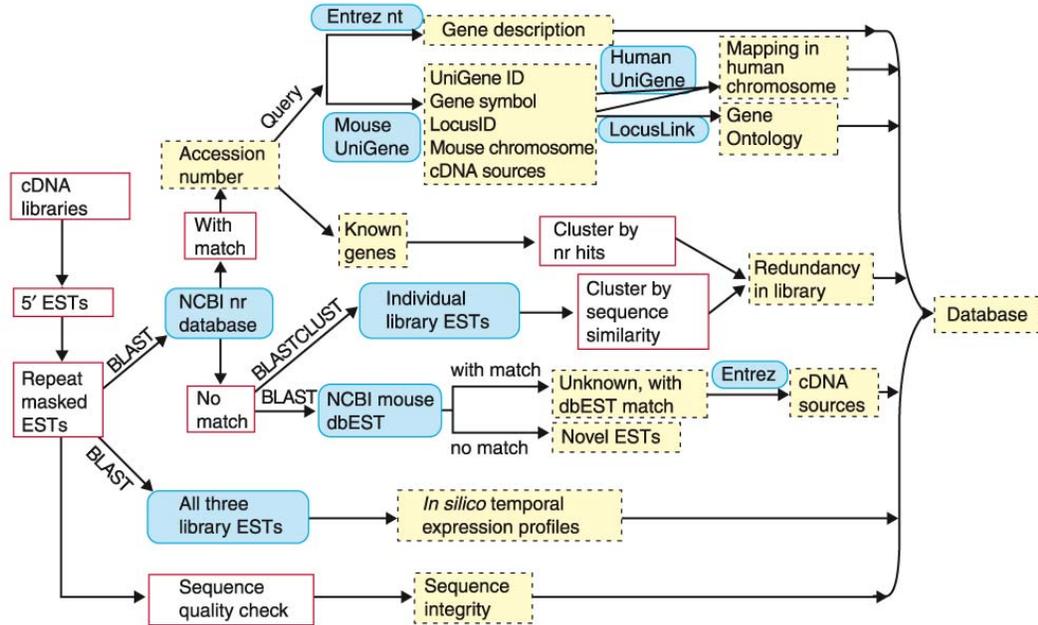


Figure 3.1. Schematic representation of the EST analysis and annotation processes. ESTs from the 5' end of cDNAs were repeat masked and checked for sequence integrity. They were BLAST-searched against the NCBI nr database, the mouse dbEST and our local database of EST sets from each library (individual library ESTs) or the entire collection of 9,638 ESTs (all three library ESTs). High-level functional annotation was achieved by searching the NCBI databases, including Entrez nt, LocusLink, mouse and human UniGene database. Databases that were BLASTed or queried are indicated by solid blue, rounded rectangles. Annotated data (highlighted by dashed, yellow rectangles) for every EST can be accessed at (<http://www.umich.edu/~igene>).

corresponding UniGene and LocusLink entries, and 48% have matching human orthologs and identified chromosomal locations.

To assist in the analyses of tissue expression patterns of the 1,319 unknown ESTs, we retrieved from the NCBI Entrez nucleotide database (http://www.ncbi.nlm.nih.gov/blast/html/blastcgihelp.html#nucleotide_databases) all dbEST entries with significant homology to our clones and determined their original tissue of isolation. Cross-comparison of every EST to the complete set of ESTs in our local database provided information regarding their redundancy and level of expression in each of the three libraries. Gene annotation, along with sequence quality assessment, was recorded in a local database with hyperlinks to NCBI Entrez nucleotide, UniGene and LocusLink databases (Figure 3.2). A complete list of all clones and corresponding annotations is available on our website (<http://www.umich.edu/~igene>).

EST clustering

To estimate the number of unique transcripts represented by our ESTs, known, unknown or novel ESTs were clustered within each library. We did not consider redundancy between libraries, as our purpose was to assess the amount of redundant sequences within an individual library and to compare EST frequency between libraries. ESTs matching known genes were placed in groups based on common nucleotide hits in the NCBI nr database. The results showed 3,604 unique genes (1,382 in M15E, 1,443 in M2PN and 779 in MRA) representing these 4,973 ESTs. Less than 50 clusters contained more than five ESTs and a majority (3,016) consisted of a single EST (Figure 3.3a). The largest cluster in the M15E, M2PN and MRA libraries contained, respectively, 45, 50 and 37 ESTs corresponding to the same gene. Table 3.2 includes highly-expressed genes of each library. Crystallin genes were excluded from this list since the inclusion of lens in the sample RNA used for library construction could bias their abundance in the M15E

Clone ID	Length	PCT_N	PCT_A	Integr.	Description
MRA-0007	319	0.00	0.28	ok	Mus musculus paired box gene 6 (Pax6), mRNA
MRA-0009	335	0.00	0.29	ok	Mouse mRNA for elongation factor 1-alpha (EF 1-alpha)
MRA-0012	314	0.00	0.21	ok	Mus musculus retinal S-antigen (Sag), mRNA
MRA-0015	162	0.00	0.22	ok	Mus musculus cyclin ania-6b gene, partial sequence
MRA-0016	270	0.00	0.21	ok	unknown, with dbEST match
MRA-0017	197	0.00	0.22	ok	Mus musculus phosphatidylethanolamine binding protein
MRA-0019	249	0.00	0.30	ok	unknown, with dbEST match
MRA-0026	220	0.00	0.23	ok	Mus spretus E6-AP ubiquitin-protein ligase (Ube3a) mRNA
MRA-0027	333	0.00	0.23	ok	Mus musculus huntington yeast partner C (Hypc) mRNA,
MRA-0028	337	0.00	0.17	ok	Mus musculus pyruvate kinase 3 (Pk3), mRNA
MRA-0029	365	0.00	0.23	ok	Mus musculus RNA polymerase II 3 (Rpo2-3), mRNA

Clone ID	GenBank	Symbol	UniGene	LocusLink	Hs. H. map	Chr.	Molecular Function
MRA-0007	NM_013627	Pax6	Mm.3608	18308	11p13	2	transcription factor,
MRA-0009	X13661	Eef1a1	Mm.196614	13627	6q14.1	19	GTP binding,translational elon
MRA-0012	NM_009118	Sag	Mm.1276	20215	2q37.1	1	calcium ion binding,
MRA-0015	AF185591						
MRA-0016	unknown, with dbEST match						
MRA-0017	U43206	Pbp	Mm.195898	23980	12q24.23	5	lipid binding,
MRA-0019	unknown, with dbEST match						
MRA-0026	AF082835						
MRA-0027	AF135440	2610317D2	Mm.102104	54614	12q	15	
MRA-0028	NM_011099	Pk3	Mm.2635	18746	15q22	9	kinase,transferase,pyruvate 1
MRA-0029	NM_009090	Rpo2-3	Mm.2186	20021	16q13-q21	11	DNA binding,transferase,DN

Clone ID	Biological Process	Cellular Components	Tissue Expressed
MRA-0007	brain development,eye morphoge	transcription factor complex,	cerebellum;eye;brain;nervous system;olfactor
MRA-0009	protein biosynthesis,translational elongation,		mammary;lung;embryo, whole embryo;kidne
MRA-0012	vision,sensory perception,signal transduction,		eye,adult-retina;eyeball;retina;pineal-glands;u
MRA-0015			
MRA-0016			
MRA-0017			testis;mammary;lung;embryo, whole embryc
MRA-0019			
MRA-0026			
MRA-0027			mammary;nervous system;lung;whole body;ba
MRA-0028	glycolysis,		mammary;lung;eye;embryo, whole embryo;k
MRA-0029	transcription,	nucleus,DNA-directed RNA po	cerebellum;embryo, whole embryo;hippocam

Figure 3.2. Screen shot of I-gene database of annotated ESTs.

Clones are hyperlinked to their sequences. GenBank, UniGene and LocusLink IDs are hyperlinked to the corresponding web page. The integrity of each sequence was assessed by sequence length, percentage of Ns (PCT_N) and percentage of As (PCT_A) contained. Arrows are used to indicate continuation of the table for each clone ID.

Table 3.2. Highly expressed genes in the M15E, M2PN and MRA libraries, excluding crystallin genes.

Accession #	Gene Name	#Occurrences	%Total
M15E Library			
NM_008218	Hemoglobin alpha, adult chain 1 (Hba-a1)	45	2.30
M22432	Protein synthesis elongation factor Tu (eEF-Tu, eEf-1-alpha)	30	1.53
NM_008220	Hemoglobin, beta adult major chain (Hbb-b1)	22	1.12
NM_011653	Tubulin alpha 1 (Tuba1)	13	0.66
NM_008084	Glyceraldehyde-3-phosphate dehydrogenase (Gapd)	9	0.46
NM_008302	Heat shock protein, 84 kDa 1 (Hsp84-1)	8	0.41
NM_008972	Prothymosin alpha (Ptma)	6	0.31
NM_025586	RIKEN cDNA 2510008H07 gene (2510008H07Rik)	6	0.31
NM_026055	RIKEN cDNA 2810465O16 gene (2810465O16Rik)	6	0.31
K01173	Dog (canine) chymotrypsin	5	0.26
NM_007393	Actin, beta, cytoplasmic (Actb)	5	0.26
NM_011664	Ubiquitin B (Ubb)	5	0.26
NM_023119	Enolase 1, alpha non-neuron (Eno1)	5	0.26
NM_023123	H19 fetal liver mRNA (H19)	5	0.26
NM_025379	cytochrome c oxidase subunit VIIb (Cox7b)	5	0.26
NM_053075	RAS-homolog enriched in brain (Rheb)	5	0.26
M2PN Library			
M22432	Protein synthesis elongation factor Tu (eEF-Tu, eEf-1-alpha)	12	0.59
NM_011664	Ubiquitin B (Ubb)	10	0.49
NM_010240	Ferritin light chain 1 (Ftl1)	9	0.44
NM_009751	Beaded filament structural protein in lens-CP94 (Bfsp1)	8	0.39
NM_009242	Secreted acidic cysteine rich glycoprotein (Sparc)	6	0.29
NM_021278	Thymosin, beta 4, X chromosome (Tmsb4x)	6	0.29
AK013100	clone:2810417D04	5	0.24
BE657894	GM700004A10F4 Gm-r1070 Glycine max cDNA clone Gm-r1070-1399	5	0.24
NM_007393	Melanoma X-actin (Actx)	5	0.24
NM_007687	Cofilin 1, non-muscle (Cfl1)	5	0.24
NM_008084	Glyceraldehyde-3-phosphate dehydrogenase (Gapd)	5	0.24
NM_018785	Formin binding protein 3 (Fnbp3)	5	0.24
MRA Library			
BC013125	Rhodopsin (Rho)	37	3.81
NM_008084	Glyceraldehyde-3-phosphate dehydrogenase (Gapd)	10	1.03
NM_008140	Guanine nucleotide binding protein, alpha transducing 1 (Gnat1)	10	1.03
NM_009118	Retinal S-antigen (Sag)	6	0.62
NM_008131	Glutamine synthetase (Glns)	5	0.52
NM_010480	Heat shock protein, 86 kDa 1 (Hsp86-1)	5	0.52
NM_011676	Unc119 homolog (C. elegans) (Unc119h)	5	0.52
X69523	R.rattus mRNA for interphotoreceptor retinoid-binding protein (Rbp3)	5	0.52
AF105711	Tubby like protein 1 (Tulp1)	4	0.41
L08075	Domesticus phosducin	4	0.41
NM_008189	Mus musculus guanylate cyclase activator 1a (retina) (Guca1a)	4	0.41
NM_008938	Mus musculus peripherin 2 (Prph2)	4	0.41

and M2PN libraries. In accordance with previous observations (Mu et al. 2001), translation factors, cell structure/cytoskeletal proteins and housekeeping genes are among the most abundant, especially in the M15E and M2PN libraries. In the MRA library, phototransduction genes are among those highly expressed, consistent with the primary functional responsibility of the mature retina. Of the 12 most abundant genes, ten are known to play important roles in retinal function; these include rhodopsin (37 occurrences), alpha-transducin 1 (Gnat1) (nine occurrences), arrestin (Sag) (six occurrences), glutamine synthetase (five occurrences), unc119 homolog (Unc119h, Hrg4) (five occurrences), interphotoreceptor retinoid-binding protein (five occurrences) and four occurrences each of tubby like protein 1 (Tulp1), phosducin, guanylate cyclase activator 1a (Guca1a) and peripherin 2 (Rds). We conclude that genes/ESTs from these three cDNA libraries are representative of their source tissues and their redundancy could be utilized to construct *in silico* expression profiles.

Unknown and novel ESTs were clustered based on sequence similarity. A total of 159 clusters, composed of 404 ESTs, were generated, while the remaining 3,256 ESTs did not cluster, resulting in a total of 3,415 (93%) non-redundant clusters for the total 3,660 ESTs (Figure 3.3b). Only five ESTs from all three libraries have redundancy that is higher than three. Therefore, the number of unique transcripts in our clone set, including all known, unknown and novel ESTs, was estimated to be up to 7,019 (81% of all clustered genes/ESTs), with 2,909 from M15E, 2,705 from M2PN and 1,405 from the MRA library. This estimation excludes mitochondrial, ribosomal and low-quality sequences. The redundancy of known ESTs in the libraries is relatively higher than that of unknown and novel ESTs, which may reflect easier identification of abundantly expressed genes, near completion of genomic sequencing in human and mouse, and recent focus on functional characterization.

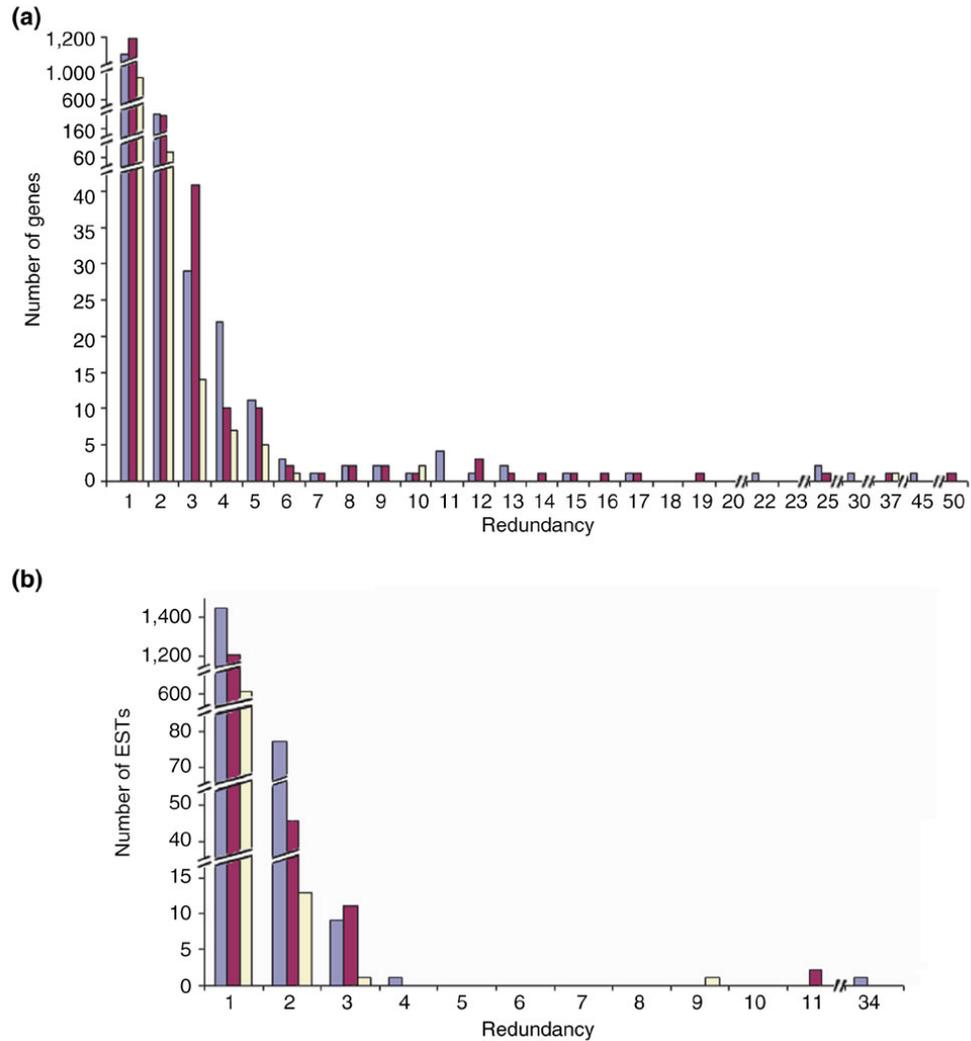


Figure 3.3. Histograms showing the number of genes or ESTs at each level of redundancy in the three libraries. Redundancy of **(a)** known genes and **(b)** unknown or novel ESTs are shown for the M15E (blue), M2PN (purple) and MRA (yellow) libraries.

Functional distribution of known ESTs

Of the 4,973 ESTs matching to known genes in the nr database, 1,142 correspond to cDNAs with unspecified function and 482 have homology with genomic sequences. The remaining 3,349 known ESTs can be divided into ten groups based on their putative functions (Figure 3.4a). The largest two groups include proteins involved in cell signaling/communication (17%) and those involved in protein expression/regulation/processing (15%). Following these were functional groups including crystallin-family genes (14%), and those involved in cell structure/motility/extracellular matrix (13%) and gene regulation/transcription factors (12%).

A detailed breakdown of functional groups by cDNA library (Figure 3.4b,c,d) demonstrated that phototransduction-specific genes are significantly more abundant in the MRA library, in concordance with functional activity at this stage. We also observed that crystallin genes are highly represented in the M15E and M2PN libraries, which is due to the fact that these two libraries were constructed from whole eye, while MRA was constructed from retina only. In general, the fraction of genes devoted to cellular functions in other categories did not deviate significantly from the overall pattern. At developmental stages, more genes appear to be involved in cell structure/motility/extracellular matrix (14%) than in mature retina (9%). There is also a slight decrease in genes for gene regulation/transcription and for protein expression/regulation/processing in the adult retina. In contrast, mature retina has more genes involved in metabolism (10%) compared to 6-7% in developmental libraries. Overall, the analysis of functional distribution of known ESTs in the three libraries is consistent with the established developmental and functional role at specific stages.

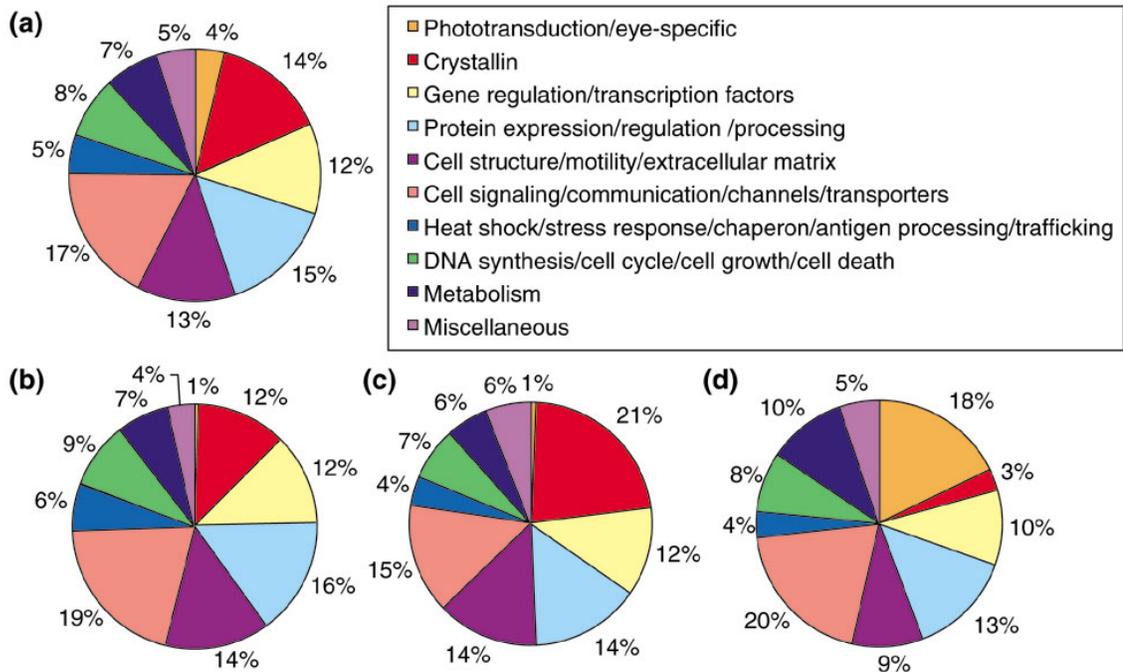


Figure 3.4. Functional categorization of ESTs with known-gene match.

(a) Categorization for the total 3,349 known ESTs across the three libraries, **(b)** the breakdown for M15E, **(c)** the breakdown for M2PN and **(d)** the breakdown for the MRA library. The number next to each category indicates the percentage of that particular group in the total ten classes.

Candidate ocular disease genes

To identify new candidate genes for ocular, in particular retinal, diseases, we have analyzed the chromosomal locations of all ESTs with known human gene matches. We first obtained the chromosomal location for each mouse gene, followed by the determination of the location of corresponding human ortholog if available. Of the 3,604 unique known genes, we were able to obtain mouse chromosomal locations for 1,964 genes, 1,522 of which have human orthologs and mapping information based on the UniGene database. Distribution by chromosome of these genes in mouse and human genome is given in Table 3.3. Expected genes per human chromosome were computed based on the Human Gene Map database (Deloukas et al. 1998). Chi-square analysis indicated that the observed frequency of ocular genes by chromosome deviates significantly from the expectation for all three libraries ($\chi^2 = 72.2, 31.9, 67.46$ for the M15E, M2PN and MRA libraries, respectively). The ratio of observed versus expected frequency by chromosomes revealed that the ocular gene density on chromosome 17 is significantly higher than expected ($p < 0.002$), but there appears to be no clustering in any specific chromosomal region. The X-chromosome has a marginally higher density of ocular genes from the M2PN and MRA libraries, whereas chromosomes 4 and 10 have marginally fewer ocular genes.

The RetNet database (<http://www.sph.uth.tmc.edu/Retnet/>) was searched to determine whether any of these genes might serve as candidates for mapped but as yet uncloned retinal disease loci. Of the 2,358 human ortholog ESTs, 641 ESTs, representing 277 non-redundant genes, were localized to one of the mapped retinal disease regions. A complete list of these genes is available in Additional data file 1 (<http://www.umich.edu/~igene>). Table 3.4 summarizes the number of genes mapping into each retinal disease locus, with one example gene listed. ESTs are localized to the cytogenetic locations of a total of 37 retinal disease loci. The RP26 region includes 142 ESTs, but this may be due to the fact

Table 3.3. Ocular gene distribution in mouse and human chromosomes.

Chr.	Mouse Genes			Human Ortholog of mouse ocular genes								
				Observed			Expected			Chi-square ²		
	M15E	M2PN	MRA	M15E	M2PN	MRA	M15E	M2PN	MRA	M15E	M2PN	MRA
1	47	49	29	53	52	26	64.2	60.7	32.7	1.95	1.24	1.38
2	61	71	29	51	51	19	46.5	44.0	23.7	0.43	1.12	0.94
3	44	39	23	36	31	16	41.5	39.3	21.2	0.74	1.74	1.26
4	61	36	21	20	19	12	30.5	28.8	15.5	<u>3.60</u>	<u>3.33</u>	0.80
5	46	35	21	27	36	7	31.5	29.8	16.1	0.65	1.30	5.11
6	36	38	19	23	28	33	39.0	36.9	19.9	6.58	2.14	8.64
7	66	61	26	27	28	6	32.9	31.1	16.8	1.05	0.30	<u>6.90</u>
8	40	36	21	20	25	15	24.9	23.5	12.7	0.95	0.10	0.43
9	39	48	23	25	19	10	25.7	24.3	13.1	0.02	1.16	0.74
10	39	41	30	19	19	4	28.3	26.7	14.4	<u>3.04</u>	2.23	<u>7.52</u>
11	85	62	42	44	35	19	36.2	34.2	18.4	1.69	0.02	0.02
12	24	16	14	33	35	21	32.7	30.9	16.7	0.00	0.55	1.14
13	21	29	12	14	10	5	14.5	13.7	7.4	0.02	1.00	0.77
14	24	33	16	22	22	18	21.6	20.4	11.0	0.01	0.13	<u>4.46</u>
15	28	28	23	24	23	15	21.2	20.0	10.8	0.37	0.44	1.63
16	20	16	14	31	16	7	17.5	16.5	8.9	10.40	0.02	0.41
17	41	31	25	48	34	25	26.0	24.6	13.3	18.51	<u>3.58</u>	10.36
18	16	14	3	3	7	2	10.8	10.2	5.5	<u>5.61</u>	1.00	2.22
19	33	26	18	25	28	18	23.0	21.7	11.7	0.18	1.82	<u>3.38</u>
20	NA	NA	NA	18	19	7	15.6	14.8	8.0	0.36	1.22	0.12
21	NA	NA	NA	13	7	5	6.3	5.9	3.2	<u>7.14</u>	0.19	1.00
22	NA	NA	NA	21	15	9	11.6	11.0	5.9	<u>7.52</u>	1.45	1.58
X	27	30	18	23	27	17	18.0	17.0	9.2	1.38	<u>5.84</u>	<u>6.65</u>
Total	798	739	427	620	586	316	620	586	316	72.20	<u>31.90</u>	67.46

² Chi-square tests with 1 degree of freedom ($df = 1$) for each chromosome and $df = 22$ for total were applied to compare the observed distribution of ocular genes with expected distribution calculated based on Human Gene Map. Significant Chi-square values with $p < 0.002$ are indicated in bold. Marginally significant Chi-square values with $p < 0.1$ were underscored, if genes from all 3 libraries showed a similar trend.

Table 3.4. Candidate ocular disease genes identified in three mouse retina / eye cDNA libraries.

Disease	Location	#ESTs in interval	#Unique Genes	Example Accession #	Gene Name
AA	11p15	60	19	NM_015814	Dickkopf homolog 3 (Xenopus laevis)
ARMD1	1q25-q31	17	8	BC023001	Regulator of G-protein signaling 2
AXPC1	1q31-q32	18	9	NM_008131	Glutamine synthetase (Glns)
BBS5	2q31	8	4	AK019448	Procollagen, type III, alpha 1
BCD	4q35-qter	2	1	U27315	Adenine nucleotide translocase-1 (Ant1)
CACD	17p13	26	16	BC008093	Eukaryotic translation initiation factor 5A Fragile X mental retardation syndrome 1
COD2, XLPCD	Xq27	2	1	NM_008031	homolog (Fmr1)
CORD1	18q21.1-q21.3	2	1	NM_009190	Vacuolar protein sorting 4b (yeast)
CORD7	6q	83	29	NM_007865	Delta-like 1 (Drosophila) SEC22 vesicle trafficking protein-like 1 (S. cerevisiae)
CORD8	1q12-q24	30	20	NM_011342	cerevisiae)
CORD9	8p11	5	3	NM_031158	Ankyrin 1, erythroid
CYMD	7p21-p15	20	10	AK002910	Chromobox homolog 3 (Drosophila HP1 gamma)
EVR3	11p13-p12	9	6	NM_013627	Paired box gene 6 (Pax6) C. elegans ceh-10 homeo domain containing
LCA3	14q24	11	8	NM_007701	homolog
LCA5	6q11-q16	45	6	NM_009945	Cytochrome c oxidase, subunit VIIa 3 (Cox7a3)
MCDR1	6q14-q16.2	41	4	AK019500	NS1-associated protein 1
MRST	15q24	12	6	NM_024431	Testis expressed gene 189
OPA2	Xp11.4-p11.2	26	10	NM_009457	Ubiquitin-activating enzyme E1, Chr X
PRD	Xp11.3-p11.23	19	8	NM_013680	Synapsin I
RCD1	6q25-q26	8	3	AF147785	Zinc finger protein ZAC1 (Zac1)
RNANC	10q21	3	2	X16461	Cell division cycle 2 homolog A (S. pombe)
RP17	17q22	10	4	NM_009296	Suppressor of Ty 4 homolog (S. cerevisiae)
RP22	16p12.1-p12.3	6	4	NM_007672	Cerebellar degeneration-related 2
RP24	Xq26-q27	4	2	NM_016697	Glypican 3
RP25	6cen-q15	45	6	NM_009945	Cytochrome c oxidase, subunit VIIa 2
RP26	2q31-q33	142	20	NM_022988	Ngg1 interacting factor 3-like 1 (S. pombe)
RP28	2p11-p16	26	17	NM_009837	Chaperonin subunit 4 (delta) (Cct4)
RP29	4q32-q34	4	3	NM_025436	Sterol-C4-methyl oxidase-like (Sc4mol)
STGD4	4p	9	7	NM_013457	Adducin 1 (alpha) DEAD/H (Asp-Glu-Ala-Asp/His) box
USH1A, USH1	14q32	13	6	NM_020494	polypeptide 13 (RNA helicase A)
USH1E	21q21	2	1	BC013562	Similar to E4tf1-60 transcription factor SMT3 (supressor of mif two, 3) homolog 2 (S. cerevisiae)
USH1G	17q24-q25	39	26	AK012619	Ubiquitin-conjugating enzyme E2E 1, UBC4/5
USH2B	3p24.2-p23	4	3	NM_009455	homolog (yeast)
USH2C	5q14-q21	7	5	NM_010151	Nuclear receptor subfamily 2, group F, member 1 Ras and a-factor-converting enzyme 1 homolog
VRNI	11q13	45	24	NM_023131	(S. cerevisiae) (Rce1)
WFS2	4q22-q24	13	9	NM_007917	Eukaryotic translation initiation factor 4E (Eif4e)
WGN1, ERVR	5q13-q14	13	9	AY037837	Single-stranded DNA binding protein 2

that several crystallin genes with high abundance in our libraries are localized in this region.

***In silico* expression profiling**

Since the cDNA clones were randomly selected for sequencing, their relative abundance in an unamplified library may represent the corresponding transcript levels.

Temporal expression profiles: To determine expression level of ESTs in the M15E, M2PN and MRA libraries, each sequence was queried against our entire collection of 9,638 sequences using a local BLAST database. For every query, the number of matching ESTs from the three libraries was recorded. Relative frequencies of ESTs were computed by totaling the number of matching genes/ESTs within this library (number of occurrences) and dividing it by the total number of ESTs sequenced (excluding low-quality sequences) for this library. *In silico* temporal expression profiles were then constructed for each EST based on their relative frequencies in the three libraries. In Table 3.5, we selectively show a number of unknown ESTs with age-restricted patterns of expression. As expected, phototransduction genes are highly represented in the MRA library (data not shown). Rhodopsin and its homologs represent 1.3% of MRA clones, but are absent in the other two libraries. We observed a number of genes/ESTs that are highly expressed in both the M2PN and MRA libraries and are restricted to these postnatal stages. These include translation repressor NAT1 (0% in M15E, 2.0% in M2PN and 1.4% in MRA), hairy and enhancer of split 6 (0% in M15E, 2.0% in M2PN and 1.5% in MRA; data not shown), and a number of unknown ESTs (Table 3.5). These genes may be specifically relevant to postnatal ocular/retinal functions. Many members of crystallin genes are highly expressed in the M2PN library. Considering both M15E and M2PN were constructed from eye tissues, the greater percentage of these genes in M2PN may reflect the development of lens at PN2.

Table 3.5. *In silico* temporal expression profiles of selected ESTs.

Clone ID	Accession Number	Number in M15E ³	Number in M2PN	Number in MRA	% in M15E	% in M2PN	% in MRA
M2PN-0339	CB844781	0	74	26	0	2.00	1.45
M2PN-0448	CB844886	1	74	25	0.02	2.00	1.39
M2PN-0649	CB845079	0	74	26	0	2.00	1.45
M2PN-0722	CB845148	0	74	26	0	2.00	1.45
M2PN-0316	CB844758	0	73	27	0	1.97	1.51
M2PN-0376	CB844818	0	73	27	0	1.97	1.51
M2PN-0391	CB844833	0	73	27	0	1.97	1.51
M2PN-0438	CB844876	0	73	27	0	1.97	1.51
M2PN-0544	CB844978	0	73	27	0	1.97	1.51
MRA-0314	CB848782	0	61	39	0	1.65	2.18
MRA-0545	CB848901	0	66	34	0	1.78	1.90
MRA-0572	CB848921	0	68	32	0	1.84	1.78
MRA-0841	CB849156	0	68	32	0	1.84	1.78
MRA-0096	CB850344	0	69	31	0	1.86	1.73
MRA-0121	CB850363	0	70	30	0	1.89	1.67
MRA-0322	CB848790	0	70	30	0	1.89	1.67
MRA-0912	CB849219	0	70	30	0	1.89	1.67
MRA-1033	CB849321	0	70	30	0	1.89	1.67
MRA-1085	CB849361	0	70	30	0	1.89	1.67
MRA-0157	CB850381	0	71	29	0	1.92	1.62
MRA-1648	CB849895	0	0	1	0	0	0.06
M15E-2659	CB842026	6	0	0	0.14	0	0
M15E-2778	CB842141	6	1	1	0.14	0.03	0.06
M2PN-1529	CB845822	0	5	0	0	0.14	0
M2PN-2316	CB846570	0	5	0	0	0.14	0
M2PN-2533	CB846779	0	5	0	0	0.14	0
M2PN-3030	CB847254	0	5	0	0	0.14	0
MRA-1021	CB849309	0	0	17	0	0	0.95
MRA-1029	CB849317	0	0	16	0	0	0.89
MRA-1028	CB849316	0	0	7	0	0	0.39
MRA-1408	CB849664	0	0	5	0	0	0.28

³ Number of occurrence of each EST in the three libraries are indicated. Percentage of every EST in each library was estimated as the number of occurrence divided by the total number of high-quality ESTs generated from that library.

