

Open Research Online

The Open University's repository of research publications and other research outputs

A cell-specific analysis of gene expression in the neural crest of zebrafish.

Thesis

How to cite:

Van Roessel, Peter J. (1998). A cell-specific analysis of gene expression in the neural crest of zebrafish. MPhil thesis. The Open University.

For guidance on citations see [FAQs](#).

© 1998 Peter J. Van Roessel

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

UNRESTRICTED

A CELL-SPECIFIC ANALYSIS OF GENE
EXPRESSION IN THE NEURAL CREST OF
ZEBRAFISH

A THESIS PRESENTED BY

PETER VAN ROESSEL

TO THE

OPEN UNIVERSITY

MILTON KEYNES, ENGLAND

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTER OF PHILOSOPHY

MAX PLANCK INSTITUTE FOR DEVELOPMENTAL BIOLOGY
TÜBINGEN, GERMANY

SEPTEMBER 1998

Sponsoring Establishment:
Medical Research Council
National Institute for Medical Research

ProQuest Number:27696795

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 27696795

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

SUMMARY

This thesis presents a new approach to the rapid and systematic identification of genes expressed differentially between specific cell types. The approach is based upon the hybridization of amplified tissue or cell-specific complex cDNA probes to robotically arrayed cDNA libraries and the computer-assisted analysis ('electronic subtraction') of the resulting hybridization patterns.

This approach was applied to two populations of cultured zebrafish neural crest cells, as well as to tissue from ventral trunk neural tube. Sets of cDNA clones were assembled that putatively represent genes expressed in early-migrating trunk neural crest cells (390 clones), late-migrating trunk neural crest cells (422 clones), or both neural crest cell populations (292 clones), but not in ventral trunk neural tube.

The procedures and techniques used were tested for reproducibility and initial conclusions are made about the feasibility and validity of the approach. A critical assessment of the techniques applied suggests aspects of the approach which might be improved in its future application.

ACKNOWLEDGEMENTS

The work presented here was carried out in the research group of Camila Esguerra and Alexander Crawford in the department of Friedrich Bonhoeffer in the Max Planck Institute for Developmental Biology in Tübingen, Germany. Additional supervision and direction was provided by Suresh Jesuthasan, also of the Max Planck Institute for Developmental Biology. Work was performed in collaboration with Matthew Clark of the department of Hans Lehrach of the Max Planck Institute for Molecular Genetics in Berlin-Dahlem, Germany.

Research was performed under the auspices of the German Human Genome Project (HGF) and was funded primarily by the German Federal Ministry for Education, Science, Research and Technology (BMBF), with additional research funding supplied by the Max Planck Society. Personal support was provided by a Graduate Research Fellowship from the National Science Foundation of the USA.

Friedrich Bonhoeffer is thanked for his support of the research group, and Suresh Jesuthasan for his generous assistance and guidance in the design and pursuit of this project. Matthew Clark, Franco Weth and Barbara Grunewald are gratefully acknowledged for their patient assistance and technical advice, and Chi-Bin Chien, Stephan Neuhauss and others of the Abteilung I fish group are similarly acknowledged for many helpful suggestions and discussions of the work. Darren Gilmour and Henry Roehl are thanked for enthusiastic discussion of the project, and they are additionally thanked with Dongya Gao and G.M. Kelly for providing cDNA markers used in these experiments. Karin Lehmann and the members and staff of Abteilung I have provided a terrific and terrifically helpful working environment, as have Nils Agster, Johannes Zander, Ivillesse Esguerra, and Ya Xian Wang. Jagan Srinivasan and Sylvie LeGuyader are thanked with particular gratitude for their extreme patience, helpful advice and invaluable company. The MRC National Institute for Medical Research in London, England, most especially Rodney King and Kristine Cronhelm, are thanked for enabling this work towards the M.Phil.

Final thanks are due to, and most gratefully given, Camila Esguerra and Alexander Crawford, whose generous support, patient teaching, bountiful enthusiasm, and creative approach to the challenges presented by this project have made work under their guidance a great pleasure.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1 OVERVIEW OF THE APPROACH.....	2
1.2 TECHNIQUES FOR STUDYING DIFFERENTIAL GENE EXPRESSION.....	3
1.2.1 <i>Gene expression analysis at the level of single cells</i>	4
1.2.2 <i>The differential expression analysis of cDNA pools</i>	5
1.2.3 <i>The analysis of arrayed cDNA libraries</i>	6
1.3 THE NEURAL CREST AS A MODEL SYSTEM FOR DIFFERENTIAL EXPRESSION ANALYSIS.....	8
1.3.1 <i>Justification as a model system</i>	8
1.3.2 <i>Distinctions between subpopulations of the neural crest</i>	11
2. MATERIALS AND METHODS.....	14
2.1 NEURAL CREST RT-PCR.....	15
2.1.1 <i>Development and Control of PCR Amplification</i>	15
2.1.2 <i>Experimental tissue and cell sources</i>	17
2.1.3 <i>RT-PCR cDNA amplification and probe synthesis</i>	19
2.1.4 <i>Slot blot analysis of amplified cDNA</i>	22
2.2 LIBRARY HYBRIDIZATIONS.....	23
2.2.1 <i>Probe labeling</i>	23
2.2.2 <i>Library characteristics</i>	25
2.3 HYBRIDIZATION PROTOCOLS.....	25
2.3.1 <i>Complex hybridizations and marker hybridizations</i>	25
2.3.2 <i>Vector hybridizations</i>	26
2.4 HYBRIDIZATION ANALYSIS AND CLONE SELECTION.....	27
2.4.1 <i>Hybridization scoring</i>	27
2.4.2 <i>'Electronic subtraction' parameters</i>	27
2.5 SCREEN OF CLONES BY IN SITU HYBRIDIZATION.....	28
2.5.1 <i>Generation of labeled antisense-RNA probe</i>	28
2.5.2 <i>In situ hybridization</i>	28
3. RESULTS.....	30
3.1 NEURAL CREST RT-PCR.....	30
3.1.1 <i>Development and control testing of PCR amplification</i>	31
3.1.2 <i>Global RT-PCR amplification of neural tube and neural crest cell-specific cDNAs</i>	35
3.2 COMPLEX HYBRIDIZATIONS.....	42
3.2.1 <i>Hybridizations</i>	44
3.2.2 <i>Filter scoring procedure and positive clones scored</i>	47
3.3 MEASURES OF REPRODUCIBILITY: THE SIMILARITY OF SIMILAR HYBRIDIZATIONS.....	50
3.3.1 <i>Similarity of filters hybridized in duplicate</i>	50
3.3.2 <i>Similarity of filters hybridized with probes made from parallel PCRs</i>	50
3.3.3 <i>Similarity of filters hybridized with probes from parallel cDNA syntheses</i>	54
3.3.4 <i>Overall similarity in the late neural crest hybridizations</i>	54
3.4 MARKER HYBRIDIZATION CONTROLS.....	56
3.4.1 <i>Ubiquitous markers</i>	56
3.4.2 <i>Neural crest markers</i>	58
3.4.3 <i>Vector oligonucleotide hybridizations</i>	60
3.5 CLONE SELECTION AND SCREENING.....	62
3.6 IN SITU SCREENING OF SELECTED CLONES.....	65
4. DISCUSSION.....	67
4.1 RT-PCR SENSITIVITY.....	68
4.2 cDNA SYNTHESIS TECHNIQUES.....	69
4.2.1 <i>Random priming</i>	71
4.2.2 <i>Homopolymer tailing</i>	71
4.2.3 <i>Oligonucleotide bridging</i>	73
4.2.4 <i>Second-strand replacement</i>	73

4.3 CONTAMINATION AND AMPLIFICATION OF NON-CDNA.....	74
4.4 HYBRIDIZATION ANALYSIS.....	76
4.5 REPRODUCIBILITY OF HYBRIDIZATION.....	77
4.5.1 <i>Filter reproducibility</i>	78
4.5.2 <i>Probe-labeling reproducibility</i>	79
4.5.3 <i>Modeling the error incurred</i>	82
4.6 ASSESSING THE VALIDITY OF THE RESULTS.....	83
4.6.1 <i>Identity and specificity of the source tissue</i>	83
4.6.2 <i>Efficacy of the subtraction</i>	84
4.6.3 <i>Selection of markers as positive control</i>	85
4.6.4 <i>In situ hybridization</i>	86
4.7 PRACTICAL APPLICATION OF THE HYBRIDIZATION DATA.....	86
5. CITATIONS.....	88

TABLE OF FIGURES

FIGURE 1. TRUNK NEURAL CREST CELLS MIGRATE ALONG TWO DISTINCT PATHWAYS.....	11
FIGURE 2. PCR AMPLIFIES ADAPTED TEMPLATE PRESENT AT LOW COPY NUMBERS, AND YIELDS NO UNSPECIFIC PRODUCT	33
FIGURE 3. DIG-LABELED pBLUESCRIPT PROBE HYBRIDIZES ADAPTED TARGET TEMPLATE AMPLIFIED WITHIN A COMPLEX POOL OF ADAPTED ZEBRAFISH CDNA	34
FIGURE 4. THE ANATOMY OF A DOUBLE-STRANDED CDNA MOLECULE	36
FIGURE 5. NEURAL CREST CELLS MIGRATE IN CULTURE.....	37
FIGURE 6. REPRESENTATIVE GEL-ELECTROPHORETIC PROFILES OF AMPLIFIED CDNA FROM MINIMAL TISSUES.....	39
FIGURE 7. SLOT BLOTS CONFIRM AMPLIFICATION OF ZEBRAFISH CDNA	41
FIGURE 8 A TYPICAL ARRAYED LIBRARY HYBRIDIZATION	43
FIGURE 9. CUSTOM SOFTWARE ASSISTS IN THE MANUAL SCORING OF HYBRIDIZATION PROFILES	48
FIGURE 10. THE OUTPUT OF THE VISUAL GRID SOFTWARE CAN BE REPRESENTED BY X-Y COORDINATE VALUES AND INTENSITY SCORES	49
FIGURE 11. OVERLAP AND SIMILARITY BETWEEN FILTERS HYBRIDIZED WITH THE SAME PROBE.....	51
FIGURE 12. OVERLAP AND SIMILARITY BETWEEN FILTERS HYBRIDIZED WITH PROBE MADE FROM THE SAME CDNA TEMPLATE, YET DIFFERENT PCRS AND LABELING REACTIONS.....	52
FIGURE 13. OVERLAP AND SIMILARITY BETWEEN FILTERS HYBRIDIZED WITH PROBES MADE FROM THE SAME CELL TYPE, YET DIFFERENT CDNA SYNTHESSES, PCRS, AND LABELING REACTIONS	55
FIGURE 14. THE SET OF <i>TOTAL</i> CLONES SCORED IN THE FOUR LNCA AND LNCB HYBRIDIZATIONS.....	55
FIGURE 15. EF1 α PROBE HYBRIDIZES LIBRARY cDNAs ENCODING EF1 α	57
FIGURE 16. RADIOLABELED OLIGONUCLEOTIDE HYBRIDIZES PLASMID VECTOR OF cDNAs	61
FIGURE 17. THE SET OF CLONES SELECTED AS NEURAL-CREST SPECIFIC	64
FIGURE 18. IN SITU HYBRIDIZATION INDICATES NON-SPECIFIC HYBRIDIZATION BETWEEN ARNA PROBES AND DAY-OLD ZEBRAFISH EMBRYOS	66
FIGURE 19. SECOND-STRAND SYNTHESIS BY RANDOM HEXAMER PRIMING.....	70
FIGURE 20. SECOND-STRAND CDNA SYNTHESIS BY HOMOPOLYMER TAILING.....	70
FIGURE 21. SECOND-STRAND PRIMING USING AN OLIGONUCLEOTIDE 'BRIDGE'	70

TABLE OF TABLES

TABLE 1. CDNA SYNTHESSES PERFORMED, WITH INFORMATION REGARDING THE SOURCE AND AMOUNT OF TISSUE USED	46
TABLE 2. PROBES USED FOR ARRAYED LIBRARY HYBRIDIZATION, WITH INFORMATION REGARDING CDNA SOURCE, EXTENT OF PCR POOLING, AND THE AMOUNT OF mRNA REPRESENTED BY THE TOTAL STARTING CDNA TEMPLATE.....	46
TABLE 3. PROBE, FILTER NUMBER, AND POSITIVES SCORED.....	53
TABLE 4. KNOWN-MARKER CLONE POSITIONS SCORE POSITIVELY PRIMARILY IN NEURAL CREST-HYBRIDIZED FILTERS	59

1. **Introduction**

INTRODUCTION

1.1 Overview of the approach

This thesis documents a project with dual intent: the development of an improved technique for the analysis of differential gene expression between tissues or cell types, and the application of this technique to discover genes specifically expressed in the developing neural crest of the zebrafish.

In overview, the technique described is based upon the hybridization of complex cDNA probes, amplified from different specific tissue samples, to multiple filter copies of an arrayed cDNA library. Computer-assisted comparison of the resulting hybridization profiles allows identification of cDNA clone sets that encode mRNAs differentially expressed between the tissues analyzed. The approach described is based on two technologies: the first is the development of methods for the ordered and reproducible high-density printing of cDNA libraries on nylon filters; the second is the development of methods for the global PCR amplification of cDNAs synthesized from small-tissue or single-cell quantities of mRNA.

As part of the German Human Genome Project, and through collaboration with the laboratory of Hans Lehrach of the Max Planck Institute for Molecular Genetics in Berlin, gridded cDNA libraries representing approximately 25,000 different cDNAs derived from late-somitogenesis stage zebrafish have been made available to our research group. Described here is the development of a protocol for the global amplification of cDNAs from small tissue samples, and the use of this protocol to amplify cDNAs specific to two populations of zebrafish neural crest cells. By hybridizing labeled probes from these neural crest-specific cDNA pools to copies of the available gridded library, and by performing 'electronic subtractions' between these results and those of a third probe, derived from ventral neural tube, sets of cDNAs were isolated that putatively encode genes specifically expressed in individual cell populations of the neural crest.

INTRODUCTION

Additionally, as most hybridizations were performed in duplicate, and as steps in the process were repeated as controls for reproducibility, some conclusions are drawn about the advantages, limitations, and overall efficacy of the approach described, as well as aspects of this approach which might be improved in its further applications.

1.2 *Techniques for studying differential gene expression*

Although previous approaches toward the identification of differentially expressed genes, such as subtractive cloning (Wang and Brown, 1991), differential display (Liang and Pardee, 1992), representational difference analysis (Lisitsyn *et al.*, 1993) and suppression subtractive hybridization (Diatchenko *et al.*, 1996), have improved our understanding of a select number of significant genes, no previously-employed technology has permitted the monitoring of more than a small fraction of genes in any given cell type. The development of comprehensive and high-throughput approaches to expression analysis with single-cell resolution, however, is mandated by the new era of genomic biology (Lander, 1996). As a more global knowledge of model genomes is acquired, with the full genomic sequencing of several species to be completed in the near future, correspondingly more comprehensive approaches toward the analysis of gene expression need to be taken. Hybridization of radiolabeled single-cell cDNA probes to high-density arrayed cDNA libraries, as described here, should with future optimization allow the simultaneous and quantitative determination of gene expression state for any given cell type. This ultimate goal has not yet been reached, yet the technologies and approach described here are a preliminary step in this future direction.

The approach to differential gene expression analysis described here relies heavily on recent technological developments in cDNA library generation and analysis. Much recent progress has been made to improve these component techniques; to our knowledge, however, they have not been assembled into the full protocol as described here. No group has yet reported the complex hybridization of an arrayed library with amplified probes from single cells or small groups of cells.

INTRODUCTION

1.2.1 Gene expression analysis at the level of single cells

The current level of spatial resolution at which expression analysis can easily be carried out is too low for many applications in biology. The need to assess differential gene expression in extremely small tissue samples, and often in single cells— as is the case for many problems in neurobiology, for example— has driven the recent development of methodologies for synthesizing and amplifying cDNA pools from extremely limited quantities of mRNA.

The amplification of known genes using gene-specific primers for cDNA synthesis and PCR presents a straightforward technical challenge. This challenge has both encouraged the development of more sensitive PCR protocols and led to improved techniques for isolating individual cells or cell cytoplasm for analysis. Single-cell reverse transcription and PCR (RT-PCR), first attempted by electrophysiologists interested in the expression states of various ion channel subunits in neurons, generally involves the isolation of cell cytoplasm by aspiration with patch-clamp microelectrodes. Such attempts often follow prior electrophysiological characterization of the cell (Monyer and Lambolez, 1995; Sucher and Deitcher, 1995). The advent of techniques for globally amplifying complex cDNA pools from limited amounts of RNA (Belyavsky *et al.*, 1989; Van Gelder *et al.*, 1990; Froussard, 1992), however, and recognition of the possibilities for differential expression screening these techniques provide, has now expanded interest in single-cell gene expression analysis beyond the realm of electrophysiologists.

The acceptance of the approach by researchers without extensive electrophysiological tools has led to the development of further techniques for the isolation of single cells or tissues. In addition to the isolation of specific cells by dissociation and morphology-based selection (Dulac, 1995), targeted expression of fluorescent protein in cells of interest has allowed the collection of specific tissue by dissociation and subsequent FACS separation (Amrein and Axel, 1997). Biochemical approaches have been

INTRODUCTION

improved to allow the synthesis of cDNA from fixed cells in paraffin sections (Crino *et al.*, 1996), and recent technological developments for laser-assisted capture of single fixed cells (Emmert-Buck *et al.*, 1996; Bonner *et al.*, 1997) should lead to the further application of sectioning and staining in the identification and isolation of target cells.

Aside from the difficulties of obtaining experimental tissue samples, the greater challenge of reproducibly amplifying cDNAs representing the full complement of mRNAs present in a small tissue sample has led to the development of several sensitive cDNA synthesis and amplification protocols. The pioneering work of Van Gelder *et al.* (1990) has led to the development of protocols based on successive linear amplifications with RNA polymerase (Eberwine, 1996; Philips and Eberwine, 1996), and that of Belyavsky *et al.* (1989) to protocols based on PCR (Dulac, 1998, Karrer *et al.*, 1995). These basic approaches, and modified versions thereof, have since been applied by various research groups to diverse biological questions. The inherent problem of assessing the success of a global amplification, however— when no direct measurements of the source material can be made— continues to cloud objective judgement on the benefits and limitations of these different techniques.

1.2.2 The differential expression analysis of cDNA pools

Karrer *et al.* (1995) and Brady *et al.* (1995) amplified complex cDNA pools from individual plant cells and from cultured hemopoietic progenitor cells, respectively, and analyzed the results by slot-blot hybridization with known marker genes. Such experimentation has confirmed the possibility of 'expression mapping' individual cells, as well as put forth the general approach as a feasible tool for the study of cell state and lineage diversification. The significant amount of work involved in slot blot analysis, however, limits the utility of this approach. Additionally, the necessarily limited perspective provided by working with a small set of known genes constrains judgements on the success of such attempts at global cDNA amplification.

INTRODUCTION

The differential screening approach of Dulac and Axel (1995), while similarly providing little insight into the success or failure of the RT-PCR approach used, has gained recognition for its success in achieving an experimental goal: the identification of genes putatively encoding mammalian pheromone receptors. By hybridizing cDNA amplified from one vomeronasal receptor-neuron to that of another, as well as the complementary procedure, Dulac and Axel identified cDNAs unique to each of the original cDNAs pools. The success of this approach perhaps most reflects the optimal characteristics of the system under investigation, however. The genes that were sought after are hypothesized to comprise 1% of the total neuronal mRNA and to be among the only differences between generally identical cell types. This demonstrates the importance of thoughtfully matching experimental technique with the biological question at hand.