Ribosomal proteins are found to be more abundant in the M15E library, consistent with previous observations in the E14.5 library (Mu et al. 2001). Examples of ESTs preferentially expressed at specific time-points are listed in Table 3.5, while *in silico* temporal expression profiles of all ESTs are available at our website (<http://www.umich.edu/~igene>).

Tissue expression profiles: To identify potential retina-enriched genes, cDNA sources of known ESTs and tissue origins of homologs of unknown ESTs were examined. Of 3,189 known ESTs (1,290 from M15E, 1,259 from M2PN and 638 from MRA) with cDNA sources annotated, we have identified 110 unique genes that are expressed preferentially in eye, retina, pineal and brain tissues, including 28 from M15E, 30 from M2PN and 52 from MRA. Roughly half (26) of the retina-enriched genes from the MRA library are known to be involved in retinal function. Table 3.6 lists these potential retina-enriched genes. This table excludes known phototransduction genes and crystallins.

In the case of unknown ESTs, those matching predominantly to ESTs from ocular tissues were considered as eye-enriched (Table 3.7). We examined the feasibility of this approach using sequences for rhodopsin, unc119h and Rlbp1 (Cralbp) (cellular retinaldehyde-binding protein 1). The rhodopsin sequence has 250 homology ESTs in the NCBI mouse dbEST database, all of which were isolated from retina-related tissues. We also observed that 34 out of 68 of unc119h and 19 out of 23 of Rlbp1 homology ESTs were identified in ocular tissues. A total of 100 unknown ESTs from our study were found to be ocular specific, as over 70% of their homologous ESTs in the NCBI nucleotide database were originally isolated from ocular tissue. A complete list of tissue expression patterns of unknown ESTs is available at our website (<http://www.umich.edu/~igene>).

Table 3.6. Retina/eye-enriched genes, excluding crystallin and phototransduction genes.

Clone ID	Gene Description	GenBank
M15E-4178	adult male testis cDNA, RIKEN clone:4930517J16 cat eye syndrome chromosome region, candidate 6 homolog (human)	AK015815
M15E-2565	(Cecr6)	NM_033567
M15E-2026	ciliary neurotrophic factor receptor (Cntfr)	NM_016673
M15E-0149	major intrinsic protein of eye lens fiber (Mip)	NM_008600
M15E-1025	silver (Si)	NM_021882
M15E-5154	sine oculis-related homeobox 6 homolog (Drosophila) (Six6)	NM_011384
M15E-2513	SNRPN upstream reading frame (Snurf)	NM_033174
M15E-3173	testis protein TEX13 (Tex13)	AF285576
M15E-1676	troponin C, cardiac/slow skeletal (Tncc)	NM_009393
M2PN-2684	adult male testis cDNA, RIKEN clone:4930579P15	AK016334
M2PN-0941	adult retina cDNA, RIKEN clone:A930007D18	AK020824
M2PN-4223	angiopoietin 4 (Agpt4)	NM_009641
M2PN-2236	beaded filament structural protein in lens-CP94 (Bfsp1)	NM_009751
M2PN-3855	calpain 12 (Capn12)	NM_021894
M2PN-3442	forkhead box containing protein N4 (Foxn4)	AF323488
M2PN-1767	lymphocyte antigen 96 (Ly96)	NM_016923
M2PN-0532	myogenic factor 6 (Myf6)	NM_008657
M2PN-1826	zinc finger protein 97 (Zfp97)	NM_011765
M2PN-2535	Similar to zinc finger protein 97, clone MGC:18740 IMAGE:3986622	BC011426
MRA-0178	RIKEN adult male spinal cord cDNA clone A330074E19	BB192844
MRA-0213	RIKEN 7 days neonate cerebellum cDNA clone A730095K19	BB260718
MRA-0392	RIKEN adult retina cDNA clone A930041F21	BB282979
MRA-0624	RIKEN 16 days embryo head cDNA clone C130040H24	BB368057
MRA-0365	RIKEN 12 days embryo eyeball cDNA clone D230035C09	BB470735
MRA-0334	RIKEN adult retina cDNA clone A930002B01	BB591375
MRA-2067	13 days embryo head cDNA, RIKEN clone:3110045I18	AK014181
MRA-0913	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2 (Pfkfb2)	NM_008825
MRA-1771	adult male hippocampus cDNA, RIKEN clone:2900046P06	AK013659
MRA-1770	adult male testis cDNA, RIKEN clone:1700110E17	AK018950
MRA-1154	adult male testis cDNA, RIKEN clone:4930438A20	AK015332
MRA-1937	adult retina cDNA, RIKEN clone:A930004D18	AK020807
MRA-1280	adult retina cDNA, RIKEN clone:A930029D14:guanine nucleotide binding protein (G protein), gamma 1 subunit	AK020903
MRA-0210	cGMP-gated cation channel protein	M84742
MRA-0623	TWEAK mRNA, partial cds	AF030100
MRA-1629	Mus musculus, clone MGC:28867 IMAGE:4512579	BC017153
MRA-1185	Mus musculus, clone MGC:38855 IMAGE:5361063	BC023051

Table 3.7. Retina/eye-enriched unknown ESTs.

Clone ID	Accession Number	Number of eye EST matches⁴	Total EST match	% Eye EST
M15E-0732	CB840441	4	4	100
M15E-3645	CB842968	162	239	68
M15E-4248	CB843444	136	210	65
M15E-4840	CB843901	6	7	86
M2PN-0590	CB845024	56	66	85
M2PN-1110	CB845520	6	6	100
M2PN-1211	CB845617	10	13	77
M2PN-2132	CB846398	116	214	54
M2PN-4105	CB848283	197	230	86
MRA-0135	CB850369	8	8	100
MRA-0274	CB848744	21	21	100
MRA-0291	CB848760	16	22	73
MRA-0304	CB848772	5	7	71
MRA-0376	CB848841	7	11	64
MRA-0671	CB849009	7	10	70
MRA-0729	CB849066	15	22	68
MRA-1155	CB849425	13	21	62
MRA-1259	CB849524	4	4	100
MRA-1408	CB849664	7	10	70
MRA-1648	CB849895	2	2	100
MRA-1656	CB849903	4	4	100
MRA-1706	CB849949	14	18	78
MRA-1722	CB849965	4	4	100
MRA-1792	CB850035	9	12	75
MRA-1858	CB850097	6	6	100

⁴ Homology ESTs of each clone were obtained from mouse dbEST and designated as ‘eye EST match’ if it was originally isolated from ocular tissue library. ‘% eye EST’ was calculated by dividing the number of homology ESTs from ocular tissue by the total number of homologs.

qRT-PCR expression profiling of selected clones

To validate *in silico* expression profiles for different developmental stages and tissues, we performed qRT-PCR analysis on a few selected ESTs. Clone MRA-1648 has two homologous sequences in the mouse dbEST and both of these ESTs were identified in mouse retina. We therefore hypothesized that MRA-1648 is a potential retina-enriched EST. qRT-PCR analysis shows that MRA-1648 is highly enriched in retina, with a four-fold higher expression in retina than brain and lung, and over eight-fold higher expression than other tissues (Figure 3.5a). Similar qRT-PCR analysis confirmed *in silico* temporal expression profiles of several ESTs obtained from E15.5 eye, PN2 eye and adult retinas (Figure 3.5b).

3.5 Discussion

EST analysis provides a powerful and rapid means of reconstructing the transcriptome of specific tissues and cell types and for identification of differentially expressed genes. In this study, 11,057 clones were isolated and sequenced, yielding 9,638 high-quality ESTs. The accumulation of low-quality sequences in the first 24 plates of clones sequenced from the M15E library indicated initial technical issues in sequencing, rather than significant differences between this and the other two libraries. We have characterized the transcriptional profiles of mouse E15.5 eye, PN2 eye and adult retina, by further annotation and analyses of 4,144, 3,701 and 1,793 ESTs, respectively. About half of the identified high-quality ESTs represented known genes. As ESTs have previously been generated from several adult retina libraries, the higher number of known genes in our MRA library was not unexpected. Until the recent large-scale transcription analyses of E14.5 retina (Mu et al. 2001), ESTs have not been generated from embryonic eyes. Not surprisingly, 29% of ESTs from the M15E library corresponded to novel transcripts, suggesting that EST generation remains a useful approach for gene discovery.

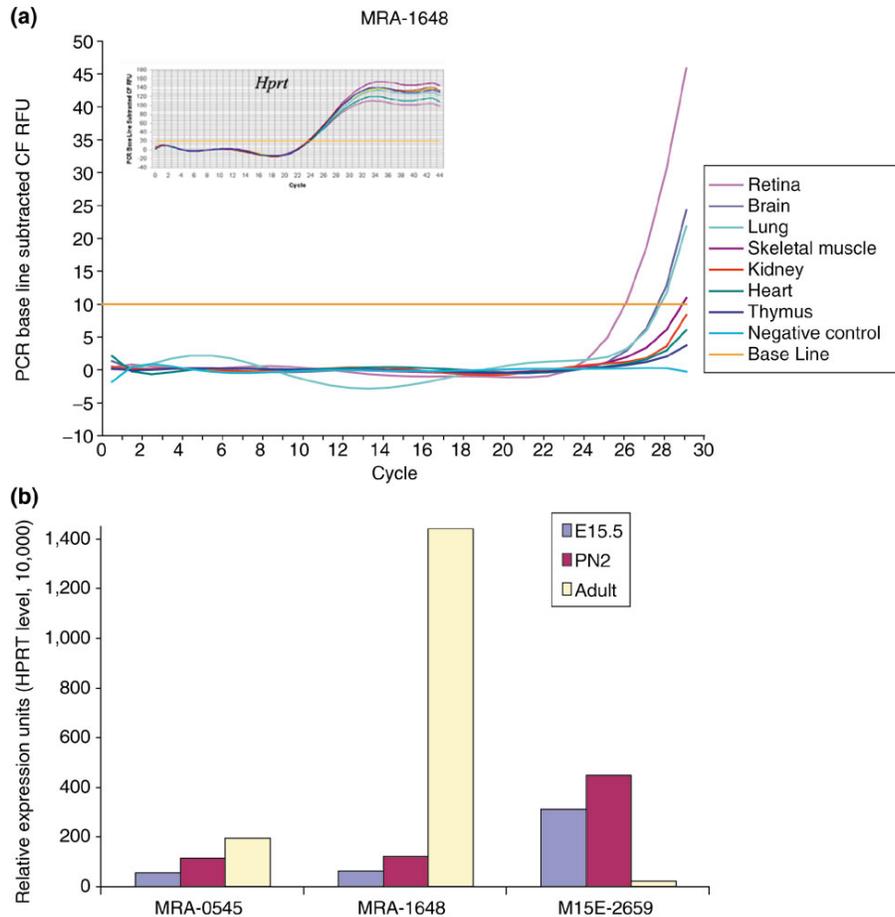


Figure 3.5. qRT-PCR validation of *in silico* expression profiles.

(a) Clone MRA-1648 showed a four-fold higher expression in retina than in brain or lung. Its expression in other tissues is at least eight-fold less than in the retina. One negative control curve is shown to demonstrate that samples are free of genomic contamination. The inset figure shows the qRT-PCR curves of *Hprt*, all of which cross the baseline at 24 ± 0.5 cycles, indicating equal input of cDNA in different tissue samples. (b) The relative expression units of three clones, MRA-0545, MRA-1648 and M15E-2659, in E15.5 eyes (blue), PN2 eyes (purple) and adult retina (yellow) were calculated by their cycle differences from *Hprt* during qRT-PCR experiments. *Hprt* was arbitrarily assigned a value of 10,000. The high transcript level of MRA-0545 and MRA-1648 and low level of M15E-2659 in adult retina confirmed the *in silico* profiles shown in Table 3.7.

Despite the large number of random cDNA clones analyzed, we did not observe much redundancy. Only 0.4% of the ESTs were detected six or more times, not considering mitochondrial and ribosomal genes (7-10%). Approximately 51% of M15E, 58% of M2PN and 67% of MRA ESTs are identified only once (singletons). This suggests that additional sequencing or *in silico* retinal EST mining may be desired for a more comprehensive transcriptional profiling, especially of the adult retina. It should be noted that the true number of unique clusters will almost certainly be lower since in the absence of full-length sequences it is not always obvious whether two ESTs represent non-overlapping segments of the same gene or two different genes. This issue is minimized by obtaining an average 500 bp or longer sequence from the 5' end, given that the coding regions of most transcripts are less than 1,000 bp in length (Lander et al. 2001).

Functional categorization of ESTs with known-gene matches underlines general differences in expression profiles of different tissues at distinct developmental stages. The data presented here suggest global changes in gene expression during eye development. For instance, the developing eye exhibited more transcripts involved in cell structure, gene regulation and protein expression, whereas the adult retina had more transcripts representing phototransduction and metabolism. Although our data are limited by the fact that the M15E and M2PN libraries were constructed from whole eye while MRA was from retina only, the overall findings are in good agreement with the expected roles of these tissues at corresponding stages. Similar to results from E14.5 retinal cDNA sets (Mu et al. 2001), the most abundant transcripts in developing eyes are those of translation factors, housekeeping genes and cell structure/cytoskeletal proteins. Both the M15E and M2PN libraries have higher numbers of clones encoding elongation factor Tu, ubiquitin B, Gapd and several structural proteins. Excluding crystallins, hemoglobin alpha adult chain 1 transcript is the most abundant gene in the M15E library. This may reflect active neovascularization of embryonic eyes and dynamic transcriptional activity in the nuclei of immature red blood cells. Levels of hemoglobin transcripts drop

dramatically to only six copies in the M2PN library and zero in MRA, perhaps reflecting the gradual maturation of red blood cells and the vascular system. Adult retina plays an important role in vision, as revealed by high expression of phototransduction genes. Although deeper sequencing of these libraries may change the frequencies of individual genes, the overall patterns of gene expression should remain invariant.

In addition to the comparison of global patterns of gene expression, ESTs were utilized to construct *in silico* temporal transcriptional profiles of individual genes in E15.5, PN2 eyes and adult retinas. Relative expression of a single gene in each library was estimated from the number of corresponding ESTs representing this gene normalized by the total number of high-quality ESTs. This approach assumes random sequencing and that the level of activity of a given gene may be inferred from the number of corresponding ESTs obtained. It is intrinsically limited in detecting posttranscriptional regulatory effects and is insufficient in identifying under-expressed genes. However, similar issues are shared by other methods for estimating gene expression, including microarray hybridization. Furthermore, high numbers of analyzed ESTs would undoubtedly increase the sensitivity of this approach. Cross-comparison between our three cDNA libraries identified a number of genes showing a restricted temporal expression. For instance, phototransduction genes are highly and restrictively expressed in the adult retina, whereas crystallin transcript levels are greatly elevated in the E15.5 and PN2 eyes. In addition to providing a list of differentially regulated genes, this approach has enabled us to infer the potential function of unknown ESTs based on their temporal expression patterns. ESTs showing striking differences in the level of expression in the three libraries may play significant roles in eye development or function of mature retina. *In silico* temporal expression profiles should also provide valuable insights into the cellular function of individual genes and may become a useful guide for gene discovery.

It is believed that a tissue selectively expresses a specific set of genes based on its functional needs. Such preferentially expressed genes are generally of importance and may have disease-causing effects. Mutations in almost all genes that are preferentially expressed in the retina or during eye development have been associated with eye or retinal diseases. Identification of eye or retina-enriched genes is therefore a promising approach to discovering potential disease genes. This is strengthened by the fact that over half of the retina-enriched genes from the MRA library are known phototransduction genes. Similarly, functional annotation of a number of crystallin genes revealed a restricted expression in ocular tissues including eye or retina. Crystallins have been long known to be major protein components of lens. Recent studies have advocated their role as molecular chaperones (Graw and Loster 2003; Kurita et al. 2003) and similar function is predicted in the retina (Xi et al. 2003). We have validated by qRT-PCR the retina-enriched expression of an EST (MRA-1648) that has been identified only from retina libraries. Our list of known genes or unknown ESTs occurring preferentially in the eye libraries may provide valuable information for gene discovery.

We have identified the human orthologs of 47% of known ESTs identified from our EST set. The chromosomal distribution of ocular genes differs significantly from the observed distribution of 30,000 human genes reported previously (Deloukas et al. 1998). A higher density of ocular genes was found in chromosomes 17 and X, confirming earlier studies that indicated an overrepresentation of retinal disease loci on these chromosomes (Blackshaw et al. 2001; Bortoluzzi et al. 2000). Chromosomes 4 and 10 were estimated to have lower ocular gene density. A similar trend was demonstrated in a previous mapping of 3,152 genes from adult human retina (Bortoluzzi et al. 2000). It is noteworthy that our mapping information is highly dependent on the accuracy of UniGene data, and increasing the number of mapped genes will definitely add to the power of this test. We have also observed 277 unique genes localized within the chromosomal intervals of

mapped genetic loci for ocular diseases. These genes therefore may be considered as valid candidates for mutation screening.

The identification of ESTs will greatly facilitate in the investigation of the complete transcriptome of specific tissues. ESTs annotated in this study will be a valuable complement to currently archived sequences from ocular tissues and should facilitate eye transcriptome analyses. A comprehensive collection of retina/eye ESTs is an important genomic resource for the positional cloning of disease genes, for large-scale expression studies and for other functional genomic studies. Mouse cDNA microarrays containing over 6,000 retina/eye ESTs have been generated in our laboratory for the transcriptional analysis of ocular tissues during development and from knockout and transgenic mice or mutants (Farjo et al. 2002; Yu et al. 2002). The high-level annotation of eye and retinal ESTs, presented here, will greatly facilitate *in silico* expression profiling and experimental approaches utilizing slide microarrays of eye-expressed genes.

3.6 References

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F. et al. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**: 1651-1656.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J Mol Biol* **215**: 403-410.
- Bernstein, S.L., Borst, D.E., Neuder, M.E., and Wong, P. 1996. Characterization of a human fovea cDNA library and regional differential gene expression in the human retina. *Genomics* **32**: 301-308.
- Bernstein, S.L., Borst, D.E., and Wong, P.W. 1995. Isolation of differentially expressed human fovea genes: candidates for macular disease. *Mol Vis* **1**: 4.
- Blackshaw, S., Fraioli, R.E., Furukawa, T., and Cepko, C.L. 2001. Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell* **107**: 579-589.
- Bortoluzzi, S., d'Alessi, F., and Danieli, G.A. 2000. A novel resource for the study of genes expressed in the adult human retina. *Invest Ophthalmol Vis Sci* **41**: 3305-3308.
- Buraczynska, M., Mears, A.J., Zarepari, S., Farjo, R., Filippova, E., Yuan, Y., MacNee, S.P., Hughes, B., and Swaroop, A. 2002. Gene expression profile of native human retinal pigment epithelium. *Invest Ophthalmol Vis Sci* **43**: 603-607.
- Cheng, G. and Porter, J.D. 2002. Transcriptional profile of rat extraocular muscle by serial analysis of gene expression. *Invest Ophthalmol Vis Sci* **43**: 1048-1058.
- Deloukas, P., Schuler, G.D., Gyapay, G., Beasley, E.M., Soderlund, C., Rodriguez-Tome, P., Hui, L., Matise, T.C., McKusick, K.B., Beckmann, J.S. et al. 1998. A physical map of 30,000 human genes. *Science* **282**: 744-746.
- Farjo, R., Yu, J., Othman, M.I., Yoshida, S., Sheth, S., Glaser, T., Baehr, W., and Swaroop, A. 2002. Mouse eye gene microarrays for investigating ocular development and disease. *Vision Res* **42**: 463-470.
- Gieser, L. and Swaroop, A. 1992. Expressed sequence tags and chromosomal localization of cDNA clones from a subtracted retinal pigment epithelium library. *Genomics* **13**: 873-876.
- Graw, J. and Loster, J. 2003. Developmental genetics in ophthalmology. *Ophthalmic Genet* **24**: 1-33.

- Jasper, H., Benes, V., Atzberger, A., Sauer, S., Ansorge, W., and Bohmann, D. 2002. A genomic switch at the transition from cell proliferation to terminal differentiation in the *Drosophila* eye. *Dev Cell* **3**: 511-521.
- Kan, Z., States, D., and Gish, W. 2002. Selecting for functional alternative splices in ESTs. *Genome Res* **12**: 1837-1845.
- Katsanis, N., Worley, K.C., Gonzalez, G., Ansley, S.J., and Lupski, J.R. 2002. A computational/functional genomics approach for the enrichment of the retinal transcriptome and the identification of positional candidate retinopathy genes. *Proc Natl Acad Sci U S A* **99**: 14326-14331.
- Kawamoto, S., Yoshii, J., Mizuno, K., Ito, K., Miyamoto, Y., Ohnishi, T., Matoba, R., Hori, N., Matsumoto, Y., Okumura, T. et al. 2000. BodyMap: a collection of 3' ESTs for analysis of human gene expression information. *Genome Res* **10**: 1817-1827.
- Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y. 2001. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res* **11**: 281-289.
- Kurita, R., Sagara, H., Aoki, Y., Link, B.A., Arai, K., and Watanabe, S. 2003. Suppression of lens growth by alphaA-crystallin promoter-driven expression of diphtheria toxin results in disruption of retinal cell organization in zebrafish. *Dev Biol* **255**: 113-127.
- Lander, E.S. Linton, L.M. Birren, B. Nusbaum, C. Zody, M.C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- Maubaret, C., Delettre, C., Sola, S., and Hamel, C.P. 2002. Identification of preferentially expressed mRNAs in retina and cochlea. *DNA Cell Biol* **21**: 781-791.
- Mu, X., Zhao, S., Pershad, R., Hsieh, T.F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. et al. 2001. Gene expression in the developing mouse retina by EST sequencing and microarray analysis. *Nucleic Acids Res* **29**: 4983-4993.
- Sharon, D., Blackshaw, S., Cepko, C.L., and Dryja, T.P. 2002. Profile of the genes expressed in the human peripheral retina, macula, and retinal pigment epithelium determined through serial analysis of gene expression (SAGE). *Proc Natl Acad Sci U S A* **99**: 315-320.
- Shimizu-Matsumoto, A., Adachi, W., Mizuno, K., Inazawa, J., Nishida, K., Kinoshita, S., Matsubara, K., and Okubo, K. 1997. An expression profile of genes in human retina and isolation of a complementary DNA for a novel rod photoreceptor protein. *Invest Ophthalmol Vis Sci* **38**: 2576-2585.

- Sinha, S., Sharma, A., Agarwal, N., Swaroop, A., and Yang-Feng, T.L. 2000. Expression profile and chromosomal location of cDNA clones, identified from an enriched adult retina library. *Invest Ophthalmol Vis Sci* **41**: 24-28.
- Sohocki, M.M., Malone, K.A., Sullivan, L.S., and Daiger, S.P. 1999. Localization of retina/pineal-expressed sequences: identification of novel candidate genes for inherited retinal disorders. *Genomics* **58**: 29-33.
- Strachan, T., Abitbol, M., Davidson, D., and Beckmann, J.S. 1997. A new dimension for the human genome project: towards comprehensive expression maps. *Nat Genet* **16**: 126-132.
- Swaroop, A. and Zack, D.J. 2002. Transcriptome analysis of the retina. *Genome Biol* **3**: REVIEWS1022.
- VanBuren, V., Piao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin, P.R., Stagg, C.A., Bassey, U.C., Aiba, K., Hamatani, T. et al. 2002. Assembly, verification, and initial annotation of the NIA mouse 7.4K cDNA clone set. *Genome Res* **12**: 1999-2003.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* **270**: 484-487.
- Venter, J.C. Adams, M.D. Myers, E.W. Li, P.W. Mural, R.J. Sutton, G.G. Smith, H.O. Yandell, M. Evans, C.A. Holt, R.A. et al. 2001. The sequence of the human genome. *Science* **291**: 1304-1351.
- Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. 1996. A large family of putative transmembrane receptors homologous to the product of the *Drosophila* tissue polarity gene *frizzled*. *J Biol Chem* **271**: 4468-4476.
- Waterston, R.H. Lindblad-Toh, K. Birney, E. Rogers, J. Abril, J.F. Agarwal, P. Agarwala, R. Ainscough, R. Alexandersson, M. An, P. et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562.
- Wistow, G. 2002. A project for ocular bioinformatics: NEIBank. *Mol Vis* **8**: 161-163.
- Xi, J., Farjo, R., Yoshida, S., Kern, T.S., Swaroop, A., and Andley, U.P. 2003. A comprehensive analysis of the expression of crystallins in mouse retina. *Mol Vis* **9**: 410-419.
- Xu, Q., Modrek, B., and Lee, C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* **30**: 3754-3766.
- Yu, J., Othman, M.I., Farjo, R., Zarepari, S., MacNee, S.P., Yoshida, S., and Swaroop, A. 2002. Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays. *Mol Vis* **8**: 130-137.

CHAPTER 4

GENERATION OF EYE-GENE MICROARRAYS

This chapter has been published in:

R. Farjo^{*}, **J. Yu^{*}**, M.I.Othman, S. Yoshida, S. Sheth, W. Baehr, A. Swaroop. *Mouse I-gene microarrays for investigating ocular development and disease*. Vision Res. 2002 Feb;42(4):463-70. ***Co-authors with equal contribution**

4.1 Abstract

Microarray technology can facilitate simultaneous expression analysis of thousands of genes and assist in delineating cellular pathways involved in development or disease pathogenesis. Since public databases and commercial cDNA microarrays have an under-representation of eye-expressed genes, I have generated over 10,000 expressed sequence tags from three unamplified mouse eye/retinal cDNA libraries. Of these, 3,000 were initially printed onto glass slides to generate custom cDNA microarrays of eye-genes. Methodology for printing of slides, hybridization, scanning and data analysis has been optimized. The I-gene microarrays will be useful for establishing expression profiles of the mouse eye/retina and provide a resource for defining molecular pathways involved in development, aging and disease.

4.2 Introduction

Molecular insights into pathogenesis of retinal diseases and systematic design of therapeutic strategies are dependent, to a large extent, on animal models (mostly mouse) and our understanding of the basic biological processes underlying development and maintenance of specialized visual functions (e.g., phototransduction). Pioneering studies

in vertebrate systems have demonstrated that cell fate determination and neuronal differentiation of retinal progenitor cells is guided by intrinsic genetic programs, inductive cell–cell interactions and extrinsic factors (Cepko 1996; Levine et al. 2000), which lead to differential expression of genes at specific stages of development. A significant number of transcription factors and signaling molecules that are expressed in developing retina have been identified (Freund et al. 1996; Jean et al. 1998) and the role for at least two such proteins in causing human eye disease delineated (Bessant et al. 1999; Freund et al. 1997; Swain et al. 1997; Swaroop et al. 1999). Nevertheless, cellular targets for the molecular regulators and signaling pathways leading to the specification of retinal neurons are poorly understood. Similarly, although a number of retinal macular disease genes have been identified (RetNet, <http://www.sph.uth.tmc.edu/RetNet/>) and animal models of human retinal disease (generally in mouse) generated, the pathways by which mutations in a particular gene (retina-specific or widely expressed) specifically lead to photoreceptor degeneration are poorly understood.

Recent advances in genomics and microarray technology provide an excellent opportunity to examine changes in retinal gene expression profiles during development and disease. Until recently, most studies have focused on the characterization of a limited number of genes or proteins. The advent of microarrays has revolutionized the pace of investigations in functional genomics (Hughes et al. 2000; Schena et al. 1995; Shalon et al. 1996; Young 2000). It is now possible to simultaneously study thousands of genes that are altered at a particular stage of development or during disease pathogenesis. Among others, the microarrays have been successfully applied to the generation of expression profiles of cell cycle (Cho et al. 2001), hematopoietic stem cells (Phillips et al. 2000), various cancers (Alizadeh and Staudt 2000; Bittner et al. 2000; DeRisi et al. 1996; Golub et al. 1999), and normal/diseased brain (Lee et al. 2000; Mirnics et al. 2000; Sandberg et al. 2000). Generation of such gene profiles is expected to lead to better insights of cellular pathways and should have significant impact on rational drug design. During the

last few years, several laboratories have generated expressed sequence tags (ESTs) from retina (Bernstein et al. 1996; Malone et al. 1999; Sinha et al. 2000; Swanson et al. 1997) and used small-scale microarrays for comparative studies (Livesey et al. 2000). However, systematic investigations for generating expression profiles of eye tissues, particularly retina, have not been attempted.

Although public gene databases provide a wealth of information, the genes expressed in the mouse (and human) eye/retina are under-represented. Furthermore, as much as 30% of the cDNA clones procured from vendors may contain multiple inserts or are mislabeled and must be re-sequenced (Halgren et al. 2001). Therefore, the goals of our studies are to construct cDNA libraries from mouse eye/retina, isolate sequence-tagged cDNAs, and produce custom eye gene microarrays for developing gene expression profiles. At this stage, we have analyzed over 3000 cDNA sequences and are printing slides with almost 2000 eye-expressed genes. Experiments have also been initiated for optimization of conditions and studies on mutant mice retina.

4.3 Methods

RNA isolation

Mouse embryonic day (E) 15.5 and postnatal day (PN) two eye tissues were dissected from CD-1 mice (Charles River Laboratory). Adult mouse retina was dissected from C57BL/6 mice. Total RNA was isolated using TRIzol reagent (Life Technologies) and purified by RNeasy kit (Qiagen). Purity and integrity of RNA was evaluated by absorbance at 260 and 280 nm and by agarose gel electrophoresis.

cDNA library construction

Total RNA (40 mg) was used for generating directional cDNA libraries with SuperScript™ Plasmid System (Life Technologies). Three libraries were constructed in the pSPORT1 vector according to the manufacturer's protocol. Briefly, first-strand cDNA was synthesized with a Not I-oligo(dT) primer-adaptor. After second-strand synthesis and ligation of Sal I adaptors, the cDNAs were digested with Not I, which generated cDNAs with Sal I sites at the 5' end and Not I sites at the 3' end. cDNAs were size-fractionated and those of 0.5 to 2 kb were digested with Sal I and Not I and ligated to pSPORT1 vector. The ligated sample was used to transform *E. coli* (ElectroMax DH-5 α) by electroporation.

DNA sequencing

DNA sequencing was performed with a high throughput automatic Sequencer (Applied Biosystems, Inc). An average of 600 bp nucleotide sequence was obtained from the 5' end of 3188 clones. BLAST search (<http://www.ncbi.nlm.nih.gov/BLAST>) was performed against GenBank and dbEST for each clone and a collection of ESTs was assembled. Information regarding these clones is stored in a Microsoft Excel database.

Printing of microarray slides

Individual clones were cultured in 96-well plates containing ampicillin. After overnight incubation, these cultures were converted into glycerol stocks that were also used for direct amplification of cDNA inserts. PCR reactions in a total volume of 100 μ l (2.5 mM MgCl₂ and 2.5 units of Amplitaq Gold (AP Biosystems)) were performed as follows: 95°C 12 min, (35 cycles: 94°C 30 s, 63°C 30 s, 72°C 45 s), 72°C 15 min. A 96-well plastic replicator (Incyte Genomics) was used to seed PCR reactions with bacteria from the glycerol stocks. PCR products were purified using the multiscreen PCR system

(Millipore). Samples were then electrophoresed on agarose gels to determine relative DNA concentrations as well as the success rate for PCR amplification. Four 96-well plates of cDNA inserts were then converted to a single 384-well plate. These samples were evaporated under vacuum and resuspended in either 4 μ l of 3X sodium chloride/sodium citrate (SSC) or 50% dimethylsulfoxide. The samples were arrayed onto CMT-GAPS slides (Corning) with a SDDC-2 arrayer (Virtek ESI). Printed slides are then stratalinked and stored in a dust-free and lightproof container until use.