1.2.3 The analysis of arrayed cDNA libraries

Advantageous aspects of both analysis methods described above are combined in the approach presented in this thesis. By using the cell-specific complex pool itself as a hybridization probe, like Dulac, yet screening reproducibly arrayed libraries of approximately 25,000 identified clones, the comprehensive screening potential of Dulac and Axel's protocol is given the systematicity of Brady *et al.*'s extensive slot-blotting. As with the latter approach, data generated with every hybridization creates an 'expression map' for each arrayed cDNA clone. With an extremely high capacity for cDNA screening, however, the future use of complex probes hybridized to arrayed libraries should allow the creation of 'expression atlases', containing information encompassing not only small numbers of known markers, but significant portions of all expressed genes in the organism studied.

The future possibility of such comprehensive analysis has been advanced by a number of recent developments in the high-density arraying of cDNA libraries. The most promising of these technologies involves the synthesis of short specific

INTRODUCTION

oligonucleotide sequences onto 1-2 cm glass chips. These chips allow small-volume hybridizations and return precise and sensitive fluorescent hybridization signals, detectable by confocal laser microscopy (Chee *et al.*, 1996; Lockhart *et al.*, 1996). These technologies, however, require sequence information for each oligonucleotide before it can be synthesized on the chip. While this is a possibility for large expressed sequence tag (EST) sets (Okubo *et al.*, 1992), the current paucity of EST libraries for many model systems, as well as the high price involved in the production of such chips, makes them an unrealistic option for most researchers at present. Related technologies for the liquid-jet printing of PCR products directly onto glass chips have also been developed (Schena *et al.*, 1995; Lashkari *et al.*, 1997), but these processes are similarly subject to prohibitive cost and an even more limited density of printing than oligonucleotide-based approaches.

The robotic printing of cDNA-carrying bacterial clones onto large nylon membranes, followed by growth, lysis, and fixation of DNA to the membrane— the technological forerunner to glass chip technology— remains a viable and relatively inexpensive alternative for generating arrayed cDNA libraries. This arraying technology (Lennon and Lehrach, 1991) has now been adapted to a high-throughput scale and offered as a free service to member groups of the German Human Genome Project (<http://www.rzpd.de/>). These cDNA-library 'colony filters' are generally produced for hybridization with specific probes as a resource for researchers interested in specific genes. The use of arrayed colony filters for complex hybridization, however, as performed by Gress *et al.* (1992), as well as groups working with smaller sets of cDNA clones (Nguyen *et al.*, 1995; Piétu *et al.*, 1996), has furthered the use of this arraying technology for differential expression analysis and has demonstrated the significant promise of this approach.

1.3 *The neural crest as a model system for differential expression analysis*

1.3.1 Justification as a model system

Analysis of zebrafish neural crest— specifically the comparison of late- and early-migrating neural crest— was selected as an appropriate project on which to test the technologies developed here based on two major considerations: the first was that the subject be well-suited to the technique; the second was that the technique be well-suited to the subject.

Underlying the first assertion— that the neural crest is appropriate for experimentation of the sort described— is that the neural crest provides a clearly defined system of very similar cells which are amenable to experimental isolation and culture. As described, an important consideration for any differential gene expression assay is the overall difference between subject populations. While estimations of transcriptional difference between cell types are most often only speculative, empirically the best results from differential expression screens have come from researchers focusing on well-defined cell types with very few, yet very strong, differences (Dulac and Axel, 1995, Brady *et al.*, 1995; Amrein and Axel, 1997). Accessibility of the subject populations is a necessary consideration as well. The study of neural crest, and zebrafish neural crest in particular, is advantageous in this regard. The zebrafish neural crest includes fewer and larger cells than its counterparts in chick or *Xenopus* (Raible *et al.*, 1992), and these cells are easily observed and accessed. A simple system for zebrafish cell culture has been previously described which provides a direct way to isolate specific cell populations (Jesuthasan, 1996). Methods for neural crest cell aspiration from live embryos (originally used in transplantation experiments), though technically more challenging, also present an established means of isolating specific cells (Raible and Eisen, 1996). A further advantage of neural crest study in zebrafish is that the optical clarity of the zebrafish embryo allows the lucid visualization of crest cells *in vivo* (Raible and Eisen, 1994). In fixed animals as well, the proximity of crest-cell

INTRODUCTION

populations to the embryo surface allows for the easy assay of gene expression by whole mount *in situ* hybridization. Such analysis is required to confirm the results of a differential gene expression screen as presented here. A final pragmatic consideration in choosing the cellular system to investigate was that the processes of neural crest development occurs in the time window represented by one of the gridded cDNA libraries available from the Resource Center of the German Human Genome Project (<http://web.rzpd.de/cgi-bin/newlib>).

Underlying the second line of reasoning— that the technique be appropriate to the subject— is that specific neural crest markers found by comparative analysis of gene expression would be of particular value to the study of the neural crest. While there exists a strong tradition of embryological experimentation in neural crest development, relatively little is known about the molecular mechanisms underlying the observed embryological phenomena. As few specific molecular markers for neural crest cells exist, an immediate value for neural crest-specific genes might come from their utility in screening mutant populations for mutations in genes affecting neural crest development (Henion *et al.*, 1996), and in analyzing the effects of existing mutations affecting neural crest development (Kelsh *et al.*, 1996; Neuhauss *et al.*, 1996). Zebrafish neural crest-specific genes are of particular value in this regard, as zebrafish is perhaps the preeminent vertebrate forward-genetic model organism (Hafter *et al.*, 1996; Driever *et al.*, 1996). Though existing markers, such as AP2 (Furthauer *et al.*, 1998)), Fkd6 (Odenthal and Nusslein Volhard, 1998)), and Snail 2 (Thisse *et al.*, 1995)) are amenable for this purpose, markers with well-characterized sub population-specific expression might allow more subtle analysis of new and existing mutants. Even without reference to mutant screens, many questions in neural crest developmental biology might be answered in part by appropriate marker genes. For example, simply knowing that pre-migratory differences in gene expression exist between neural crest sub-populations would itself inform the current debate about the role of intrinsic specification in neural crest-cell fate choice.

INTRODUCTION

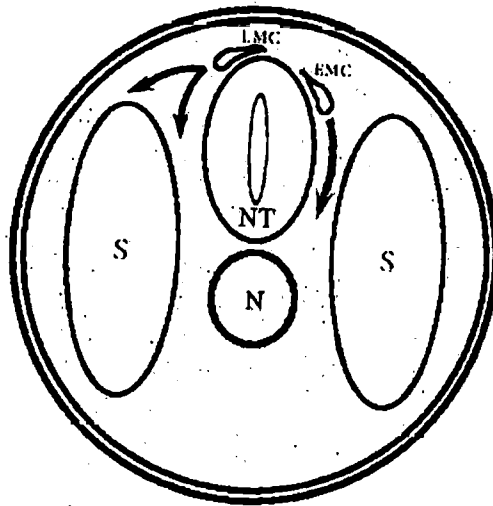


Figure 1. Trunk neural crest cells migrate along two distinct pathways. Early-migrating cells (EMC, purple) migrate ventrally between neural tube (NT) and somite (S). Late-migrating cells (LMC, red) take the ventral pathway, but can also migrate along a lateral pathway, between somite and dorsal ectoderm. N, notochord. After Raible, *et al.* (1992).

INTRODUCTION

1.3.2 Distinctions between subpopulations of the neural crest

The neural crest itself is a multipotent population of precursor cells that migrate away from the dorsal-most aspect of the neural tube, giving rise to a broad array of progeny: cell types of known neural crest origin include neurons, glial cells, endocrine cells, pigments cells, cartilage cells, and bone cells (Le Douarin, 1982; Weston, 1991; Bronner-Fraser and Fraser, 1991; Le Douarin et al., 1993; Eisen and Weston, 1993; Raible and Eisen, 1994). One of the most intriguing challenges of neural crest developmental biology is that of deciphering the mechanisms by which these apparently similar cells achieve such a startling diversity of fates. Among the basic issues inviting investigation are the nature and relative contributions of environmental stimuli and cell-intrinsic determinants to fate choice, the time-course of fate restriction and competency change, and the mechanisms controlling neural crest cell migration, which itself potentially determines which environmental signals the cells receive.

One of the least ambiguous approaches to understanding what role, if any at all, intrinsic signals play in guiding neural crest-cell fate choice is to search for differences present between subgroups of neural crest cells prior to migration. Such cells should still be subject to the same environmental influences. Any such difference that correlates with later differences in cell fate may provide indications about the causal mechanisms of lineage divergence and fate choice.

A particularly well-characterized distinction between subpopulations of neural crest cells in zebrafish trunk is one made between 'early' and 'late'-migrating cells. As described in Figure 1, the first crest cells to migrate start from a relatively more lateral position on the neural tube and migrate exclusively along a 'ventral pathway' between the neural tube and somite; those cells with origins more toward the dorsal midline, however, leave the neural tube at a later time and migrate not only along the ventral pathway, but also along a 'lateral' pathway between the somite and ectoderm (Raible and Eisen, 1994). These distinctions in migratory behavior correlate with differences in

INTRODUCTION

cellular progeny. While early-migrating trunk neural crest cells have neurogenic potential (giving rise to dorsal root ganglion neurons, among others), late-migrating cells— even those that take the same ventral pathway— never become neurons (Raible and Eisen, 1996). Consistent with this observation, neuronal derivatives of head neural crest are found in more lateral positions of the pharyngeal arches, where a more lateral position correlates with the early onset of migration (Schilling and Kimmel, 1994).

The degree to which fates are restricted in these two subpopulations of neural crest, as well as the broader question as to whether intrinsic differences between these cell types exist, has been the subject of much recent investigation. One key observation relevant to this debate is that late-migrating cells— even if placed in an 'early' environment— are not competent to become neurons (Raible and Eisen, 1996). This thus suggests that there *are* cell-intrinsic differences in the properties of late and early neural crest— that late migrating cell fates are somehow restricted before migrating from the neural tube. Recent work in chick by Baker *et al.* (1997), suggesting that late and early migrating neural crest populations have equivalent potential, however, illustrates the continuing controversy surrounding this issue.

In an intriguing— if still speculative— twist, it has been suggested that the differences between cell populations are not differences in their competency to respond to signals originating from the environment of the ventral pathway, but rather to signals that originate with the early migrating crest cells themselves. A caveat to the observation by Raible and Eisen (1996) that late migrating cells placed in an early environment do not become neurons, for example, is that these cells *can* become neurons when the early migrating cells are laser-ablated from the ventral pathway. The authors provide additional evidence for the existence of regulative cell-cell interactions between crest-cell populations by observing that following ablation of laterally located crest cells, medially-located crest cells (putatively destined for late migration) shift to a more lateral position and migrate early. They further parallel potential lateral-inhibitory processes involved to similar processes in drosophila neurogenesis; no evidence has

INTRODUCTION

yet been published, however, that the molecules involved in insect lateral inhibition, such as those of the Delta-Notch signaling pathway, are active in vertebrate neural crest.

2. **Materials and Methods**

2.1 *Neural crest RT-PCR*

2.1.1 **Development and Control of PCR Amplification**

2.1.1.1 *Universal primer sequences*

Two 18-base and one 20-base primer sequences were designed for use in global RT-PCR experiments. Primers were designed to be short, to contain sites for digestion with restriction enzymes, and to have a relatively lower annealing temperature at the 3' end than at the 5' end. Primer sequences were checked with oligonucleotide analysis software (Oligo[®] v.5.0) for possible dimerization and hairpinning, and primers used in the same reactions were checked to assure minimal base-pairing interactions.

Primer 1, containing an *Xba* I restriction site, has the sequence

5'-AGCGGCACTCTAGATACA-3'.

Primer 2, containing an *Eco* RI restriction site, has the sequence

5'-TCCCCAACGAATTCAGAC-3'.

Primer 3, containing a *Pst* I restriction site, has the sequence

5'-CTGAACAGTTCTGCAGCTGT-3'.

2.1.1.2 *Control templates*

2.1.1.2.1 For the assay of PCR sensitivity

To generate a directional target template for amplification with two distinct primers, plasmid Bluescript II KS - was restricted with *Rsa* I to yield a blunt 1192 bp fragment. This fragment was ligated to Primer 2 (annealed to its complement with a 3' TGTG overhang), PCR amplified and gel purified. The resulting 1218 bp product was then restricted with *Nae* I, after which Primer 3 (annealed to its complement with a 3' ATAG overhang) was ligated to the blunt ends of the resulting 447 bp and 781 bp fragments.

MATERIALS AND METHODS

The ligation products were PCR amplified, and the 465 bp amplicon was gel purified, quantified, and serially diluted for use as a directional PCR template.

2.1.1.2.2 For detection of a target amplified in a complex pool

Plasmid Bluescript II KS - (Stratagene) was cut with *Pvu* II to yield a 492 bp blunt-ended fragment. Primer 1 (annealed to its complement with a 3' CACA overhang) was ligated to the 492 bp fragment, and the resulting 522 bp fragment was PCR amplified, gel-purified, quantified, and serially diluted for use as a target PCR template.

2.1.1.3 *Optimized PCR Reaction*

2.1.1.3.1 Chemical mix

Template to be amplified was seeded into a 50 μ l mix containing: 1x OptiPrime No.5 PCR Buffer (Stratagene); 0.7 mM HCl; 0.2 mM each dATP, dCTP, dGTP, and dTTP (Pharmacia); 0.4 μ g Primer 2; 5% Formamide (Sigma); 0.13 mg/ml BSA (NEB); and 2.5 units *Taq* Polymerase (Pharmacia).

2.1.1.3.2 Thermocycling

Thermocycling was performed as follows on a Hybaid Touchdown Thermocycler using simulated tube mode and 0.5 ml Safe-Lock reaction tubes (Eppendorf).

Stage 1 (initial denaturation)
94°C 3 min.

Stage 2 (30 cycles)
94°C 15 sec.
60—45,5°C 2 min. (0.5° cooler each cycle)
Ramp to 72°C 2 min. (ramp 0.3°/sec.)

Stage 3 (10 cycles)
94°C 15 sec.
50°C 2 min.
Ramp to 72°C 2 min (ramp 0.3°/sec.)

Stage 4 (final extension)
72°C 8 min.

MATERIALS AND METHODS

2.1.1.4 *Southern blot hybridization to detect target in a complex pool*

Southern Blots to detect pBluescript amplification product in a complex pool of PCR products were modeled after a standard protocol for capillary Southern transfer (Sambrook *et al.*, 1989) as modified by Weth (1993). Downward transfer with 1M Ammonium Acetate pH 9.3 was used to transfer approximately 300 ng of PCR product (10 µl of PCR reaction), run out in a 1% TAE Agarose gel (Gibco), onto positively-charged nylon membranes (Boehringer Mannheim). After UV crosslinking, membranes were hybridized with the product of a DIG-High Prime labeling reaction (Boehringer Mannheim) performed according to the manufacturer's instructions with 20 ng of PCR target fragment as template. Hybridization was performed as for complex filter hybridizations (see below) in 10 mL of hybridization buffer at 68°C. The binding of conjugate Anti-DIG-Alkaline Phosphatase antibody (Boehringer Mannheim) and a color reaction with CSPD (Boehringer Mannheim) were carried out as per manufacturer's directions. The membrane was exposed to standard autoradiographic film (Kodak) for two hours.

2.1.2 **Experimental tissue and cell sources**

Zebrafish (Tübingen Wildtype) were kept under standard conditions in institute facilities (Brand *et al.*, 1995). Embryos, once born, were kept at 28.5°C in embryo medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 MgCl₂) until use. Developmental stages used for culture were determined with reference to Raible *et al.* (1992).

2.1.2.1 *Ventral neural tube (VNT):*

A single embryo was kept until the eight-somite stage (~13 hrs (Kimmel *et al.*, 1995)), after which the epidermis was removed using watchmaker's forceps, the body axis was separated from the yolksac, and the section of trunk corresponding to the

MATERIALS AND METHODS

middle six somites was isolated. This section of trunk was incubated for several minutes in a 1:1 dilution of Trypsin/EDTA (Sigma):Ringer solution (Normal Ringer Solution; Westerfield, 1993) at room temperature and digestion was stopped in ringer with approximately 0.5% (w/v) BSA. Somites, notochord, pronephrous, and other tissues adjacent to the neural tube were removed with tungsten needles and gentle passaging through the opening of a pulled glass pipette. The isolated neural tube was then longitudinally sectioned into dorsal and ventral halves, the ventral of which was sectioned again into three parts of equivalent length but varying origin along the anteroposterior axis. These three sections, each on the order of a hundred cells, were used to seed the three VNT-probe cDNA syntheses.

2.1.2.2 *Early neural crest (ENC):*

Neural tubes adjacent to the second to eighth somites of eight-somite embryos were isolated as described above. The neural tubes were placed in sterile tissue culture dishes and cultured in a modified L-15 medium (containing BSA, 1 mg/ml ; gentamycin, 0.01 mg/ml; and buffered with 20 mM HEPES pH 7.3) for six to seven hours at 28.5°C. By the end of the culture period, migratory cells were visible and adhered to the floor of the dish. These were aspirated from the dish using pulled borosilicate capillaries (Clark Electromedical Instruments) and were seeded directly into the lysis mixture (see below) on ice.

2.1.2.3 *Late neural crest (LNC):*

Neural crest cell culture was performed as with early crest, but the neural tube used was adjacent to the first eight somites of a twenty-somite (19 hr) embryo. Cultures were incubated for seven to eight hours before cells were isolated.

MATERIALS AND METHODS

2.1.3 RT-PCR cDNA amplification and probe synthesis

The protocol developed and used here is based on that of Dulac (1998) and to a lesser extent on the manufacturer's instructions for the SMART™ PCR cDNA Synthesis Kit (Clontech). Table 1 presents a summary of the cDNA syntheses performed. Table 2 summarizes the number of amplifications of these cDNA templates pooled to make each probe, the number of cells represented by each probe (cell equivalents), and the filters to which the probe was hybridized.

2.1.3.1 *Ventral neural tube*

Three ventral neural tube cDNA syntheses (VNT1, VNT2, and VNT3) were performed with the following steps.

Cells chosen, with approx. 0.5-1.5 μ l culture medium, were added to 4 μ l of a mix containing 125 ng/ μ l Primer 2-poly(dT)₁₈, 0.625% IPEGAL (Sigma), and, when RNase Inhibitor was used, 0.5 units Prime RNase Inhibitor (5 Prime->3 Prime) and 20 units placental RNase inhibitor (Boehringer Mannheim) in nuclease-free, reverse-osmotically purified water (Sigma). The mixture was centrifuged down immediately after the addition of cells and kept on ice until all samples were collected. Cells were then lysed by incubation at 70°C for five minutes, after which the mixture was cooled to room temperature for two minutes to allow annealing of primer to mRNA. The mixture was then chilled on ice and spun down.

First-strand (antisense) cDNA was synthesized by adding 5 μ l of a mixture containing 2x 1° cDNA synthesis buffer (Gibco), 0.2 M DTT (Gibco), 1 mM each dATP, dCTP, dGTP, and dTTP (Pharmacia) and 100 units SuperScript II Reverse Transcriptase (Gibco), and incubating at 37° for one hour.

Second-strand (sense) cDNA was synthesized by adding 66 μ l of a mixture containing 1.14x 2° strand cDNA synthesis buffer (Gibco), 0.23 mM each dATP,

MATERIALS AND METHODS

dCTP, dGTP, and dTTP (Pharmacia), 200 units T4 DNA Ligase (NEB), 20 units *E. coli*. DNA Polymerase (NEB), and 2 units RNase H (Gibco), and incubating at 16°C for two hours.