Direct labeling of the target RNA

A mixture of total RNA (10 μ g) and oligo-dT (2 μ g) (in a total volume of 22 μ l) was heated to 70°C for 10 min and quickly chilled on ice. The following were then added sequentially: Cy3TM-dCTP (12.5 μ M) or Cy5TM-dCTP (25 μ M) (Amersham), 1X first-strand buffer, DTT (10 mM) (Life Technology), dNTP mix (0.5 mM each dATP, dGTP, dTTP, and 0.25 mM dCTP), 40 units RNase inhibitor and 400 units SuperScriptTM II reverse transcriptase (Life Technology) in a total volume of 40 μ l. The reaction was incubated at 42°C for 2 h to generate Cy3- or Cy5-labeled target cDNA. Starting RNA was eliminated by adding 2 units of RNase H and 10 μ g of RNase A for 15 min at 37 °C. Target cDNAs were purified using QIAquickTM PCR purification kit (Qiagen) and concentrated to 10 μ l.

Indirect labeling of the target RNA

Total RNA (16 μ g) and oligo-dT (5 μ g) were combined in a total volume of 15.5 μ l, incubated at 70°C for 10 min and quickly chilled on ice. The following were added to a total volume of 30 μ l: 1X first-strand buffer, 10 mM DTT (Life Technology), aminoallyl-dUTP/dNTP mix (500 μ M dATP, 500 μ M dCTP, 500 μ M dGTP, 300 μ M dTTP, 200 μ M aa-dUTP), and 380 units SuperScriptTM II reverse transcriptase (Life Technology). Each sample was incubated at 42°C for 2 h. 10 μ l of 1N NaOH and 10 μ l of 0.5 M EDTA were added and the reaction was incubated at 65°C for 15 min to hydrolyze starting RNA. The reaction was neutralized by adding 25 μ l of 1 M Tris-HCl, pH 7.5, and

cDNAs were purified using GFX™ columns (Amersham) and vacuum dried. Each cDNA pellet was resuspended in 4.5 µl of H₂O. Cy3 or Cy5 monoreactive dye (Amersham) was each resuspended in 4.5 µl of 0.1 M sodium bicarbonate buffer (pH 9.0). cDNAs and Cy3 or Cy5 were mixed and incubated at room temperature in the dark for 1 h to allow coupling of the dyes. To stop the reaction and prevent cross-coupling, 4.5 µl of 4 M hydroxylamine was added to each reaction, followed by 15 min incubation in the dark. Samples were then purified using QIAquick™ PCR purification kit (Qiagen) and vacuum dried.

Reference RNA

Reference RNA for microarray experiments was generated by combining total RNAs from several mouse tissues and cell lines. A mixture of 17 mg of total RNA was obtained, with 7 mg from mouse retina at different stages of development (E14-16, PN2-3, PN10-12 and adult), 3 mg from P19 embryonic carcinoma cells, another 3 mg from P19 cells induced to differentiate into neuronal and glial cells in the presence of retinoic acid, and 4 mg from neuroblastoma stem cells N1E-115.

Microarray hybridization with labeled targets

Microarray slides were prehybridized in prehybridization buffer (5X SSC, 1% Bovine serum albumin) and 0.1% Sodium dodecyl sulfate (SDS) at 42°C for 1 h and washed five times in H₂O for 2 s each. The slides were then rinsed with isopropanol and centrifuged at 1900 rpm for 2 min.

An equal volume of 2X hybridization buffer (50% formamide, 10X SSC, 0.2% SDS) was added to the target mixture, consisting of Cy3 and Cy5 labeled target cDNAs, 1 µg poly(A) RNA (Sigma), 2 µg mouse Cot-1 DNA (Life Technology), 1 µg yeast tRNA (Life Technology), and 10 µg salmon sperm DNA (Life Technology). The

hybridization mixture was applied to microarray slides. For indirect labeling, dye-labeled cDNAs were resuspended in 45 μ l of GlassHyb (Clontech), heated to 95°C for 5 min and centrifuged briefly before applying to the slides. After putting a 22 \times 40 mm coverslip (Grace Bio-Lab), each slide was placed in a hybridization chamber (Corning Microarray Technology). Droplets of DEPC water were placed in the two reservoirs at either end of the chamber, which was then sealed and placed in a 42°C water bath overnight for 16–20 h. Slide microarrays were removed from the chamber and immersed into 2X SSC until the coverslip moved away freely. Slides were then sequentially washed twice in 0.1X SSC, 0.1% SDS and 0.1X SSC. After rinsing in DEPC water for 5 s, the microarrays were centrifuged at 1900 rpm for 2 min to dry.

Slide scanning and data analysis

After hybridization and washing, slides were scanned using an Affymetrix 428 scanner. Images were acquired for Cy3 and Cy5 channels in a 16-bit TIFF format and then analyzed using Jaguar 2.0 (Affymetrix). Laser power and the gain on the photomultiplier tube (PMT) were kept constant during scanning of each individual slide. The PMT setting was chosen so that the highest intensity values lied below the saturation point. Both channels were normalized against each other using Jaguar software. Spot intensity and background signals were quantified for each channel and used to generate spreadsheet data for Microsoft Excel.

4.4 Results

Schematic representation of microarray steps

The steps involved in the microarray strategy that we have used are shown in Figure 4.1. Clones from three unamplified mouse eye libraries were sequence-tagged and

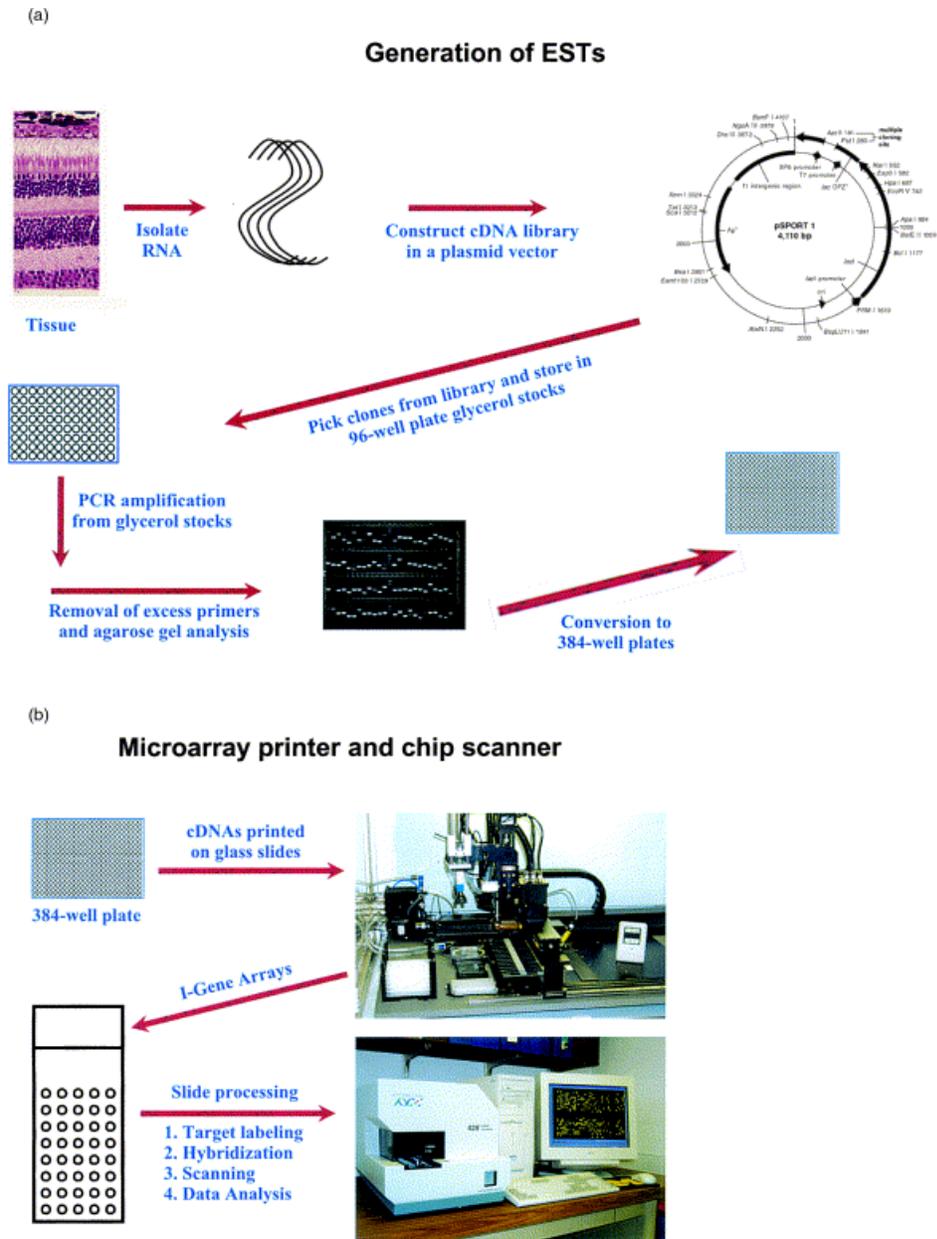


Figure 4.1. A schematic representation of the microarray strategy.

(A) The generation of clones and cDNAs. (B) The printing and hybridization of microarray slides.

analyzed to determine their identity. Inserts from these cDNAs were isolated and purified prior to printing. I-Gene microarrays were printed at 22–24°C with a relative humidity of 55–65%. Cy-labeled cDNAs were used as targets to hybridize the slides. Images of Cy3 and Cy5 were acquired by laser scanning and analyzed using Jaguar 2.0.

Characterization of the mouse cDNA libraries

Unamplified libraries from mouse E15.5 and PN2 eyes, and from adult retina were generated in pSPORT1 plasmid vector. More than 90% of the clones had an average insert size of 1.5 kb. Almost 2500 clones from each of the three libraries have been isolated and stored in triplicate as glycerol stocks in 96-well plates. Of these, 3188 have been sequenced at the 5' end and the sequences analyzed (Figure 4.2). 1952 cDNA clones show strong homology to known genes or ESTs, whereas 886 clones are classified as novel ESTs. A total of 99 clones matched ribosomal genes, 152 matched mitochondrial genes, and 99 clones gave poor sequencing results.

Preparation of DNA inserts for printing

Two rounds of PCR are performed on each glycerol stock in order to increase the yield of cDNA inserts required for printing of slides. After purification, the average yield is approximately 3 µg per clone with a PCR success rate of >90%. Purified PCR products are vacuum dried and resuspended in 3X SSC at an average concentration of 200 ng/ l.

Printing

Spots with 100 µm diameter are printed with Stealth pins (Telechem) using SDDC-2 arrayer (Virtek ESI). Samples are pre-printed eight times onto a blotting glass before the slides are printed. The pins are washed with H₂O in four cycles of 6 s each and dried between samples.

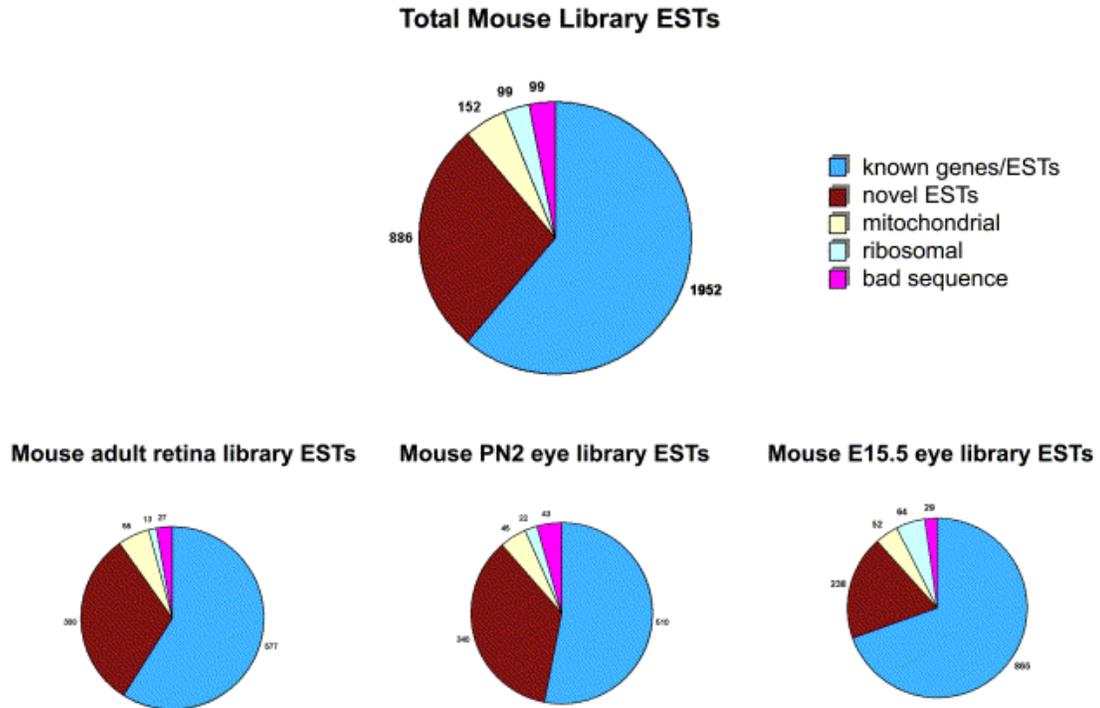


Figure 4.2. Pie charts representing different clone types identified in our mouse eye libraries.

Currently a total of 3188 cDNA clones have been sequenced and analyzed.

Hybridization and analysis of slides

Slides containing over 2000 clones (printed in duplicate) have been utilized to optimize protocols involved in printing, arraying, dye labeling, hybridization, and scanning. We have performed multiple hybridizations using an identical RNA target (from a P19 embryonic carcinoma cell line or N1E-115 neuroblastoma stem cells) labeled with both Cy3 and Cy5 respectively. Experiments were performed to compare hybridizations of P19 (Cy3) vs. P19 (Cy5) and N1E-115 (Cy3) vs. N1E-115 (Cy5) targets. Since the targets in each experiment are same, hybridizations intensities from each channel should be identical. However, Cy3 and Cy5 have different levels of incorporation in RNA during the labeling step. This is corrected by normalizing the Cy3 and Cy5 channels so that the total measured fluorescence intensity is equal between the two channels (Hegde et al. 2000; Quackenbush 2001). The total integrated intensities across all spots in one channel should be equal to the other channel, especially when the RNA targets are the same. Spots are classified as outliers if they have intensity values that are less than the background or close to 0, or if the spot has very high intensity. These outliers were removed prior to data analysis. Jaguar 2.0 was used to generate synthetic images for each channel and calculate intensity ratios for each experiment (Figure 4.3). Methods for direct and indirect labeling were tested and used in multiple experiments. After scanning and data analysis, 90–95% of the spots have a ratio between -2 and +2, demonstrating relatively equal amounts of hybridization at both the Cy3 and Cy5 channels (Figure 4.4). Repetitive experiments illustrate that false positives are recognizable. Some spots appear to show a change greater than two-fold between the two samples; however, when multiple experiments are performed, the average ratio of these spots no longer reveals a significant change in the expression level. As for labeling methods, our data indicates that indirect labeling may be a more effective and

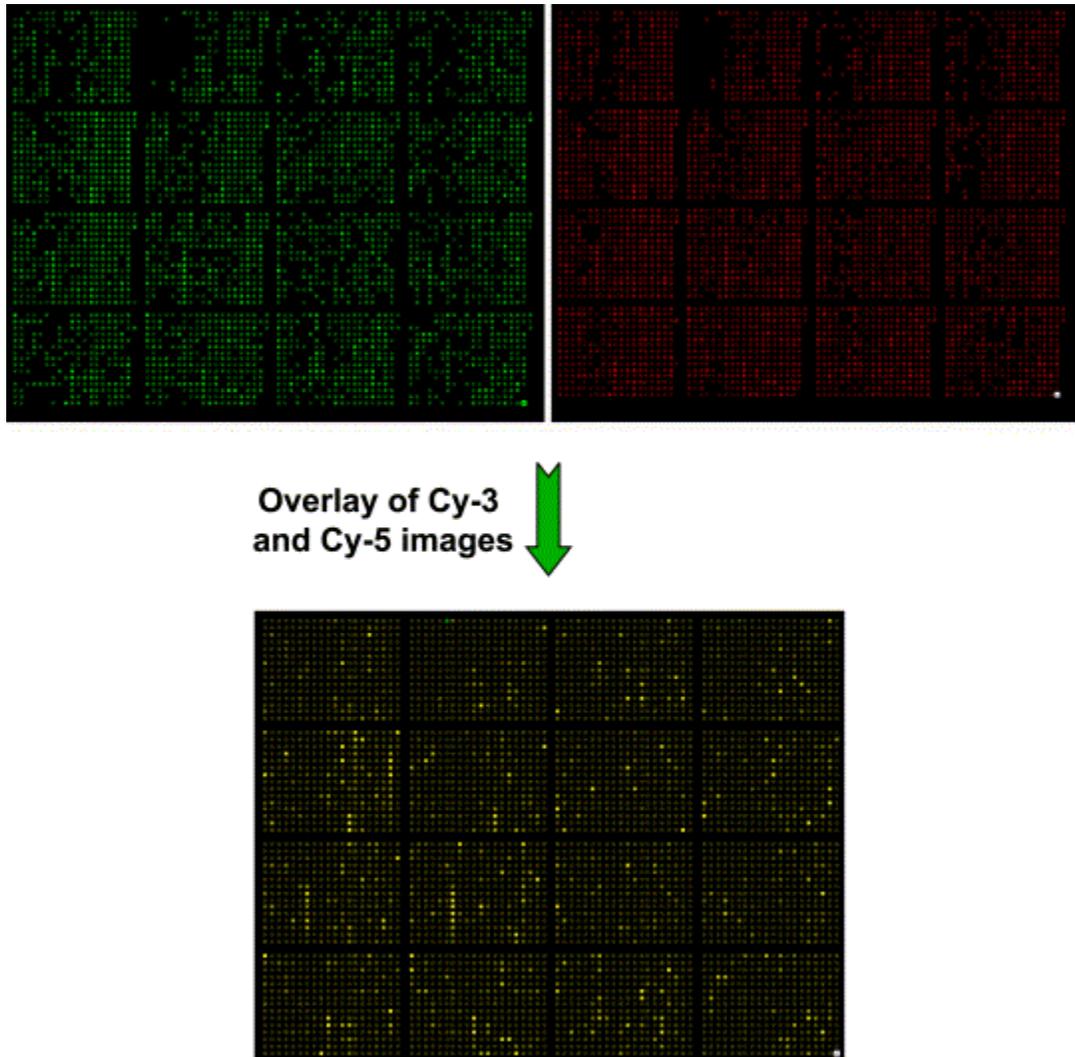


Figure 4.3. A typical microarray experiment with identical target RNAs. Jaguar 2.0 software (Affymetrix) calculates pixel intensity minus background and generates a synthetic image for both the Cy3 and Cy5 channels. These images are then overlaid to generate a composite synthetic image that can be used for qualitative analysis. Yellow spots have an equal amount of hybridization from the Cy3 and Cy5 targets.

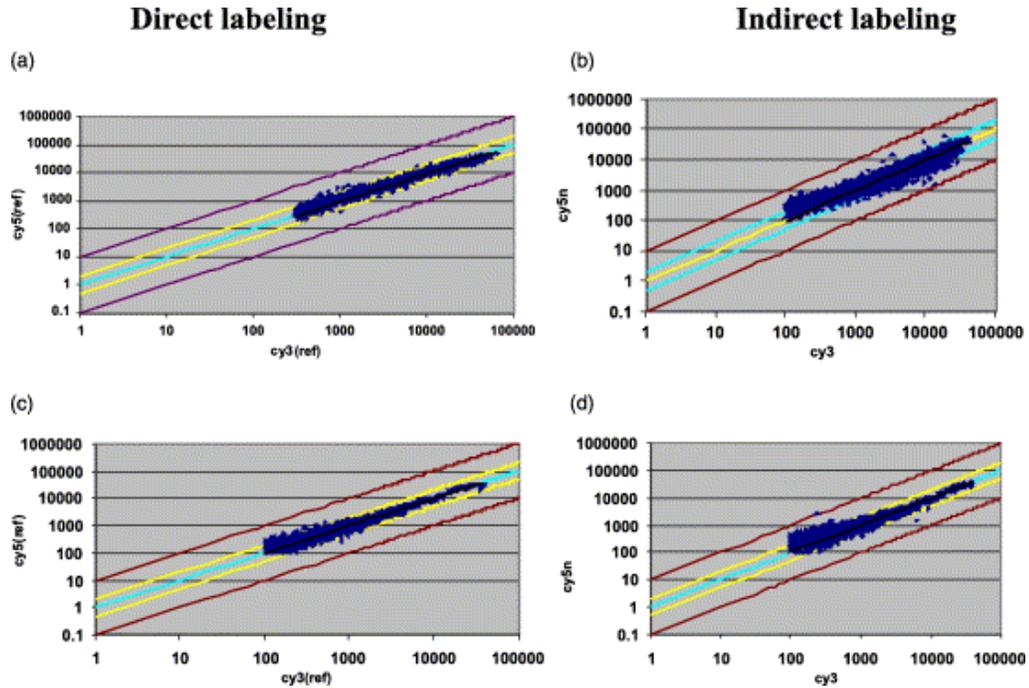


Figure 4.4. Scatter plot analysis of microarray hybridization experiments with identical target RNAs.

The graphs on the left represent data from experiments utilizing direct labeling and graphs on the right from indirect labeling. Cy3 and Cy5 channels were normalized against each other. As expected, most genes have a fold change between -2 and +2 between the identical targets. Linear regression of data from direct labeling experiments (a and b) show R^2 (correlation of determination) values of 0.828 and 0.929, respectively. R^2 values of data from amino-allyl indirect labeling (c and d) are 0.972 and 0.969, respectively.

reproducible method and provide better data. With indirect labeling, the slope of a scatter plot produced by the data set is also closer to one and the signal/noise ratio was significantly higher.

4.5 Discussion

Generation of eye-expressed ESTs

We have isolated 7500 clones from mouse eye/retina cDNA libraries; of these, 3188 have been sequenced so far. Almost 30% of these sequences do not show significant homology to any known gene or EST in the public databases, further confirming an under-representation of eye genes. Many cDNAs reveal homology to the genomic sequence (continuously being deposited as part of the Human Genome Project), which can be used to elucidate their gene structure and chromosomal location. This can assist in the identification of potential candidate genes for retinal/eye diseases. A non-redundant set of clones is currently being organized into new stocks, which will be used to print the I-Gene arrays. ESTs from additional mouse eye libraries may be added to augment this non-redundant set of clones.

I-Gene arrays

Currently, slides are being printed with >6000 clones. These slides have permitted us to optimize the hybridization methodology and perform several experiments. Although our current protocols are generating acceptable data, additional standardizations may be necessary to further reduce the inter- and intra-slide variability.

Multiple experiments must be performed when working with microarrays to minimize differences between tissue samples that can lead to false positives and to obtain statistically significant data. It is also necessary to use identical methodology to reduce

experimental variations. Although microarray experiments provide researchers with semi-quantitative data, it is possible to develop algorithms that will convert microarray ratios into true fold changes in gene expression. In any event, the experimental results obtained from microarray analysis must be confirmed using Northern blots and/or quantitative reverse transcription-PCR.

Microarray technology still needs better methods for image acquisition and data analysis. Inter-experiment and inter-personnel variations are a significant problem. At times, investigators may also need to directly compare data generated by different laboratories. Utilization of a common target RNA (i.e., reference RNA) can provide better normalization of results for studying changes in gene expression (Bassett et al. 1999). Even where the experimental design involves a straightforward comparison of two RNA species, use of reference RNA will permit better normalization of spot intensities and provide tools for statistical analysis of the data. The utilization of identical mouse I-gene cDNA microarrays and standard methodologies by different laboratories should permit additional cluster analysis of microarray data and lead to better insights into cellular pathways involved in eye/retinal development and disease pathogenesis.

4.6 References

- Alizadeh, A.A. and L.M. Staudt. 2000. Genomic-scale gene expression profiling of normal and malignant immune cells. *Curr Opin Immunol* **12**: 219-225.
- Bassett, D.E., Jr., M.B. Eisen, and M.S. Boguski. 1999. Gene expression informatics--it's all in your mine. *Nat Genet* **21**: 51-55.
- Bernstein, S.L., D.E. Borst, M.E. Neuder, and P. Wong. 1996. Characterization of a human fovea cDNA library and regional differential gene expression in the human retina. *Genomics* **32**: 301-308.
- Bessant, D.A., A.M. Payne, K.P. Mitton, Q.L. Wang, P.K. Swain, C. Plant, A.C. Bird, D.J. Zack, A. Swaroop, and S.S. Bhattacharya. 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nat Genet* **21**: 355-356.
- Bittner, M., P. Meltzer, Y. Chen, Y. Jiang, E. Seftor, M. Hendrix, M. Radmacher, R. Simon, Z. Yakhini, A. Ben-Dor, N. Sampas, E. Dougherty, E. Wang, F. Marincola, C. Gooden, J. Lueders, A. Glatfelter, P. Pollock, J. Carpten, E. Gillanders, D. Leja, K. Dietrich, C. Beaudry, M. Berens, D. Alberts, and V. Sondak. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**: 536-540.
- Cepko, C.L. 1996. The patterning and onset of opsin expression in vertebrate retinæ. *Curr Opin Neurobiol* **6**: 542-546.
- Cho, Y.S., M.K. Kim, C. Cheadle, C. Neary, K.G. Becker, and Y.S. Cho-Chung. 2001. Antisense DNAs as multisite genomic modulators identified by DNA microarray. *Proc Natl Acad Sci U S A* **98**: 9819-9823.
- DeRisi, J., L. Penland, P.O. Brown, M.L. Bittner, P.S. Meltzer, M. Ray, Y. Chen, Y.A. Su, and J.M. Trent. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* **14**: 457-460.
- Freund, C., D.J. Horsford, and R.R. McInnes. 1996. Transcription factor genes and the developing eye: a genetic perspective. *Hum Mol Genet* **5 Spec No**: 1471-1488.
- Freund, C.L., C.Y. Gregory-Evans, T. Furukawa, M. Papaioannou, J. Looser, L. Ploder, J. Bellingham, D. Ng, J.A. Herbrick, A. Duncan, S.W. Scherer, L.C. Tsui, A. Loutradis-Anagnostou, S.G. Jacobson, C.L. Cepko, S.S. Bhattacharya, and R.R. McInnes. 1997. Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell* **91**: 543-553.

- Golub, T.R., D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, H. Coller, M.L. Loh, J.R. Downing, M.A. Caligiuri, C.D. Bloomfield, and E.S. Lander. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**: 531-537.
- Halgren, R.G., M.R. Fielden, C.J. Fong, and T.R. Zacharewski. 2001. Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones. *Nucleic Acids Res* **29**: 582-588.
- Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J.E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush. 2000. A concise guide to cDNA microarray analysis. *Biotechniques* **29**: 548-550, 552-544, 556-562.
- Hughes, T.R., M.J. Marton, A.R. Jones, C.J. Roberts, R. Stoughton, C.D. Armour, H.A. Bennett, E. Coffey, H. Dai, Y.D. He, M.J. Kidd, A.M. King, M.R. Meyer, D. Slade, P.Y. Lum, S.B. Stepaniants, D.D. Shoemaker, D. Gachotte, K. Chakraborty, J. Simon, M. Bard, and S.H. Friend. 2000. Functional discovery via a compendium of expression profiles. *Cell* **102**: 109-126.
- Jean, D., K. Ewan, and P. Gruss. 1998. Molecular regulators involved in vertebrate eye development. *Mech Dev* **76**: 3-18.
- Lee, C.K., R. Weindruch, and T.A. Prolla. 2000. Gene-expression profile of the ageing brain in mice. *Nat Genet* **25**: 294-297.
- Levine, E.M., S. Fuhrmann, and T.A. Reh. 2000. Soluble factors and the development of rod photoreceptors. *Cell Mol Life Sci* **57**: 224-234.
- Livesey, F.J., T. Furukawa, M.A. Steffen, G.M. Church, and C.L. Cepko. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. *Curr Biol* **10**: 301-310.
- Malone, K., M.M. Sohocki, L.S. Sullivan, and S.P. Daiger. 1999. Identifying and mapping novel retinal-expressed ESTs from humans. *Mol Vis* **5**: 5.
- Mirnics, K., F.A. Middleton, A. Marquez, D.A. Lewis, and P. Levitt. 2000. Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. *Neuron* **28**: 53-67.
- Phillips, R.L., R.E. Ernst, B. Brunk, N. Ivanova, M.A. Mahan, J.K. Deanehan, K.A. Moore, G.C. Overton, and I.R. Lemischka. 2000. The genetic program of hematopoietic stem cells. *Science* **288**: 1635-1640.
- Quackenbush, J. 2001. Computational analysis of microarray data. *Nat Rev Genet* **2**: 418-427.

- Sandberg, R., R. Yasuda, D.G. Pankratz, T.A. Carter, J.A. Del Rio, L. Wodicka, M. Mayford, D.J. Lockhart, and C. Barlow. 2000. Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc Natl Acad Sci U S A* **97**: 11038-11043.
- Schena, M., D. Shalon, R.W. Davis, and P.O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.
- Shalon, D., S.J. Smith, and P.O. Brown. 1996. A DNA microarray system for analyzing complex DNA samples using two- color fluorescent probe hybridization. *Genome Res* **6**: 639-645.
- Sinha, S., A. Sharma, N. Agarwal, A. Swaroop, and T.L. Yang-Feng. 2000. Expression profile and chromosomal location of cDNA clones, identified from an enriched adult retina library. *Invest Ophthalmol Vis Sci* **41**: 24-28.
- Swain, P.K., S. Chen, Q.L. Wang, L.M. Affatigato, C.L. Coats, K.D. Brady, G.A. Fishman, S.G. Jacobson, A. Swaroop, E. Stone, P.A. Sieving, and D.J. Zack. 1997. Mutations in the cone-rod homeobox gene are associated with the cone- rod dystrophy photoreceptor degeneration. *Neuron* **19**: 1329-1336.
- Swanson, D.A., C.L. Freund, J.M. Steel, S. Xu, L. Ploder, R.R. McInnes, and D. Valle. 1997. A differential hybridization scheme to identify photoreceptor-specific genes. *Genome Res* **7**: 513-521.
- Swaroop, A., Q.L. Wang, W. Wu, J. Cook, C. Coats, S. Xu, S. Chen, D.J. Zack, and P.A. Sieving. 1999. Leber congenital amaurosis caused by a homozygous mutation (R90W) in the homeodomain of the retinal transcription factor CRX: direct evidence for the involvement of CRX in the development of photoreceptor function. *Hum Mol Genet* **8**: 299-305.
- Young, R.A. 2000. Biomedical discovery with DNA arrays. *Cell* **102**: 9-15.

CHAPTER 5

EVALUATION AND OPTIMIZATION OF PROCEDURES FOR TARGET LABELING AND HYBRIDIZATION OF CDNA MICROARRAYS

This chapter has been published in:

J. Yu, M.I. Othman, R. Farjo, S. Zarepari, S.P. MacNee, S. Yoshida, A. Swaroop.
Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays. Mol. Vis. 2002 Apr 26;8:130-7.

5.1 Abstract

The purpose of this study was to evaluate and optimize methods of target labeling and microarray hybridization using I-gene microarrays. Standardized protocols that consistently produce low background and high intensity hybridization with small amounts of starting RNA are needed to extract differentially expressed genes from a pool of thousands of unaltered genes. Two identical aliquots of RNA from P19 cell line were labeled with Cy3 or Cy5 dyes using four different methods and self-self hybridization was performed on mouse I-gene microarrays. The validity and reproducibility of these protocols were further examined using target RNAs isolated from wild-type ($Nrl^{+/+}$) and a mutant ($Nrl^{-/-}$) deficient for neural retinal leucine zipper (*Nrl*) gene. Hybridizations were also carried out on microarray slides of human genes with varying amounts of starting RNA from human retina. Using self-self hybridization, we optimized the protocols for direct labeling (R-square = 0.93), aminoallyl indirect labeling (R-square = 0.97), Genisphere 3DNA labeling (R-square = 0.96), and for microarray hybridization and washing. Although small amounts of initial RNA can be used in TSA method, inconsistent labeling was encountered under our experimental conditions. When retinal RNA targets from $Nrl^{+/+}$ and $Nrl^{-/-}$ mice were tested by direct and aminoallyl indirect

labeling protocols, both produced varying hybridization results with low intensity spots and non-uniform backgrounds. However, the Genisphere 3DNA labeling procedure consistently yielded strong hybridization and R-square values of 0.92 or higher. Furthermore, expression profiles were compatible with prior knowledge of this mouse model. Serial analysis of hybridizations with various starting amounts of RNA showed that the Genisphere 3DNA protocol could produce reliable signal intensity with 3 μg of total RNA. These studies should encourage further use of microarray technology for gene profiling during eye development and in retinal diseases.

5.2 Introduction

Microarray technology is widely utilized for disease diagnostics (Pomeroy et al. 2002; Sorlie et al. 2001), candidate gene identification (Bulyk et al. 2001; McDonald and Rosbash 2001), expression profiling (Eisen et al. 1998; Pollack et al. 1999), and pathway constructions (Iyer et al. 1999; Pilpel et al. 2001). Applications of this relatively new approach in vision research are rapidly growing with exciting prospects. Downstream targets of the photoreceptor homeobox gene *Crx* have been described by applying a small set of cDNA microarrays on *Crx*^{+/+} and *Crx*^{-/-} mouse models (Livesey et al. 2000). Expression profiles of specific eye tissues (Jun et al. 2001), retinal diseases (Joussen et al. 2001), or biological processes (Cavallaro et al. 2001; Shelton et al. 1999) have been examined. To enlarge the repertoire of eye (particularly retina) expressed genes and to enhance the potential use of cDNA microarray technology in vision research, a large amount of ESTs expressed in eye and retina have been sequenced (Buraczynska et al. 2002; Gieser and Swaroop 1992; Mu et al. 2001; Sinha et al. 2000) and eye gene microarrays generated (Farjo et al. 2002).

Owing to its high-throughput nature, cDNA microarray technology is vulnerable to systematic variations introduced during experimental processes (Kerr et al. 2000).

Although a number of statistical algorithms have been developed to normalize microarray data and to control experimental variations (Kerr and Churchill 2001; Tseng et al. 2001; Waterston et al. 2002), high quality input images are still the prerequisite for obtaining significant new output. This requires reproducible procedures for labeling of cDNA targets, prehybridization, hybridization, and washing of slides to consistently generate high intensity and low background images (high signal-to-noise ratios). A number of protocols, including direct and indirect labeling of cDNA targets, have been utilized in different laboratories (Yang et al. 2002) or by various vendors (Genisphere, Hatfield, PA; Perkin-Elmer, Boston, MA). However, a careful and systematic evaluation of these protocols has not been described, especially using slide microarrays. Furthermore, many of these techniques require over 10 μg of starting RNA, which makes it difficult to perform multiple microarray experiments when using eye tissues.

We have produced mouse eye and human retinal pigment epithelium (RPE) gene microarrays at the Sensory Gene Microarray Node, [Kellogg Eye Center](#), University of Michigan. Using several batches of slide microarrays, we have evaluated available protocols for labeling, hybridization and washing. In addition, we investigated the minimum amount of starting RNA needed to yield reproducible results using the preferred Genisphere 3DNA labeling method.

5.3 Methods

Generation of eye gene microarrays

Two cDNA libraries were constructed from mouse eyes at embryonic day 15.5 and postnatal day 2.5, respectively, and a third from adult retinas. The cDNA clones were isolated, amplified, and printed onto Corning CMT-GAPS slides (Corning Inc., Corning, NY), as previously described (Farjo et al. 2002). Two sets of mouse eye gene arrays,

M2500 and M6000, were generated, containing nearly 2500 or 6000 cDNAs, respectively, that were randomly printed in duplicate. M2500 arrays were used to optimize microarray procedures by self-self hybridization, while M6000 arrays were hybridized with two different RNA targets to validate the optimized methods in identifying differential expression. To generate human microarrays, ESTs were obtained from two cDNA libraries constructed from native human RPE (Buraczynska et al. 2002). Slide arrays containing over 2500 of these clones (called H2500) were printed in duplicate as replicated super-grids. A range of starting amounts of RNA targets were hybridized to H2500 slides to estimate the minimum amount of RNA required for use with the Genisphere 3DNA labeling method.

Cell culture, tissue preparation and RNA isolation

P19, a teratocarcinoma cell line derived from an embryonic carcinoma induced in a C3H/He strain mouse, was cultured in alpha MEM (Sigma, St. Louis, MO) with ribonucleosides and deoxyribonucleosides adjusted to contain 1.5 g/L sodium bicarbonate, supplemented with 10% fetal bovine serum and 0.5 μ M retinoic acid (Sigma) at 37°C in a humidified atmosphere of 5% CO₂. Cells in tissue culture plates were washed with ice-cold PBS, and homogenized in Trizol (Invitrogen, Carlsbad, CA) for RNA isolation.

Retinas from Nrl^{+/+} and Nrl^{-/-} mice (Mears et al. 2001) were dissected at postnatal day 21. Animals utilized in this study were handled as approved by the University Committee on Use and Care of Animals (UCUCA, Ann Arbor, MI). Dissected retinas were immediately frozen on dry ice and kept in -80°C freezer until use.

Two pairs of human eyes, one from a 46-year-old donor and the other from a 24-year-old donor, were obtained few hours after death from the Michigan Eye Bank (Ann Arbor, MI) and the National Disease Research Interchange (Philadelphia, PA),

respectively. The donor eyes were acquired for research purposes with family consent and processed in compliance with University of Michigan regulations. The retina was dissected from the eye, rapidly frozen on dry ice, and kept at -80°C .

Total RNA was isolated using Trizol reagent (Invitrogen) and further purified by RNeasy kit (Qiagen, Valencia, CA). Purity and RNA integrity were evaluated by absorbance at 260 nm and 280 nm, and by denaturing formaldehyde agarose gel electrophoresis. High quality RNAs with A260/A280 ratio over 1.9 and intact ribosomal 28S and 18S RNA bands were utilized for the microarray experimentation.

Direct labeling of cDNA targets

The direct labeling methods for fluorescent cDNA targets were reported previously (Farjo et al. 2002) (Figure 5.1A) and are briefly summarized as follows. A mixture of 10 μg total RNA and 2 μg oligo-dT in a total volume of 22 μl was heated to 70°C for 10 min and chilled on ice for 4 min. A reverse transcription labeling mixture of 18 μl was added to RNA to provide a final concentration of 25 μM dATP, 25 μM dGTP, 25 μM dTTP, 12.5 μM dCTP, 10 mM DTT, 1X first-stand buffer, 400 U SuperScript II, 40 U RNase inhibitor, and 12.5 μM Cy3-dCTP or 25 μM Cy5-dCTP (Amersham, Piscataway, NJ). The reaction was incubated at 42°C for 2 h to generate fluorescent-labeled cDNA. Starting RNA template was removed by adding 2 U RNase H and 10 μg RNase A, followed by incubation at 37°C for 15 min. Cy3 or Cy5-labeled cDNA targets were mixed together, purified using QIAquick PCR purification kit (Qiagen, Valencia, CA) and concentrated to 10 μl . Blocking reagents, including 1 μg poly(A) RNA (Sigma), 2 μg mouse Cot-1 DNA (Invitrogen), 1 μg yeast tRNA (Invitrogen), and 10 μg salmon sperm DNA (Invitrogen) were added to the labeled cDNA, followed by the addition of an equal volume of 2X hybridization buffer (50% formamide, 10X SSC, 0.2% SDS).

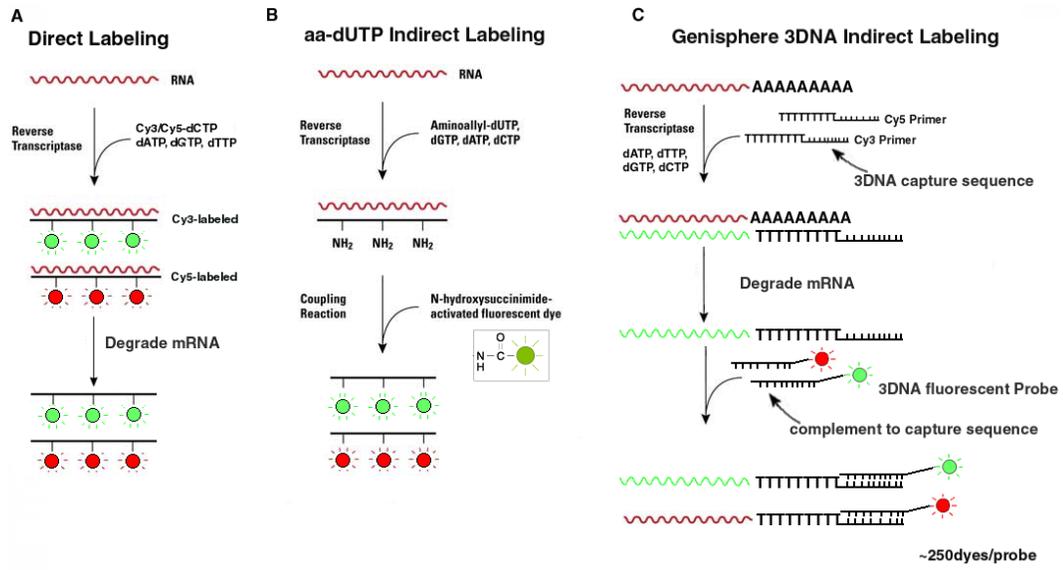


Figure 5.1. Target labeling procedures.

Schematic representations of experimental procedures. **A:** Direct labeling. **B:** Aminallyl indirect labeling. **C:** Genisphere 3DNA labeling.

Aminoallyl indirect labeling of cDNA targets

A number of protocols for aminoallyl indirect labeling (Figure 5.1B) were evaluated. The method we used is as follows. Briefly, 10 µg total RNA and 5 µg oligo-dT primers were mixed to a final volume of 18.4 µl, incubated at 70°C for 10 min and snap-cooled in ice. Reverse transcription labeling mixture (11.6 µl) was then added to RNA to obtain a labeling reaction, containing 0.5 mM dATP, dCTP, dGTP, 0.3 mM dTTP, 0.2 mM aminoallyl-dUTP (aa-dUTP), 400 U SuperScript II, 10 mM DTT, and 1X first strand buffer. This mixture was incubated at 42°C for 3 h or overnight to generate aminoallyl-labeled cDNA. To hydrolyze RNA template, 10 µl 1 M NaOH and 10 µl 0.5 M EDTA were added to the reaction and incubated at 65°C for 15 min. The reaction was neutralized by 25 µl 1 M Tris-HCl (pH 7.5). Unincorporated aa-dUTP and free amines were removed by QIAquick PCR purification kit (Qiagen) and the sample was then vacuum dried. Aminoallyl-cDNA pellet was resuspended in 4.5 µl 0.1 M sodium carbonate buffer (pH 9.0) and coupled with Cy3 or Cy5 monoreactive dye (Amersham) prepared in DMSO for 1 h at room temperature in the dark. Uncoupled dyes were removed by QIAquick PCR purification kit (Qiagen). Cy3 and Cy5 labeled cDNA targets were mixed, vacuum dried and resuspended in 45 µl GlassHyb (Clontech, Palo Alto, CA).

MICROMAX Tyramide Signal Amplification (TSA) labeling of cDNA targets

This labeling was carried out using MICROMAX TSA Labeling and Detection Kits (Perkin-Elmer, Boston, MA) as previously described (Yoshida et al. 2002), except that 2 µg total RNA was used for each of the biotin or dinitrophenyl labeling.

Genisphere 3DNA indirect labeling of cDNA targets

Labeling of total RNA using this method is outlined in Figure 5.1C, and was performed using 3DNA Submicro Expression Array Detection kit according to manufacture's protocol (Genisphere, Hatfield, PA). Briefly, total RNA was reverse transcribed using reverse transcription (RT) primers tagged with either Cy3 or Cy5 specific 3DNA capture sequence. The synthesized tagged cDNAs were then fluorescently labeled by Cy3-3DNA or Cy5-3DNA based on the complementarity of capture sequence with 3DNA capture reagents.

Microarray hybridization and signal detection

Microarray slides were prehybridized in buffer containing 5X SSC, 1% Bovine Serum Albumin and 0.1% SDS at 42-50°C for 1 h and washed by dipping five times in distilled water. The slides were then dipped in isopropanol for 1 s and centrifuged at 1000 rpm for 2 min to dry in 50 ml un-capped centrifuge tubes.

The direct or aminoallyl indirect labeled targets were heated at 95°C for 5 min, snap-cooled on ice for 30 s, and applied to prehybridized slide in a CMT-Hybridization chamber (Corning Inc., Corning, NY). Genisphere 3DNA labeled targets were incubated at 75-80°C for 10 min, followed by 50°C for 20 min before applied to prehybridized slide. A 22 x 60 mm coverslip (Grace Bio-Lab, Bend, OR) was cleaned with compressed air and then gradually placed on the slide to form a thin layer of labeled targets. To maintain humidity inside the chamber, 10 µl DEPC water was added to the two reservoir wells. The chamber was then tightly sealed and incubated at 42-50°C water bath overnight for 16-20 h. Slide was then removed from the chamber, washed for 10 min sequentially in 2X SSC/0.2% SDS buffer, 2X SSC buffer and 0.2X SSC buffer, rinsed in distilled water for 5 s, and dried by centrifugation at 1000 rpm for 2 min.

The hybridized slides were scanned with Affymetrix 428 scanner (Affymetrix, Santa Clara, CA) using appropriate gains on the photomultiplier tube (PMT) to obtain the highest intensity without saturation. A 16 bit TIFF image was generated for each channel, Cy3 and Cy5.

Image extraction and data analysis

Scanned images for Cy3 and Cy5 were then overlaid with GLEAMS software (NuTec, Atlanta, GA). This software utilizes auto-segmentation and edge detection to calculate spot intensities and backgrounds. Signal-to-noise ratios (SNR) were calculated as the mean pixel intensity over the intensity standard deviation for all pixels in a spot. For each slide, an Excel-type spreadsheet was generated for further analyses. Spots with background-subtracted intensity lower than 100 in either Cy3 or Cy5 channel were filtered out. Global normalization was then applied to correct artifacts caused by different dye incorporation rates or scanner settings for two dyes. Scatter plots in log scale were performed to visualize fold changes between two channels by plotting background-subtracted Cy5 intensity against Cy3, with parallel fold lines across data points. A linear regression (using the data without the logarithmic transformation) trend-line with intercept at origin was applied to the scatter of background-subtracted Cy5 and Cy3 intensities and the coefficient of determination (R^2 value) was calculated to indicate how well Cy5 and Cy3 intensities fit in this linear relationship.

5.4 Results and Discussion

Optimizing microarray procedures using self-self hybridization

In order to perform self-self hybridization, total RNA isolated from P19 cell line (Jones-Villeneuve et al. 1982) was divided into 2 aliquots that were labeled with Cy3 and

Cy5 dyes, respectively. Since both dye-labeled samples were identical, these hybridizations should ideally produce similar intensities in both channels for every spot. Global normalization was applied to normalize Cy5 intensity against Cy3, based on the assumption that total intensity of Cy5 channel is equal to that of Cy3. This approach was adapted to account for different dye incorporation ratios, various scanning scales, and other systematic variations. The normalized data should produce a Cy3 and Cy5 overlaid false color image of primarily yellow and a scatter plot with the majority of spots having Cy3 intensity similar to Cy5. The four labeling methods were tested multiple times using M2500 slides and modified repeatedly to achieve the best possible results with slide microarrays.

The TSA method gave inconsistent labeling with either Cy3 or Cy5 in 3 trials of self-self hybridizations. We opted not to use this method because of time consuming post-hybridization manipulations and bias generated by signal amplification that may cause inconsistency.

A number of direct cDNA fluorescent labeling protocols, including those developed by Microarrays Inc. (Nashville, TN), Corning microarray technology (Corning Inc.), Amersham (Piscataway, NJ), and The Institute for Genomic Research (TIGR, Rockville, MD) (Hegde et al. 2000), were evaluated by hybridizing P19 (Cy3) against P19 (Cy5) RNA targets using M2500 slides. Various blocking reagents, hybridization buffers, and washing conditions were also tried to produce optimal hybridization with highest intensity and lowest background. Overlaid images of Cy3 and Cy5 showed mostly yellow spots, although for low intensity spots preferential incorporation of Cy5 was observed (Figure 5.2A). The scatter plots demonstrated that 95% of spots lie within -2 and +2 fold lines (Farjo et al. 2002).

Since Cy3 and Cy5 have different incorporation rates in direct labeling methods, aminoallyl indirect labeling protocols were examined. A protocol from TIGR consistently yielded low background and high intensity hybridizations in 4 self-self experiments using

M2500 arrays; however, for low hybridization spots, signals from Cy5 channel were still slightly higher than Cy3 (Figure 5.2B). Scatter plot analysis of these slides showed regression with R^2 of 0.97 (Farjo et al. 2002).

3DNA indirect fluorescent labeling method utilizes DNA dendrimer probes that include a "capture sequence," which is complementary to 5'-end sequence of either Cy3 or Cy5 tagged RT primers. Self-self hybridizations using cDNA targets labeled by 3DNA method produced mostly yellow overlaid images of Cy3 and Cy5 (Figure 5.2C) with tight scatter plots (Figure 5.3). Furthermore, for low hybridization spots, it produced low signals for both channels, reflecting equal incorporation of dyes (Figure 5.2C).

Validation of labeling methods using two different RNA targets

Optimized protocols for direct, aminoallyl indirect, and Genisphere 3DNA methods were used to label retinal RNAs from *Nrl*^{+/+} mice by Cy3 and *Nrl*^{-/-} mice by Cy5. The labeled targets were then hybridized to M6000 arrays to identify differentially expressed genes. Since only a specific set of genes are altered in the *Nrl*^{-/-} mice retina relative to *Nrl*^{+/+} (Mears et al. 2001), a tight scatter plot with few outlier spots is expected. Scatter plot analysis was applied to normalized data from three hybridizations utilizing the same batch of RNA labeled with different techniques (Figure 5.4A,C,E). Both direct and aminoallyl indirect labeled targets generated hybridization results with uneven background and scatter plots showed regression with R^2 less than 0.9. On the other hand, Genisphere 3DNA labeling protocol consistently produced results with 0.92 R^2 values. Furthermore, outlier spots generated by this method identified genes that were shown to be differentially expressed in the two RNAs (Mears et al. 2001). Flip-over hybridizations, with RNA from *Nrl*^{+/+} mice retina labeled by Cy5 and RNA from *Nrl*^{-/-} mice retina by Cy3, showed that both direct and aminoallyl indirect labeling have signals biased to Cy5 (Figure 5.4B,D). Dye flip-over experiments using Genisphere 3DNA

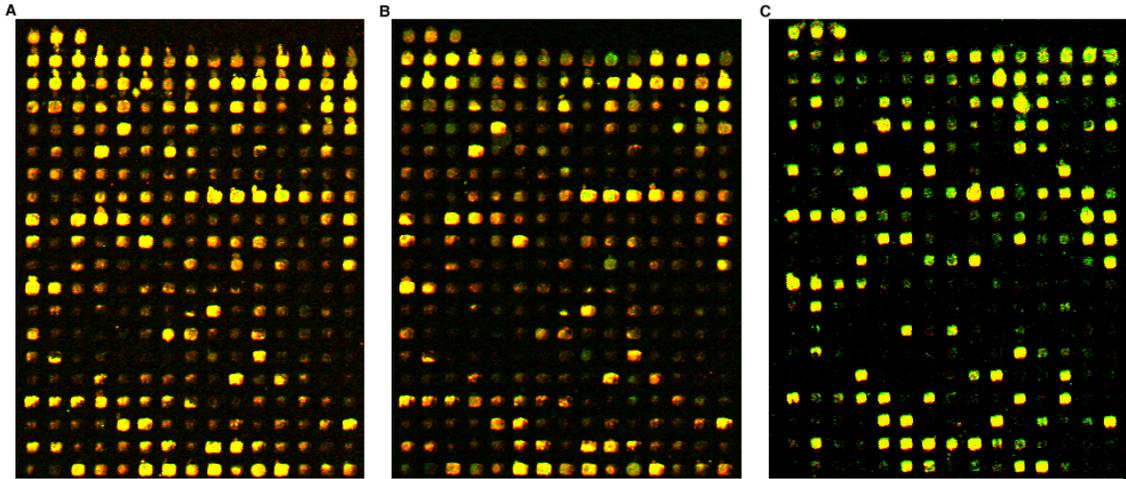


Figure 5.2. False color overlaid images

Cy3 (green) and Cy5 (red) images of a sub-grid of M2500 slides after self-self hybridization were overlaid to show relative expression of each spots in both channels. False color overlaid images of Cy3 (green) and Cy5 (red) channels of a sub-grid from self-self hybridization using M2500 slides. In both **A** (direct labeling) and **B** (aminoallyl indirect labeling), higher intensity in Cy5 was detected for spots that have low hybridization. Genisphere 3DNA indirect labeling (**C**) produced primarily yellow overlaid images for abundant genes and low signal in both channels for low hybridization spots, indicating equal incorporation of dyes.

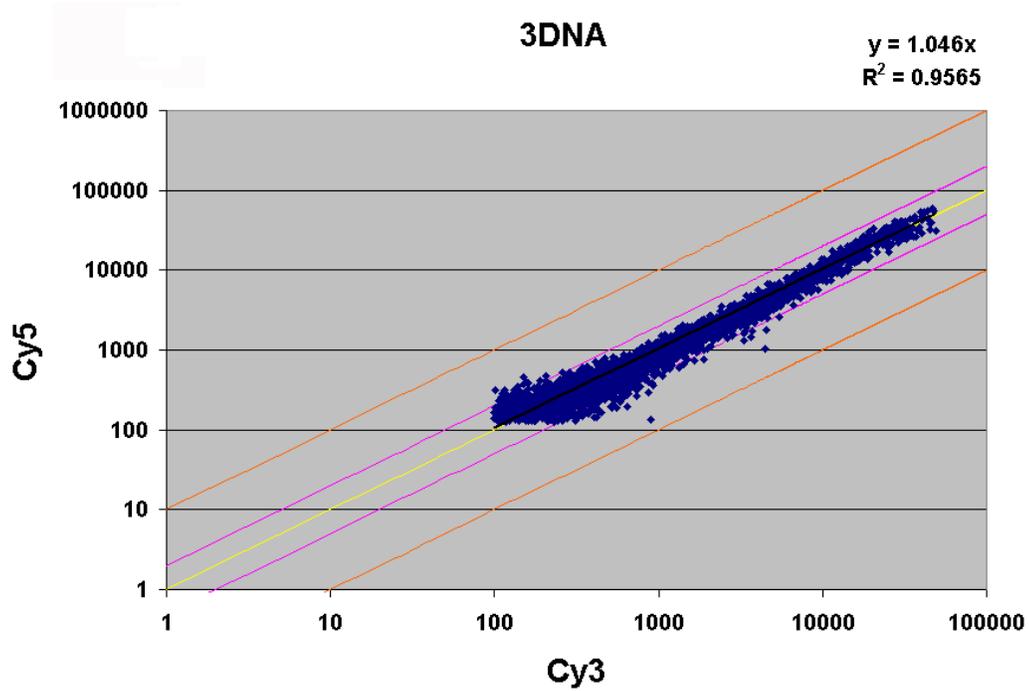


Figure 5.3. Scatter plot of self-self hybridization

Genisphere 3DNA labeled targets were hybridized to M2500 slides, containing over 5000 spots. The scatter plot indicates a majority of spots lie within 2-fold lines. Linear regression of data shows an R^2 value of 0.96.

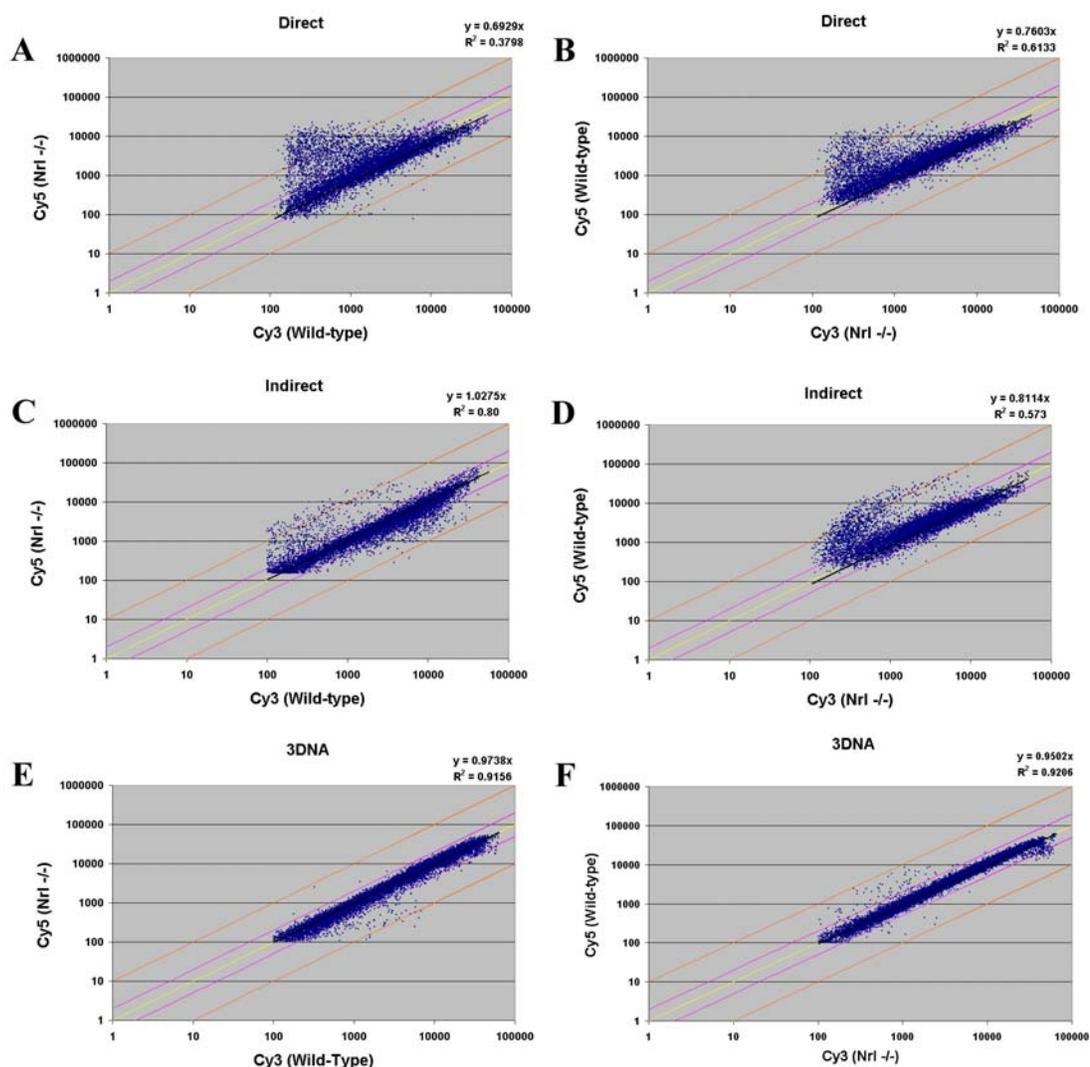


Figure 5.4. Scatter plots of two targets hybridization

Two different targets, labeled using RNA either from $Nrl^{+/+}$ (Wild-type) or $Nrl^{-/-}$ mice retina, were hybridized to M6000 slides, which contains over 13000 spots. The graphs on the left represent data generated by labeling RNA from $Nrl^{+/+}$ mouse retina by Cy3 and RNA from $Nrl^{-/-}$ mice retina by Cy5, while the graphs on the right was created by dye flip-over hybridizations. Similar scatter plots were observed between **A** and **B** (direct), and **C** and **D** (aminoallyl indirect), although a reverse pattern was expected for dye flip-over experiments (as shown in **E** and **F** with Genisphere 3DNA method). Furthermore, the Genisphere 3DNA method results in a higher R^2 value of 0.92 in both **E** and **F**, compared to 0.4 in **A**, 0.6 in **B**, 0.8 in **C**, and 0.6 in **D**.

method generated flip-over scatter plots indicating equal incorporation of Cy3 and Cy5 dyes (Figure 5.4F).

Hybridization with different amounts of total RNA

To empirically estimate the lowest amount of RNA required for high quality hybridization, a series of self-self hybridizations were performed using H2500 slides with 0.25, 0.5, 1, 2, 3, and 5 μg starting RNA isolated from two pairs of human retinas. The target RNAs were labeled by the Genisphere 3DNA method. In spite of the six different RNA amounts used, high R^2 scatter plots were consistently obtained indicating equal incorporation of Cy3 and Cy5 dyes. However, hybridizations with less than 2 μg of RNA produced relatively low and variable spot intensities. In addition, the use of 3 μg or more total RNA resulted in higher signal-to-noise ratios (Figure 5.5).

In summary, we have evaluated and optimized protocols for fluorescent labeling of cDNA targets and hybridization conditions for cDNA microarray experiments. Four major labeling techniques, including direct, aminoallyl indirect, TSA, and 3DNA method, were examined using mouse eye arrays or human RPE arrays. The Genisphere 3DNA labeling method produced superior and consistent results in both self-self and $\text{Nrl}^{+/+}$ versus $\text{Nrl}^{-/-}$ mice retina RNA hybridizations. This procedure was found to be less time consuming and more robust. We believe that these protocols can serve as templates for researchers that intend to use slide microarrays for investigating expression changes during eye development and disease.

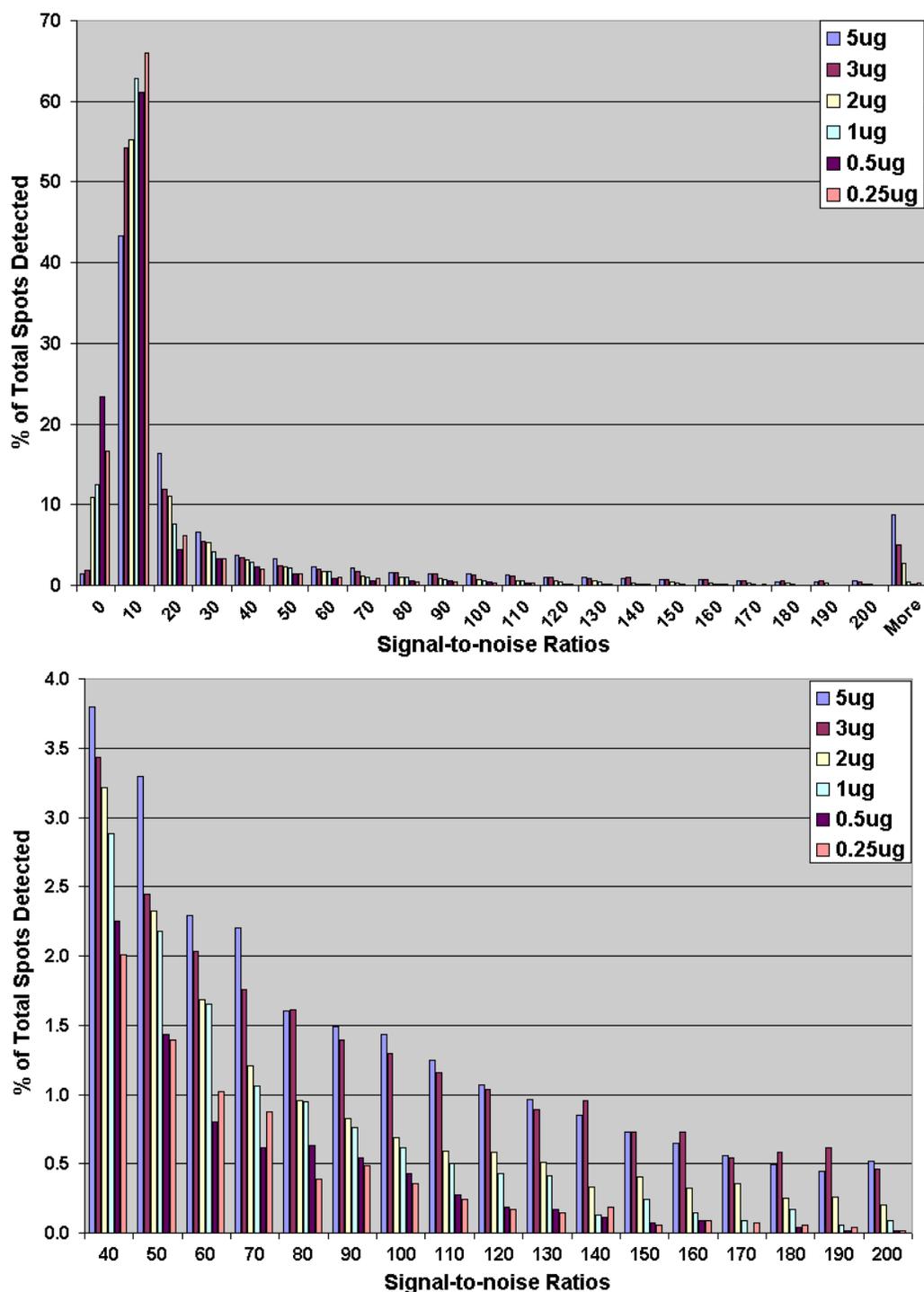


Figure 5.5. Signal-to-noise ratio (SNR) for serial hybridization

Histogram of SNR were constructed for all spots detected in a series of hybridizations performed with different amounts of starting RNA: 5, 3, 2, 1, 0.5, and 0.25 μg . A: With higher amounts of RNA, fewer spots have SNR lower than 20. B: Targets derived from 3-5 μg of total RNA resulted in more spots with higher SNR

5.5 References

- Bulyk, M.L., X. Huang, Y. Choo, and G.M. Church. 2001. Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc Natl Acad Sci U S A* **98**: 7158-7163.
- Buraczynska, M., A.J. Mears, S. Zarepari, R. Farjo, E. Filippova, Y. Yuan, S.P. MacNee, B. Hughes, and A. Swaroop. 2002. Gene expression profile of native human retinal pigment epithelium. *Invest Ophthalmol Vis Sci* **43**: 603-607.
- Cavallaro, S., B.G. Schreurs, W. Zhao, V. D'Agata, and D.L. Alkon. 2001. Gene expression profiles during long-term memory consolidation. *Eur J Neurosci* **13**: 1809-1815.
- Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**: 14863-14868.
- Farjo, R., J. Yu, M.I. Othman, S. Yoshida, S. Sheth, T. Glaser, W. Baehr, and A. Swaroop. 2002. Mouse eye gene microarrays for investigating ocular development and disease. *Vision Res* **42**: 463-470.
- Gieser, L. and A. Swaroop. 1992. Expressed sequence tags and chromosomal localization of cDNA clones from a subtracted retinal pigment epithelium library. *Genomics* **13**: 873-876.
- Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J.E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush. 2000. A concise guide to cDNA microarray analysis. *Biotechniques* **29**: 548-550, 552-544, 556-562.
- Iyer, V.R., M.B. Eisen, D.T. Ross, G. Schuler, T. Moore, J.C. Lee, J.M. Trent, L.M. Staudt, J. Hudson, Jr., M.S. Boguski et al. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* **283**: 83-87.
- Jones-Villeneuve, E.M., M.W. McBurney, K.A. Rogers, and V.I. Kalnins. 1982. Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. *J Cell Biol* **94**: 253-262.
- Joussen, A.M., S. Huang, V. Poulaki, K. Camphausen, W.D. Beecken, B. Kirchhof, and A.P. Adamis. 2001. In vivo retinal gene expression in early diabetes. *Invest Ophthalmol Vis Sci* **42**: 3047-3057.
- Jun, A.S., S.H. Liu, E.H. Koo, D.V. Do, W.J. Stark, and J.D. Gottsch. 2001. Microarray analysis of gene expression in human donor corneas. *Arch Ophthalmol* **119**: 1629-1634.