The cDNA was blunted by adding 6 units T4 DNA Polymerase (NEB) and incubating for five minutes at 16°, after which the reaction was stopped by incubation at 70° for 25 minutes.

The cDNA was adapted for PCR amplification by ligating Primer 2 (annealed to its phosphorylated complement with a 3' TGTG overhang) to both blunt ends. This was accomplished by adding 24 µl of a mixture containing 4.16x Ligase Buffer (NEB), 1 µg pre-annealed Primer 1 : cPrimer 1-TGTG-3' and 400 units T4 DNA Ligase (NEB), followed by 20 hours of incubation at 16 °C.

PCR amplification of the probe was performed using 3 µl of the ligation product (template cDNA) as template in a 50 µl PCR reaction with a composition similar to that described for the PCR sensitivity tests above. The thermocycler and tubes used were identical. The previously-optimized PCR thermocycling program was used, but was modified to increase polymerization time, as this was empirically shown to increase the upper range of amplified cDNA lengths (data not shown). The program used was as follows:

Stage 1 (initial denaturation)
94°C, 3 min.

Stage 2 (30 cycles)
94°C 15 sec.
60—45.5°C 2 min. (0.5° cooler each cycle)
Ramp to 72°C 4 min. (ramp 0.3°/sec.)

Stage 3 (10 cycles)
94°C 15 sec.
50°C 2 min.
Ramp to 72°C 4-10 min (ramp 0.3°/sec.; each cycle 40 sec.longer)

Stage 4 (final extension)
72°C 8 min.

MATERIALS AND METHODS

PCR products were analysed by standard gel electrophoresis in TAE buffer (Sambrook *et al.*, 1989) on a 1,2% agarose (Gibco) gel. For further separation of DNA bands, TAE gels of 1% agarose (Gibco) and 1% Low melting-point agarose (NuSieve®/FMC-Biozym) were used.

To produce linearly amplified antisense cDNA for later random-primed labeling, the amplicons of three parallel PCR amplifications of VNT3 cDNA (a cDNA pool made without the addition of RNase inhibitors) were pooled together, spectrophotographically quantified, and purified using QIAquick Spin Columns (Qiagen) according to the manufacturer's instructions. 200 ng of this purified PCR product was used as template in a 200 µl reaction with a composition identical to that of the PCR mixture, save the addition of 1.7 µg 5'-biotinylated Primer 3 instead of Primer 2, and a total of 10 units of *Taq* Polymerase.

Cycling parameters were as follows:

Stage 1 (initial denaturation)
94°C, 3 min.

Stage 2 (40 cycles)
94°C 15 sec.
50°C 2 min.
Ramp to 72°C 4-10 min (ramp 0.3°/sec.; 12 sec. longer each cycle)

The product of this linear amplification reaction was purified with QIAquick columns and quantified by spectrophotometry.

2.1.3.2 *Early neural crest*

Two early neural crest cDNA probes (ENC1 and ENC2) were created similarly to those from Ventral Neural Tube, with the following modifications: Primer 2, modified with a 3' poly(dT)₁₈ and a 5' ACACACACAC spacer sequence, was used to prime first-strand cDNA synthesis, and Primer 1, annealed to its complement with a 3' ATAG overhang was ligated on to adapt the cDNA for PCR with this primer; no RNase Inhibitor was used; 5 units *E. coli*. DNA Ligase were used instead of T4 DNA Ligase

MATERIALS AND METHODS

in the second strand synthesis; and 10 PCR amplicons from ENC1 cDNA were pooled for use as template in the linear antisense-strand amplification.

2.1.3.3 *Late neural crest*

Three late neural crest cDNA probes (LNC1, LNC2, and LNC3) were created similarly to the early neural crest probes, with the exception of the primers used, which were identical to those used for the VNT syntheses, and the cDNA amplicon pooling steps to make linear amplification template. For the probes used on filters 39-1-207 and 39-1-208 (LNC A) 16 amplicons were pooled— three from LNC1 cDNA, three from LNC2 cDNA, and ten from LNC3 cDNA; for the probes used on filters 39-1-209 and 39-1-210 (LNC B) the same mix of cDNAs was pooled; and for the probes used on filter 39-2-224 (LNC X) and on filter 39-2-225 (LNC Y), ten amplicons were pooled each from LNC1 and LNC2 cDNA, respectively.

2.1.4 Slot blot analysis of amplified cDNA

2.1.4.1 *Blotting of cDNA*

Slot blots were performed by suction-blotting 550 ng of alkali denatured cDNA (incubated for thirty minutes at 37°C in 0,2 M NaOH/ 2mM EDTA, followed by neutralization with 0,3 M Sodium Acetate) per slot onto a positively charged nylon membrane (Boehringer Mannheim), using a Gibco slot-blotting apparatus, followed by UV-crosslinking. The cDNA amplicons blotted were created as described earlier, but the number of amplifications used for each slot varied. For each cDNA pool blotted, the following number of PCR amplicons were mixed: 3 PCRs each for VNT1, VNT2, VNT3, ENC Mock and VNT Mock; 10 PCRs each for ENC1 and ENC2; 3 PCRs of whole embryo cDNA (created using a similar process but approximately 1 ng of purified, poly (dA)-tail-enriched, late-somitogenesis stage zebrafish RNA as template (Esguerra, unpubl.)) were pooled for southern hybridizations as a positive control.

2.1.4.2 *Hybridization of probe to blots*

Radiolabeled hybridization probes were generated from cDNAs encoding zebrafish EF1- α (Gao *et al.*, 1997) and zebrafish β -actin (Kelly *et al.*, 1998) by restricting with *Hind* III and *Hinc* II, respectively, gel purifying the resulting fragments, and using approximately 40 ng as templates for a random hexamer priming reaction (T7 Quick Prime Kit, Pharmacia) according to the manufacturer's directions. Reaction product was diluted in 7.5 mL hybridization buffer and denatured before hybridization. The pretreatment and hybridization of the slot-blotting membrane filter was similar to that of the gridded library hybridizations (see below), with the exception of hybridization buffer volumes, here 7.5 mL per filter.

2.1.4.3 *Exposure*

After washing, again using the same conditions as for the hybridization of gridded libraries, the hybridized blots were exposed to standard autoradiographic film (Kodak) for four hours at -70° C.

2.2 *Library hybridizations*

2.2.1 *Probe labeling*

2.2.1.1 *complex probes*

For all tissue-specific antisense cDNA pools (with exception noted below), 32 P-labeled probe for hybridization to the gridded libraries was created using the following procedure: 400 ng of biotinylated single-stranded antisense cDNA was diluted in 68 μ l H₂O, denatured at 95°C for five minutes, and placed on ice. 20 μ l T7 QuickPrime reagent mix (Pharmacia), 10 μ l α ³²P dCTP (ca. 100 μ Ci) (Amersham) and 2 μ l (20 units) T7 DNA Polymerase were added, and the mixture was incubated for thirty minutes at 37°C. Reaction products were purified using QIAQuick spin columns

MATERIALS AND METHODS

(Qiagen) according to the manufacturers protocol. The purified reaction product was incubated with 0.5 mg Streptavidin-bound paramagnetic beads (M280 Streptavidin, Dynal) in 150 μ l of 0.66 M NaCl for twenty minutes at 37°C. To denature radiolabeled probe from the Streptavidin-bound template, the mixture was incubated for five minutes at 95°C, placed on ice, briefly centrifuged, and magnetically concentrated. The supernatant was saved and the procedure was repeated after re-suspending the beads in 50 μ l of 10 mM Tris-HCl pH 8.5 buffer, after which the second supernatant was added to the first. The combined supernatants were added to 0.25 mg of prewashed and concentrated Dynabeads Oligo-(dT)₂₅ (Dynal) and incubated for 20 minutes at 25°C. After magnetic concentration, the supernatant was removed and saved as probe, and the beads were washed three additional times with 100 μ l 2x SSC; all wash supernatants were added to the final probe. 10 μ l of the final probe (ca. 550 μ l) were removed for scintillation counting, and 500 ng of oligo(dA)₄₀ was added to the remainder, followed by five minutes incubation at room temperature. For the probes used on only one filter (hybridizations *LNCX* and *LNCY*, (table 2)), all components of the initial hexamer-priming reaction were reduced by half.

2.2.1.2 *Known marker probes*

For the neural crest-specific marker hybridizations (filter 39-1-211), probes were made by restricting AP2 and B220 (Furthauer, 1998; M. Clark, unpubl.) with both Not I and Sal I, and by restricting Snail2 (Thisse, 1995) with *Eco* RI. Fragments correlating to cDNA inserts were gel-purified and quantified, and approximately 40 ng of each was used as template in a T7 random-priming reaction with α -³²P dCTP.

For the hybridization of EF1- α to filter 39-1-211, EF1- α probe was prepared as described for use in the slot-blot analysis of the cDNA pools (see above).

MATERIALS AND METHODS

2.2.1.3 Oligonucleotide vector probe

To create a probe for the plasmid vector carrying all cDNA sequences, 100 picomoles (335 ng) of vector oligonucleotide (5'-GCACGCGTACG-3') was added to a 50 μ l Polynucleotide Kinase (PNK) reaction using 50 units T4 PNK (NEB) and 250 μ Ci γ -³²P-ATP. The entire reaction, after 30 minutes of incubation at 37°C, was added to the oligonucleotide hybridization buffer (described below).

2.2.2 Library characteristics

The library array used for complex hybridization analysis represented 42% of a late-somitogenesis zebrafish (*Danio rerio*) cDNA library created by Matt Clark of the Max Planck Institute for Molecular Genetics (unpubl.). The full library contains approximately 65,800 separate clones, which, after 'oligo-fingerprint' hybridization with more than 200 seven to 12-base nucleotide sequences, is thought to contain about 20,000 clusters, or groups of cDNAs which hybridize to similar sets of these oligonucleotides. The library was made from mRNA purified directly from late-somitogenesis stage (24 hour) zebrafish. Library cDNAs are directionally cloned into the pSport vector between *Not* I and *Sal* I restriction sites with the SP6 promoter controlling transcription of antisense RNA.

Bacterial clones of this library are robotically arrayed on 22 cm x 22 cm nylon filters. These are grown in place and lysed to deposit their cDNA at the given filter locus. The filter should have approximately 27,600 colonies printed in duplicate (~55,300 clones), arranged in a 48 x 48 grid of 5 x 5 blocks, each containing 24 bacterial clones, or 12 separate cDNAs in duplicate.

2.3 Hybridization protocols

2.3.1 Complex hybridizations and marker hybridizations

2.3.1.1.1 Pretreatment

MATERIALS AND METHODS

Filters were prewashed in 5x SSC, 0,5% SDS and 1mM EDTA for one hour at 50°C. Filters were then prehybridized in 50 mL hybridization buffer (1% BSA; 7% SDS; 0,5 M Na₂HPO₄; 1 mM EDTA) for four to twelve hours at 67°C.

2.3.1.1.2 Hybridization

For hybridization of the filters, the prehybridization buffer was discarded and replaced with the radiolabeled probe dissolved in 20 ml fresh hybridization buffer. As recommended by the filter provider, probe was diluted to approximately 1-2 million cpm/ml. Hybridization was carried out for 12-16 hours at 67°C in a rotating 10 cm-diameter glass hybridization flask (Oncor/Appligene).

2.3.1.1.3 Washing and exposure

After hybridization, the filters were washed two times 30 min. each in 2x SSC, 0.1% SDS, and then once in 0.2x SSC, 0.1% SDS at 67°C. Following wash steps, filters were exposed against a phosphor screen (Molecular Dynamics) for 7-12 hours and imaged at both 100 and 200 µm resolution on a Storm[®] phosphorimager (Molecular Dynamics).

2.3.2 Vector hybridizations

Oligonucleotides were hybridized to all arrayed library copies with the method of Hoheisel *et al.* (1994) and these hybridizations were done in parallel in one vessel. After 30 minutes prehybridization in 400 ml oligonucleotide hybridization buffer (4x SSC, 7.2% Sodium N-lauroylsarcosine), labeled probe was added to the container and hybridized to the filters for three hours at 4°C. Filters were washed once in fresh hybridization buffer, 30 minutes at 25°C, and exposed against phosphor screens from 45 minutes to 1.5 hours.

2.4 *Hybridization analysis and clone selection*

2.4.1 *Hybridization scoring*

Complex hybridizations were scored using Visual Grid software (MPI for Molecular Genetics/ Genome Pharmaceuticals Corp.) running on Windows 95™ or Windows NT™ operating systems. Program output was manipulated on an institute server running Digital UNIX® V3.2D-1 Worksystem Software (Rev. 41).

2.4.2 *'Electronic subtraction' parameters*

The following formulae describe the parameters used for selection of filter loci. "*ENC and ENC'*" refers to "any clone found positive in both hybridizations *ENC* and *ENC'*"; "*ENC or ENC'*" refers to "any clone found positive in either *ENC* or *ENC'*", etc.

2.4.2.1 *Early neural crest specific*

= (*ENC and ENC'*) not (*LNCA, LNCA', LNCB, LNCB', VNT, or VNT'*)

2.4.2.2 *Both neural crest specific*

= [(*ENC or ENC'*) and (*LNCA, LNCA', LNCB, or LNCB'*)] not *VNT or VNT'*

2.4.2.3 *Late neural crest specific*

= [(*LNCA and LNCA'*), (*LNCA and LNCB*), (*LNCA and LNCB'*), (*LNCA' and LNCB*), (*LNCA' and LNCB'*), or (*LNCB and LNCB'*)] not (*ENC, ENC', VNT, or VNT'*)

2.5 Screen of clones by in situ hybridization

2.5.1 Generation of labeled antisense-RNA probe

XL1 Blue bacteria were obtained as requested from the RZPD library service carrying cDNAs cloned into a Plasmid pSport (Gibco) vector between Not I and Sal I cloning sites, with the SP6 promoter controlling transcription of antisense RNA. Single bacterial colonies were picked from streaked cultures (on TB-agarose plates with 50 µg/ml ampicillin) and cultured 16 hours in 3 ml of 2x TY medium containing ampicillin (Sambrook *et al.*, 1989). Plasmids were purified in parallel using a QIAprep8 plasmid purification kit (QIAGEN) as per instructions, and were subsequently linearized with Sma I (NEB).

In vitro transcription was performed after Thisse and Thisse, 1998 and Thisse *et al.*, 1993. 500 ng linearized plasmid in restriction buffer (125ng/µl) were used as template in a 10 µl *in vitro* transcription reaction consisting of: 1x Transcription Buffer (Boehringer Mannheim); 1 mM each ATP, CTP, and GTP (Pharmacia); 0,65 mM UTP (Pharmacia); 0,35 mM DIG-UTP (Boehringer Mannheim); 20 units RNase Inhibitor (Boehringer Mannheim); and 10 units Sp6 RNA Polymerase (Boehringer Mannheim). Reaction Products were precipitated by adding 40 µl EtOH:4M LiCl (75:2.5) and resuspended in 10 µl H₂O.

2.5.2 In situ hybridization

2.5.2.1 Embryo preparation and hybridization

Hybridization, antibody binding, and color reaction are based on protocols of Strahle *et al.* (1993), Lanfear *et al.* (1993) and Thisse and Thisse (1998). Fish embryos held at 28.5°C were collected at approximately 24 hours of age and incubated in fixative (PBS; 4% paraformaldehyde) for four hours at room temperature. Following multiple washes in PTW (PBS; 0,2% Tween 20[®] (Sigma)) the fish were dehydrated in 100% Methanol and held at -30°C for several hours. Fish were then gradually rehydrated to PTW, rinsed in hybridization buffer (50% Formamide; 5X SCC; 0,1% Tween 20[®]; tRNA 500 µg/ml (Sigma); and heparin 50 µg/ml (Sigma)) and prehybridized for one hour at 68°C

MATERIALS AND METHODS

in 400 µl hybridization buffer. To hybridize, 1 µl RNA probe diluted in 100 µl hybridization buffer was added to the prehybridization mix and the fish were incubated for a further sixteen hours at 68°C.

2.5.2.2 Washing and detection

After hybridization, the fish were washed at 68°C for ten minutes each with hybridization buffer:2x SSC at ratios of 75:25, 50:50, 25:75, and then with 100% 2x SSC. Two thirty-minute, 68°C washes in 0.2x SSC then followed, and these again were followed by five-minute room-temperature washes with 2x SSC:PBT (PTW; 0,2% BSA) in ratios of of 75:25, 50:50, and 25:75. Following a final wash in 100% PBT, the fish were left at room temperature to block in MABT buffer (Maleic acid 100 mM; NaCl 150 mM; Tween 0.1%; pH 7,5) with 2% Blocking Reagent (Boehringer Mannheim). Antibody binding was achieved by addition of Anti-DIG-AP Fab fragments (Boehringer Mannheim), pre-incubated with fixed zebrafish embryos, to a final antibody dilution of 1:8000 and incubation overnight at 4°C.

Before staining, the fish were washed five times in PBT for lengths of five, ten, twenty, thirty, and sixty minutes. After two five-minute washes in staining buffer (NaCl 100 mM; MgCl₂ 50 mM; Tris-HCl 100 mM; Tween 0.1%; pH 9.5) fish were stained by the addition of Nitro Blue Tetrazolium (Sigma) and 5-Bromo 4-Chloro 3-Indolyl Phosphate (Sigma) at concentrations of 0.275 mg/ml and 0.125 mg/ml, respectively.

Staining was stopped at an appropriate time by rinsing several times with PTW, and fish were mounted for microscopy in 90% glycerol. Digital photography was done using a CCD camera (Sony) with Nomarski (DIC) optics at 10x magnification on a Zeiss Axiophot Microscope.

3. Results

RESULTS

3.1 *Neural crest RT-PCR*

3.1.1 Development and control testing of PCR amplification

All reproducible approaches to gene expression analysis of limited tissue amounts are highly dependent on the amplification technique they use. As an initial step in developing such a technique, a PCR reaction was extensively optimized for reagent, buffer conditions, and cycling parameters, such that it could reproducibly detect only a few molecules of a plasmid sequence adapted with specifically-designed primer annealing sites.

In order to make an amplifiable template of known sequence and quantity, annealed primer adapter sequences were ligated to the ends of a fragment of pBluescript DNA. As shown in lanes 3-6 of Figure 2, the adapted 465 bp fragment could be reproducibly amplified to approximately 50 ng from starting template amounts totaling only five attograms (theoretically equivalent to four double-stranded template molecules). This represents an approximately 10 billion-fold amplification, or, for a forty-round PCR, an efficiency per cycle of approximately 80% ($k= 0.83$, where starting template is amplified by 2^{40k}).

As a further test of the sensitivity and specificity of the optimized PCR, a similarly-adapted 522 bp plasmid fragment was amplified from within a complex pool of zebrafish cDNA. This was done for cases in which the complex pool was not adapted for PCR amplification (see Figure 2, lanes 7-11) and in cases where the cDNA pool was adapted with the same primer binding sites and was concurrently amplified (top panel, Figure 3). In the latter case, Southern blot hybridizations were carried out to detect the amplified fragment within the PCR amplicon (Figure 3, lower panel). A digoxigenin-labeled probe made from this pBluescript fragment hybridized to a Southern-blotted PCR reaction containing 750 pg adapted zebrafish cDNA and as little as 400 ag target fragment. As the detection limit of the DIG system is stated by the

RESULTS

manufacturer to represent approximately 0.3 pg target DNA, and as about 0.3 μg of PCR product was loaded onto the gel (1 million-fold more than the detected limit), this result indicates that the abundance profile of the initial template is very roughly preserved.