- Kerr, M.K. and G.A. Churchill. 2001. Statistical design and the analysis of gene expression microarray data. *Genet Res* **77**: 123-128.
- Kerr, M.K., M. Martin, and G.A. Churchill. 2000. Analysis of variance for gene expression microarray data. *J Comput Biol* **7**: 819-837.
- Livesey, F.J., T. Furukawa, M.A. Steffen, G.M. Church, and C.L. Cepko. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. *Curr Biol* **10**: 301-310.
- McDonald, M.J. and M. Rosbash. 2001. Microarray analysis and organization of circadian gene expression in *Drosophila*. *Cell* **107**: 567-578.
- Mears, A.J., M. Kondo, P.K. Swain, Y. Takada, R.A. Bush, T.L. Saunders, P.A. Sieving, and A. Swaroop. 2001. Nrl is required for rod photoreceptor development. *Nat Genet* **29**: 447-452.
- Mu, X., S. Zhao, R. Pershad, T.F. Hsieh, A. Scarpa, S.W. Wang, R.A. White, P.D. Beremand, T.L. Thomas, L. Gan et al. 2001. Gene expression in the developing mouse retina by EST sequencing and microarray analysis. *Nucleic Acids Res* **29**: 4983-4993.
- Pilpel, Y., P. Sudarsanam, and G.M. Church. 2001. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet* **29**: 153-159.
- Pollack, J.R., C.M. Perou, A.A. Alizadeh, M.B. Eisen, A. Pergamenschikov, C.F. Williams, S.S. Jeffrey, D. Botstein, and P.O. Brown. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23**: 41-46.
- Pomeroy, S.L., P. Tamayo, M. Gaasenbeek, L.M. Sturla, M. Angelo, M.E. McLaughlin, J.Y. Kim, L.C. Goumnerova, P.M. Black, C. Lau et al. 2002. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415**: 436-442.
- Shelton, D.N., E. Chang, P.S. Whittier, D. Choi, and W.D. Funk. 1999. Microarray analysis of replicative senescence. *Curr Biol* **9**: 939-945.
- Sinha, S., A. Sharma, N. Agarwal, A. Swaroop, and T.L. Yang-Feng. 2000. Expression profile and chromosomal location of cDNA clones, identified from an enriched adult retina library. *Invest Ophthalmol Vis Sci* **41**: 24-28.
- Sorlie, T., C.M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98**: 10869-10874.

- Tseng, G.C., M.K. Oh, L. Rohlin, J.C. Liao, and W.H. Wong. 2001. Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res* **29**: 2549-2557.
- Waterston, R.H. K. Lindblad-Toh E. Birney J. Rogers J.F. Abril P. Agarwal R. Agarwala R. Ainscough M. Alexandersson P. An et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562.
- Yang, Y.H., S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai, and T.P. Speed. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**: e15.
- Yoshida, S., B.M. Yashar, S. Hiriyanna, and A. Swaroop. 2002. Microarray analysis of gene expression in the aging human retina. *Invest Ophthalmol Vis Sci* **43**: 2554-2560.

CHAPTER 6

**ALTERED EXPRESSION OF GENES OF THE BMP/SMAD AND
WNT/CALCIUM SIGNALING PATHWAYS IN THE CONE-ONLY NRL-
KNOCKOUT MOUSE RETINA, REVEALED BY GENE PROFILING USING
CUSTOM CDNA MICROARRAYS**

This chapter has been submitted for publication as:

J. Yu, S. He, J. S. Friedman, A. J. Mears, D. Ghosh, D. Hicks, A. Swaroop. *Altered expression of genes of the Bmp/Smad and Wnt/Ca²⁺ signaling pathways in the cone-only Nrl-knockout mouse retina, revealed by gene profiling using custom cDNA microarrays.*

6.1 Summary

Many mammalian retinas are rod-dominant, and hence our knowledge of cone photoreceptor biology is relatively limited. To gain insights into molecular and functional differences between rods versus cones, we examined gene expression profiles of wild type and *Nrl*^{-/-} (Neural retina leucine zipper knockout) mouse retina, which unlike the former has no rods but a large number of functional cones. Our analysis, using custom microarrays of eye-expressed genes, provided equivalent data using either direct or reference-based experimental designs, and confirmed differential expression of rod- and cone-specific genes in the *Nrl*^{-/-} retina. We detected altered expression of several genes that encode cell signaling or structural proteins. Prompted by these findings, additional real-time PCR analysis revealed that genes belonging to the Bmp/Smad and Wnt/Ca²⁺ signaling pathways are expressed in the mature wild type retina and that their expression is significantly altered in the *Nrl*^{-/-} retina. Chromatin immunoprecipitation

studies suggest that Bmp4 and Smad4, two genes down-regulated in the $Nrl^{-/-}$ retina, may be direct transcriptional targets of Nrl. Activation of the Smad-mediated Bmp pathway by Nrl may provide a mechanism for integrating cell signaling networks in photoreceptors. We propose that Bmp/Smad and Wnt/ Ca^{2+} pathways participate in cell-cell communication in the mature retina and expression changes observed in the $Nrl^{-/-}$ retina reflect their biased utilization in rod versus cone homeostasis.

6.2 Introduction

In mammals, vision is initiated in the retina, a highly structured part of the brain consisting of over 50 types of neurons that are organized in three distinct layers (Dowling 1987; Masland 2001). The outer nuclear layer consists exclusively of two types of photoreceptors, rods and cones, responsible for detection and transduction of light energy. Rods function under low ambient light and form the major photoreceptor population of many mammals, including humans (95%) and mice (97%). Cones are responsible for phototransduction in bright light, providing high acuity and color vision (Dowling 1987). Cones are present in higher proportion in the central region of the retina, and their survival is essential for maintaining central vision, which is compromised in several retinal and macular diseases (Curcio et al. 2000; Hicks and Sahel 1999). Rods and cones possess distinct subsets of proteins involved in the phototransduction cascade (Molday 1998; Nathans 1999), but despite significant neuroanatomical and physiological advances (Leskov et al. 2000; Marc and Jones 2003; Masland 2001), little progress has been made towards delineating the molecular mechanisms that underlie functional distinctions between the two photoreceptor types, their communication with other neurons, and their maintenance, survival or remodeling in response to extrinsic or intrinsic insults.

One approach to dissect systematically the regulatory networks and molecules associated with rod or cone photoreceptor function is to take advantage of animal models that exhibit preferential utilization of one or the other photoreceptor subtype. Unfortunately, many species with cone-rich retinas (e.g., ground squirrel, chick) present difficulties with respect to experimentation and are less amenable to genetic manipulations. Because of a large number of naturally-occurring and experimentally-generated mutants (<http://jaxmice.jax.org/info/index.html>), the mouse offers a unique opportunity to examine this complex question. The *Nrl*-knockout (*Nrl*^{-/-}) mouse, recently generated in our laboratory, exhibits a unique cone-only retinal phenotype (Mears et al. 2001). *Nrl*, a basic motif leucine zipper protein of the Maf-subfamily, is specifically and highly expressed in the rod photoreceptors (Swain et al. 2001; Swaroop et al. 1992) and pineal gland (Mears, A.J. and Swaroop, A., unpublished data). It interacts with other transcription factors, such as the homeodomain protein *Crx*, and regulates (either alone or synergistically) the expression of several rod genes (Chen et al. 1997; Lerner et al. 2001; Mitton et al. 2000; Mitton et al. 2003; Rehemtulla et al. 1996). Missense mutations in the human *NRL* gene are associated with autosomal dominant retinitis pigmentosa (Bessant et al. 1999; DeAngelis et al. 2002). Targeted deletion of *Nrl* (*Nrl*^{-/-}) in mouse results in a retina with no rod photoreceptors and a concomitant increase in functional S-opsin expressing cones (Mears et al. 2001). This apparent functional switching of photoreceptor subtypes (from rods to cones) has been validated by histology, electrophysiology, biochemical and molecular analysis (Nikonov, S.S., Daniele, L., Mears, A.J., Swaroop, A. and Pugh, E.N., Jr. J., unpublished data) (Mears et al. 2001; Zhu et al. 2003).

Microarray-based global profiling of gene expression, in combination with bioinformatic tools, can yield valuable insights into cell- or tissue-specific functions. Expression profiling of tissues from mice deficient in a transcription factor gene can point to downstream regulatory targets, provide candidates for functional studies, and

facilitate positional cloning of human disease loci (DeRyckere and DeGregori 2002; Horton et al. 2003; Livesey et al. 2000; Mu et al. 2004). Analysis of the retinal transcriptome during development and aging and in mouse models of retinal dysfunction has been a subject of intense investigation (Blackshaw et al. 2001; Mu et al. 2004; Swaroop and Zack 2002; Yoshida et al. 2002). Since commercially available microarrays do not have adequate representation of genes transcribed in developing and mature eye/retina, several groups have produced custom slide microarrays of eye/retina-expressed genes (Chowers et al. 2003; Mu et al. 2004; Swaroop and Zack 2002). For gene profiling, we isolated and sequenced cDNAs from mouse eye/retina libraries, annotated over 10,000 expressed sequence tags (ESTs), and produced cDNA microarrays (called I-gene microarrays) (Farjo et al. 2002; Yu et al. 2003).

We have used the $Nrl^{-/-}$ mouse model to delineate molecular distinctions between rods and cones. We generated an expression profile of the mature retina from the cone-only $Nrl^{-/-}$ mice using I-gene microarrays and compared it to the gene profile of the rod-dominated wild type mouse retina. We demonstrate differential expression of several genes, encoding phototransduction, structural and signaling proteins, in the $Nrl^{-/-}$ retina. Of particular importance are the findings that genes encoding components of Bmp/Smad and Wnt/ Ca^{2+} signaling pathways are expressed in mature retina and their expression is altered in the $Nrl^{-/-}$ retina. Expression analysis by real-time PCR and chromatin immunoprecipitation (ChIP) studies suggest that the activity of the Smad-mediated Bmp signaling pathway is modulated by Nrl in the mature retina. We propose that rods and cones exhibit selective bias in the utilization of different signaling pathways for cell-cell communication and controlling intracellular functions.

6.3 Materials and Methods

Tissue preparation, RNA isolation and Northern analysis

All procedures involving mice were approved by the University Committee on Use and Care of Animals (UCUCA) of the University of Michigan. Retinas were dissected from the wild type and $Nrl^{-/-}$ mice at postnatal day (P) 21 and snap-frozen on dry ice. Total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA) and purified by RNeasy kit (Qiagen, Valencia, CA). RNA integrity was verified by denaturing formaldehyde-agarose gels. Total RNA samples with 260 nm to 280 nm absorbance ratio of greater than 1.9 were used for studies. Northern analysis was performed as described (Swaroop et al. 1992).

Reference RNA for retinal gene profiling

To generate the reference RNA for microarray hybridizations, total RNA was pooled from the following tissues and cell lines: mouse eye or retina at embryonic day (E) 14-16, P 2-3, P 10-12, and adult (5.5 mg); mouse adult brain (1.5 mg); P19 embryonic carcinoma cells (McBurney et al. 1982) (3 mg); retinoic acid-induced P19 cells that differentiated into neuronal and glial like cells (3 mg); and neuroblastoma cell line N1E-115 (Amano et al. 1972) (4 mg). The pooled RNA was divided into aliquots (at a concentration of 1.28 $\mu\text{g}/\mu\text{l}$) and stored at -80°C until use.

Target labeling and microarray hybridization

Mouse I-gene cDNA microarray slides (Farjo et al. 2002) contained PCR-amplified products from over 6,500 cDNAs, printed in duplicate; cDNAs were isolated from libraries constructed from E 15.5 eyes, P 2 eyes, and adult retinas, sequenced, and annotated (Yu et al. 2003). Target RNA (10 μg of total RNA) was labeled using 3DNA

Submicro Expression Array Detection Kit (Genisphere, Hatfield, PA) and hybridized to microarray slides, as described (Yu et al. 2002). The slides were scanned using Affymetrix 428 scanner (Affymetrix, Santa Clara, CA) to obtain highest intensity of signal, without reaching saturation for a maximum of 10 out of the 13,440 spots.

Image processing and data analysis

Scanned array images without major defects (such as scratches or blobs) were analyzed using AnalyzerDG (MolecularWare Inc, Cambridge, MA) in a batch mode. A “contour shape” was utilized to detect spots for intensity calculations, while a “cell method” was set to calculate background on an individual basis in a local square region centered on the spot. A data file containing spots’ intensities and annotations was exported for each array in the tab-delimited text format and then imported into the statistical package, R (<http://www.r-project.org/>). Intensities of Cy3 and Cy5 channels for each spot were calculated separately by subtracting corresponding median background from the mean signal intensity. Genes with negative background-corrected intensities in either channel were filtered out.

Normalization of intensities in Cy3 and Cy5 channels is an important step to reduce systematic variation caused by experimental procedures (such as dye effects). The most commonly used technique is the global normalization, which calculates a single normalization factor that equalizes the average/median intensities in two channels. By contrast, normalization based on a locally weighted linear squares (LOWESS) procedure (Cleveland 1979) balances two channels piecewise; different scale factors are calculated for spots within different ranges of intensities. To avoid the effects of dramatically altered genes on these scale factors, in this study, a data-driven LOWESS normalization was applied to individual datasets using a cluster of least-altered genes on the array identified by a rank-based algorithm. Briefly, let (R_j, G_j) denote the measurements for the

j^{th} gene in the red and green channels, respectively, $j = 1, \dots, m$. We calculate the ranks in each channel separately, take the difference in ranks between the two channels, and fit a three-component normal mixture model to the difference in ranks for the m genes. There will be three classes of genes to consider: those for which the rank in the red channel is significantly higher than that in the green channel, those for which the rank in the green channel is significantly higher than that in the red channel, and those for which the ranks do not substantially change between the two channels.

The genes whose ranks did not change between the two channels were used to perform a slide-dependent LOWESS normalization. This allowed us to normalize genes and redefine a new horizontal zero axis that was used to compute log-ratios of intensities in two channels. For indirect comparisons, log-ratios of the wild type slide were subtracted from the $\text{Nr1}^{-/-}$ slide to obtain reference-corrected values for each gene. An Empirical Bayes method was then applied to the replicated arrays to obtain a B statistic for each gene (Lonnstedt and Speed 2002). This was done by first calculate the estimates for the differences between the wild type and the $\text{Nr1}^{-/-}$ samples using standard linear model. Based on these estimates, the B statistic is obtained as the posterior log odds of differential expression. Compared to t-test, this method is more reliable if the variance of log-ratios of a gene is unusually small (Lonnstedt and Speed 2002).

Quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

To validate gene expression changes, qRT-PCR analysis was performed as described (Yu et al. 2003). Briefly, total RNA (2.5 μg) treated with RQ1 RNase-Free DNase (Promega, Madison, WI) was subjected to reverse transcription using oligo-d(T) primers and with (+RT sample) or without (-RT) SuperScript II (Invitrogen). For each gene, real time PCR reactions from wild type and $\text{Nr1}^{-/-}$ samples were performed in triplicate with SybrGreen I (Molecular Probes, Eugene, OR) and analyzed with the

iCycler IQ real-time PCR Detection system (Bio-Rad, Hercules, CA). The average threshold cycle (C_t) differences between the two samples were normalized against Hprt (control) in the corresponding cDNA preparation.

Immunohistochemistry

Anti-Gnb3 polyclonal antibodies were generated in rabbit against the peptide (ADITLAELVSGLEVV) and affinity-purified (Invitrogen). Due to the low number of cones in wild type mouse retina (3% of photoreceptors), we performed immunohistochemical analyses on frozen sections of adult pig retina, which contains approximately 15-20% cones (Hendrickson and Hicks 2002). Cryostat sections of pig retina were permeabilized with 0.1% Triton X-100 (5 min), and then pre-incubated in blocking buffer (PBS supplemented with 0.2% bovine serum albumin, 0.1% Tween 20, 5% normal rabbit serum and 0.1% NaN₃) for 30 min. The sections were then incubated overnight at 4°C in anti-Gnb3 polyclonal antibodies (diluted 1:200 in blocking buffer) combined with rho-4D2 anti-rod opsin monoclonal antibody (1 µg/ml) (Hicks and Molday 1986) or with biotinylated peanut agglutinin (PNA, from Vector Laboratories; 10 µg/ml). After extensive washing in PBS, sections were incubated in a mixture of goat anti-rabbit IgG-Alexa488 (Molecular Probes Inc., Eugene, OR; 2 µg/ml), and either rabbit anti-mouse IgG- or streptavidin-Alexa594 (each 2 µg/ml) in blocking buffer for 2 hr. Slides were washed extensively, mounted, and viewed by laser scanning confocal microscopy (Zeiss LSM 510 v2.5 scanning device with Zeiss Axiovert 100 inverted microscope). Control experiments were performed by omitting the primary antibody.

Chromatin Immunoprecipitation (ChIP)

A commercially available assay kit (Upstate Biotechnologies, Charlottesville, VA) was used for ChIP studies. Briefly, four snap-frozen retinas from wild type mice

were crosslinked for 15 min at 37°C with 1% formaldehyde in PBS containing proteinase inhibitors. The retinas were washed four times in ice-cold PBS with proteinase inhibitors and then incubated on ice for 15 min. The tissue was then sonicated on ice eight times using 20-second pulses. The remaining steps were essentially performed as described by the manufacturer, using anti-NRL polyclonal antibody (Swain et al. 2001).

Putative promoter regions (5'-upstream of the transcription start site) for Rho, Bmp4, Smad4 and Bmpr1a were determined *in silico* (<http://www.ncbi.nlm.nih.gov/mapview>). Each DNA sequence was analyzed using MatInspector (<http://www.genomatix.de/index.html>). PCR primers were designed to flank the putative Nrl binding sites (Nrl response element, NRE) (Rehemtulla et al. 1996) either predicted manually or by MatInspector (Table 6.1). The sequence closest to the transcription start site was chosen. ChIP DNAs from two independent experiments were used for PCR using equal amounts for input, with antibody, and no antibody reactions.

6.4 Results

We chose to generate gene expression profiles of P21 retinas from rod-dominated wild type and rodless (cone only) *Nrl*^{-/-} mice. At this stage, the differentiation and laminar organization of retinal neurons are complete, and phototransduction pathways in the retina (from photoreceptors to ganglion cells) are fully functional.

High concordance between direct and indirect microarray comparisons

We examined two different experimental designs, direct and indirect comparison (Yang and Speed 2002), for their ability to identify differentially expressed genes. Two direct comparison experiments (2 slides), in which Cy3-labeled wild type and Cy5-labeled *Nrl*^{-/-} retinal RNA targets were hybridized simultaneously to the same slide, were performed (Figure 6.1A). Five indirect comparisons were carried out with the reference

Table 6.1. Putative NRE and the PCR primers used for ChIP enrichment tests.

Gene	Putative NRE (base pairs upstream of the transcription start site)	ChIP PCR Primers
Bmp4	<u>GACAGTGACGCAGGGAATC</u> AA (-1495)	Forward 5'-AACCTGCTATGGGAGCACAG-3' Reverse 5'-GGAATGTCAGGTTGGAAGGA-3'
Smad4	<u>CAATTTTGATGACAATGAAG</u> AATTG (-1299)	Forward 5'-GCATCCAAAGGGTCATGAGT-3' Reverse 5'-GGGTAAGCCAAAGGGACAAT-3'
Bmpr1a	<u>ATGTATGACTGTGCATCACA</u> T (-1684)	Forward 5'-GGTGGATATGAGGGAATGGA-3' Reverse 5'-TGGTGGTCCACTACCATCTG-3'
Rhodopsin	<u>GGATGCTGAATCAGCCTCT</u> (-70)	Forward 5'-GATGGGATAGGTGAGTTCAGGA-3' Reverse 5'-GAGAAGGGCACATAAAAATTGG-3'

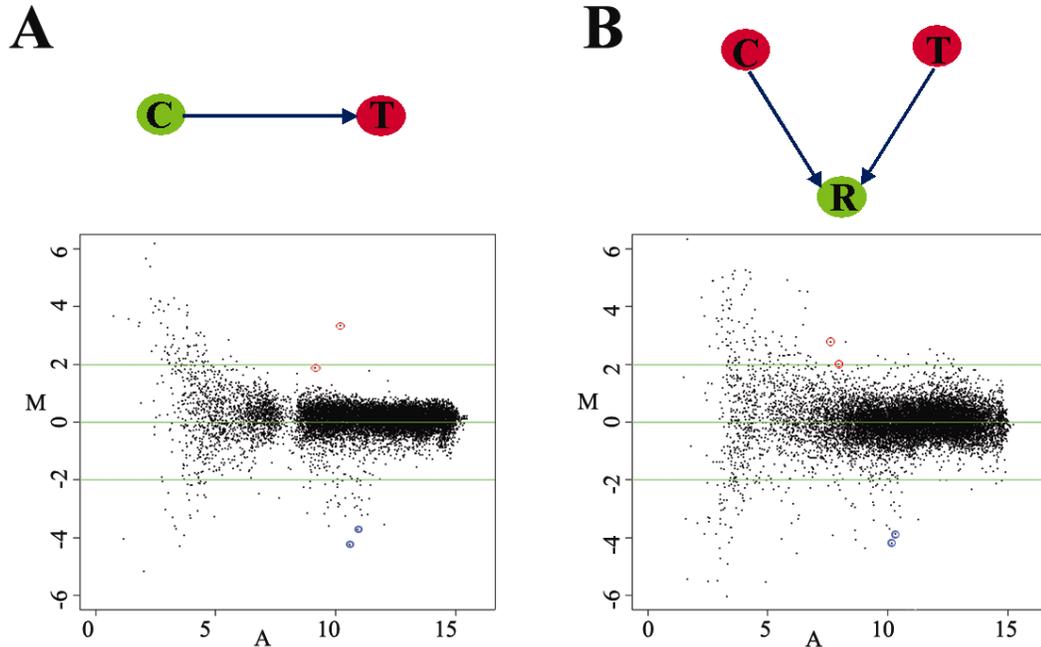


Figure 6.1. Representative scatter-plots showing differential expression of two copies each of S-opsin (circled spots above 0 axis) and Rhodospin (circled spots below 0 axis).

(A) Direct comparisons, where control (C) and treatment (T) are labeled by Cy3 (green) or Cy5 (red) and compared directly on one slide hybridization. MA-plot of wild type (G) and $Nrl^{-/-}$ (R) mouse retinal gene expression in one direct comparison experiment, where $M = \log_2(R/G)$ and $A = (\log_2(R \times G))/2$. (B) Indirect comparisons, where C and T are labeled by Cy5 (red) and compared to a common reference sample (R), labeled by Cy3 (green) in separate slide hybridization. Samples from the wild type (R_1) retina and the common reference (G_1) were hybridized to one slide, whereas samples from the $Nrl^{-/-}$ retina (R_2) and the reference (G_2) were hybridized to another slide. Expression levels of the wild type and $Nrl^{-/-}$ retinas were compared by removing the signal from common reference after data normalization. X-axis, $A = (\log_2(R_1 \times G_1) + \log_2(R_2 \times G_2))/2$. Y axis, $M = \log_2(R_1/G_1) - \log_2(R_2/G_2)$.

RNA labeled by Cy3 and hybridized in conjunction with Cy5-labeled either wild type or *Nrl*^{-/-} retinal RNA (total of 10 slides) (Figure 6.1B). Of the 13,440 spots on the I-gene microarray slide, 97.4% (13,092 spots) showed higher spot intensity than the background (i.e., were considered detected) in both direct comparison slides, whereas 91.6% (12,307 spots) spots were detected in all 10 slides with indirect comparisons. Scatter-plot analysis of signal intensities of the wild type and *Nrl*^{-/-} retinas revealed high similarity between direct and indirect comparisons (Figure 6.1). A majority of genes have log₂ ratios centered at 0, indicating no change between the two tested samples. Only few spots displayed over 4-fold change in both methods, although direct comparison showed a tighter scatter within the 4-fold lines. The variance of indirect comparison is 1.8 of that of direct comparison and is mainly due to the use of the extra reference sample. Both methods were able to successfully identify the duplicate spots of S-opsin and rhodopsin as the most up- or down-regulated genes in the *Nrl*^{-/-} retina, respectively.

To compare the power of direct versus indirect methods in identifying differential gene expression, similar statistical analysis protocols were applied to both datasets to rank the genes based on their odds of being altered. The top 40 genes identified by direct comparison included 31 genes that were also ranked in top 40 by indirect comparison. Furthermore, 37 of the top 40 genes revealed by indirect experiments were present in the top 75 genes with direct comparison (Table 6.2). This indicates a possible false positive rate of 3/40 in the indirect comparison, assuming that only genes ranked in the top 75 by both methods are truly positive.

Differential expression of genes in the *Nrl*^{-/-} retina

Based on the greatest odds of difference identified by both direct and indirect comparisons, a non-redundant set of 74 highly ranked genes was selected; of these, 50 genes are known (Table 6.3), whereas 24 cDNAs only show homology to sequences in

Table 6.2. The number of genes commonly identified by both direct and indirect comparisons as having the highest odds of differential expression in the top 40, 50 and 70 lists.

common in both		indirect comparison		
		top 40	top 50	top 70
direct comparison	top 40	31	31	31
	top 50	34	35	37
	top 70	37	44	52

the EST or genomic databases (Table 6.4). Of the 50 known differentially expressed genes, only eight demonstrate higher expression in the $Nrl^{-/-}$ retina as compared to the wild type. The majority of these genes can be divided into the following functional groups: phototransduction, transcriptional regulation, signaling pathways, and structural or membrane-associated proteins. One-third of the altered genes (24/74) represent proteins of unknown function (Table 6.4).

To determine whether gene profiling could validate previously reported patterns of gene expression (Mears et al. 2001), we first examined the phototransduction genes that are represented on I-gene microarrays. Except for the two genes, cone arrestin 3 (*Arr3*) and *Pde6b*, which did not exhibit expression change in the $Nrl^{-/-}$ retina by microarray analysis, our data confirmed previous results (see Table 6.3). To further establish the fidelity of the microarray data, we examined the overlaid microarray images of spots corresponding to three genes: *Gnb3* (cone transducin beta subunit) and its counterpart in rods - *Gnb1*, and one random EST (AC021049) (corresponding to the clone MRA-1648, with partial sequence homologous to cold shock domain protein A). The overlaid images of these three genes were in agreement with their microarray data (Figure 6.2A). Northern blot and qRT-PCR analysis further confirmed the expression differences detected by microarray experiments (Figure 6.2B, C). Consistent with these observations, *Gnb3* immunoreactivity was specific to cones: intense within outer segments, and moderate in cone cell bodies and synaptic pedicles (Figure 6.2Da,b). The staining within the outer segments was distinct from that observed for rod opsin (Figure 6.2Dc), but co-localized precisely with PNA (Figure 6.2Dd,e), a commonly used marker of cone outer matrix sheaths, demonstrating the cone specificity of *Gnb3* immunolabeling.

Table 6.3. The top 50 differentially expressed genes. Microarray and qRT-PCR columns show fold-change, where (-) and (+) indicate lower and higher expression in *Nrl*^{-/-} retina, respectively.