RESULTS

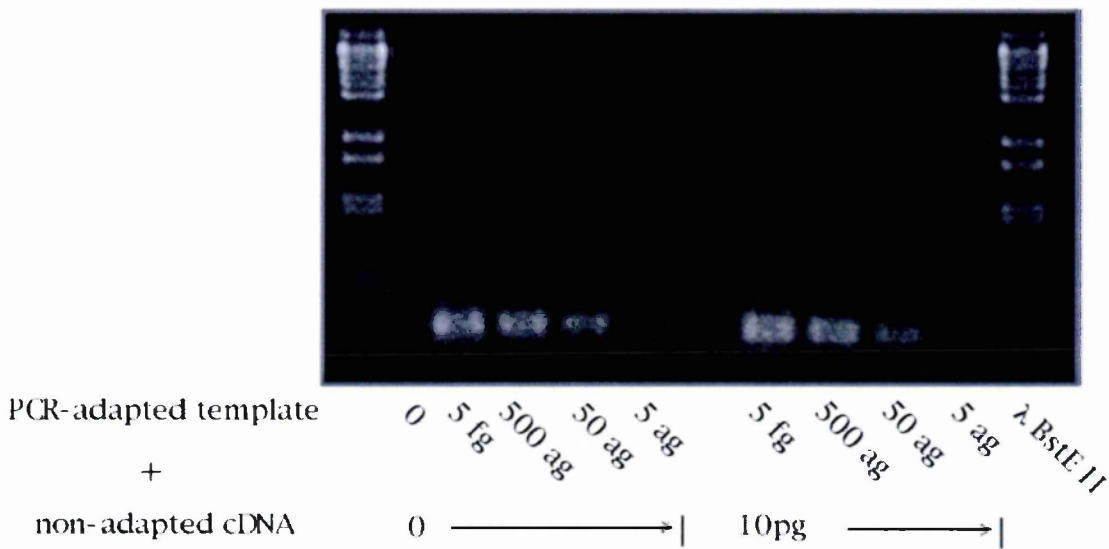


Figure 2. PCR amplifies adapted template present at low copy numbers, and yields no unspecific product. A 465 bp fragment of plasmid DNA, adapted with specific primer sequences, is amplified by a forty-round PCR. Lanes 3-6 are 10-fold serial dilutions of the target template, lanes 7-11 are 10-fold dilutions of target template plus non-adapted zebrafish cDNA. Lanes 2 and 6 are negative controls to check PCR contamination, and lanes 1 and 11 are molecular weight ladder (λ BstE II digest, 0.5 μ g).

RESULTS

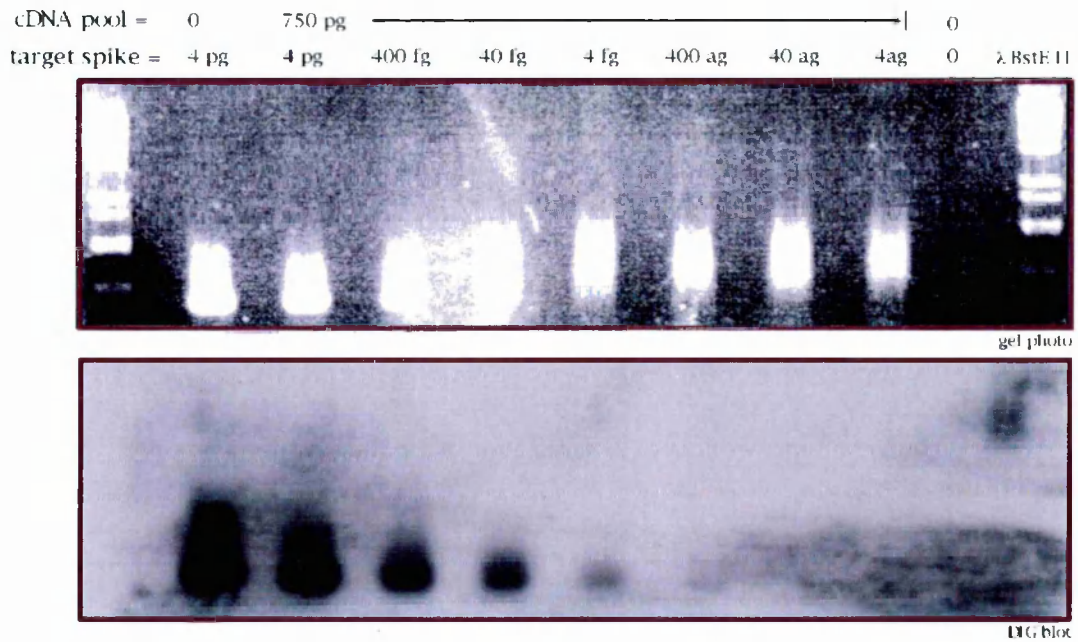


Figure 3. DIG-labeled pBluescript probe hybridizes adapted target template amplified within a complex pool of adapted zebrafish cDNA. Top panel shows gel electrophoresis of amplified product: lane 2, 4 pg target template alone; lanes 3-9, 10 fold dilution series of target template plus 750 pg adapted cDNA; lane 10, negative control with no template; lanes 1 and 11, molecular weight ladder (λ BstE II digest, 0.5 μ g). Bottom panel shows Southern blot hybridization of the same samples, chemiluminescently detected.

RESULTS

3.1.2 Global RT-PCR amplification of neural tube and neural crest cell-specific cDNAs

3.1.2.1 Tissue collection and cDNA synthesis

After experimentation to choose an appropriate method for efficient and reproducible cDNA synthesis (see discussion for details), a method based on traditional second-strand replacement (Okayama and Berg, 1982; Gubler and Hoffman, 1983) modified for PCR amplification (Esguerra, unpubl.; Figure 4) was applied to ventral neural tube tissue and to cultured neural crest cells. In order to harvest neural crest cells potentially corresponding to the early and late-migrating neural crest sub-populations, distinct cell cultures were performed with trunk neural tube from similar anteroposterior location but different developmental stages.

To isolate putative early-migrating neural crest ('ENC'), anterior-trunk neural tubes from eight-somite embryos were cultured for six to seven hours in serum-free medium (Jesuthasan, 1996; unpubl.). Examples of such a culture can be seen in Figure 5. When cells were fully separated from neural tube, they were aspirated from culture and seeded into cDNA synthesis reactions. Two such cDNA syntheses were performed on early neural crest, each seeded with approximately ten cells. To isolate putative late-migrating neural crest ('LNC'), neural tubes from the same axial position were cultured from 20-somite embryos. Three syntheses were carried out from late neural crest, each also starting with approximately ten cells. Three cDNA syntheses were also performed on small pieces of ventral neural tube ('VNT'), microdissected from the anterior trunk of an eight-somite embryo. A summary of these cDNA syntheses is presented in Table 1.

RESULTS



Figure 4. The anatomy of a double-stranded cDNA molecule. Primer 1 (green), with an 18-base 3' stretch of poly(dT), primes the synthesis of first-strand antisense cDNA (blue) from the 3' poly(dA) tail of the template mRNA. After limited RNase digestion, second-strand sense cDNA (red) is primed by remaining RNA fragments. After blunt-ending the double stranded cDNA, Primer 2, annealed to its complement (orange), is ligated to both ends. cDNA is amplified by PCR with Primer 2, while Primer 1 confers directionality and is used to generate single-stranded antisense cDNA for a radiolabeling procedure.

RESULTS

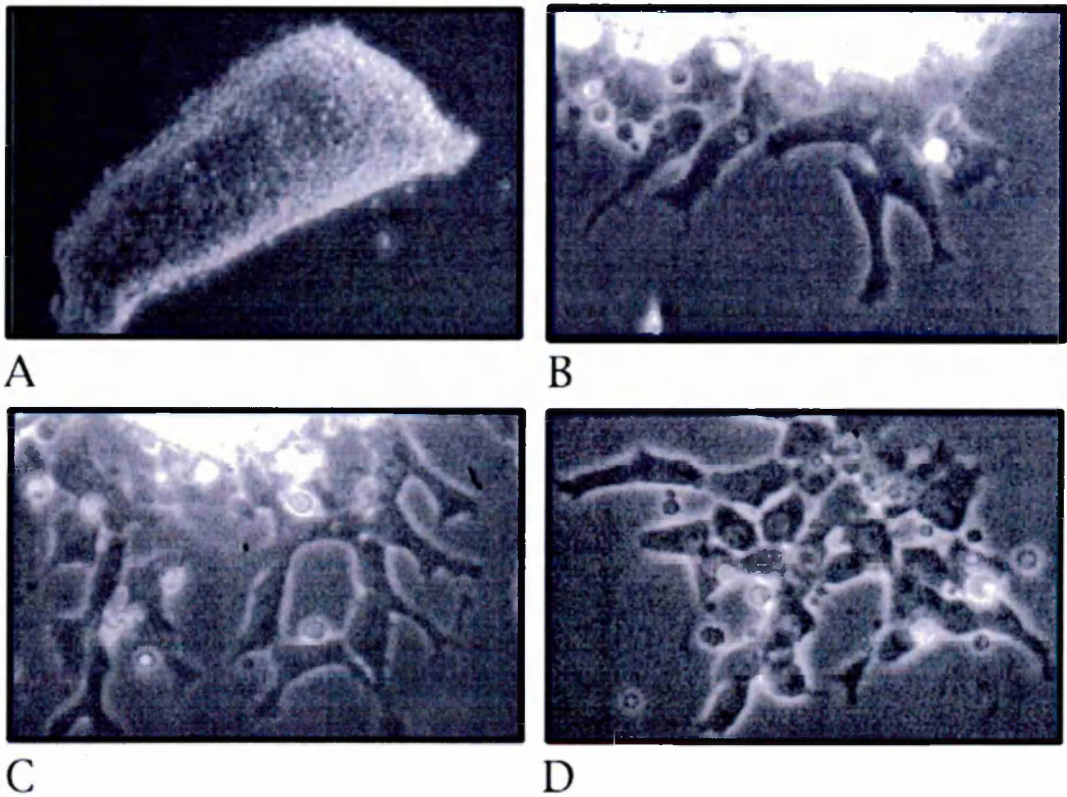


Figure 5. Neural crest cells migrate in culture. A piece of neural tube, dissected from the trunk of an eight-somite embryo (A) is shown after 5 hours (B), 7 hours (C), and at eight hours, after washing away the neural tube tissue in preparation for aspiration (D).

RESULTS

3.1.2.2 *Size distribution and banding patterns of amplified pools*

Portions of the cDNA synthesized from neural crest and neural tube tissues were amplified using an extended version of the optimized, forty-round PCR protocol. Following amplification, amplified cDNA pools were analyzed by agarose gel electrophoresis. Selected examples of these cDNA pools are shown in Figure 6. In panel A amplifications from the three ventral neural tube syntheses are shown. The first two of these syntheses, VNT1 and VNT2, were performed using RNase inhibitors. These RNase inhibitors appear to have an inhibitory effect on the synthesis process. Several amplifications from one cDNA synthesis of early neural crest cells are shown in panel B and amplifications of one cDNA synthesis from late neural crest cells are shown in panel C. In the late neural crest cDNA amplicons faint banding patterns can be seen. When these amplicons were run out on an electrophoretic gel with a higher percentage of agarose, the results visible in panel D were obtained. The variable banding indicates that individual PCR amplifications of different aliquots of cDNA from one cDNA synthesis reaction can produce inconsistent results. Such banding was not seen, however, for VNT and ENC pools.

RESULTS

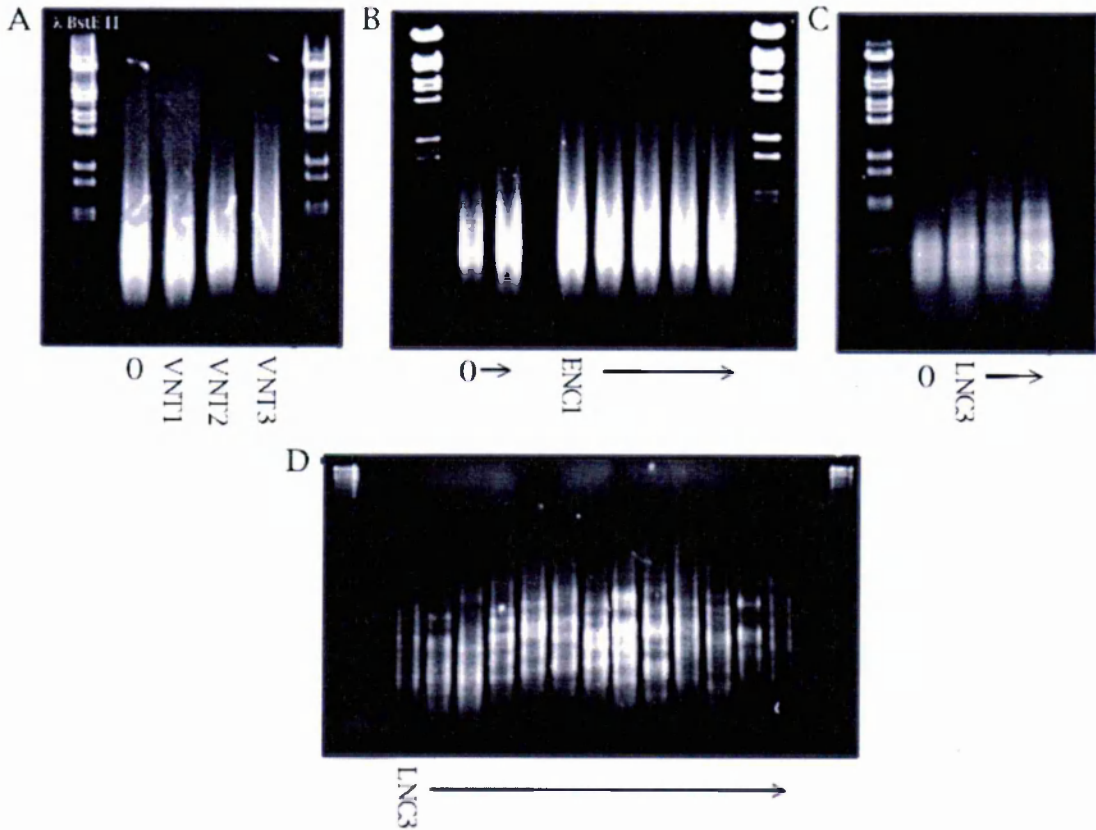


Figure 6. Representative gel-electrophoretic profiles of amplified cDNA from minimal tissues. Amplicon products from amplifications of ventral neural tube (A), early neural crest (B), and late neural crest (C) are shown run on 1% TAE gels. Negative controls cDNA syntheses with no cells (lanes marked '0') yield DNA product. LNC pools exhibit banding; this was checked on a 2% low-melting point TAE gel (D). Banding appears to vary between PCRs.

RESULTS

3.1.2.3 *Slot-blot testing of cDNA pools with β -actin and EF1- α*

One phenomenon common to the three sets of cDNA syntheses performed, as well as to similar protocols for single-cell RT-PCR amplification (Dulac, 1998), is that the negative controls made without cells produced amplifiable cDNA. Although the upper range of amplified DNA lengths were shorter in the negative control pools than in the experimental pools used to make complex probes, control slot-blot hybridizations were carried out to ascertain that the amplified experimental DNA was zebrafish cDNA, and that DNA amplified from the negative control was not zebrafish DNA. ^{32}P -labeled random-primed probe from zebrafish EF1- α (Gao *et al.*, 1997) and from β -actin (Kelly *et al.*, 1997) was hybridized to membranes blotted with cDNA pools from all experimental ventral neural tube (VNT) syntheses, from both early neural crest (ENC) syntheses, from the negative control (mock) syntheses for each, and, as positive control, from a previously generated cDNA pool made from poly(dA)-purified mRNA. The results of these slot-blot hybridizations are presented in Figure 7. No signal was obtained from the blotted control pools, suggesting that the DNA obtained was non-zebrafish cDNA. Signal was obtained from all experimental pools, although weaker signals in experimental cDNA pools derived from fewer cells suggests that a portion of the DNA in these pools was also non-cDNA contaminant. It is possible, however, that the weak signal obtained reflects the expression state of these genes in the early neural crest cell population sampled.

RESULTS

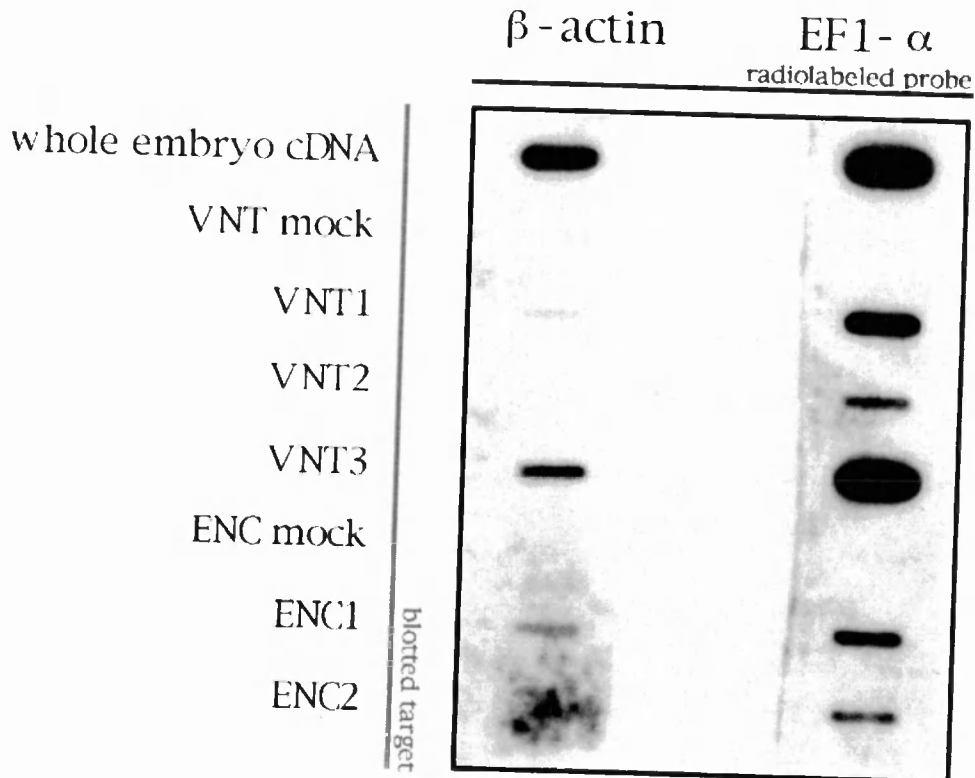


Figure 7. Slot Blots confirm amplification of zebrafish cDNA. Horizontals represent similar blotted cDNA pools; verticals represent hybridization with radiolabeled zebrafish β -actin (left) and EF1- α (right). Each slot is blotted with 545 ng cDNA. Blotted pools are: whole embryo cDNA (see text); VNT mock, 3 pooled amplifications of a no-tissue control synthesis for VNT probes; VNT1, VNT2, and VNT3, 3 pooled amplifications each of three separate cDNA syntheses from ventral neural tube; ENC mock, 3 pooled amplifications of a no-tissue control synthesis for ENC probes; ENC1 and ENC2, 10 pooled amplifications each of two separate cDNA syntheses from early neural crest cells. A faint β -actin signal in the VNT 2 pool is not visible in this printed reproduction.

RESULTS

3.2 *Complex hybridizations*

To counteract possible variability in PCR amplification, PCR amplicons were combined after many parallel amplifications (Table 2). For each pool of this amplified and combined cDNA from a given tissue type, single-stranded antisense cDNA was created by a linear amplification reaction using the internal primer sequence of the cDNA (Figure 4, green). This single-stranded template was used in a random-primed ³²P-labeling reaction. After removal of the template cDNA and measures designed to reduce the possibility of poly(dA)-poly(dT) background hybridization, the labeled probe was hybridized to a robotically-arrayed cDNA library colony filter putatively representing more than 26,000 separate zebrafish cDNAs. The filters used in this study were printed with a set of bacterial colonies representing approximately 40% of a 'late-somitogenesis cDNA library' made from 26-somite zebrafish embryos (Resource Center of the German Human Genome Project (RZPD); Matt Clark, unpubl.). Filters are given identification numbers by their creators (RZPD) and these numbers indicate the library printed, the set of clones printed from that library, and the date of printing. An exposure of a representative filter is shown in Figure 8, which shows a filter hybridized with a probe from early neural crest, as well as an enlarged detail in pseudocolor.

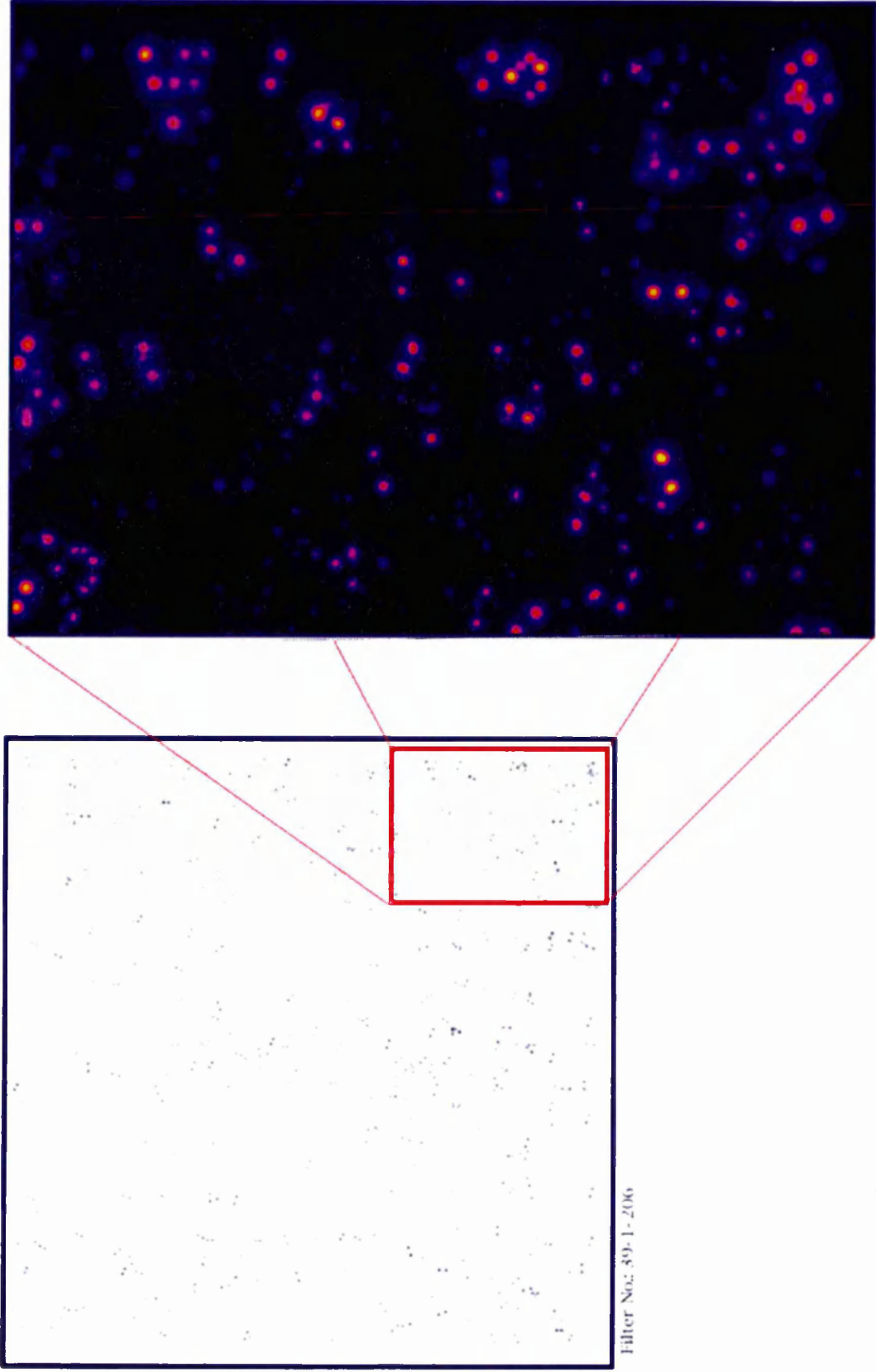


Figure 8. A typical arrayed library hybridization profile. Early Neural Crest probe hybridized to filter 39-1-206 (*ENC'*), with enlarged, pseudocolored view. Positively scored cDNA clones total 1221.

RESULTS

3.2.1 Hybridizations

3.2.1.1 *Ventral neural tube: two hybridizations*

Two filters, *VNT* and *VNT'*, were hybridized with ^{32}P -labeled probe made from ventral neural tube cDNA. Antisense template was made from pooled cDNA amplicons of three PCR reactions (Table 2). The final amplified and pooled cDNA used for probe synthesis should represent a tenth of the original cDNA synthesis reaction, or an estimated 10-100 cell equivalents.

3.2.1.2 *Early neural crest: two hybridizations*

Filters *ENC* and *ENC'* were hybridized with one probe made from early neural crest cDNA. Antisense template was made from pooled cDNA amplicons of 10 PCR reactions (Table 2). This probe represents the amplification of about 3.5 cell equivalents.

3.2.1.3 *Late neural crest: six hybridizations*

Filters *LNCA* and *LNCA'* were hybridized with one probe made from late neural crest cDNA. Antisense probe template was made from pooled cDNA amplicons of sixteen PCRs (Table 2). This probe represents approximately five cell equivalents.

As a control for PCR amplification and labeling procedures, 16 additional PCRs were performed on a set of original cDNA template sources identical to those used for the first LNC probe (Table 2). After linear amplification to make antisense cDNA, filters *LNCB* and *LNCB'* were hybridized with one probe made from this second batch of late neural crest cDNA.

As a control for the cDNA synthesis procedure, ten PCRs of LNC cDNA were amplified each from two different cDNA synthesis reactions, made from two different populations of late-migrating neural crest cells (Table 1; Table 2). cDNA amplicons from the same cDNA syntheses were pooled and radiolabeled probes were generated.

RESULTS

Each probe should represent approximately 3.5 cell equivalents. These probes were hybridized to filters *LNCX* and *LNCY*.

Table 1. cDNA syntheses performed, with information regarding the source and amount of tissue used. *Syntheses VNT1 and VNT2 were not used for further study, as RNase Inhibitor appeared to inhibit cDNA syntheses. **Ages given are at the time of neural tube isolation: cells were harvested 6-8 hours later from a culture at 28.5°C.

cDNA Syntheses	Tissue Used	Starting Tissue Amounts	RNase Inhibitor Used?
VNT1	Trunk ventral neural tube from 8-somite embryo	100-300 cells each	yes*
VNT 2	"	"	yes*
VNT3	"	"	no
ENC 1	Cultured neural crest cells from 8-somite embryo**	6-10 cells each	no
ENC2	"	"	no
LNC1	Cultured neural crest cells from 20-somite embryo**	8-12 cells each	no
LNC2	"	"	no
LNC3	"	"	no

Table 2. Probes used for arrayed library hybridization, with information regarding cDNA source, extent of PCR pooling, and the amount of mRNA represented by the total starting cDNA template. *These filters were printed with a different set of clones from the late-somitogenesis zebrafish cDNA library than the others used in this study; hence these hybridizations were used solely for purposes of experimental control.

Probe	Number PCR Amplifications Pooled	cDNA Template Source	Starting mRNA represented	Filters Hybridized
VNT	3	VNT3	10 to 30 cell equivalents	39-2-209; 39-2-211
ENC	10	ENCI	approx. 3.5 cell equivalents	39-1-206; 39-2-210
LNCA	16	LNC1; LNC2; LNC3 (3; 3; and 10 PCRs)	approx. 5 cell equivalents	39-1-207; 39-1-208
LNCB	16	LNC1; LNC2; LNC3 (3; 3; and 10 PCRs)	"	39-1-209; 39-2-210
LNCX	10	LNC1	approx. 3.5 cell equivalents	39-2-224*
LNCY	10	LNC2	"	39-2-225*

RESULTS

3.2.2 Filter scoring procedure and positive clones scored

Custom-designed software assisted the manual scoring of hybridization signals and assignment of subjective intensity values for each positive signal. A screen projection of the software used is presented in Figure 9. Briefly described, after exposing hybridized filters to phosphor screen, the software was used to layer an electronic gridwork over the phosphorimager-generated image file. This gridwork represents the pattern of 48 x 48 blocks in which sets of bacterial colonies are printed. 12 bacterial clones are printed in duplicate within each of these blocks, filling 24 nodes in a five-by-five array where the center position is blank. Duplicate positions are arranged as to make identification of any duplicate pair unambiguous; moreover, as only paired signals are scored, the scoring of background is reduced. The software allows these duplicate pairs, or block patterns, to be scored and given a subjective intensity value, ranging from one (weak) to three (strong). The software catalogues the scoring input as a list of x-y coordinates, 384-well plate freezer addresses, and intensity values. This is illustrated in Figure 10, where a raster presentation of x-y coordinates with intensity information creates a schematic of the scored filter, and where a list output of this data is presented as it is used in subsequent data processing. An institute server running the UNIX[®] operating system was then used to analyze the output lists. Freezer addresses (serving as unique names for every clone) were stripped from the original file and sorted. These were then counted, compared, and manipulated using UNIX commands for list processing, word searching, and list comparison.

Scoring yielded on the order of 1000 positive signals per filter. The exact number of positives scored per filter is given in Table 3. A first indication that filters did not return entirely reproducible results was that filters with identification numbers beginning with '39-2' (printed on a different day than those labeled '39-1') tended to return fewer positively-scored clones.

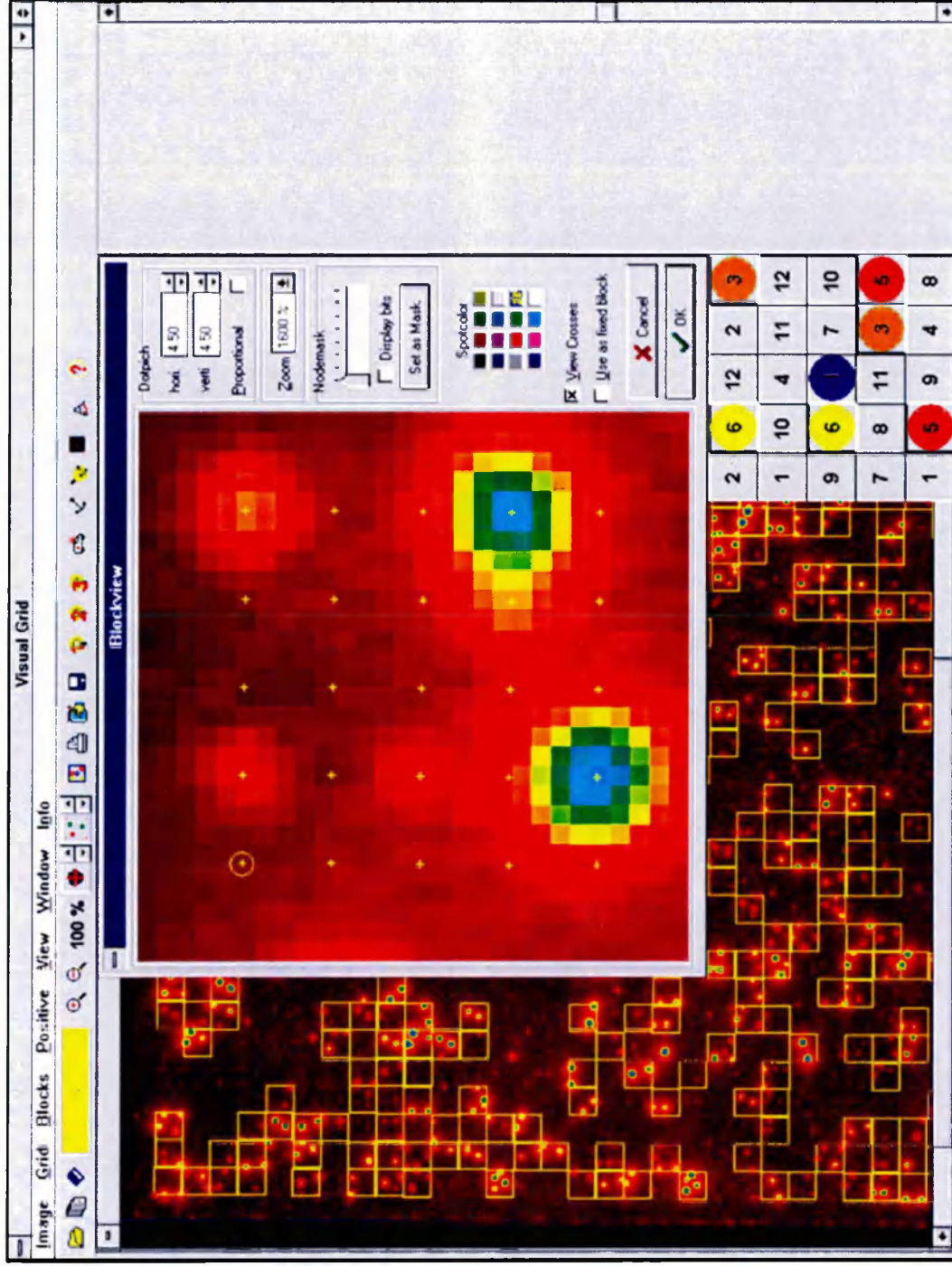
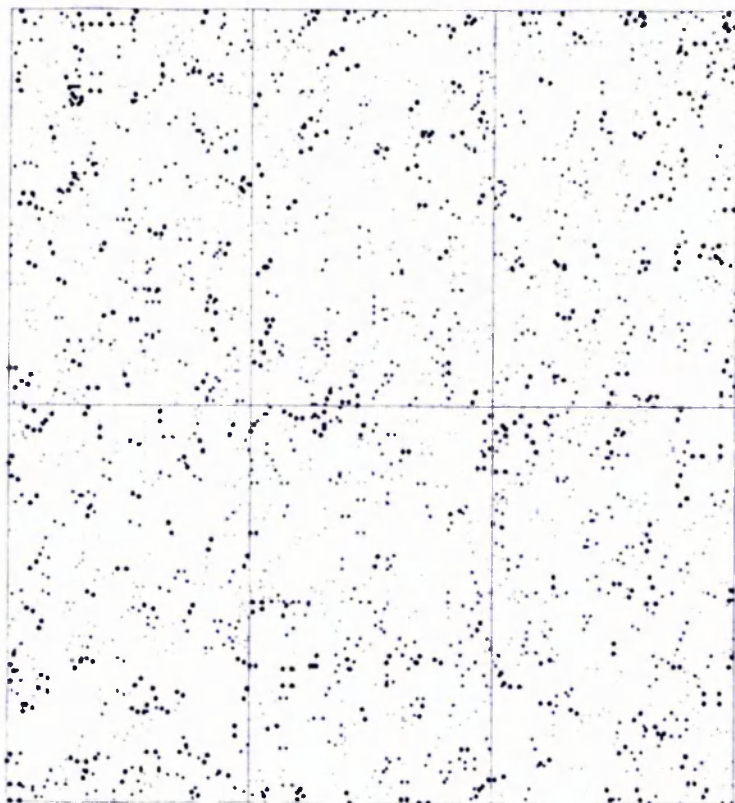


Figure 9. Custom software assists in the manual scoring of hybridization profiles. A screen image of the analysis software depicts the scoring process. A grid is overlaid on the phosphorimage of the filter being scored, and scored blocks (yellow) are cataloged by selecting the positive clone patterns on the button-field at bottom right. The 'Blockview' window shows an enlarged block with 24 printed cDNA nodes, representing 12 bacterial clones printed in duplicate (the center node is empty in all blocks). The block selected was scored positive for three clones: pattern 6 (intensity 1), pattern 3 (intensity 2), and pattern 5 (intensity three).



```
Session: Name:gen.mpib-tuebat  
roessel@gen: ~/current >more  
2 12 58N24 3  
3 13 58N24 3  
4 47 58G24 3  
5 48 58C24 3  
6 62 64D24 1  
7 64 64D24 1  
8 67 58C24 2  
9 68 58C24 2  
10 78 52A24 1  
11 80 52A24 1  
12 86 7024 2  
13 86 7024 2  
14 86 7024 2  
15 97 55M24 1  
16 98 55M24 1  
17 97 61M24 1  
18 99 61M24 1  
19 137 55F24 2  
20 133 55F24 2  
21 156 31A24 3  
22 158 31A24 3  
23 191 71J24 3  
24 192 71J24 3  
25 201 71H24 1  
26 202 71H24 1  
27 213 53F24 3  
28 215 53F24 3  
29 19 28M23 2  
30 20 28M23 2  
31 21 16L23 1  
32 24 16L23 1  
33 26 16K23 1  
34 29 16K23 1
```

Figure 10. The output of the Visual Grid software can be represented by x-y coordinate values and intensity scores. Left, A Visual Grid output raster image of scored clones on filter 39-1-206 (Early Neural Crest probe). Right, output data representing a hybridized filter with x and y coordinates, 96-well plate freezer address, and intensity score.

RESULTS

3.3 Measures of reproducibility: the similarity of similar hybridizations

3.3.1 Similarity of filters hybridized in duplicate

Hybridizations were carried out in duplicate in order to assess the reproducibility of hybridization and filter-scoring methods. For ventral neural tube, early neural crest, and two late neural crest duplicate hybridizations, labeled probes were created and split between two hybridization flasks with theoretically identical filters. Filters were scored independently and the results were compared to assess their similarity. Circular diagrams depicting overlap between these duplicate sets of scored clones can be seen in Figure 11. They present the full *unique* set of clones scored and show the proportion of this set present in each individual filter and the proportion present in both. In addition to these diagrams, a 'similarity' value (expressed as a percentage) is presented for each filter pair. This value reflects the proportion of clones that are present in duplicate in the *total* (not unique) set of scored clones.

3.3.2 Similarity of filters hybridized with probes made from parallel PCRs

A similar comparison was then carried out between filters hybridized with different probes which were both amplified from the same pools of cDNA. It was intended that these comparisons assess the reproducibility of the PCR and probe-labeling procedures. These comparisons made were made for each of the four possible A-B pairs of *LNCA* and *LNCB* hybridizations, and are presented in Figure 12.