Accession #	Symbol	Gene name	Microarray	qRT-PCR
Phototransduction				
BC022793	<i>Gnat1</i>	Transducin-alpha 1	-8.7	-256
BC013125	<i>Rho</i>	Rhodopsin	-5.0	-256
NM_008142	<i>Gnb1</i>	Guanine nucleotide binding protein, beta 1	-3.5	-128
AF105711	<i>Tulp1</i>	Tubby like protein 1	-2.8	-8
NM_010316	<i>Gng3</i>	Guanine nucleotide binding protein, gamma 3	-2.2	-5
BC014826	<i>Opn1mw</i>	M-opsin	-1.6	
NM_007378	<i>Abca4</i>	ATP-binding protein transporter	-1.5	
NM_024458	<i>Rpr1</i>	Rod photoreceptor 1	-1.5	-256
NM_012065	<i>Pde6g</i>	cGMP phosphodiesterase 6 gamma	-1.5	-16
M24086	<i>Sag</i>	Retinal S-antigen	-1.5	
NM_007538	<i>Opn1sw</i>	S-opsin	+8.9	+256
NM_013530	<i>Gnb3</i>	Guanine nucleotide binding protein, beta 3	+3.2	+10
NM_008189	<i>Gcap1</i>	Guanylate cyclase activator 1a (retina)	+2.5	+4.6
Transcriptional regulation				
NM_007840	<i>Ddx5</i>	DEAD box polypeptide 5	-5.8	-2
NM_013708	<i>Nr2e3</i>	Nuclear receptor subfamily 2, group E, member 3	-3.0	-256
NM_008736	<i>Nrl</i>	Neural retina leucine zipper gene	-1.7	-256
NM_007905	<i>Edr</i>	Early development regulator	-1.5	
AJ252060	<i>Zranb1</i>	Zinc finger, RAN-binding domain containing 1	+1.5	
Signaling pathways				
NM_007553	<i>Bmp2</i>	Bone morphogenetic protein 2	-5.4	-4
NM_009427	<i>Tob1</i>	Transducer of ErbB-2.1	-4.8	-3
NM_007554	<i>Bmp4</i>	Bone morphogenetic protein 4	-3.7	-12
M84819	<i>Rxrg</i>	Retinoid X receptor gamma	-3.7	+5
NM_022992	<i>Arl6ip5</i>	ADP-ribosylation-like factor 6 interacting protein 5	-2.7	
NM_008714	<i>Notch1</i>	Notch homolog 1	-1.7	-1.6
NM_009758	<i>Bmpr1a</i>	Bone morphogenetic protein receptor 1A	-1.5	-32
NM_007595	<i>Camk2b</i>	Calmodulin-dependent protein kinase II beta	+2.2	+3
Structural and Membrane Proteins				
AF010137	<i>Jag2</i>	Jagged 2	-8.1	-3
NM_008663	<i>Myo7a</i>	Myosin VIIa	-3.7	-5
NM_009837	<i>Cct4</i>	Chaperonin subunit 4 (delta)	-3.7	-2
NM_005277	<i>GPM6A</i>	Glycoprotein M6A	-3.4	-2
XM_126191	<i>Unc119h</i>	Unc119 homolog	-2.9	-8
NM_010837	<i>Mtap6</i>	Microtubule-associated protein 6	-2.7	+8
M11686	<i>Krt1-18</i>	Keratin complex 1, acidic, gene 18	-2.6	-2
NM_080556	<i>Tm9sf2</i>	Transmembrane 9 superfamily member 2	-1.9	
NM_019634	<i>Tm4sf2</i>	Transmembrane 4 superfamily member 2	-1.8	
NM_008938	<i>Prph2</i>	Peripherin 2	-1.8	-3
AF044312	<i>Epb4.1l2</i>	Erythrocyte protein band 4.1-like 2	+1.8	

NM_027324	<i>Sfxn1</i>	Sideroflexin 1	+1.5	
Other functions				
NM_016764	<i>Prdx4</i>	Peroxiredoxin 4	-8.5	-7
NM_012053	<i>Rpl8</i>	Ribosomal protein L8	-2.1	
NM_008898	<i>Por</i>	P450 (cytochrome) oxidoreductase	-2.1	
J00378	<i>Cryba1</i>	Beta-A3/A1-crystallin	-2.0	
NM_010421	<i>Hexa</i>	Hexosaminidase A	-1.7	
NM_008084	<i>Gapd</i>	Glyceraldehyde-3-phosphate dehydrogenase	-1.7	
X00654	<i>Cryga</i>	Gamma-4 crystallin	-1.6	
X56974	<i>Rn18s</i>	External transcribed spacer	-1.6	
NM_076608	<i>Hsp-25</i>	Heat shock protein	-1.6	
X52379	<i>Eno1</i>	Enolase 1, alpha non-neuron	-1.4	
NM_008131	<i>Glns</i>	Glutamine synthetase	-1.4	
NM_023695	<i>Crybb1</i>	Crystallin, beta B1	+1.4	

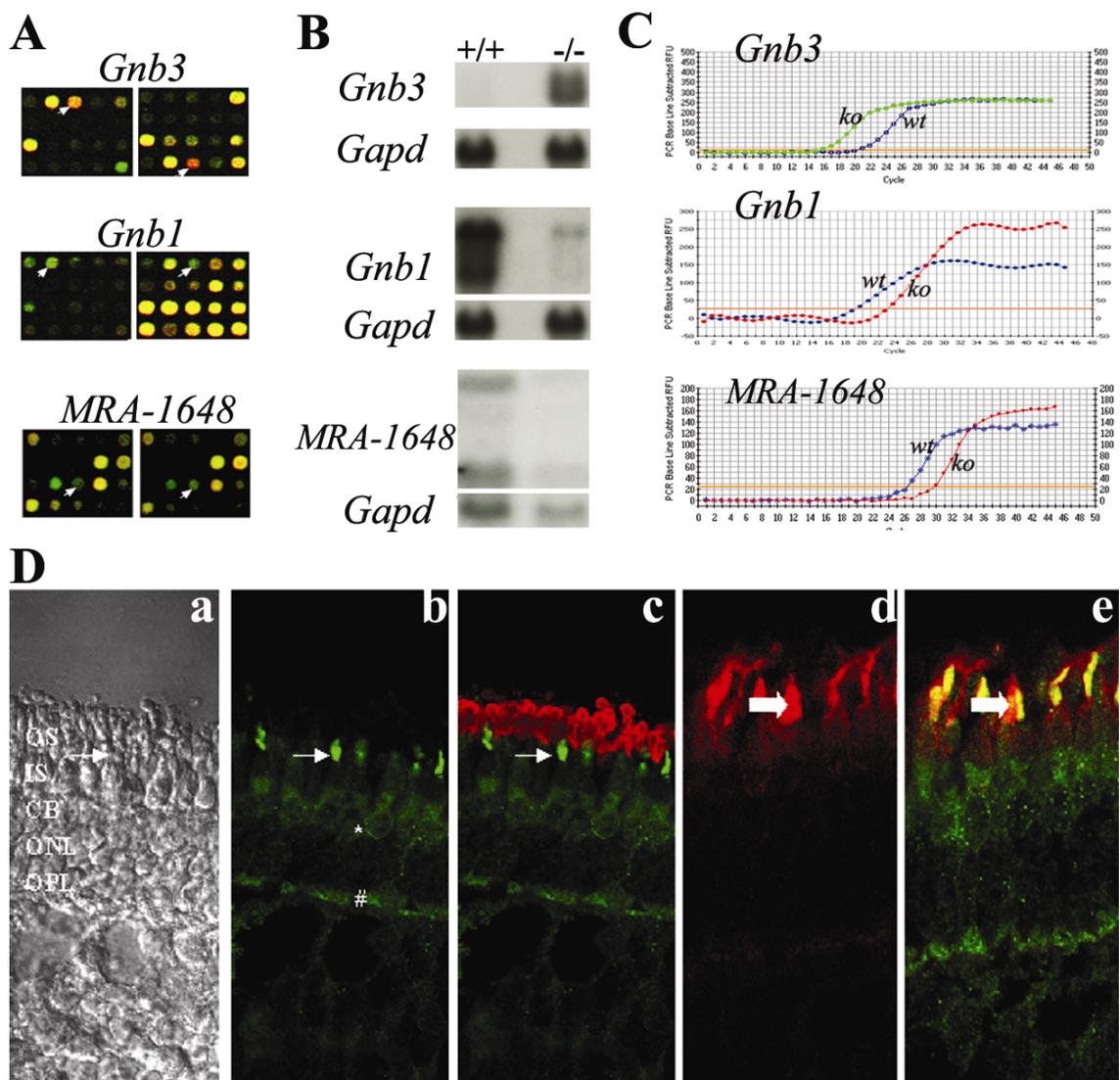


Figure 6.2. Differential expression of three genes: *Gnb3*, *Gnb1* and MRA-1648.

The overlaid images of red ($Nrl^{-/-}$) and green (wild type) channels showing microarray spots corresponding to these 3 genes (white arrows). The spots appear as green if expression is higher in wild type, red if higher in knockout, and yellow if equivalent expression is observed. (B) Northern analysis comparing total RNA levels in wild type (+/+) and $Nrl^{-/-}$ (-/-) retinas. *Gapd* was utilized to normalize the amount of loaded samples. (C) qRT-PCR analysis of transcript levels in wild type (wt) and $Nrl^{-/-}$ (ko) retinas. (D) Immunohistochemical localization of *Gnb3* in adult porcine retina: confocal images of retinal sections viewed by normal illumination to show outer layers (a), and with anti-*Gnb3* alone (b), *Gnb3* merged with rho-4D2 rod opsin staining (c), PNA alone (d), or anti-*Gnb3* merged with PNA staining (e). Intense *Gnb3* immunoreactivity was observed in cone outer segments (arrow, a-c), as well as in the cytoplasm of cone cell bodies (*, b) and within cone synaptic pedicles (#, b). Labeling with rod opsin antibody (red) and *Gnb3* antibody (green) shows the separation of the two patterns (c). Cone outer segments were outlined by specific staining with PNA (d), and merging of signals for PNA (red) and *Gnb3* (green) shows overlap of the two stains (wide arrows in both d and e). Abbreviations: CB, cone cell bodies; IS, photoreceptor inner segments; ONL, outer nuclear layer; OPL, outer plexiform layer; OS, photoreceptor outer segments. Scale bar in panel J = 20 μ m.

Table 6.4. The top 24 differentially expressed unknown genes / ESTs. Microarray and qRT-PCR columns show fold-change, where (-) and (+) indicate lower and higher expression in *Nrl*^{-/-} retina, respectively.

Clone ID	Accession #	Gene name	Microarray qRT-PCR	
MRA-1749	CB849992	EST	-6.5	-1.8
MRA-1928	AC104161	Human chromosome 3 clone RP11-32J15	-7.6	-1.9
MRA-1917	NM_021432	RIKEN cDNA 1110020M21Rik	-5.8	-3.5
MRA-1076	NM_153594	RIKEN cDNA 5330414D10Rik	-3.2	
MRA-1753	NM_025335	RIKEN cDNA 0610041E09Rik	-3.0	-2.1
MRA-1648	AC021049	Human 12p12-21.8-27.2 BAC RP11-904M10	-2.0	-6.5
MRA-1090	NM_025993	RIKEN cDNA 2510025F08Rik	-2.0	-1.6
MRA-1037	AC111007	Chromosome 3 clone RP23-226I10	-1.9	
MRA-1102	CB849372	EST	-1.8	
MRA-1865	CB850104	EST	-1.8	
MRA-1458	NM_023168	RIKEN cDNA 1110025J15Rik	-1.6	
MRA-1679	CB849926	EST	-1.5	
MRA-2092	CB850310	EST	-1.5	
MRA-1485	CB849741	EST	-1.5	
MRA-0674	AC002397	Mouse chromosome 6 BAC-284H12	+2.9	
MRA-0560	CB848912	EST	+2.5	
M15E-0368	CB840173	EST	+2.4	
MRA-0358	CB850400	EST	+1.9	
M15E-1859	CB841340	EST	+1.8	
MRA-1285	CB849550	EST	+1.8	
MRA-1484	CB849740	EST	+1.6	
MRA-1009	NM_172385	RIKEN cDNA 9630010P11Rik	+1.5	
MRA-1592	AC079044	Mouse Chromosome 16 clone C57BL/6J	+1.5	
MRA-1032	CB849320	EST	+1.5	

QRT-PCR validation of differentially expressed genes

Due to variability inherent in microarray technology, qRT-PCR analysis was applied to validate expression data (Table 6.3 and 6.4). For 33 out of 35 genes examined (95%), qRT-PCR data were in good agreement with the microarray results with respect to the direction of observed changes (i.e. qualitative). For 21 genes, microarray data underestimated the expression change. Two genes, *Rxrg* and *Mtap6* (also called *Stop*), showed a discrepancy between the two methods. Microarray data indicated slight reduction whereas qRT-PCR demonstrated clear augmentation of their expression in the mutant retina. This discrepancy may be the result of differential expression of various isoforms (generated by alternative splicing) encoded by the two genes. While microarray method is hybridization-based and may reflect cumulative effects, qRT-PCR using gene-specific primers generates expression data for specific isoform(s).

Repression of Bmp/Smad signaling pathway in the *Nrl*^{-/-} retina

Microarray analysis showed that the expression of four genes belonging to the Bmp/Smad pathway (*Bmp2*, *Bmp4*, *Bmpr1a* and *Tob1*) was reduced in the *Nrl*^{-/-} retina. Since the expression and role of Bmp/Smad proteins have not been documented in mature retina, we decided to explore this further by qRT-PCR analysis of other pathway components that were not represented on I-gene microarray. All genes of the Smad-mediated Bmp signaling pathway exhibited lower expression in the cone-only *Nrl*^{-/-} retina (Figure 6.3A). The expression levels of *Bmp4*, *Bmp2*, and *Bmpr1a* were reduced by 12-, 4-, and 32-fold, respectively, in the *Nrl*^{-/-} retina, whereas R-Smads (*Smad 1* and *Smad 5*) and co-Smad (*Smad 4*) transcripts were marginally decreased. The expression of anti-Smads (*Smad 6* and *Smad 7*) and *Tob1* was also reduced in the *Nrl*^{-/-} retina.

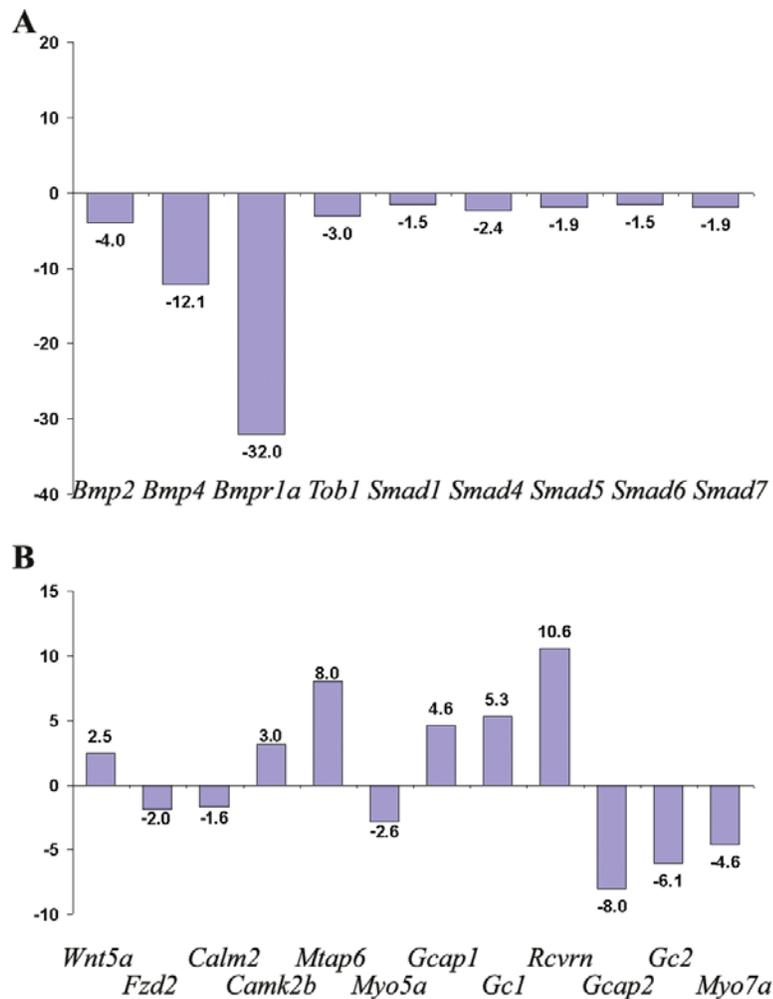


Figure 6.3. Expression analysis by qRT-PCR.

The number associated with each bar indicates fold-change in the *Nrl*^{-/-} retina relative to the wild type. (A) Reduced expression of genes associated with the Bmp/Smad pathway in the *Nrl*^{-/-} retina. (B) Altered expression of genes associated with Wnt/Ca²⁺ signaling and Ca²⁺ homeostasis in the *Nrl*^{-/-} retina. Expression changes in genes encoding Wnt5a ligand, Fzd2 receptor, and ten Ca²⁺ regulated proteins are shown.

Altered expression of the components of Wnt/Ca²⁺ signaling pathway in the Nrl^{-/-} retina

Altered expression of several Ca²⁺/Calmodulin-regulated genes, including Myo5a, Myo7a, and Mtap6, prompted us to examine the transcript levels of calmodulin (Calm2) and Camk2b by qRT-PCR. Both genes demonstrated altered expression in the Nrl^{-/-} retinal RNA as compared to the wild type (Figure 6.3B). We also detected higher expression of the Wnt5a ligand by qRT-PCR, indicating the overall increased activity and utilization of this pathway in the Nrl^{-/-} retina.

ChIP analysis

To investigate if the reduced expression of Bmp/Smad genes in the Nrl^{-/-} retina is a direct result of the lack of Nrl, we performed ChIP assays that permit *in vivo* analysis of the binding of a transcription factor to its cognate cis-regulatory sequence elements (Boyd and Farnham 1999; Wells and Farnham 2002). DNA fragments bound to Nrl were isolated from mouse retina using an anti-Nrl antibody (Swain et al. 2001). Although this antibody can cross-react with a P45 protein, this protein is present only in the cytoplasmic and not in the nuclear fraction, and shall not bind to chromatin (Swain et al. 2001). The ChIP DNA was examined for the enrichment of putative promoter regions of the three Bmp/Smad pathway genes (Bmp4, Smad4 and Bmpr1a) (Figure 6.4). PCR amplification of the NRE region in the rhodopsin promoter (Chen et al. 1997; Rehemtulla et al. 1996) was used as positive control for ChIP assay. We were able to amplify the genomic DNA fragments including the putative Nrl binding sites upstream of transcription start sites of Bmp4 and Smad4; however, no enrichment of the promoter region of Bmpr1a was observed.

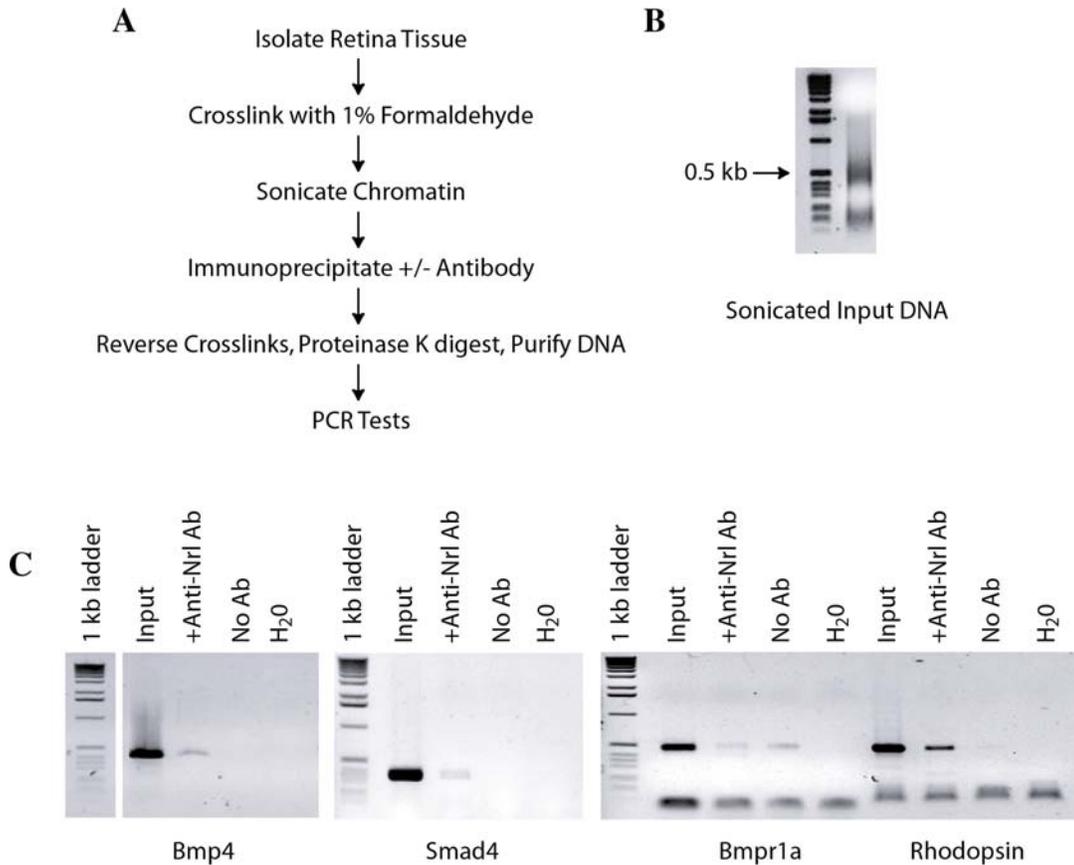


Figure 6.4. Chromatin immunoprecipitation (ChIP)

(A) A flowchart of the ChIP protocol. (B) An example of the sonicated input DNA used for ChIP PCR. (C) Enrichment of Bmp4 and Smad4, but not Bmpr1a, promoter region sequences in ChIP DNA obtained with anti-Nrl antibody. Amplification of rhodopsin promoter sequence in Nrl-ChIP DNA establishes the validity of the procedure. Sonicated input DNA and H₂O lanes serve as positive and negative controls, respectively.

6.5 Discussion

Our knowledge of signaling molecules and cell-cell communication pathways that specify physiological divergence between rods and cones and their homeostasis is currently limited. The availability of the *Nrl*^{-/-} mouse, which exhibits a complete lack of rods and significantly enhanced cone function, permits dissection of molecular differences between the two types of photoreceptors and associated pathways. We took advantage of the advances in microarray technology and utilized custom I-gene microarrays to compare the gene expression profiles of mature P21 retinas of the rod-dominant wild type mouse with that of the *Nrl*^{-/-} mouse. In addition to a large number of differentially expressed genes encoding known proteins, we identified 24 novel ESTs that might serve as candidate genes for human retinal diseases but would require additional molecular and functional analyses. The major findings from expression profiling presented in this report are that: (i) components of Bmp/Smad and Wnt/Ca²⁺ signaling pathways are expressed in the mature retina, suggesting their physiological relevance in regulating cell-cell communication and/or homeostasis; (ii) Bmp/Smad and Wnt/Ca²⁺ pathways are reduced in the cone-only *Nrl*^{-/-} retina, indicating their differential utilization in rod and cone associated inter- and intra-cellular signal transduction. (iii) At least two Bmp/Smad pathway genes (*Bmp4* and *Smad4*) may be direct targets of regulation by *Nrl* in mature rod photoreceptors.

Global gene expression profiling using microarrays is a powerful methodology to generate molecular signatures of specific cell-types, yielding valuable data for investigating complex biological processes. However, a careful experimental design is critical; for cDNA microarrays, issues related to direct versus indirect comparisons have been repeatedly emphasized (Churchill 2002; Yang and Speed 2002). Indirect comparison is more advantageous since it allows researchers to cross-compare studies performed under different conditions. Recent advances have greatly minimized

systematic sources of variations by improved technique (Yu et al. 2002), appropriate normalization (Quackenbush 2002; Tseng et al. 2001; Yang et al. 2002) and sufficient replication (Lee et al. 2000; Pan et al. 2002). In this study, we used I-gene microarrays and evaluated the performance of indirect method using a reference RNA target relative to the direct comparison of wild type and *Nrl*^{-/-} samples. Using the reference, we were able to obtain signals for >97% of spots on a single slide and about 91% of spots on all 12 slides studied. This high representation of retinal genes in the reference sample provides a basis for future indirect comparisons using I-gene microarrays and meta-analysis of the data. The percentage of positive spots on I-gene microarrays is substantially higher than that obtained with commercial microarray slides (52%) (Yoshida et al. 2002) or Affymetrix GeneChips (approximately 60%²) using the retinal tissue. Although the precise fold change was not identical, similar patterns of gene expression were revealed by direct and indirect comparisons. As one would expect, increasing the number of independent replicates decreased the number of false positives. Despite systematic variation introduced by the use of a reference sample, we found that the indirect method provided comparable results to direct comparison.

Gene profiling of whole retina from wild type and *Nrl*^{-/-} mice would yield expression differences that reflect the morphological and physiological changes associated with the loss of *Nrl* function. As predicted, the lack of rods and their replacement by cones resulted in higher expression of cone phototransduction genes with a concomitant reduction in rod gene transcripts. A majority of the differentially-expressed genes showed lower expression in the *Nrl*^{-/-} retina, reflecting the bias towards rod genes in the I-gene microarrays (containing cDNAs derived from rod-dominant wild type retinal tissue). Gene profiling suggested a remodeling of the *Nrl*^{-/-} retina, as evident by altered expression of many genes encoding structural proteins (e.g., *Prph2*, *Myo5a*, *Myo7a*, *Krt1-18*, *Cct4* and *Mtap6*) (see Table 6.3). This is consistent with the collapse of the sub-retinal space, shortening of photoreceptor outer segments and rewiring of neurons

in the outer plexiform layer (Strettoi, E., Mears, A.J. and Swaroop, A., unpublished data) (Mears et al. 2001). It should be noted that these expression changes probably reflect secondary rather than primary effects of the loss of Nrl.

Proteins of Wnt (wingless) and transforming growth factor- β (TGF β) families initiate a diverse array of signal transduction pathways that control cell-cell communication and play prominent roles during development (Derynck and Zhang 2003; Kuhl et al. 2000; Miyazawa et al. 2002; Shi and Massague 2003; Wrana 2000). Our studies demonstrated altered expression of components of Wnt/Ca²⁺ and Bmp-Smad pathways in the mature Nrl^{-/-} retina (Figure 6.5). In the Wnt/Ca²⁺ pathway, Wnt5a binds to its receptor Frizzled 2 (Fzd2), stimulating intracellular Ca²⁺ release and activating Ca²⁺/calmodulin-dependent kinase II beta (Camk2b) in a G-protein dependent manner (Kuhl et al. 2000). The gene profiling data suggested alterations in Ca²⁺ homeostasis, as indicated by increased expression of three (Gcap1, Gc1 and Rcvrn) and reduced expression of two (Gcap2 and Gc2) genes encoding Ca²⁺-regulated proteins. Furthermore, we observed altered expression of several Ca²⁺ dependent genes, such as Calm2, Camk2b, Mtap6, Myo5a and Myo7a. Since these genes are expressed in both rods and cones (Ebrey and Koutalos 2001; Yang and Garbers 1997), we hypothesize that the expression changes may reflect differences in Ca²⁺ homeostasis in rods versus cones (Krizaj and Copenhagen 2002; Rebrik and Korenbrot 1998). We suggest that the Wnt/Ca²⁺ pathway may modulate intracellular Ca²⁺ concentration and may have more direct control on cone function or homeostasis. A regulatory role of the Wnt/Ca²⁺ pathway in photoreceptor activity has been indicated in previous studies. Wnt5a and Fzd2 are known to signal through G α t, which is highly and specifically expressed in photoreceptors, and are altered in mouse retinal mutants (Liu et al. 1999; Yang-Snyder et al. 1996). In addition, two Fzd2 homologous genes, sFRP2 and Mfrp, known to inhibit the Wnt/Ca²⁺ signaling pathway, were shown to have altered transcript levels in retinitis pigmentosa patients and in the rd6 mouse (Jones et al. 2000a, b; Kameya et al. 2002).

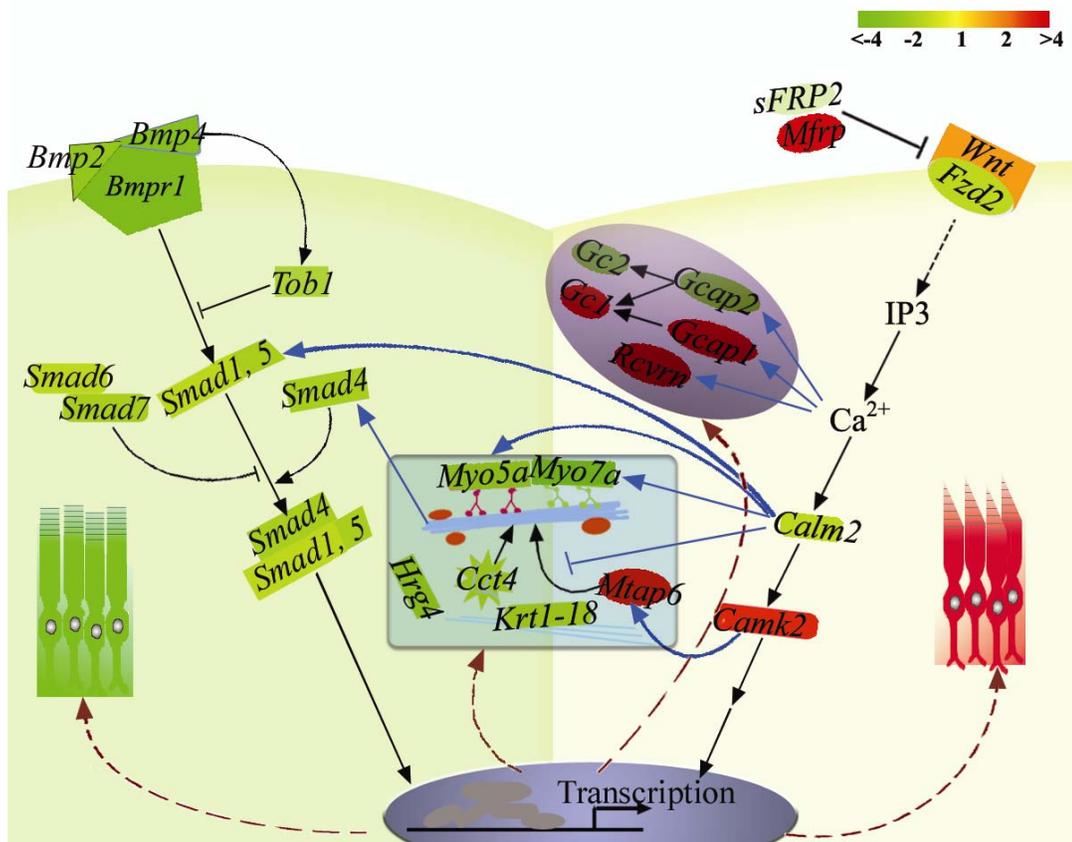


Figure 6.5. A schematic representation of signaling molecules and corresponding pathways altered in the *Nrl*^{-/-} mouse retina.

Expression changes were identified in several genes that could be classified as components of the Bmp/Smad pathway (Miyazawa et al. 2002) and/or the Wnt/Ca²⁺ signaling pathway (Kuhl et al. 2000). Proposed cross-talk (blue arrows) between these two signaling pathways and their regulation of cytoskeletal (light-blue rectangle box) or Ca²⁺-dependent proteins (purple ellipse) are indicated. Genes with increased expression in the *Nrl*^{-/-} are represented in red, whereas those with decreased expression are shown in green. The color bar indicates the fold change in expression.

Regulated expression of synaptic proteins, such as Myo5a, Myo7a and Hrg4 (Unc119h), is in good agreement with the known roles of the Wnt/Ca²⁺ signaling in the synthesis and release of neurotransmitters (Braun and Schulman 1995).

Bone morphogenetic proteins (Bmps) are members of TGF β superfamily of secreted proteins that primarily transduce signals through Smad proteins (Derynck and Zhang 2003; Miyazawa et al. 2002; Shi and Massague 2003; Wrana 2000). Bmp signaling is essential for lens development and is involved in retinal differentiation (Belecky-Adams et al. 1999; Faber et al. 2002; Furuta and Hogan 1998; Sakuta et al. 2001). Although a role for Bmps has been suggested in adult retinal pigment epithelium (Mathura et al. 2000), little is known about Bmp and Smad pathways in the mature retinal photoreceptors. We postulate that reduced expression of genes of Bmp/Smad pathway has relevance to the lack of rod function in the Nrl^{-/-} retina. ChIP data strongly suggests that expression of Bmp4 and Smad4 may be directly regulated by Nrl in the mature retina (see Figure 6.4). The absence of enrichment in Bmpr1a suggests that Nrl does not directly act on its promoter in the region tested. Modulation of Smad-mediated Bmp signaling pathway by Nrl (a transcription factor expressed in rod photoreceptors and not in other retinal neurons) argues in favor of its role in rod-mediated cell-cell communication. Activated Smad-complexes can positively or negatively regulate transcription of numerous target genes and modulate cellular response. It is tempting to speculate that in addition to directly controlling the expression of specific phototransduction genes in mature rods (Lerner et al. 2001; Mitton et al. 2000; Mitton et al. 2003; Rehemtulla et al. 1996), Nrl can influence rod function and maintenance by modulating and integrating discrete signaling events.

6.6 References

- Amano, T., Richelson, E., and Nirenberg, M. 1972. Neurotransmitter synthesis by neuroblastoma clones (neuroblast differentiation-cell culture-choline acetyltransferase-acetylcholinesterase-tyrosine hydroxylase-axons-dendrites). *Proc Natl Acad Sci U S A* **69**: 258-263.
- Belecky-Adams, T.L., Scheurer, D., and Adler, R. 1999. Activin family members in the developing chick retina: expression patterns, protein distribution, and in vitro effects. *Dev Biol* **210**: 107-123.
- Bessant, D.A., Payne, A.M., Mitton, K.P., Wang, Q.L., Swain, P.K., Plant, C., Bird, A.C., Zack, D.J., Swaroop, A., and Bhattacharya, S.S. 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nat Genet* **21**: 355-356.
- Blackshaw, S., Fraioli, R.E., Furukawa, T., and Cepko, C.L. 2001. Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell* **107**: 579-589.
- Boyd, K.E. and Farnham, P.J. 1999. Coexamination of site-specific transcription factor binding and promoter activity in living cells. *Mol Cell Biol* **19**: 8393-8399.
- Braun, A.P. and Schulman, H. 1995. The multifunctional calcium/calmodulin-dependent protein kinase: from form to function. *Annu Rev Physiol* **57**: 417-445.
- Chen, S., Wang, Q.L., Nie, Z., Sun, H., Lennon, G., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Zack, D.J. 1997. Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* **19**: 1017-1030.
- Chowers, I., Gunatilaka, T.L., Farkas, R.H., Qian, J., Hackam, A.S., Duh, E., Kageyama, M., Wang, C., Vora, A., Campochiaro, P.A., et al. 2003. Identification of novel genes preferentially expressed in the retina using a custom human retina cDNA microarray. *Invest Ophthalmol Vis Sci* **44**: 3732-3741.
- Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* **32 Suppl**: 490-495.
- Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* **74**: 829-836.
- Curcio, C.A., Owsley, C., and Jackson, G.R. 2000. Spare the rods, save the cones in aging and age-related maculopathy. *Invest Ophthalmol Vis Sci* **41**: 2015-2018.
- DeAngelis, M.M., Grimsby, J.L., Sandberg, M.A., Berson, E.L., and Dryja, T.P. 2002. Novel mutations in the NRL gene and associated clinical findings in patients with dominant retinitis pigmentosa. *Arch Ophthalmol* **120**: 369-375.