RESULTS

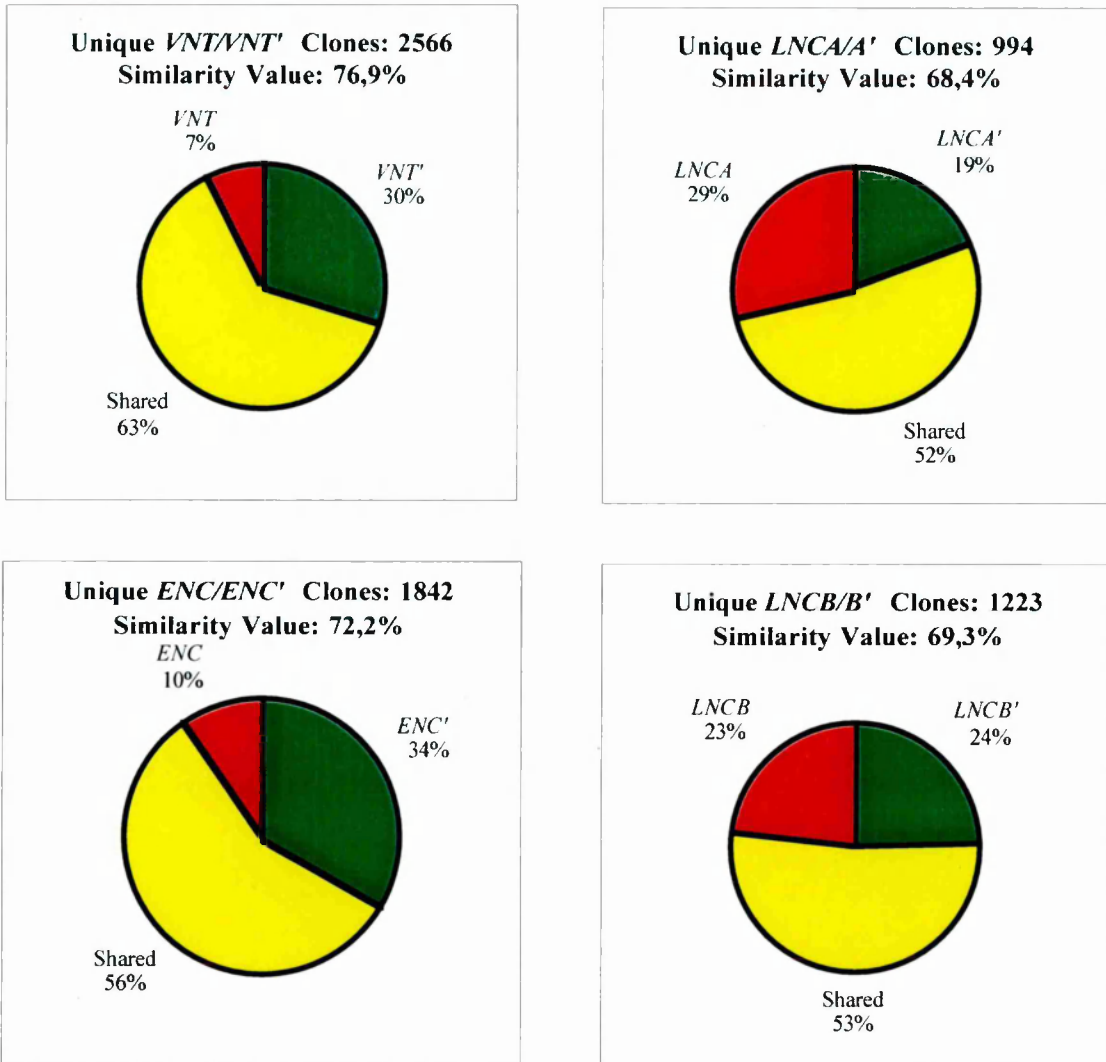


Figure 11. Overlap and similarity between filters hybridized with the same probe. Pie charts depict the *unique* set of clones scored in each filter pair, and show overlap between scored filters. Red and green wedges represent clones present in one filter. Yellow wedge represents clones present in both filters. Similarity Value is the percentage of *total* clones present in duplicate, and reflects the percentage of clones on one filter present in the other.

RESULTS

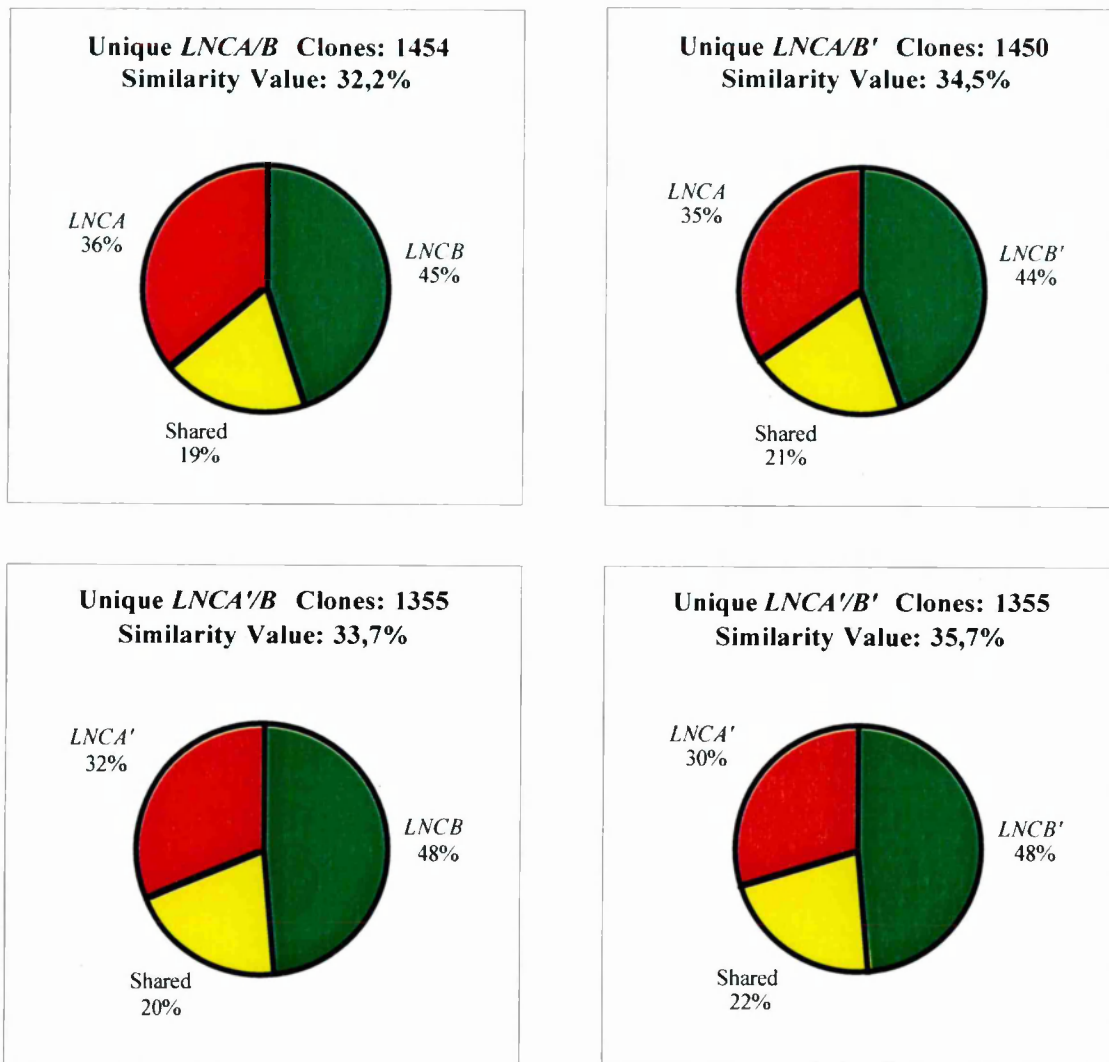


Figure 12. Overlap and similarity between filters hybridized with probe made from the same cDNA template, yet different PCRs and labeling reactions. Pie charts and Similarity Values as in Figure 6.

RESULTS

Table 3. Probe, filter number, and positives scored for all complex hybridizations. '39-2' filters have less printed cDNA, and score fewer clones.

Hybridization	Filter	Clones scored
<i>ENC</i>	39-1-206	1221
<i>ENC'</i>	39-2-210	1661
<i>VNT</i>	39-2-209	1791
<i>VNT'</i>	39-2-211	2377
<i>LNCA</i>	39-1-207	807
<i>LNCA'</i>	39-1-208	704
<i>LNCB</i>	39-1-209	926
<i>LNCB'</i>	39-1-210	945
<i>LNCX</i>	39-2-224	1470
<i>LNCY</i>	39-2-225	1336

RESULTS

3.3.3 Similarity of filters hybridized with probes from parallel cDNA syntheses

In an attempt to test the reproducibility of probe making and hybridization a further step back in the procedure, set overlap and a similarity value were determined for two additional Late Neural Crest hybridizations (*LNCX* and *LNCY*). The two probes used to hybridize these filters were amplified from different cDNA pools made from similar cells. This comparison is illustrated in Figure 13. The filters used for this hybridization, however, represent a different portion of the original late-somitogenesis cDNA library, and are arrayed with a different set of clones. The results of these hybridizations were thus not used for any further functional analysis. Indicating that PCR and cDNA synthesis are not the sole sources of experimental error, the similarity between these two hybridizations is in fact greater than those of filter pairs hybridized with probes from the same cDNA but different PCRs. A tendency is observed, however, suggesting similarity correlates with the number of clones scored.

3.3.4 Overall similarity in the late neural crest hybridizations

As the similarity values returned in these comparisons were much lower than had been expected, the four Late Neural Crest (with *LNCA* and *LNCB* probes) hybridization clone sets were further electronically analyzed to determine the frequency with which scored clones are positive in one, two, three, or four out of four hybridizations (Figure 14). This data suggests that random background alone (false positives) cannot account for the error between duplicate filters. Hybridization results must to some degree incompletely represent the set of sequences present in a given probe (see discussion).

RESULTS

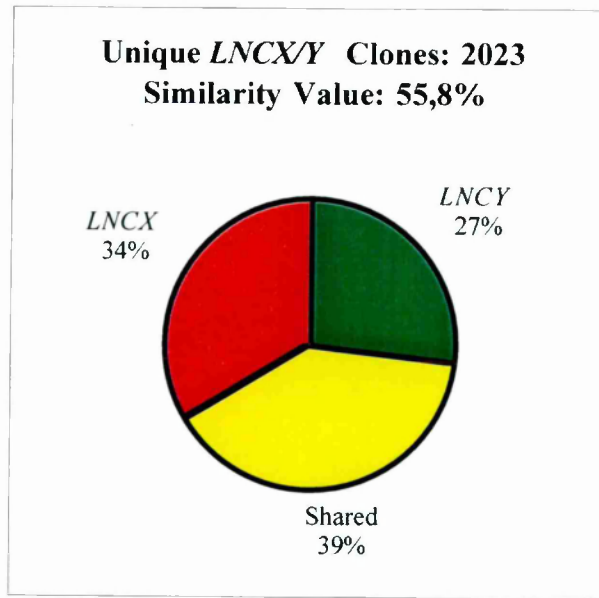


Figure 13. Overlap and similarity between filters hybridized with probes made from the same cell type, yet different cDNA syntheses, PCRs, and labeling reactions. Pie chart format and Similarity Value as in Figure 6.

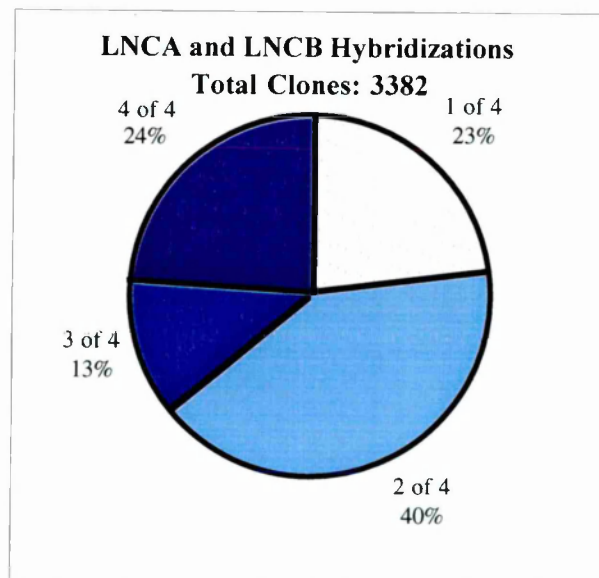


Figure 14. The set of *total* clones scored in the four LNCA and LNCB hybridizations. Wedges show the percentage of clones positive in one, two, three, and four of four hybridizations.

RESULTS

3.4 *Marker hybridization controls*

For further use as experimental controls, known markers were hybridized to the gridded cDNA libraries: an ubiquitous marker was used to check the probes for consistency with slot-blot hybridization results; specific markers allow the probes to be checked for neural crest specificity; and a vector-specific oligonucleotide was used to assess the distribution of cDNA on the filters.

3.4.1 *Ubiquitous markers*

The arrayed library was probed with labeled cDNA of zebrafish EF1- α , a ubiquitously and highly expressed gene, to determine corresponding filter locations. The EF1- α hybridization is shown in Figure 15. Additionally in this figure, an enlarged detail of the EF1- α hybridization is compared with an analogous detail from an early neural crest hybridization. Surprisingly, these filter loci were scored positively faintly and sporadically, if at all, in the complex probe hybridizations. The EF1- α marker hybridization was not electronically scored, but was confirmed by eye to accord with EF1- α hybridizations performed and scored previously by Matthew Clark (pers. comm.). This lack of signal from abundant genes assumed to be present in the original cDNA pools (as indicated in the slot blots described previously, Sec. 3.1.2.3) was also seen when the complex hybridizations were assessed for scoring of β -actin, cytochrome c, and other abundant genes of known position in the arrayed library.

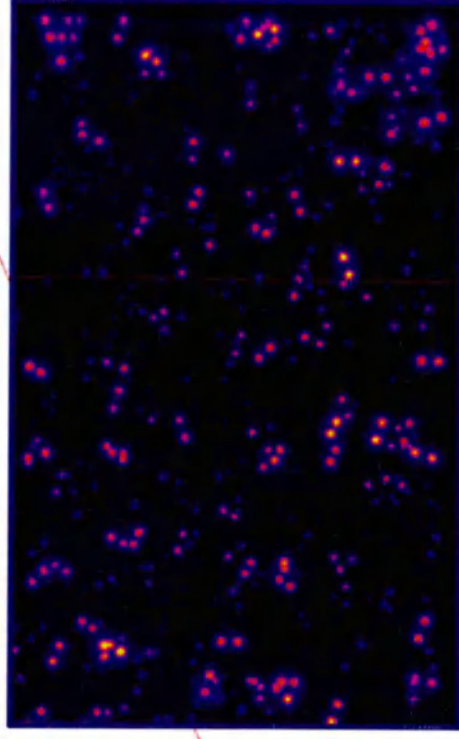
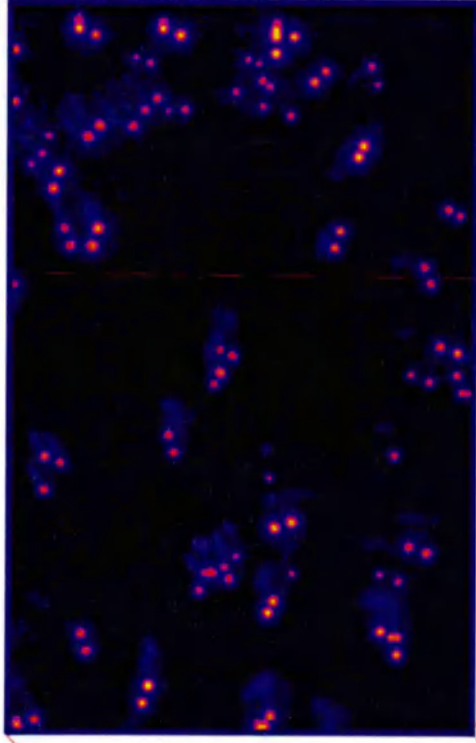
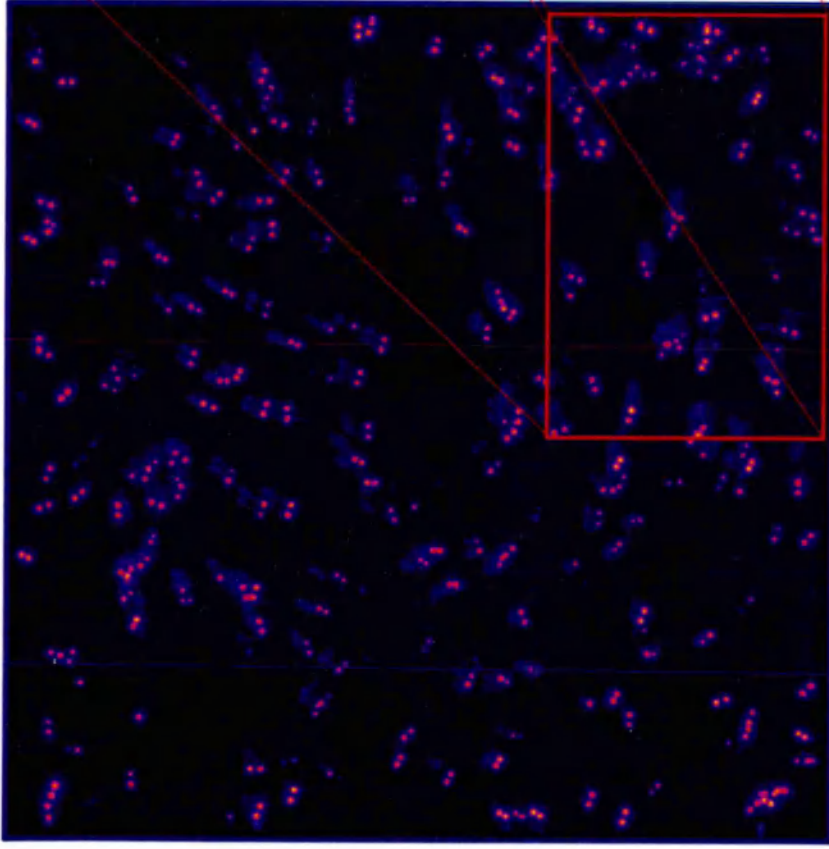


Figure 15. EF1 α probe hybridizes library cDNAs encoding EF1 α . These clones however, do not appear to give a hybridization signal on filters hybridized with neural crest of ventral neural tube probes. EF1 α is shown by slot blot to be present in neural crest and neural tube cDNA pools; some filtering of highly-abundant clones may occur during radiolabeled probe creation process (see discussion). Left—EF1 α probe hybridized to filter 39-1-211. Upper right—enlarged detail of filter 39-1-211. Lower right—analogue detail of filter 39-1-206, hybridized with probe from Early Neural Crest.

RESULTS

3.4.2 *Neural crest markers*

To find the filter locations of cDNAs encoding mRNAs known to be expressed differentially in trunk neural crest, filter 39-1-211 was hybridized with radioabeled probe generated using fragments of *snail2* (Thisse *et al.*, 1995), *AP2* (Furthauer *et al.*, 1998), and *B220* (M. Clark, C. Thisse, and B. Thisse, unpubl.) cDNAs as template. These cDNAs all encode transcription factors known to be expressed in the developing neural crest of zebrafish. The marker hybridization using *snail2*, *AP2*, and *B220* yielded nine scored positive signals, although distinction between very weak 'positives' and background hybridization was somewhat difficult. One of these nine had been previously identified as *B220* (Matt Clark, pers. comm.); the other eight have not been identified.

To check the validity of the complex hybridizations, library clones that hybridized with the markers were checked to see if they scored positive after hybridization with complex probes. When filters hybridized with neural tube or neural crest complex probes were assessed for the scoring of these marker positions, the pattern presented in Table 4 was obtained. All but two of the marker positions were scored in at least one neural crest filter, yet many markers (especially those with weak signals in the marker hybridization) scored positive in only one filter, rendering them non-selected by the parameters used for choosing neural crest specific clones (see below). Two others were not selected by the final selection parameters because they scored positively in the filters hybridized with probes from ventral neural tube.

RESULTS

Hybridization	VNT	VNT'	ENC	ENC'	LNCA	LNCA'	LNCB	LNCB'
N.C. marker	(39-2-209)	(39-2-211)	(39-1-206)	(39-2-210)	(39-1-207)	(39-1-208)	(39-1-209)	(39-1-210)
Position 1	-	-	2	3	-	1	-	-
Position 2	1	-	2	2	-	-	-	-
Position 3	-	-	-	-	-	-	-	-
Position 4	-	-	-	-	-	-	-	1
Position 5	-	-	-	-	-	-	1	-
Position 6	-	-	-	-	1	-	-	-
Position 7	-	-	-	-	-	-	-	1
Position 8	1	1	3	3	2	2	-	-
B220	-	-	-	-	-	-	-	-

Table 4. Known-marker clone positions score positively primarily in neural crest filters. cDNA clones putatively encoding AP2, Snail2, and B220— all transcription factors expressed in migrating neural crest— score positive in filters hybridized with neural crest specific probes. 1 (yellow), 2 (orange), and 3 (red) refer to weak, medium, and strong hybridization signal intensities respectively. Only Position 1 is pulled out by screening parameters as 'neural crest specific'. Others are either not reproducibly scored (one or fewer signals), or scored positive in ventral neural tube hybridizations.

RESULTS

3.4.3 Vector oligonucleotide hybridizations

Inconsistencies in the distribution of cDNA on the arrayed filters presents a significant possible hindrance to differential hybridization analysis. To assess this distribution, an 11-base radiolabeled oligonucleotide complementary to a sequence of the plasmid used as a cDNA vector was hybridized to all filters in a single parallel hybridization. The results of this hybridization for one typical filter (39-1-206) are shown in Figure 16. Though the technical goal in library gridding is to create reproducible filters with consistent cDNA amounts and distribution, inconsistencies in the amount of cDNA plasmid per spot are clearly visible. Also in Figure 16, enlarged details of this filter and three others are compared. Presenting a more serious hindrance the goals of this project, the inconsistencies between printed spots appear themselves inconsistent between filters.

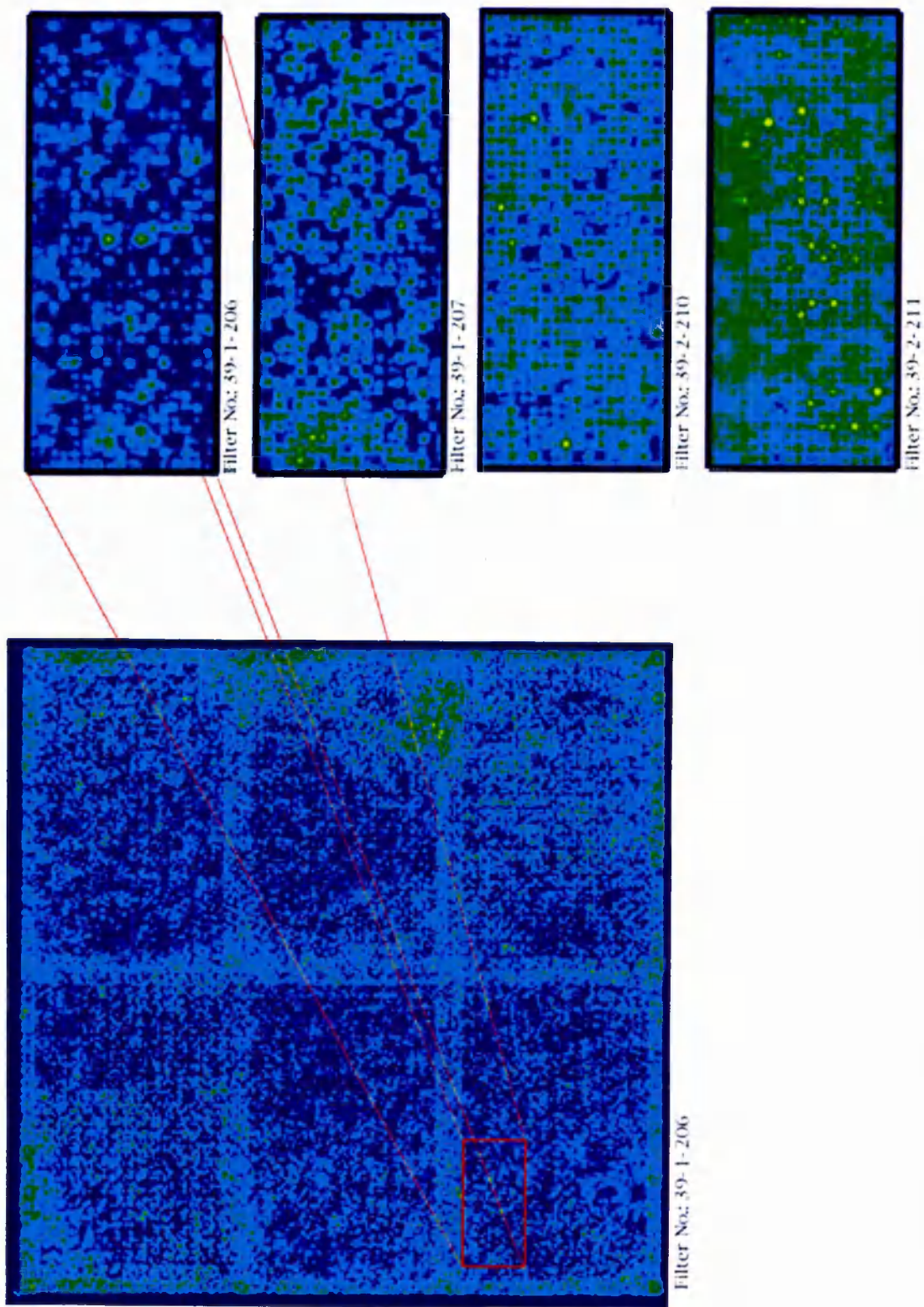


Figure 16. Radiolabeled oligonucleotide hybridizes plasmid vector of cDNAs. Left— filter 39-1-206, hybridized by a vector-specific oligo, shows the distribution of cDNAs on the filter. Right— comparison of filter 39-1-206 detail and three analogous arrayed filter details shows the variation in cDNA distribution on 'duplicate' filters.

RESULTS

3.5 Clone selection and screening

By electronically manipulating data sets, cDNA filter loci were identified that met defined criteria for selection as neural-crest specific. The selection parameters were designed to reduce false positives at the expense of set size. Maintaining the biochemical subtraction terminology of a 'tracer' pool from which 'driver' sequences are subtracted, tracer clones had to be positive in at least two hybridizations to be considered, while driver clones present in even only one hybridization would be removed from the tracer set. These 'electronic subtractions' were as follows:

3.5.1.1.1 Early neural crest

To isolate cDNAs encoding mRNAs specific to the early neural crest but not present in late neural crest or ventral neural tube, a 'tracer' pool was defined as the set of all clones scored positive in both filters hybridized with early neural crest probe. Subtracted from this tracer pool was the 'driver' set of all clones scored positive in one or more of the filters hybridized with probes from late neural crest or ventral neural tube. 390 clones were selected.

3.5.1.1.2 Late neural crest

To isolate cDNAs encoding mRNAs present in the late neural crest but not present in early neural crest or ventral neural tube, a 'tracer' pool was defined as the set of all clones scored positively in two or more of the four filters hybridized with late neural crest probe. Subtracted from this tracer set was the 'driver' set of all clones scored positively in one or more of the filters hybridized with probes from early neural crest or ventral neural tube. 422 clones were selected.

3.5.1.1.3 Both neural crest

To isolate cDNAs encoding mRNAs present in both early and late neural crest but not present in ventral neural tube, a 'tracer' pool was defined as the set of all clones

RESULTS

scored positively in either or both of the filters hybridized with early neural crest probe and in one or more of the filters hybridized with late neural crest probe. Subtracted from this tracer pool was the 'driver' set of all clones scored positively in one or both of the filters hybridized with probes from ventral neural tube. 292 clones were selected.

RESULTS

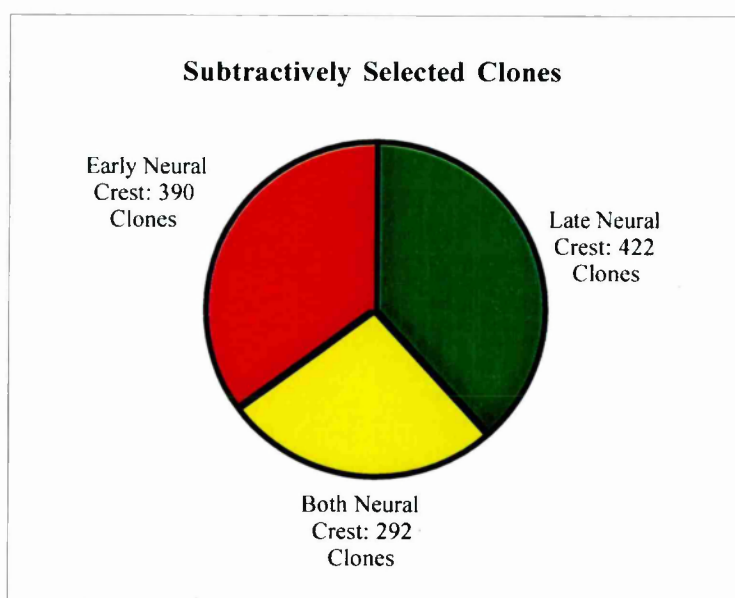


Figure 17. The set of clones selected as neural-crest specific. Chart shows overlap between clones scored for ENC and LNC. Red represents clones positive in ENC hybridizations only, green represents those positive in LNC hybridizations only, and yellow represents those positive in both.

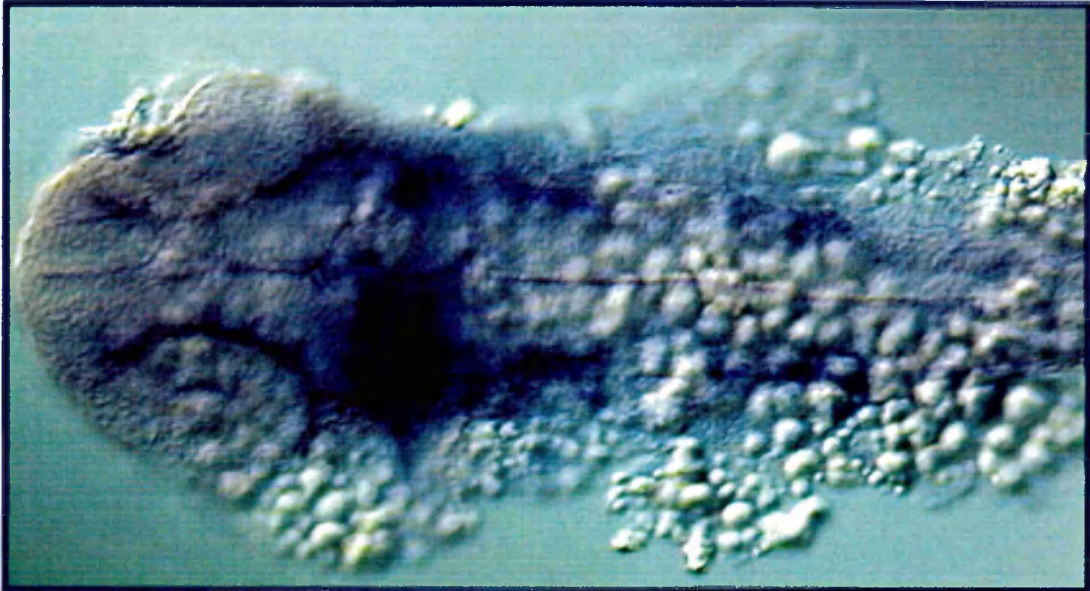
RESULTS

3.6 *In situ screening of selected clones*

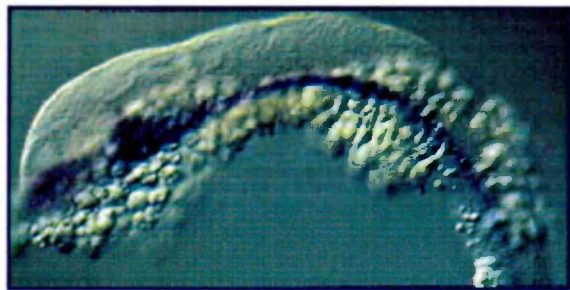
In situ screening of the selected clones has been unsuccessful. Although more than fifty separate cDNAs were screened in this manner, the same pattern of non-specific hybridization was observed for all antisense RNA probes. It is suspected that a sequence of plasmid multiple cloning site, present between the transcription initiation point and the start of cDNA coding sequence, is responsible for this signal-obscuring background.

To control for technical aspects of the *in situ* hybridization protocol, positive control *in situs* were carried out by hybridizing krox 20- and sonic hedgehog- encoding antisense RNA to 24 hour zebrafish embryos in the same manner used for hybridization of the unknown sequences selected in this screen. Both probes returned the expected signals: staining in rhombomeres 3 and 5 for krox 20 (Lanfear al., 1991; data not shown); and staining of the notochord and prechordal plate for sonic hedgehog (Krauss *et al.*, 1993; Figure 18). Additional diagnostic controls performed included *in situ* hybridizations without labeled RNA (data not shown), after which fish returned no color signal— indicating that the antibody binding and color reaction do not lead to the background observed— and *in situ* hybridization with B220, a known neural crest specific cDNA isolated from this library and cloned into the same plasmid vector. No neural crest-specific signal was obtained from hybridizations with B220, giving additional indication that the *in situ* pattern obtained reflects an experimental artifact and not the true expression pattern of the transcripts selected.

RESULTS



in situ: ESC-specific pool cDNA



in situ: sonic hedgehog cDNA

Figure 18. In situ hybridization indicates non-specific hybridization between aRNA probes and day-old zebrafish embryos. Above: *in situ* hybridization of an aRNA probe made from a cDNA clone putatively encoding an early migrating neural crest specific mRNA. Similar patterns were obtained from all library cDNA clones. Lower right: *in situ* positive control hybridization. Sonic Hedgehog aRNA probe produces the expected staining of notochord and prechordal plate.

4. **Discussion**

DISCUSSION

4.1 ***RT-PCR sensitivity***

As stated previously, a primary concern for any protocol involving the synthesis and comparison of amplified cDNAs is that the amplification procedure used be reproducible. As estimated in Alberts *et al.* (1994), a typical mammalian cell's mRNA population contains very many transcripts of a very few genes and very few transcripts of very many genes. At the low end of this abundance profile are approximately 12,000 genes transcribed at less than ten copies per cell. If, as the model suggests, the typical cell contains mRNAs encoding approximately 15,000 genes, then more than two-thirds of the transcriptome's (expressed genome's) complexity might come from genes present at less than one copy in 100,000. While these numbers are necessarily coarse estimates, due to a lack of appropriate technology for measurement, they effectively illustrate the requirement for sensitivity faced by any amplification procedure used. An amplification procedure missing the requisite sensitivity to detect rare mRNAs might return extremely variable results upon repeated experiments.

Such variability of PCR result has been described before by Karrer *et al.* (1994), where it is called the 'Monte Carlo effect' and presented as an 'inherent limitation' of single cell-based RT-PCR protocols. Perhaps this phenomenon, however, characterized by large variation in the amount of PCR product amplified from small amounts of starting template, should not be considered an inherent phenomenon, but simply symptomatic of an insufficiently-sensitive PCR paradigm. The authors do, however, draw conclusions about ways in which this effect might be overcome— by increasing, for example, the amount of cDNA template available for PCR amplification by initial linear RNA amplification rounds (eg. Eberwine *et al.*, 1992).

Inefficient PCR amplification should affect experimental outcome for rare molecules in the initial stages of amplification. A PCR efficiency of 0.5 (as previously defined, Sec. 3.1.1), for example, could easily leave a single molecule unamplified over two rounds ($p= 0.25$), whereas a pool of thousands of molecules should quite reproducibly

DISCUSSION

double in size. In this respect initial linear amplification of cDNA— creating larger starting pools for rare molecules before exponential PCR amplification— should reduce the variability of PCR results. Pooling multiple PCR products, as was done in this study, should also alleviate the effects of such error. The same consideration emphasizes the need for a highly efficient cDNA synthesis protocol. cDNA synthesis, generally performed once without any thermocycling, has few chances to occur. If this fails for a molecule present only at a few copies in a pool, this could again cause experimental outcomes which disproportionately represent rare mRNA species.

4.2 ***cDNA synthesis techniques***

Many variants of a few basic protocols were attempted before choosing a method for the synthesis of cDNA from limiting amounts of mRNA. While similar in the techniques used to generate first-strand cDNA, the various protocols offer different methods of priming the second-strand cDNA. The three basic paradigms of protocols attempted and rejected included random priming of the second strand (Froussard, 1992) terminal-transferase tailing of the first strand with complementary homopolymer priming of the second strand (Tam *et al.*, 1989), and ligating or polymerizing primer sites at the 3' end of first strand cDNA, using annealed oligonucleotides as template or 'bridges' for ligation. These protocols were all investigated for their capability to produce directional cDNA for PCR amplification; they all exhibited limitations, however, making them unsuitable for our application without a great degree of further experimentation and optimization.

DISCUSSION

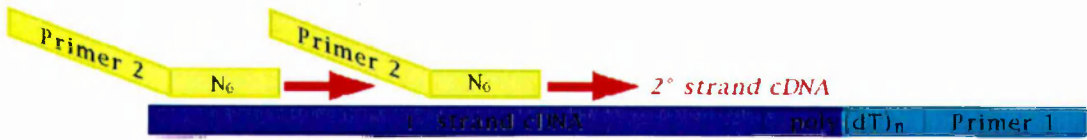


Figure 19. Second-strand synthesis by random hexamer priming. Primer 2, with random 3' hexamers (yellow), recognizes complementary sites in the first-strand cDNA (blue) and primes second-strand cDNA (red).



Figure 20. Second-strand cDNA synthesis by homopolymer tailing. Terminal Transferase is used to add a sequence of a single nucleotide (orange) to the end of the first-strand cDNA (blue). A primer with a complementary 3' sequence (yellow) is used to prime the second strand (red).



Figure 21. Second-strand priming using an oligonucleotide 'bridge'. A known sequence (a homopolymer tail, or the three G's added by MMLV reverse transcriptases, for example; orange), is recognized by a complementary 3' sequence of the oligo bridge (yellow) allowing the ligation of complementary Primer 2 (red). This leaves the first-strand cDNA (blue) ready for PCR amplification using Primer 1 (green) and Primer 2. cPrimer 2 can also be synthesized in place by DNA polymerase.

DISCUSSION

4.2.1 **Random priming**

Initial attempts at priming synthesis of the second strand by a specific oligonucleotide coupled to a 3' N₆ randomer (figure 19) were generally unsuccessful. Theoretical considerations also suggest the unsuitability of this approach when dealing with particularly low amounts of mRNA. The annealing of primer with template is a second-order reaction, dependent on the concentrations of both nucleotide sequences. Using randomers as primer reduces the effective concentration of complementary priming sequence for any given template sequence. When the concentration of one component is already only barely sufficient, the overall reaction is severely hindered, requiring a much higher concentration of random primer in the reaction mix than is feasible or economical. An advantage of random priming (though outweighed by this far more relevant disadvantage) is that it produces cDNAs whose length is independent of gene sequence or mRNA length. This is an ideal characteristic for cDNA that will be amplified when the preservation of abundance profile is desired. As PCR favors short fragments that polymerize quickly, if cDNA size correlates with transcript size, short transcripts will be relatively enriched in the amplified cDNA pool.

4.2.2 **Homopolymer tailing**

Homopolymer tailing with terminal transferase, a procedure by which a chain of a single nucleotide (generally A or T) is polymerized at the 3' end of the first strand, was the cDNA synthesis method adapted for use by Catherine Dulac (1995). Following such tailing, a second-strand cDNA can be synthesized using a primer coupled to a 3' stretch of the complementary nucleotide (figure 20). For Dulac's application, this method of priming the second strand was combined with a 'limited' first-strand cDNA synthesis such that the resultant cDNAs would be of a uniformly short length, alleviating the loss of accuracy in abundance representation potentially conferred by PCR. Terminal transferase tailing, however, has inherent limitations when used to create directional cDNAs, as has been our goal. Tailing with dATP and priming the

DISCUSSION

second strand with a specific primer-poly(dT)_n, as Dulac does, runs the risk of creating non-directional contaminant by mispriming events associated with hairpinning during the second-strand synthesis and subsequent PCR. Tailing with dTTP again allows the possibility of mispriming, and tailing with either dGTP or dCTP runs the risk of creating high-melting temperature cDNA stretches that would lead to hybridization artifacts. Despite these drawbacks, many separate protocols for priming a second strand of DNA on a control template of DNA were developed in our search for a reliable and efficient second-strand priming method.