- DeRyckere, D. and DeGregori, J. 2002. Identification and characterization of transcription factor target genes using gene-targeted mice. *Methods* **26**: 57-75.
- Derynck, R. and Zhang, Y.E. 2003. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425**: 577-584.
- Dowling, J.E. 1987. *The retina: an approachable part of the brain*. Harvard University Press, Cambridge, MA.
- Ebrey, T. and Koutalos, Y. 2001. Vertebrate photoreceptors. *Prog Retin Eye Res* **20**: 49-94.
- Faber, S.C., Robinson, M.L., Makarenkova, H.P., and Lang, R.A. 2002. Bmp signaling is required for development of primary lens fiber cells. *Development* **129**: 3727-3737.
- Farjo, R., Yu, J., Othman, M.I., Yoshida, S., Sheth, S., Glaser, T., Baehr, W., and Swaroop, A. 2002. Mouse eye gene microarrays for investigating ocular development and disease. *Vision Res* **42**: 463-470.
- Furuta, Y. and Hogan, B.L. 1998. BMP4 is essential for lens induction in the mouse embryo. *Genes Dev* **12**: 3764-3775.
- Hendrickson, A. and Hicks, D. 2002. Distribution and density of medium- and short-wavelength selective cones in the domestic pig retina. *Exp Eye Res* **74**: 435-444.
- Hicks, D. and Molday, R.S. 1986. Differential immunogold-dextran labeling of bovine and frog rod and cone cells using monoclonal antibodies against bovine rhodopsin. *Exp Eye Res* **42**: 55-71.
- Hicks, D. and Sahel, J. 1999. The implications of rod-dependent cone survival for basic and clinical research. *Invest Ophthalmol Vis Sci* **40**: 3071-3074.
- Horton, J.D., Shah, N.A., Warrington, J.A., Anderson, N.N., Park, S.W., Brown, M.S., and Goldstein, J.L. 2003. Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proc Natl Acad Sci U S A* **100**: 12027-12032.
- Jones, S.E., Jomary, C., Grist, J., Stewart, H.J., and Neal, M.J. 2000a. Altered expression of secreted frizzled-related protein-2 in retinitis pigmentosa retinas. *Invest Ophthalmol Vis Sci* **41**: 1297-1301.
- Jones, S.E., Jomary, C., Grist, J., Stewart, H.J., and Neal, M.J. 2000b. Modulated expression of secreted frizzled-related proteins in human retinal degeneration. *Neuroreport* **11**: 3963-3967.

- Kameya, S., Hawes, N.L., Chang, B., Heckenlively, J.R., Naggert, J.K., and Nishina, P.M. 2002. Mfrp, a gene encoding a frizzled related protein, is mutated in the mouse retinal degeneration 6. *Hum Mol Genet* **11**: 1879-1886.
- Krizaj, D. and Copenhagen, D.R. 2002. Calcium regulation in photoreceptors. *Front Biosci* **7**: d2023-2044.
- Kuhl, M., Sheldahl, L.C., Park, M., Miller, J.R., and Moon, R.T. 2000. The Wnt/Ca²⁺ pathway: a new vertebrate Wnt signaling pathway takes shape. *Trends Genet* **16**: 279-283.
- Lee, M.L., Kuo, F.C., Whitmore, G.A., and Sklar, J. 2000. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A* **97**: 9834-9839.
- Lerner, L.E., Gribanova, Y.E., Ji, M., Knox, B.E., and Farber, D.B. 2001. Nrl and Sp nuclear proteins mediate transcription of rod-specific cGMP- phosphodiesterase beta-subunit gene: involvement of multiple response elements. *J Biol Chem* **276**: 34999-35007.
- Leskov, I.B., Klenchin, V.A., Handy, J.W., Whitlock, G.G., Govardovskii, V.I., Bownds, M.D., Lamb, T.D., Pugh, E.N., Jr., and Arshavsky, V.Y. 2000. The gain of rod phototransduction: reconciliation of biochemical and electrophysiological measurements. *Neuron* **27**: 525-537.
- Liu, X., Liu, T., Slusarski, D.C., Yang-Snyder, J., Malbon, C.C., Moon, R.T., and Wang, H. 1999. Activation of a frizzled-2/beta-adrenergic receptor chimera promotes Wnt signaling and differentiation of mouse F9 teratocarcinoma cells via Galphao and Galphat. *Proc Natl Acad Sci U S A* **96**: 14383-14388.
- Livesey, F.J., Furukawa, T., Steffen, M.A., Church, G.M., and Cepko, C.L. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. *Curr Biol* **10**: 301-310.
- Lonnstedt, I. and Speed, T. 2002. Replicated microarray data. *Statistica Sinica* **12**: 31-46.
- Marc, R.E. and Jones, B.W. 2003. Retinal remodeling in inherited photoreceptor degenerations. *Mol Neurobiol* **28**: 139-147.
- Masland, R.H. 2001. Neuronal diversity in the retina. *Curr Opin Neurobiol* **11**: 431-436.
- Mathura, J.R., Jr., Jafari, N., Chang, J.T., Hackett, S.F., Wahlin, K.J., Della, N.G., Okamoto, N., Zack, D.J., and Campochiaro, P.A. 2000. Bone morphogenetic proteins-2 and -4: negative growth regulators in adult retinal pigmented epithelium. *Invest Ophthalmol Vis Sci* **41**: 592-600.

- McBurney, M.W., Jones-Villeneuve, E.M., Edwards, M.K., and Anderson, P.J. 1982. Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. *Nature* **299**: 165-167.
- Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., and Swaroop, A. 2001. Nrl is required for rod photoreceptor development. *Nat Genet* **29**: 447-452.
- Mitton, K.P., Swain, P.K., Chen, S., Xu, S., Zack, D.J., and Swaroop, A. 2000. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. *J Biol Chem* **275**: 29794-29799.
- Mitton, K.P., Swain, P.K., Khanna, H., Dowd, M., Apel, I.J., and Swaroop, A. 2003. Interaction of retinal bZIP transcription factor NRL with Flt3-interacting zinc-finger protein Fiz1: possible role of Fiz1 as a transcriptional repressor. *Hum Mol Genet* **12**: 365-373.
- Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T., and Miyazono, K. 2002. Two major Smad pathways in TGF-beta superfamily signalling. *Genes Cells* **7**: 1191-1204.
- Molday, R.S. 1998. Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases. The Friedenwald Lecture. *Invest Ophthalmol Vis Sci* **39**: 2491-2513.
- Mu, X., Beremand, P.D., Zhao, S., Pershad, R., Sun, H., Scarpa, A., Liang, S., Thomas, T.L., and Klein, W.H. 2004. Discrete gene sets depend on POU domain transcription factor Brn3b/Brn-3.2/POU4f2 for their expression in the mouse embryonic retina. *Development* **In press**.
- Nathans, J. 1999. The evolution and physiology of human color vision: insights from molecular genetic studies of visual pigments. *Neuron* **24**: 299-312.
- Pan, W., Lin, J., and Le, C.T. 2002. How many replicates of arrays are required to detect gene expression changes in microarray experiments? A mixture model approach. *Genome Biol* **3**.
- Quackenbush, J. 2002. Microarray data normalization and transformation. *Nat Genet* **32 Suppl**: 496-501.
- Rebrik, T.I. and Korenbrot, J.I. 1998. In intact cone photoreceptors, a Ca²⁺-dependent, diffusible factor modulates the cGMP-gated ion channels differently than in rods. *J Gen Physiol* **112**: 537-548.
- Rehemtulla, A., Warwar, R., Kumar, R., Ji, X., Zack, D.J., and Swaroop, A. 1996. The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proc Natl Acad Sci U S A* **93**: 191-195.

- Sakuta, H., Suzuki, R., Takahashi, H., Kato, A., Shintani, T., Iemura, S., Yamamoto, T.S., Ueno, N., and Noda, M. 2001. Ventroptin: a BMP-4 antagonist expressed in a double-gradient pattern in the retina. *Science* **293**: 111-115.
- Shi, Y. and Massague, J. 2003. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**: 685-700.
- Swain, P.K., Hicks, D., Mears, A.J., Apel, I.J., Smith, J.E., John, S.K., Hendrickson, A., Milam, A.H., and Swaroop, A. 2001. Multiple phosphorylated isoforms of NRL are expressed in rod photoreceptors. *J Biol Chem* **276**: 36824-36830.
- Swaroop, A., Xu, J.Z., Pawar, H., Jackson, A., Skolnick, C., and Agarwal, N. 1992. A conserved retina-specific gene encodes a basic motif/leucine zipper domain. *Proc Natl Acad Sci U S A* **89**: 266-270.
- Swaroop, A. and Zack, D.J. 2002. Transcriptome analysis of the retina. *Genome Biol* **3**: REVIEWS1022.
- Tseng, G.C., Oh, M.K., Rohlin, L., Liao, J.C., and Wong, W.H. 2001. Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res* **29**: 2549-2557.
- Wells, J. and Farnham, P.J. 2002. Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods* **26**: 48-56.
- Wrana, J.L. 2000. Regulation of Smad activity. *Cell* **100**: 189-192.
- Yang, R.B. and Garbers, D.L. 1997. Two eye guanylyl cyclases are expressed in the same photoreceptor cells and form homomers in preference to heteromers. *J Biol Chem* **272**: 13738-13742.
- Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**: e15.
- Yang, Y.H. and Speed, T. 2002. Design issues for cDNA microarray experiments. *Nat Rev Genet* **3**: 579-588.
- Yang-Snyder, J., Miller, J.R., Brown, J.D., Lai, C.J., and Moon, R.T. 1996. A frizzled homolog functions in a vertebrate Wnt signaling pathway. *Curr Biol* **6**: 1302-1306.
- Yoshida, S., Yashar, B.M., Hiriyanna, S., and Swaroop, A. 2002. Microarray analysis of gene expression in the aging human retina. *Invest Ophthalmol Vis Sci* **43**: 2554-2560.

- Yu, J., Farjo, R., MacNee, S.P., Baehr, W., Stambolian, D.E., and Swaroop, A. 2003. Annotation and analysis of 10,000 expressed sequence tags from developing mouse eye and adult retina. *Genome Biol* **4**: R65.
- Yu, J., Othman, M.I., Farjo, R., Zarepari, S., MacNee, S.P., Yoshida, S., and Swaroop, A. 2002. Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays. *Mol Vis* **8**: 130-137.
- Zhu, X., Brown, B., Li, A., Mears, A.J., Swaroop, A., and Craft, C.M. 2003. GRK1-dependent phosphorylation of S and M opsins and their binding to cone arrestin during cone phototransduction in the mouse retina. *J Neurosci* **23**: 6152-6160.

CHAPTER 7
GENE EXPRESSION PROFILES OF RETINA DURING POSTNATAL
DEVELOPMENT AND COMPARISON TO THE $Nrl^{-/-}$ RETINA

7.1 Summary

To delineate molecular distinctions between rod and cone differentiation, I generated and compared gene expression profiles of retinas from wild-type and $Nrl^{-/-}$ mice at 5 developmental ages (corresponding to different stages of photoreceptor differentiation) using I-gene cDNA microarrays. Our analysis of the microarray data identified altered expression of 923 genes, either during development or between age-matched wild-type and $Nrl^{-/-}$ retinas. These include 304 genes that are expressed highly during early stages of development, many of which are involved in protein biosynthesis and cell structure. Additional 223 genes, represented by metabolic or catalytic enzymes and crystallins, are expressed at high levels during later stages. The $Nrl^{-/-}$ retina exhibited prolonged expression of genes involved in protein biosynthesis, whereas transcript levels of enzymes and crystallins remained unaltered. A group of 135 genes that included those encoding rod phototransduction proteins, demonstrated a gradual increase in expression during normal development and were significantly repressed in the mutant retina; this set also included molecules of Bmp/Smad pathway. Chromatin immunoprecipitation studies in the laboratory have shown that many genes in this group are direct targets of transcriptional regulation by *Nrl*. This study has also revealed a number of genes and

unknown ESTs that may play significant roles in rod- and/or cone-mediated visual function.

7.2 Introduction

Rods and cones are two major types of photoreceptors responsible for vision (Rodieck 1998). They have distinct visual functions: rods have high visual sensitivity and mediate dim light vision, whereas cones are responsive to bright light and provide color vision (Dowling 1987). Several studies have attempted to understand the molecular basis underlying these differences. A discrete set of proteins exists for rod and cone phototransduction - a chemical cascade that converts light stimuli to electrical impulses in photoreceptor cells (Dhallan et al. 1992; Molday 1998). Previous reports have indicated an important role of Ca^{2+} in phototransduction and articulated that the quantitative difference in Ca^{2+} homeostasis between rods and cones accounts for some differences in their function (Korenbrod and Rebrik 2002; Pepe 2001). By comparative transcript analysis of mature rod- and cone-dominated retinas, our study recently suggested biased utilization of Bmp/smad and Wnt/ Ca^{2+} signaling in rod- and cone-mediated visual functions (Yu, J. et al., manuscript submitted; Chapter 6). However, it is yet to be illustrated, how and at what stage of photoreceptor maturation, these pathways become dissimilar between rods and cones.

To delineate the dynamic program that distinguishes these two kinds of neurons at molecular level, large-scale transcript profiling of developing retinas was undertaken. We also generated temporal transcript profiles of the newly generated $\text{Nrl}^{-/-}$ mouse, which contains only cones in the retina (Mears et al. 2001). *Nrl* is a basic motif leucine zipper transcription factor that is preferentially expressed in the rod photoreceptors (Swain et al.

2001). Missense mutations of its human ortholog *NRL* cause retinal disease with photoreceptor degeneration (Bessant et al. 2003; Bessant et al. 1999; DeAngelis et al. 2002). *Nrl* regulates the expression of several rod genes either by itself or synergistically with other transcriptional factors (Chen et al. 1997; Mitton et al. 2000; Mitton et al. 2003; Rehemtulla et al. 1996). Studies of the *Nrl*^{-/-} retina demonstrated a morphological and functional switch of rods to cones (Mears et al. 2001). Together with the rod-dominated wild-type mouse, this mutant provides a unique opportunity for comparative studies of rods and cones.

Microarray technology has been successfully used for obtaining global knowledge of transcriptomes of many tissues under various conditions (Furlong et al. 2001; Sperger et al. 2003). Comparative gene expression analysis of wild-type and knockout mouse retinas, in particular, has led to the discovery of downstream binding targets of the deleted gene and the identification of novel retinal disease genes (Kennan et al. 2002; Livesey et al. 2000; Mu et al. 2004). Analysis of retinal gene expression patterns during embryonic development has also been lately attempted to identify differential expression along the dorsal-ventral axis (Diaz et al. 2003). With an effort to focus on eye-expressed genes, several groups have recently generated customized microarray slides containing only eye-expressed genes (Chowers et al. 2003; Farjo et al. 2002; Mu et al. 2004; Swaroop and Zack 2002).

In this study, I-gene microarrays were exploited for the transcript analysis of wild-type and *Nrl*^{-/-} retinas during 5 developmental stages: postnatal day (P) 0, 2, 6, 10 and 21. In the wild-type, the developing and mature retinas were found to be characterized, respectively, by transcripts involved in protein biosynthesis and cell

structure and those for energy metabolism and phototransduction. A group of co-expressed genes, including those involved in rod phototransduction and Bmp signaling, were greatly repressed in the *Nrl*^{-/-} mouse. This study identified a number of candidate *Nrl* target genes and indicated an important role of Bmp signaling in rod maturation and functional maintenance.

7.3 Materials and Methods

Microarray hybridization

Mouse retinas were dissected at postnatal day 0, 2, 6, 10 and 21 for RNA isolation using TRIzol reagent (Invitrogen, Carlsbad, CA) and RNeasy kit (Qiagen, Valencia, CA). All procedures involving mice were performed in accordance with the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan. Mouse I-gene cDNA microarray slides were generated as previously described (Farjo et al. 2002). The set of slides used in the present study contained, in duplicates, over 6,000 genes / ESTs (Yu et al. 2003). Total RNA was labeled using 3DNA Submicro Expression Array Detection Kit (Genisphere, Hatfield, PA) and hybridized to the slides as previously described (Yu et al. 2002).

Microarray data analysis

Scanned microarray images were analyzed using AnalyzerDG (MolecularWare Inc, Cambridge, MA) in a batch mode. Intensity values for each channel were separately calculated by subtracting background intensities from the spot intensities. Analyzed data was normalized within-slides by “printploess” methods, followed by between-slide scaling (Yang et al. 2002) using a statistical package, limma, in R (<http://www.r-project.org/>). This normalization generated log-transformed dataset with equal variance

between slides. Previous analysis has shown that these data generally follow a normal distribution (Yu et al. 2004).

To identify genes differentially expressed during the development of the wild-type retina, a one-way ANOVA (analysis of variance) model was used for each gene as

$$M_{ik} = \mu + A_i + \varepsilon_{ik}, \quad (1)$$

where μ is the average intensity over conditions tested (e.g., the 23 wild-type experiments), A_i represents the effect of the i^{th} level of factor A ($i = 1, \dots, d$ with d : the number of levels – 5 time-points here), k is the number of replicates in each time-point and ε_{ik} represents random error. Here, the age of retina was considered as a categorical variable with 5 levels.

Similarly, to identify genes with altered expression between age-matched wild-type and $Nrl^{-/-}$ mouse retinas, the following one-way ANOVA model was used for each gene at each time-point:

$$M_{jk} = \mu + B_j + \varepsilon_{jk}, \quad (2)$$

where μ is the average intensity over conditions tested (e.g., 5 wild-type and 5 $Nrl^{-/-}$ experiments of P0 retinas), B_j represents the effect of j^{th} level of factor B ($j = 1, \dots, d$ with d : the number of levels – 2 here), and k is the number of replicates. Here, the strain of mouse was considered as a categorical variable with 2 levels.

Gene X is selected for further analysis if it has average intensity over 128 and is significantly altered ($P < 0.01$) and has at least 2-fold difference between its highest and lowest expression levels. Spots corresponding to the same genes were analyzed independently, without taking the average, due to the presence of some low-quality areas on the microarrays used.

Retinal samples and selected genes were clustered by a combination of hierarchical clustering using Spotfire (<http://www.spotfire.com>) and unsupervised learning of mixture models in Matlab (Yu, J. and Hero, A., manuscript in preparation) (Figueiredo and Jain 2002). The data were visualized by Spotfire and TreeView (Eisen et

al. 1998). Hierarchical clustering in Spotfire was used for an exploratory analysis of the obvious features of the dataset. The unsupervised clustering algorithm has the advantage of automatically detecting the appropriate number of clusters in the dataset by starting with an initial number of clusters (k_{max}), for example 35, which is much larger than the optimal number of clusters in the dataset. The robustness of this method regarding initialization was tested by using different values of k_{max} . The possibility of obtaining a specific number of clusters is calculated as the fraction of runs that gives this number of clusters.

Quantitative real-time RT-PCR

To confirm the differences in gene expression obtained using cDNA microarrays, qRT-PCR were performed as described (Yu et al. 2003). For each gene, qRT-PCR reactions for cDNAs prepared from wild-type and $Nrl^{-/-}$ retina at all 5 time-points were performed simultaneously. Each sample was conducted in triplicates in the iCycler IQ real-time PCR Detection system (Bio-Rad, Hercules, CA). RT-PCR of each gene was repeated for 3 times using RNAs isolated from different mice. The average crossing-cycle differences between wild-type and age-matched $Nrl^{-/-}$ were normalized to the relative amounts of Hprt in the same cDNA preparations. To normalize the samples of different development times, either Hprt or a data-driven normalization was used. The data-driven normalization calculated scaling factors such that the expression patterns for 2 selected genes (*Opn1sw* and *Pdc*) were most consistent between qRT-PCR and microarray methods.

Chromatin Immunoprecipitation (ChIP)

ChIP was performed using the assay kit from Upstate Biotechnologies (Charlottesville, VA), as previously described (Yu, J. et al. manuscript submitted;

Chapter 6). Briefly, retinas were obtained from wild-type mice and snap frozen on dry ice. Four retinas were crosslinked for 15 min at 37°C, washed four times in ice-cold PBS with proteinase inhibitors and then incubated on ice for 15 minutes. The tissue was then sonicated on ice eight times in 20-second pulses. The remaining steps were essentially performed as described by the manufacturer using anti-NRL polyclonal antibody (Swain et al. 2001).

Putative promoter regions were determined *in silico* (<http://www.ncbi.nlm.nih.gov/mapview>). Each DNA sequence was analyzed using MatInspector (<http://www.genomatix.de/index.html>) and PCR primers were designed to flank putative AP1-like sites either predicted manually or by MatInspector. If there was more than one AP-1 like site, the element located closest to the 5' untranslated sequence was used. ChIP DNA from two independent chromatin immunoprecipitation experiments was obtained. At least one PCR was performed with each of the two ChIP DNA sets for any given primer. Equal volumes of input, with antibody and no antibody DNA was used in each PCR reaction.

7.4 Results

The *Nrl*^{-/-} mice do not develop rod photoreceptor cells, and have an increased number of cone photoreceptors (Mears et al. 2001). To examine mechanisms underlying rod and cone photoreceptor differentiation, we compared transcript profiles of *Nrl*^{-/-} retina with that of stage-matched wild-type retina. Five stages of retinal development were analyzed: postnatal day 0 (P0), P2, P6, P10 and P21 (Figure 7.1A). The formation of rod precursor cells peaks at P2 in mice (Young 1985). Rhodopsin expression, however, is not detectable until later; the post-mitotic cells become rhodopsin-positive from P4 to P13 (Morrow et al. 1998). P0 and P2 are, therefore, time-points before the onset of rhodopsin expression; genes (such as *Nrl*) regulating the expression of phototransduction

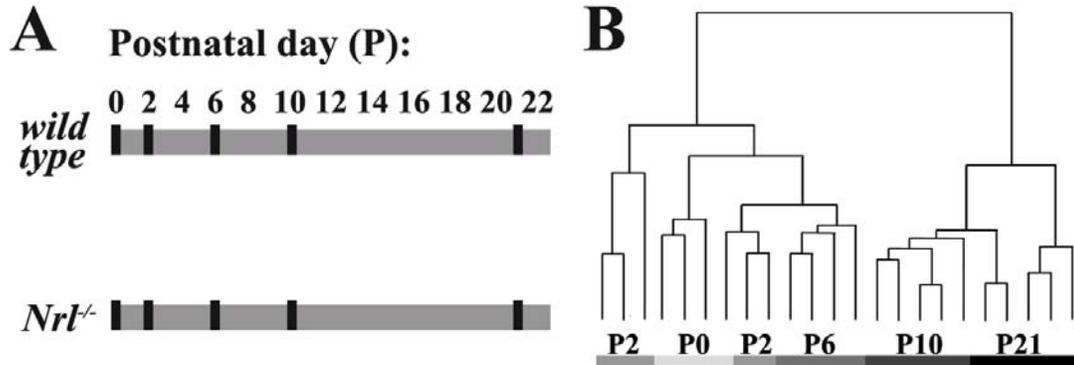


Figure 7.1. Gene expression profiling of the developing wild-type and *Nrl*^{-/-} retinas. (A) Microarray analyses were performed using retinas from wild-type and *Nrl*^{-/-} mice of matching ages: P0, P2, P6, P10 and P21. Black bars indicate the developmental stages that were sampled. (B) Hierarchical clustering of wild-type retinal samples collected at 5 developmental ages. A total of 23 microarray hybridization, including 4 replicates for each of P0 and P6 retinas and 5 replicates for each of P2, P10 and P21 retinas, were analyzed; replicated experiments of the same stage were indicated by the same color bar. This clustering was based on expression levels of the 923 selected genes. Retinas of overall similarity in the expression of these genes, such as retinas of the same age and those of neighboring time-points, are grouped together. Retinas of the 3 early developmental time-points (P0, P2 and P6) group together, whereas those of the 2 late time-points (P10 and P21) form a separate cluster.

genes are active at these stages. The middle stage, P6, is a transition point from a developing retina (P0-P6) to a mature functional retina (P10-P21). During P10 to P21, photoreceptor outer segments are actively formed and phototransduction proteins become abundant; the retina is mature by P21.

For each developmental stage, five independent replicates of microarray experiments were performed using retinas dissected from different mice; a total of 50 microarray slides were hybridized. Data from 48 of these 50 slides were used for further analysis; 2 slides with scratches were discarded. All hybridizations were performed by comparing a retinal RNA target with a common reference target, which is a mixture of RNAs isolated from retinas, eyes and brains at several developmental ages, and from 2 neuron-related cell lines (Yu et al. 2002). I-gene microarray slides, containing 13,440 spots representing, in duplicates, over 6,000 eye-expressed genes, which is a subset of the 10,000 cDNAs previously reported, were used for this study (Yu et al. 2003; Yu et al. 2002).

Two groups of genes were of particular interest: genes (807) regulated during retinal development; and those (334) altered between stage-matched wild-type and knockout retinas, including 74 genes from P0, 11 from P2, 41 from P6, 84 from P10, and 124 from P21. After eliminating redundant genes that are differentially expressed both during development and between wild-type and $Nrl^{-/-}$, we identified a total of 923 genes for further analysis. We first examined similarity in gene profiles of 23 wild-type retinal samples on the basis of the expression levels of these 923 genes. The results demonstrated high degree of reproducibility among samples at the same developmental stage (Figure 7.1B). We also obtained two well-defined clusters; one included all

replicates of P10 and P21 retinal samples, and the other all of the P0, P2 and P6 retinas. This data suggested that, during retinal development, distinct gene expression profiles characterize developing (P0 to P6) and maturing retinas (P10 to P21). This is further confirmed by hierarchical clustering of the expression profiles of the 923 genes at all stages of both wild-type and knockout retinas (Figure 7.2A). Two major groups of genes were identified: (i) genes highly expressed in P10 and P21 retinas (the “retina-late” genes); and (ii) genes highly expressed in P0, P2, and P6 retinas (the “retina-early” genes).

To determine the number of naturally occurring clusters in these 923 genes, unsupervised cluster algorithm was applied by specifying k_{max} (maximum number of clusters) ranging from 10 to 50. The possibilities of obtaining 5, 6 and 7 clusters are found to be 15%, 41% and 34%, respectively. A majority of genes were consistently grouped together despite varying k_{max} values. Using k_{max} equal to 35, this algorithm divided the 923 genes into 6 clusters (Figure 7.2B). Clusters 1, 3 and 4 contained primarily retina-early genes, whereas clusters 2, 5 and 6 were for retina-late genes. In particular, cluster 4 corresponds to genes with gradually reducing transcript levels during the development of the wild-type retina, while in the knockout, this reduction is delayed until after P6. This lag could also be observed, less obviously, in cluster 3. Both of these clusters were dominated by genes associated with protein biosynthesis and cell structure. One of these is the protein synthesis elongation factor Tu (eEF-Tu), which had 15 copies, represented by different clones on the microarray; all copies of eEF-Tu fell into cluster 4. By contrast, genes in cluster 5 had increased expression levels during the development of both the wild-type and the knockout retinas. Two major functional groups of genes in

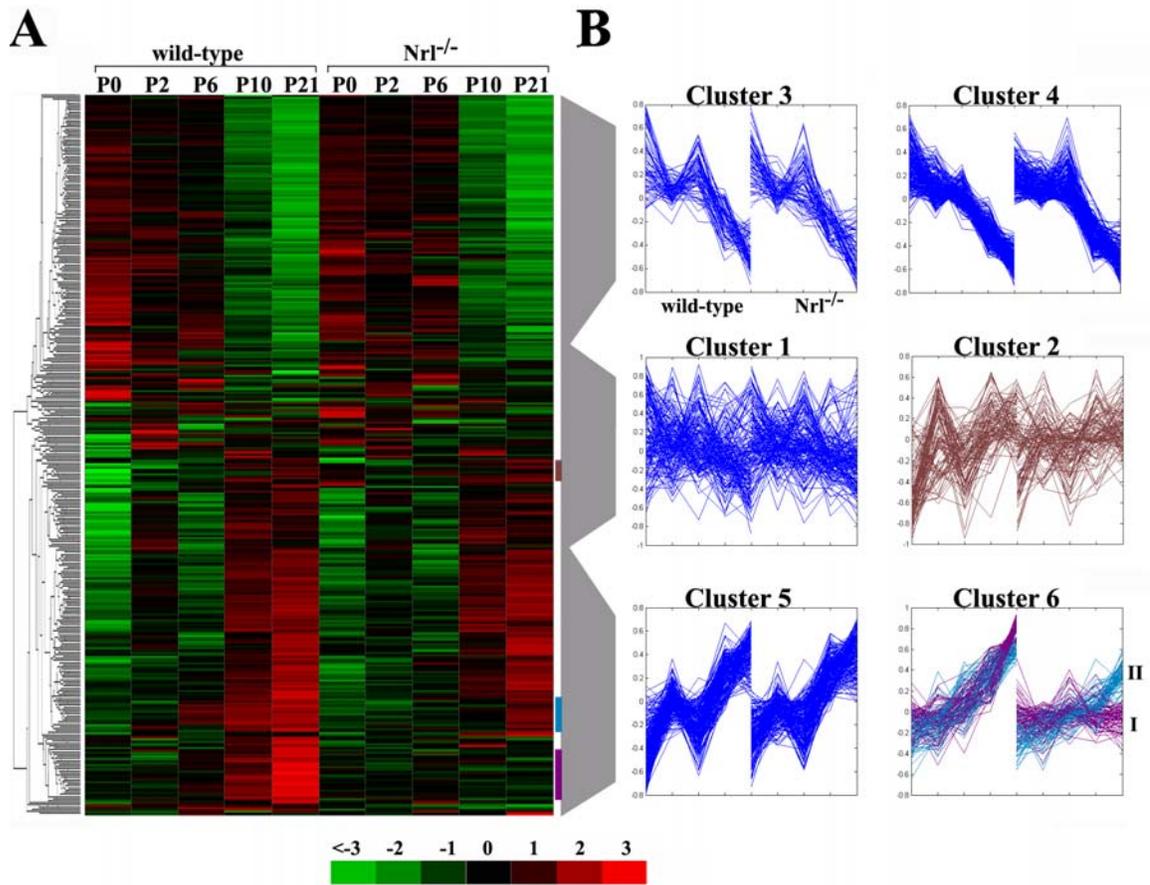


Figure 7.2. Time course of genes with altered expression during the development of the wild-type or *Nrl*^{-/-} retinas.

(A) Time course analysis of expression profiles of 923 genes identified two obvious clusters: one increases and the other decreases expression during development; expression levels are indicated by the horizontal colored scale bar. The colored vertical bars indicate genes belonging to cluster 2 and 6 with curves of matching colors. (B) Six groups of co-expressed genes, corresponding to (shown in gray) various areas in A, were identified. In each time-trajectory plot, expression of genes is ordered as in (A): 5 developmental stages of wild-type and those of *Nrl*^{-/-} and Y-axis represents relative intensity of each gene in 10 samples.

this cluster were metabolic or catalytic enzymes and crystallins. Another interesting group of retina-late genes was cluster 6, which contained a number of phototransduction genes that were expressed at dramatically lower levels (or not expressed) in the knockout retina.

The 135 cluster 6 retina-late genes were further divided into 2 groups, I and II, by K-means clustering with $K = 2$, since two obvious clusters could be observed (Figure 7.2B cluster 6, indicated by curves of different colors). The 69 group-I retina-late genes, exemplified by rhodopsin (Rho), were not expressed in all stages of the knockout retina (Figure 7.3A). They represented mature retinal genes that were not expressed in the rod-less *Nrl*^{-/-} retina. Group II contained 65 retina-late genes, which demonstrated delayed expression during development in the mutant retina, compared with that of the wild-type (Figure 7.3B). Many genes in group-II, for instance S-antigen and Phosducin (Pdc, also called rod photoreceptor 1), are also involved in phototransduction.

To determine whether the silencing of cluster 6 genes is a direct effect of the deletion of *Nrl*, we performed ChIP experiments using anti-*Nrl* antibody and evaluated the enrichment of promoter regions of several of these genes. Of the 12 genes examined, 10 were selected from group I, including Rho, *Nr2e3*, *Gnat1*, *Rxrg*, *Bmp4*, *Mtap6*, *Prdx4*, *Ddx5* and two unknown Riken cDNAs (2510025F08Rik and 1110020M21Rik), whereas 2 genes, *Bmpr1a* and *Pdc*, were from group II of cluster 6. Of the 10 group-I genes, all except *Prdx4*, are found to be positively enriched in *Nrl*-ChIP (Friedman, J. and Swaroop, A., unpublished data). ChIP analyses of *Rxrg*, *Ddx5*, *Prdx4*, 2510025F08Rik and 1110020M21Rik were shown with *Opn1mw* as negative control since its promoter does

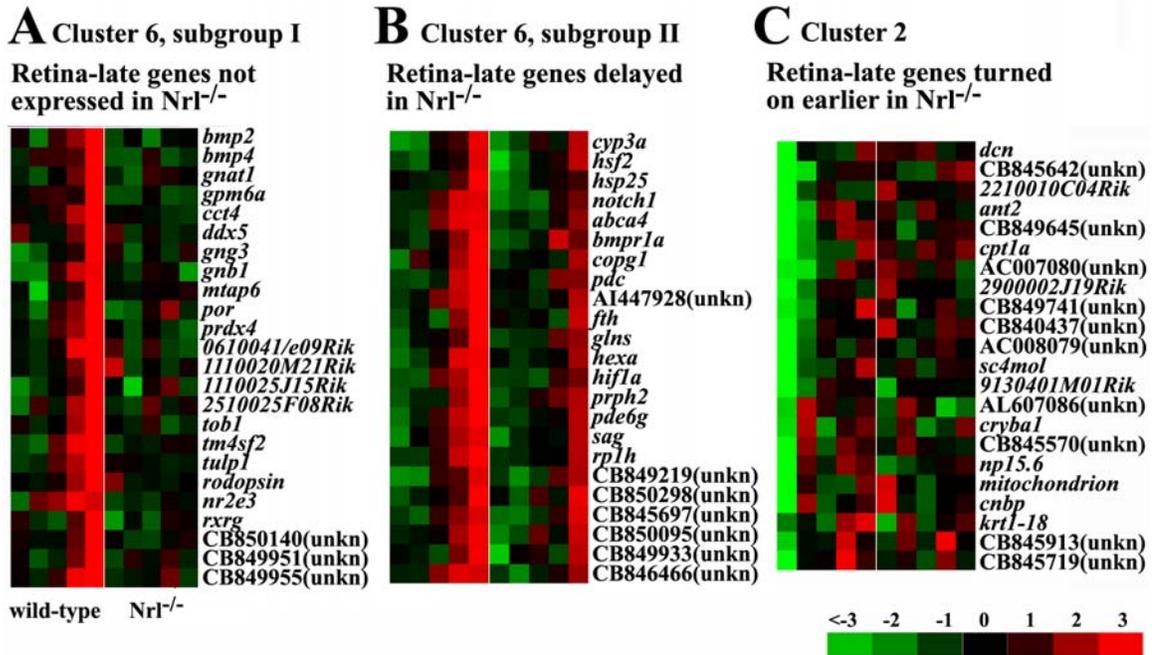


Figure 7.3. Time course of genes differentially expressed in the wild-type and *Nrl*^{-/-} retinas.

Cluster diagrams are shown, in the order of P0, P2, P6, P10 and P21 of wild-type and those of *Nrl*^{-/-} (x-axis), for genes corresponding to 3 selected clusters from Figure 7.2. (A) Retina-late genes that increase expression in the time course of wild-type retinal development, but remain lowly expressed in all stages of *Nrl*^{-/-} retinal development; (B) Retina-late genes that are induced as retina matures; this induction is delayed in the *Nrl*^{-/-} retina. (C) Retina-late genes that are repressed at early stages of wild-type retina but not at that of *Nrl*^{-/-}.

not bind Nr1 (Figure 7.4). For the 2 group-II genes, one (Pdc) was positive and the other (Bmpr1a) was negative.

A subset (22) of the cluster 2 retina-late genes was expressed early in the mutant retina. These genes were typically turned on at P2, with transcript levels maintained at later stages of the wild-type retinas. In the knockout, their expression at P0 was already at comparable levels as later stages, suggesting an earlier induction (Figure 7.3C). One member of this set was decorin (Dcn), a preproprotein capable of disrupting Bmp/Smad-dependent transcription (Abdel-Wahab et al. 2002).

To validate the temporal expression profiles detected by cDNA microarrays, qRT-PCR analyses were conducted for 12 genes. These include *Opn1sw*, *Nr2e3*, *Bmp2*, *Bmp4*, *Bmpr1a*, *Tob1*, *Camk2b*, *Rxrg*, *Cct4*, *Myo7a*, *Pdc* and *Calm2*. Two kinds of normalization, using either *Hprt* or data-driven scale factors, were performed to account for the differences in starting amount of cDNAs. Interestingly, *Hprt* normalized data showed enormous discrepancy between microarray and qRT-PCR results (Figure 7.5A). Data-driven scale factors were calculated such that the results generated by qRT-PCR were most consistent with microarray data for 2 genes: *Opn1sw* and *Pdc*, which are known phototransduction proteins. Using this normalization, the microarray and qRT-PCR results were found highly consistent. All genes, except *Rxrg*, have positive correlation coefficient with median of 0.71 (Figure 7.5B). Differential expression of over 30 genes between wild-type and mutant retinas at a single stage (P21) has been previously validated (Yu, J and Swaroop, A, manuscript submitted; Chapter 6).

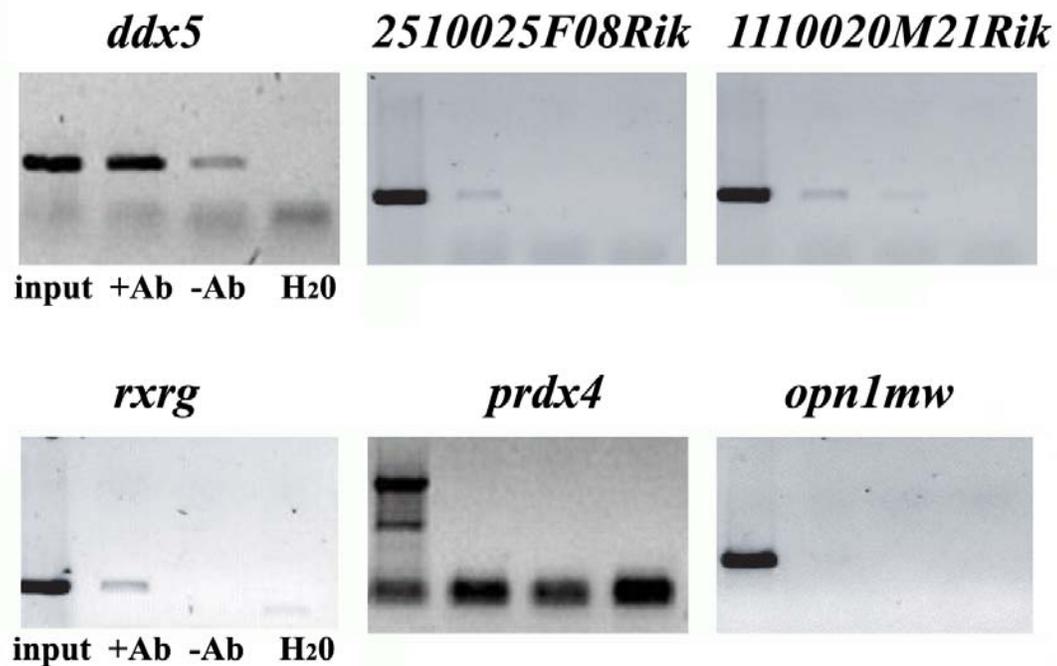


Figure 7.4. ChIP analyses of selected genes.

Chromatin immunoprecipitation (ChIP) assay was done using anti-Nrl antibody. PCR was performed using primers designed to amplify upstream promoter regions of selected genes. Parallel PCRs were conducted to amplify the extracted DNA (input), DNA samples precipitated with anti-Nrl (+Ab) or without antibody (-Ab) and H₂O. All input samples showed at least one band, indicating the existence and successful amplification of the corresponding genes in the input. An enrichment of the band intensity in +Ab sample, compared to -Ab sample, indicates that the gene is bound by Nrl. *Opn1mw* was selected as a negative control.

7.5 Discussion

Previous studies have illustrated the pivotal role of Nrl in rod photoreceptor development and function using transgenic mouse (Mears et al. 2001). Analysis of gene expression of adult Nrl^{-/-} retina, compared to that of wild-type, have identified several potential target genes of Nrl and have suggested the biased utilization of Bmp/Smad signaling pathways in rod- and cone-mediated visual function (Yu, J. et al. manuscript submitted; Chapter 6). However, the molecular and signaling mechanisms underlying photoreceptor differentiation and functional maintenance remain to be delineated. In the studies presented in this chapter, we took advantage of the I-gene cDNA microarrays and systematically analyzed gene expression during postnatal retinal development for both wild-type and Nrl^{-/-} retinas. Differential gene expressions during normal development and between the two mouse strains were identified. A major emphasis was placed on developmentally altered genes whose expression was also changed by the deletion of Nrl.

The gene profiles of control developing retina were in concordance with the expected roles of expressed genes. For example, the retina-late genes included crystallins, which are shown to localize in the nuclear layers of retina and change expressions during different biological processes (Xi et al. 2003). Several crystallins have chaperone activity and may play a protective role under stress conditions, such as apoptosis (Andley et al. 2000; Horwitz 1992; Kamradt et al. 2001; Wang and Spector 1994). Higher expression of crystallins, together with several oxidative enzymes, suggested increased functional needs during late stages, perhaps indicating stress response. The temporal expression profiles of phototransduction genes were consistent with previous studies, which showed that a majority of these genes become detectable when the outer segments are formed (Morrow et al. 1998).

Retina-early genes, represented by many protein biosynthesis and cytoskeletal genes, showed lower levels of expression during maturation. Previous studies by *in silico* EST analyses have demonstrated a similar developmental expression pattern in the retina: the developing young retina exhibits more transcripts of protein biosynthesis and cell structure, whereas the adult retina highly expresses genes for energy metabolism and phototransduction (Mu et al. 2001; Yoshida et al. 2002; Yu et al. 2003). Microarray analysis was, therefore, useful in validating previously reported gene profiles and revealing novel as yet undiscovered patterns of gene expression during postnatal retinal development.

Since wild-type and the $Nrl^{-/-}$ retina have distinct photoreceptor subtypes (Mears et al. 2001; Zhu et al. 2003), the expression of genes that are driven by photoreceptor function should change in the $Nrl^{-/-}$ retina, whereas those involved in general housekeeping functions would remain unaltered. Indeed, these two types of genes were observed in our dataset by clustering analysis. Transcripts of the group of retina-late genes represented by enzymes and crystallins became more abundant as retina matures, and this pattern did not change in the $Nrl^{-/-}$ (Figure 7.2B, cluster 5). Genes involved in protein biosynthesis, by contrast, demonstrated a significant delay in their reduction of expression in the mutant retina (Figure 7.2B, cluster 4). The mechanism(s) underlying this lag are unclear; it is, perhaps, due to a negative feedback loop since the retina is preparing to express retina-late genes, such as phototransduction proteins that were absent in the knockout.

The group of genes with temporal expression patterns similar to that of rhodopsin (cluster 6) exhibited significant difference between the wild-type and mutant retinas.

Many known rod phototransduction genes belonged to this group. Interestingly, 4 Bmp/Smad pathway proteins, which included Bmp2, Bmp4, Tob1 and Bmpr1a, demonstrated temporal profiles highly similar to that of rod phototransduction proteins in both wild-type and mutant retinas. Dcn, a gene that can disrupt Bmp/Smad signaling through Ca²⁺ signaling and activation of Camk2b (Abdel-Wahab et al. 2002), is induced in the Nrl^{-/-} (Figure 7.3C). Along with the induced expression of Camk2b (Figure 7.5), this data suggested a potential mechanism for the disruption of Bmp/Smad signaling in the development of the mutant retina. In addition, ChIP analyses showed that Nrl could directly bind to Bmp4 promoter. These findings again advocated a biased utilization of Bmp/Smad signaling by rod- or cone-dominated retinas (Yu, J. et al. manuscript submitted; Chapter 6). It also suggested a potential developmental role of Bmp signaling, controlled by Nrl, in determining whether the rod photoreceptors were able to reach a fully functional stage. This data is consistent with the essential roles of Bmp signaling during early embryonic retinal development (Derynck and Zhang 2003; Miyazawa et al. 2002; Shi and Massague 2003; Wrana 2000). However, more work is needed to evaluate the role of Bmp signaling in regulating expression of phototransduction genes.

Many co-expressed genes are reported to have similar functions and may be co-regulated (Sherlock 2000; Slonim 2002). Clustering of genes based on similar expression patterns have been repeatedly used to gain insights into functions of unknown genes (Furlong et al. 2001; Livesey et al. 2000). In this study, ChIP analyses have identified several genes, including two unknown Riken cDNAs, of cluster 6 as direct targets of Nrl. Other unknown genes in the same cluster, therefore, are also strong candidates of Nrl target genes. Since Nrl is an essential transcriptional factor specific to rods, these

unknown genes may serve important roles in rod function, and may be good candidate genes for retinal diseases. The cluster 6 genes had been further divided into two groups: group-I had genes repressed in all stages of mutant retinal development, and group-II contained genes with delayed response. Both sub-clusters contained phototransduction proteins, some of which are known targets of Nrl. More genes were shown positive, by ChIP analyses, in group I (82%) than in group II (50%). It could be hypothesized that expression of group II genes is synergistically regulated by Nrl and other transcription factors.

During qRT-PCR corroboration of microarray data, we figured out that Hprt expression might change during development. This finding is consistent with earlier reports regarding the differential expression of many housekeeping genes under various conditions (Bustin 2000; Thellin et al. 1999; Warrington et al. 2000). Therefore, normalization of quantitative qRT-PCR data, for the purpose of validating microarray results, needs to be done using a group of housekeeping genes as internal controls, as previously suggested (Vandesompele et al. 2002).

Although microarray technology is useful in systematically examine transcript levels in multiple conditions, its power is limited by the catalog of genes physically present on the array. Since cDNAs on the I-gene microarrays used for this study were isolated from wild-type retinas, rod genes are considerably more abundant than cone genes (Farjo et al. 2002; Yu et al. 2003; Yu et al. 2002). This may explain, at least in part, the lack of retina-late genes (potential cone genes), whose expression was induced in the mutant retina, in our dataset. Furthermore, since this study analyzed transcript

profiles of the whole retina, altered expression of genes may not be direct effects of the deletion of *Nrl*.

In summary, this study used I-gene cDNA microarrays to investigate gene expression patterns during the development of wild-type and *Nrl*^{-/-} retinas. High levels of transcripts associated with protein biosynthesis and cell structure marked the early developing retinas, whereas genes involved in energy metabolism and phototransduction were preferentially expressed in the adult retinas. The development of mutant retina was slightly delayed, as indicated by the prolonged expression of some retina-early genes. This study identified a number of unknown transcripts as potential *Nrl* target genes. Some of them might become retinal disease genes. For the validation of microarray results, our data elucidated the importance of using more than 1 housekeeping genes to normalize cDNA amounts. Our results illustrated an essential role of Bmp/Smad signaling in rod photoreceptor development and function. These findings are a step toward a complete description of the genetic networks of *Nrl*, and may assist in the delineation of rod and cone photoreceptor differentiation.

7.6 Reference

- Abdel-Wahab, N., S.J. Wicks, R.M. Mason, and A. Chantry. 2002. Decorin suppresses transforming growth factor-beta-induced expression of plasminogen activator inhibitor-1 in human mesangial cells through a mechanism that involves Ca²⁺-dependent phosphorylation of Smad2 at serine-240. *Biochem J* **362**: 643-649.
- Andley, U.P., Z. Song, E.F. Wawrousek, T.P. Fleming, and S. Bassnett. 2000. Differential protective activity of alpha A- and alphaB-crystallin in lens epithelial cells. *J Biol Chem* **275**: 36823-36831.
- Bessant, D.A., G.E. Holder, F.W. Fitzke, A.M. Payne, S.S. Bhattacharya, and A.C. Bird. 2003. Phenotype of retinitis pigmentosa associated with the Ser50Thr mutation in the NRL gene. *Arch Ophthalmol* **121**: 793-802.
- Bessant, D.A., A.M. Payne, K.P. Mitton, Q.L. Wang, P.K. Swain, C. Plant, A.C. Bird, D.J. Zack, A. Swaroop, and S.S. Bhattacharya. 1999. A mutation in NRL is associated with autosomal dominant retinitis pigmentosa. *Nat Genet* **21**: 355-356.
- Bustin, S.A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**: 169-193.
- Chen, S., Q.L. Wang, Z. Nie, H. Sun, G. Lennon, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, and D.J. Zack. 1997. Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* **19**: 1017-1030.
- Chowers, I., T.L. Gunatilaka, R.H. Farkas, J. Qian, A.S. Hackam, E. Duh, M. Kageyama, C. Wang, A. Vora, P.A. Campochiaro, and D.J. Zack. 2003. Identification of novel genes preferentially expressed in the retina using a custom human retina cDNA microarray. *Invest Ophthalmol Vis Sci* **44**: 3732-3741.
- DeAngelis, M.M., J.L. Grimsby, M.A. Sandberg, E.L. Berson, and T.P. Dryja. 2002. Novel mutations in the NRL gene and associated clinical findings in patients with dominant retinitis pigmentosa. *Arch Ophthalmol* **120**: 369-375.
- Derynck, R. and Y.E. Zhang. 2003. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425**: 577-584.
- Dhallan, R.S., J.P. Macke, R.L. Eddy, T.B. Shows, R.R. Reed, K.W. Yau, and J. Nathans. 1992. Human rod photoreceptor cGMP-gated channel: amino acid sequence, gene structure, and functional expression. *J Neurosci* **12**: 3248-3256.
- Diaz, E., Y.H. Yang, T. Ferreira, K.C. Loh, Y. Okazaki, Y. Hayashizaki, M. Tessier-Lavigne, T.P. Speed, and J. Ngai. 2003. Analysis of gene expression in the developing mouse retina. *Proc Natl Acad Sci U S A* **100**: 5491-5496.

- Dowling, J.E. 1987. *The retina: an approachable part of the brain*. Harvard University Press, Cambridge, MA.
- Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**: 14863-14868.
- Farjo, R., J. Yu, M.I. Othman, S. Yoshida, S. Sheth, T. Glaser, W. Baehr, and A. Swaroop. 2002. Mouse eye gene microarrays for investigating ocular development and disease. *Vision Res* **42**: 463-470.
- Figueiredo, M.A.T. and A.k. Jain. 2002. Unsupervised learning of finite mixture models. *IEEE transactions on pattern analysis and machine intelligence* **24**: 381-396.
- Furlong, E.E., E.C. Andersen, B. Null, K.P. White, and M.P. Scott. 2001. Patterns of gene expression during Drosophila mesoderm development. *Science* **293**: 1629-1633.
- Horwitz, J. 1992. Alpha-crystallin can function as a molecular chaperone. *Proc Natl Acad Sci U S A* **89**: 10449-10453.
- Kamradt, M.C., F. Chen, and V.L. Cryns. 2001. The small heat shock protein alpha B-crystallin negatively regulates cytochrome c- and caspase-8-dependent activation of caspase-3 by inhibiting its autoproteolytic maturation. *J Biol Chem* **276**: 16059-16063.
- Kennan, A., A. Aherne, A. Palfi, M. Humphries, A. McKee, A. Stitt, D.A. Simpson, K. Demtroder, T. Orntoft, C. Ayuso, P.F. Kenna, G.J. Farrar, and P. Humphries. 2002. Identification of an IMPDH1 mutation in autosomal dominant retinitis pigmentosa (RP10) revealed following comparative microarray analysis of transcripts derived from retinas of wild-type and Rho(-/-) mice. *Hum Mol Genet* **11**: 547-557.
- Korenbrodt, J.I. and T.I. Rebrik. 2002. Tuning outer segment Ca²⁺ homeostasis to phototransduction in rods and cones. *Adv Exp Med Biol* **514**: 179-203.
- Livesey, F.J., T. Furukawa, M.A. Steffen, G.M. Church, and C.L. Cepko. 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. *Curr Biol* **10**: 301-310.
- Mears, A.J., M. Kondo, P.K. Swain, Y. Takada, R.A. Bush, T.L. Saunders, P.A. Sieving, and A. Swaroop. 2001. Nrl is required for rod photoreceptor development. *Nat Genet* **29**: 447-452.
- Mitton, K.P., P.K. Swain, S. Chen, S. Xu, D.J. Zack, and A. Swaroop. 2000. The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. *J Biol Chem* **275**: 29794-29799.

- Mitton, K.P., P.K. Swain, H. Khanna, M. Dowd, I.J. Apel, and A. Swaroop. 2003. Interaction of retinal bZIP transcription factor NRL with Flt3-interacting zinc-finger protein Fiz1: possible role of Fiz1 as a transcriptional repressor. *Hum Mol Genet* **12**: 365-373.
- Miyazawa, K., M. Shinozaki, T. Hara, T. Furuya, and K. Miyazono. 2002. Two major Smad pathways in TGF-beta superfamily signalling. *Genes Cells* **7**: 1191-1204.
- Molday, R.S. 1998. Photoreceptor membrane proteins, phototransduction, and retinal degenerative diseases. The Friedenwald Lecture. *Invest Ophthalmol Vis Sci* **39**: 2491-2513.
- Morrow, E.M., M.J. Belliveau, and C.L. Cepko. 1998. Two phases of rod photoreceptor differentiation during rat retinal development. *J Neurosci* **18**: 3738-3748.
- Mu, X., P.D. Beremand, S. Zhao, R. Pershad, H. Sun, A. Scarpa, S. Liang, T.L. Thomas, and W.H. Klein. 2004. Discrete gene sets depend on POU domain transcription factor Brn3b/Brn-3.2/POU4f2 for their expression in the mouse embryonic retina. *Development* **In press**.
- Mu, X., S. Zhao, R. Pershad, T.F. Hsieh, A. Scarpa, S.W. Wang, R.A. White, P.D. Beremand, T.L. Thomas, L. Gan, and W.H. Klein. 2001. Gene expression in the developing mouse retina by EST sequencing and microarray analysis. *Nucleic Acids Res* **29**: 4983-4993.
- Pepe, I.M. 2001. Recent advances in our understanding of rhodopsin and phototransduction. *Prog Retin Eye Res* **20**: 733-759.
- Rehemtulla, A., R. Warwar, R. Kumar, X. Ji, D.J. Zack, and A. Swaroop. 1996. The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proc Natl Acad Sci U S A* **93**: 191-195.
- Rodieck, R.W. 1998. *The first steps in seeing*. Sinauer Associates, Sunderland.
- Sherlock, G. 2000. Analysis of large-scale gene expression data. *Curr Opin Immunol* **12**: 201-205.
- Shi, Y. and J. Massague. 2003. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**: 685-700.
- Slonim, D.K. 2002. From patterns to pathways: gene expression data analysis comes of age. *Nat Genet* **32 Suppl**: 502-508.
- Sperger, J.M., X. Chen, J.S. Draper, J.E. Antosiewicz, C.H. Chon, S.B. Jones, J.D. Brooks, P.W. Andrews, P.O. Brown, and J.A. Thomson. 2003. Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. *Proc Natl Acad Sci U S A* **100**: 13350-13355.

- Swain, P.K., D. Hicks, A.J. Mears, I.J. Apel, J.E. Smith, S.K. John, A. Hendrickson, A.H. Milam, and A. Swaroop. 2001. Multiple phosphorylated isoforms of NRL are expressed in rod photoreceptors. *J Biol Chem* **276**: 36824-36830.
- Swaroop, A. and D.J. Zack. 2002. Transcriptome analysis of the retina. *Genome Biol* **3**: REVIEWS1022.
- Theillin, O., W. Zorzi, B. Lakaye, B. De Borman, B. Coumans, G. Hennen, T. Grisar, A. Igout, and E. Heinen. 1999. Housekeeping genes as internal standards: use and limits. *J Biotechnol* **75**: 291-295.
- Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: RESEARCH0034.
- Wang, K. and A. Spector. 1994. The chaperone activity of bovine alpha crystallin. Interaction with other lens crystallins in native and denatured states. *J Biol Chem* **269**: 13601-13608.
- Warrington, J.A., A. Nair, M. Mahadevappa, and M. Tsyganskaya. 2000. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol Genomics* **2**: 143-147.
- Wrana, J.L. 2000. Regulation of Smad activity. *Cell* **100**: 189-192.
- Xi, J., R. Farjo, S. Yoshida, T.S. Kern, A. Swaroop, and U.P. Andley. 2003. A comprehensive analysis of the expression of crystallins in mouse retina. *Mol Vis* **9**: 410-419.
- Yang, Y.H., S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai, and T.P. Speed. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* **30**: e15.
- Yoshida, S., B.M. Yashar, S. Hiriyanna, and A. Swaroop. 2002. Microarray analysis of gene expression in the aging human retina. *Invest Ophthalmol Vis Sci* **43**: 2554-2560.
- Young, R.W. 1985. Cell proliferation during postnatal development of the retina in the mouse. *Brain Res* **353**: 229-239.
- Yu, J., R. Farjo, S.P. MacNee, W. Baehr, D.E. Stambolian, and A. Swaroop. 2003. Annotation and analysis of 10,000 expressed sequence tags from developing mouse eye and adult retina. *Genome Biol* **4**: R65.
- Yu, J., M.I. Othman, R. Farjo, S. Zarepari, S.P. MacNee, S. Yoshida, and A. Swaroop. 2002. Evaluation and optimization of procedures for target labeling and hybridization of cDNA microarrays. *Mol Vis* **8**: 130-137.

- Yu, J., Mears, A.J., Yoshida, S., Farjo, R., Carter, T.A., Ghosh, D., Hero, A., Barlow, C., and Swaroop, A. 2004. From disease genes to cellular pathways: a progress report. *Novartis Found Symp* **255**: 147-160; discussion 160-144, 177-148.
- Zhu, X., B. Brown, A. Li, A.J. Mears, A. Swaroop, and C.M. Craft. 2003. GRK1-dependent phosphorylation of S and M opsins and their binding to cone arrestin during cone phototransduction in the mouse retina. *J Neurosci* **23**: 6152-6160.

CHAPTER 8

CONCLUSIONS

8.1 Concluding Remarks

A better understanding of molecular mechanisms that lead to the degeneration of rod and cone photoreceptors is fundamental to the design of therapeutic strategies for retinal and macular dystrophies (Hims et al. 2003; Yu et al. 2004; Zack et al. 1999). The present study aimed to unravel the disparate mechanisms underlying normal rod- and cone- functions. To do so, this thesis has taken advantage of a unique animal model, the *Nrl^{-/-}* mouse, which contains only cone photoreceptors, and has used the cDNA microarrays to examine genomic scale gene expression profiling.

The key contributions of this thesis are:

- (1) Over 10,000 eye-expressed genes or ESTs were generated from three eye cDNA libraries and were functionally annotated.
- (2) Custom cDNA microarrays, containing thousands of eye-expressed genes, were produced and evaluated for profiling ocular tissue gene expression.
- (3) Conditions for microarray experimentation were optimized to reduce variance and identify significant gene expression changes.
- (4) A number of genes and novel ESTs, showing similar temporal expression patterns to genes of rod phototransduction cascade, were identified. These genes are potentially involved in rod function.

- (5) Several downstream target genes of Nrl transcriptional regulatory pathways were identified.
- (6) Bmp/Smad signaling was suggested to play a critical role in rod differentiation and mature rods, whereas Wnt/Ca²⁺ pathway may be important for cones.
- (7) Nrl was suggested to regulate the expression of genes involved in Bmp/Smad signaling.
- (8) Transcripts for protein biosynthesis and cell structure and those for energy metabolism and phototransduction were shown to be overly expressed in the developing and mature retinas, respectively.

These results implied, for the first time, the participation of Bmp/Smad and Wnt/Ca²⁺ signaling pathways in photoreceptor function and demonstrated their biased utilization in rods versus cones.

8.2 Implications

The results of this thesis study are significant for a number of reasons. The ESTs generated and annotated have greatly increased the number of sequences available for eye-expressed genes or ESTs in the public domain databases (<http://www.ncbi.nlm.nih.gov/>). The optimized microarray techniques have been used and recommended by several other researchers, either in or outside of the field of vision research (Dobbin et al. 2003; Manduchi et al. 2002). The I-gene cDNA microarrays are being widely used for transcript analysis of ocular disease and development (Xi et al. 2003; Swaroop A., unpublished data).

This study identified several candidate rod genes, the promoters of which bind to Nrl directly. In addition, several potentially cone genes were identified. The rod/cone genes identified in this study, especially those mapped in the region of known loci for

ocular diseases, are potential candidate genes for retinopathies. This study is also an important step forward for delineating the regulatory networks controlled by Nrl. The revelation of Bmp/Smad and Wnt/Ca²⁺ signaling in phototransduction will lead to a better understanding of rod and cone homeostasis.

8.3 Future Directions

Two major aspects of future investigations are: (1) the follow up of candidate rod or cone genes; and (2) the characterization the Bmp/Smad pathway in rods. The transcriptional factors, such as Ddx5 and Tm4sf2, which showed similar expression profiles as Rho, can be selected for further investigation. ChIP analyses, co-immunoprecipitation and luciferase assay can be used to determine the regulation of Nrl on these genes *in vivo*. The exact binding site of Nrl on these genes can be decided through electrophoretic-mobility shift assay and footprinting experiments. On the other hand, immunohistochemistry or in-situ hybridization will localize these genes in specific cell-types of the retina. To examine their disease-causal effects, mutation screening of these genes can be performed. Eventually, one can undertake condition loss or gain of function studies in mouse.

To elucidate the role of Bmp/Smad signaling in rods, conditional knockout mice must be generated since a majority of the genes involved in this pathway are essential for embryonic development (Winnier et al. 1995; Zhang and Bradley 1996). Recently, transgenic mice expressing Cre-recombinase in specific photoreceptor subtypes have been generated (Akimoto et al. 2004; Akimoto and Swaroop, unpublished data). Using these and other Cre-lines of mice, conditional mutants of genes involved in rods or cones may be generated.

8.4 References

- Akimoto, M., Filippova, E., Gage, P.J., Zhu, X., Craft, C.M., and Swaroop, A. 2004. Transgenic mice expressing Cre-recombinase specifically in M- or S-cone photoreceptors. *Invest Ophthalmol Vis Sci* **45**: 42-47.
- Bessant, D.A., Ali, R.R., and Bhattacharya, S.S. 2001. Molecular genetics and prospects for therapy of the inherited retinal dystrophies. *Curr Opin Genet Dev* **11**: 307-316.
- Bessant, D.A., Payne, A.M., Plant, C., Bird, A.C., Swaroop, A., and Bhattacharya, S.S. 2000. NRL S50T mutation and the importance of 'founder effects' in inherited retinal dystrophies. *Eur J Hum Genet* **8**: 783-787.
- Dobbin, K., Shih, J.H., and Simon, R. 2003. Statistical design of reverse dye microarrays. *Bioinformatics* **19**: 803-810.
- Dryja, T.P., McGee, T.L., Hahn, L.B., Cowley, G.S., Olsson, J.E., Reichel, E., Sandberg, M.A., and Berson, E.L. 1990a. Mutations within the rhodopsin gene in patients with autosomal dominant retinitis pigmentosa. *N Engl J Med* **323**: 1302-1307.
- Dryja, T.P., McGee, T.L., Reichel, E., Hahn, L.B., Cowley, G.S., Yandell, D.W., Sandberg, M.A., and Berson, E.L. 1990b. A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature* **343**: 364-366.
- Dunn, N.R., Winnier, G.E., Hargett, L.K., Schrick, J.J., Fogo, A.B., and Hogan, B.L. 1997. Haploinsufficient phenotypes in Bmp4 heterozygous null mice and modification by mutations in Gli3 and Alx4. *Dev Biol* **188**: 235-247.
- Hims, M.M., Diager, S.P., and Inglehearn, C.F. 2003. Retinitis pigmentosa: genes, proteins and prospects. *Dev Ophthalmol* **37**: 109-125.
- Manduchi, E., Scarce, L.M., Brestelli, J.E., Grant, G.R., Kaestner, K.H., and Stoeckert, C.J., Jr. 2002. Comparison of different labeling methods for two-channel high-density microarray experiments. *Physiol Genomics* **10**: 169-179.
- Winnier, G., Blessing, M., Labosky, P.A., and Hogan, B.L. 1995. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev* **9**: 2105-2116.
- Xi, J., Farjo, R., Yoshida, S., Kern, T.S., Swaroop, A., and Andley, U.P. 2003. A comprehensive analysis of the expression of crystallins in mouse retina. *Mol Vis* **9**: 410-419.
- Yu, J., Mears, A.J., Yoshida, S., Farjo, R., Carter, T.A., Ghosh, D., Hero, A., Barlow, C., and Swaroop, A. 2004. From disease genes to cellular pathways: a progress report. *Novartis Found Symp* **255**: 147-160; discussion 160-144, 177-148.

- Zack, D.J., Dean, M., Molday, R.S., Nathans, J., Redmond, T.M., Stone, E.M., Swaroop, A., Valle, D., and Weber, B.H. 1999. What can we learn about age-related macular degeneration from other retinal diseases? *Mol Vis* **5**: 30.
- Zhang, H. and Bradley, A. 1996. Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Development* 122: 2977-2986.