The best of these homopolymer-tailing protocols was abandoned, however, when finally tried on first-strand cDNA from a real synthesis: the homopolymer tailing of excess first-strand primer, present at a several million-fold excess to real first-strand cDNA, creates an optimal ~100-bp target for second-strand synthesis and subsequent PCR amplification. In practice, the amplification of these non-cDNA encoding primer pairs completely flooded out the amplification of directional cDNA. While attempts to develop a successful block to these 'empty' primer pairs were undertaken, they were quickly given up with the realization that any such block would, in principle, have to be 99,99999% effective simply to reduce a ten million-fold excess of empty primer to cDNA to a 1:1 ratio. It is not completely understood how this fails to be a more significant problem for both Belyavsky *et al.*'s and Dulac's approach to tailed-cDNA priming. Rajewski *et al.* use a precipitation step, gel electrophoresis-based size selection, a second precipitation, an additional column-based size selection, and a second electrophoretic size selection to overcome this difficulty, yet in doing so they create ample opportunity for the rare cDNA product to be lost. It is worth noting that Dulac tails her first-strand cDNA with poly(dA), and amplifies the resulting cDNA with only one primer sequence, non-directionally. It is possible that short tailed-and-primed primer sequences created in her approach, being palindromic, would form stem-loop structures and inhibit their own amplification in PCR. This explanation has been given in other contexts as well to explain the non-amplification of small sequences with potentially hairpin-forming structures (Vos *et al.*, 1995).

DISCUSSION

4.2.3 **Oligonucleotide bridging**

The final of these three unsuccessful paradigms— foregoing with the second-strand synthesis all together, ligating on or synthesizing the complement to a second strand primer at the 3' terminus of the first strand and proceeding directly to PCR— was attempted in two sets of experiments. The initial attempts utilized terminal-transferase tailing of the first strand to provide a target for the annealing of primer 'bridges' that would then allow ligation or polymerization of the primer's complement onto the first strand cDNA. This approach was eventually ruled out for the same reason that other terminal transferase approaches were abandoned. The second set of approaches utilized the existing terminal sequence of the first strand for the same purpose. The use of a bridge of nine Inosine residues (allowing promiscuous base-pairing) was attempted under several conditions yet failed. Also attempted was a PCR-based full-length synthesis kit (SMART[®], Clontech), which uses a primer ending with three guanosine residues to capitalize on the tendency of MMLV-derived reverse transcriptases to add three deoxy-cytosine residues to the end of first-strand cDNAs. While this functioned marginally for our purposes (the kit is not designed to deal with extremely low amounts of mRNA, and its primers did not function in our optimized PCR), the high cost, the unavailability of the SMART[®] oligonucleotide bridge for separate purchase, the legal barriers to modification of the technology, and the lack of information regarding the nature of this oligonucleotide— whose proprietary modifications are not available for public knowledge— lead us to reject further experimentation with the kit. An attempt was made to attach three guanosine residues to the 3' end of one of our optimized primer sequences for similar use as an oligo bridge, yet this attempt failed.

4.2.4 **Second-strand replacement**

The cDNA synthesis protocol finally chosen for use is a variant of the standard RNA replacement technique first described by Okayama and Berg (1992) and Gubler and Hoffman (1983) and as modified for PCR amplification by C. Esguerra (unpubl.). In this technique, after priming the first strand using an poly(dT)-based primer and

DISCUSSION

priming the second strand using RNA fragments remaining after limited RNase treatment, the double stranded cDNA is blunted and then adapted at both ends by ligation with a second primer. While the PCR amplification that follows uses the same primer for both strands, directionality is conferred on the final product by the presence of the internal primer sequence used to prime the first strand. The final cDNA structure obtained is diagrammed in figure 4. No effort was made to reduce the size of first strand synthesis by limiting reagent concentrations (e.g., Dulac, 1995), as when this was attempted with other cDNA synthesis protocols, the general result was a loss of reaction product— not shorter DNA, but simply less DNA.

Concerns about the efficiency of this protocol— which compromises enzyme function by serial addition of new buffers at each step, and relies on the inefficient process of blunt ended ligation— were relegated to future experimentation in view of the more pressing need to perform complex hybridizations with the available filters. Indeed such hybridizations present perhaps the most effective way to judge the success or failure of cDNA synthesis and amplification techniques. The goal of highly sensitive and reproducible cDNA amplification might be best served by refining parallel hybridization techniques and bringing them to bear on the development of an optimal protocol.

4.3 *Contamination and amplification of non-cDNA*

One characteristic of the RT-PCR procedure used is that negative controls performed with no tissue yield smears of amplified DNA. To assess the nature of the DNA amplified in the control and experimental samples, Southern blot hybridizations (slot blots) were performed to check for the presence of zebrafish EF1- α and β -actin. In addition to indicating that the DNA amplified from the probe synthesis reactions was zebrafish cDNA, the experiments suggest that the DNA amplified from the mock synthesis reactions does not appear to be contaminating zebrafish cDNA, but rather DNA of unknown origin. Such contamination was also reported by Catherine Dulac

DISCUSSION

(1998), who hypothesized that the appearance of DNA in her negative controls was due to amplification of contaminating bacterial DNA present in the enzymes and other reagents.

This contaminant probably is not limited to the negative control, as is indicated by the relatively low proportion of EF1- α or β -actin encoding sequence in the amplified cDNA prepared from extremely small amounts of tissue (relative to cDNA pools made from larger amounts of tissue (fig. 7)). Of concern to this project was the chance that the contaminating DNA might return an artifactual hybridization signal from the gridded library, and thus confound hybridization analysis. This possibility cannot be entirely discounted, as no mock probe was hybridized to an arrayed library. There are reasons to discount this possibility, however, primarily as the contaminating DNA does not seem to be the product of reverse transcription. Bacterial RNA has no 3' poly(dA) tail, and thus should not be primed by our first strand poly(dT)₁₈ primer (although internal priming of bacterial poly(dA) stretches is possible); furthermore, the addition of RNase Inhibitor, which seems to inhibit Reverse Transcriptase function (compare cDNA length-distributions of VNT2 and VNT3 syntheses, for example) increases the abundance of contaminant relative to fish cDNA, suggesting that this inhibition of enzyme function does not affect the production of contaminating template as severely as it does the synthesis of zebrafish cDNA.

It seems most probable that the contaminant is produced by the blunting of contaminating DNA sequence after second-strand synthesis and the ligation of primer adapters to these DNA ends. In this case, the contaminant should not contain the first-strand primer used to linearly amplify template, and should not produce single-stranded template for the random-primed ³²P-labeling reaction. Any double stranded contaminating DNA being radiolabeled should anneal to itself rather than to the gridded library and thus should not contribute significantly to artifactual background (see further discussion on the abundance filtering effect).

DISCUSSION

4.4 *Hybridization analysis*

Scoring of the hybridization results was done manually with the assistance of specially developed computer software operating on a Windows NT system. Cataloging of the results was made immeasurably easier by the use of such software, yet the amount of effort needed to score a single filter (more than six hours for a filter with 1500 positives) still causes the scoring process to be the rate-limiting step in the overall approach described here. While improved automatic scoring procedures are being developed, those currently available returned an unacceptably high rate of error in their identification of positive clones. The subjectivity in the scoring of intensity—spots were assessed by eye and given a value of one (low) to three (high)—combined with the high variation in cDNA amount per spot (discussed below), rules out any truly quantitative analysis of the expression levels of a given gene. While some filters gave generally low numbers of positive scores, this problem could not be addressed simply by lengthening exposure times and choosing more positives. A drawback of the high density of clones on the filters used is that longer exposures caused strong positives to obscure neighboring signals.

After computer-assisted scoring, data is stored as lists of x-y coordinates, including signal intensity and freezer address data. While unfortunately no software currently exists specifically for the manipulation of such lists, the UNIX[®] operating system is particularly amenable to this sort of data processing. By stripping output files to a single word for each clone (identifying each by its freezer address), system commands could be used to identify clones specific to given sets of files. UNIX commands function well for this purpose, yet software designed for electronic subtraction would be of tremendous assistance. It is still an enormous advantage of electronic subtraction over traditional biochemical subtraction, however, that selection parameters can be chosen and re-chosen based upon output results. The additional data analysis performed (defining and determining similarity between hybridization profiles,

DISCUSSION

searching for known markers in the output files, etc.) was also made possible by the flexibility of data processing afforded by the computer.

An additional benefit of electronically analyzing and processing the hybridization data is that each archived data set increases in value with every additional experiment done on the gridded reference library used. Further hybridizations with complex probes give an ever more-comprehensive set of clones for use as driver in further subtractions and generate an increasingly comprehensive gene expression atlas containing expression data for each clone in the library. Of particular value is all effort undertaken to characterize the clones in the library itself by sequence and cluster analysis. Oligo fingerprinting with more than 200 short (7 to 12-base) oligonucleotides has been performed on the library used by Matthew Clark, identifying 'clusters' of clones containing potentially similar sequences (as in Hoheisel *et al.*, 1994). Supplementing a current project to re-array single clones from approximately 25,000 clusters, from this library in combination with a second, gastrulation-stage cDNA library, additional research is currently aimed at generating expressed sequence tags (ESTs) for approximately 15,000 of the unique clones (M. Clark, pers. comm.). With this additional data in place, it will be possible for the results of a subtractive screen to be immediately assessed for sequence information, for overlap with the results of other hybridization probes, and even for *in situ* hybridization data on cDNAs of the set.

4.5 Reproducibility of hybridization

This approach to expression analysis has great promise, yet many practical difficulties remain to be overcome. Primary among these is the low reproducibility of probe-labeling, hybridization, and scoring procedures. As described here, hybridizations of the same probe to duplicate copies of the same arrayed library yielded significantly disparate scored results. These results were again less similar when probes made from different amplifications of the same cDNA were used— yet a better value was returned by duplicate hybridizations performed with probes derived from

DISCUSSION

different amplifications of different cDNA. These values indicate that the current technology is less than optimal, yet no immediate conclusions should be drawn about the reproducibility of the PCR amplification or cDNA synthesis. In the case of the late neural crest probes, 10 to 16 PCRs were pooled before creating labeled probe. It therefore seems unlikely that the discrepancies in results obtained could have been due solely to differences in the success of the PCR amplification step. Rather, it is likely that error in the printing of filters, as well as artifacts introduced during the radioactive labeling procedure, both described below, are primarily responsible.

4.5.1 Filter reproducibility

Though not completely accounting for the low similarity between probes produced with different PCRs, a major source of dissimilarity comes from differences in the filters hybridized. Variability introduced here, in combination with any background introduced from hybridization or scoring, must be responsible for the lack of similarity between hybridizations of different library arrays with the same probe. Figure 16 shows identical details of four filters hybridized with a labeled oligonucleotide complementary to a sequence of the cDNA plasmid vector. This 'vector oligo' should label each plasmid-containing spot on each filter hybridized. As can be inferred from this figure, the degree of variability between filters that *should* be identical replicas is one of the greatest current limitations on the approach.

The variation in cDNA printing can have many possible sources, but all relate most probably to the handling and printing of the bacterial colonies carrying the cDNA plasmids. Variation in duplicate filters may come from well-to-well contamination in the 384-well plates used for bacterial growth and storage, from residual bacterial contamination on the pins used to pick and print colonies on to the filters, from the failure of pins to make good contact with the filters during printing, from unequally distributed pressure of the printing pin-block on the filters, from poor bacterial lysis after incubation, and, most problematically, from differing filter incubation times and

DISCUSSION

varying growth rates of bacterial clones. Relying on colony filters, on which bacteria are printed, grown, and lysed, rather than on filters printed directly with controlled amounts of DNA (from PCR, for example) renders the comparison of relative gene expression levels exceedingly difficult. Thus any attempt to quantify differences in gene expression in terms more subtle than simple black-and-white analysis must rely on the extremely labor-intensive quantification of DNA per filter spot. As automatic scoring procedures do not yet give reproducible and accurate results with printed filters such as those used in this report, the possibility of hybridizing a vector-specific oligo to the filter and then quantifying the result is not feasible.

4.5.2 **Probe-labeling reproducibility**

As seen in figure 15, clones that score positive when hybridized with radiolabeled EF1- α sequence generally do not score positive when hybridized with probe derived from amplified cell-specific cDNA. It can also be seen, however, that in slot-blot hybridizations, the same EF1- α probe hybridizes strongly to the amplified tissue-specific cDNA pools. As library clones known to encode β -actin and cytochrome c also appear to return little or no signal from the cell-specific probes, we have assumed the presence of an artifact, incurred during the labeled-probe creation process, that leads to the selective loss of signal from genes abundant in the original pool of cDNA being labeled. The possible existence of such a 'low-pass' abundance-filtering process, akin to biochemical normalization (Soares *et al.* 1994) , is also evidenced by oligo-fingerprinting cluster analysis of the final hybridization results. When cluster sizes (the number of copies of a select cDNA in the printed library) are determined for clones selected by our parameters, they generally are no larger than one or, rarely, two. This suggests that the genes scored positive in our hybridization results are themselves of low abundance in the original cDNA population used to create the library.

DISCUSSION

A possible mechanism by which this artifact is created is the process of linear amplification used to create single-stranded template for random-primed ^{32}P labeling. This reaction, based on the optimized PCR protocol developed for cDNA amplification, characteristically gives a lower efficiency of amplification per cycle over forty rounds ($k \approx 0.60$, where starting product is amplified by $40k$) than does the optimized PCR ($k \approx 0.83$, as previously defined (Sec. 3.1.1)) though it uses the same buffer, enzyme, primer sequences and annealing temperature. It is a possibility that this inefficiency results from the relatively high concentration of starting template at the start of the amplification reaction, and that the amplification of abundant genes is inhibited relative to that of less-abundant genes due to the re-annealing of complementary stretches of DNA sequence during primer-annealing and extension phases of the reaction cycle. Such re-annealing would inhibit Taq polymerase function and would occur as a function of gene abundance, affecting predominantly abundant genes for which the concentration of complementary sequence in solution is relatively high.

Although such an inhibition of linear amplification should only reduce the proportion of abundant genes relative to rare ones in the DNA pool used for labeling, the elimination of abundant-gene signal in our hybridization may be due to secondary consequences of this inefficient reaction step. cDNA template remaining double-stranded and non-biotinylated going into the labeling reaction (such as that failing to be linearly amplified with the biotinylated, antisense primer) would not be removed by the streptavidinylated beads used to purify the final reaction products. Single-stranded, antisense probe created in this reaction would thus be capable of annealing to the strand of cDNA used as template in its creation, as well as to probe resulting from priming of its template's complementary strand, all of which would remain in the probe mixture following streptavidin purification. As the final probe mixture was not denatured before its addition to the hybridization flask (indeed it was incubated to encourage annealing of oligo(dA)₄₀ to possible oligo(dT)_n stretches), it remains entirely possible that most probe created from abundant template would have annealed to complementary

DISCUSSION

sequences in solution, and returned no signal from complementary cDNA sequences affixed to the gridded library filter.

The degree to which this artifact would affect the observed measures of reproducibility is unclear. If the artifact *is* incurred as hypothesized, because all linear amplification reactions were carried out with similar reaction conditions and similar starting concentrations of template, one would expect the reaction products to be similar in composition, and therefore a minimal effect of this abundance-filtering process on duplicate-reproducibility scores. An additional factor that might affect the outcome of such a process, however, would be the original abundance distribution of the different cDNA species in the different pools amplified. This was potentially variable between probes, especially considering the banding evident in the LNC probes, which probably indicates a smaller number of more abundant cDNAs in the amplified pool.

While this 'normalization' or 'low-pass' abundancy-filtering effect was unexpected and unplanned, the result could perhaps be a desired and useful one. The application of a similar process is in fact mentioned by Nguyen *et al.* (1995). Cutting the signal intensity of abundant clones might allow the scoring of rarer cDNAs that would have been missed in a non-filtered probe. With a sub-optimal RT-PCR protocol, however, this filtering should decrease the overall similarity between hybridization results from different pools. Furthermore, abundant cDNAs (with correspondingly high cluster sizes in the arrayed library) should hybridize to a large number of colonies. The presence of such abundant cDNAs in two probes should significantly increase the similarity values measured between their hybridizations; the absence of such clones in this experiment could explain a great deal of the dissimilarity obtained.

Additional experiments should be done, however, to test and characterize this effect: one should investigate the intensity of hybridization signal relative to mRNA abundance for a given set of known and well-characterized genes. Theoretically, if one were to make probes from 24- hour whole embryos, one could relate hybridization intensity to cluster size for a set of cDNAs in the library, assuming that cluster size reflects

DISCUSSION

abundance in the original library. If done for several different linear amplification conditions, the characteristics of this 'filtering' process could be determined and such filtering could be conscientiously incurred. In the meanwhile, for the purpose of differential comparison ('electronic subtraction') it seems advisable that some internal control be carried out to determine the degree of 'filtering' before comparing disparate hybridization results.

4.5.3 Modeling the error incurred

Regardless of the stages in which variability in hybridization results is introduced, this error can be thought of as consisting either of 'false positives'— clones scored positively that should not have been— or 'false negatives'— clones not scored that should have been. Determining the nature of the error in these terms is a primary challenge when trying to assess the cause of inconsistencies between filter hybridizations. Models based on both of these sorts of error can make predictions about the results of parallel hybridizations.

A model of exclusively false positives can be tested on the four Late Neural Crest hybridizations. Such a model invoking error due to high random (non-overlapping) background would predict— assuming probes LNCA and LNCB were identical— that many clones of the set of four hybridizations would be positive on all four filters, and that many clones would be present only in one. This, empirically, is not the case (figure 14). Assuming the probes were not completely identical, due to irreproducible PCR or labeling, would predict that a large proportion of clones be present in one, two, or four hybridizations, and can not explain the large proportion of positive clones empirically seen in three of the four hybridizations. This model is also contradicted by the observation that a large proportion of clones positive in only *one* of the LNCA hybridizations is also positive in one or both of the LNCB hybridizations (data not shown).

DISCUSSION

A second model, that of false negatives, assumes that only a randomly incomplete subset of sequences in a given probe returns positive signals from the arrayed library—in other words, that not all of the library clones that *could be* scored positively *are* scored positively. This model could account for clones positive in three out of four hybridizations. Estimating that 69% of a probe yields hybridization signals (figure 11; thus yielding 69% similarity between filters hybridized with the same probe, and indicating that only 48% of the clones present in the probe are scored on both filters), and estimating 34% similarity between LNCA and LNCB probe hybridizations (figure 12), this model predicts that only 13.5%, or $(1-0.69)(1-0.34)^2$, of the total clones should be present in one of four filters. This percentage is much lower than that empirically observed. Using this model to predict the number of clones positive in four of four hybridizations also returns a value much lower than that observed.

Neither of these forms of error can be assumed to dominate: a model is needed that incorporates both random background and randomly incomplete hybridization. Such a model could be designed to fit the empirical results. Positing the presence of random background should increase the expected number of clones scored positively one and four times out of four, and some degree of incomplete hybridization might explain the small proportion of clones scored in three out of four hybridizations. As a basis for such a model, the similarity between duplicate filters (S) could be expressed as $S = (1-x)p$, where ' x ' represents the proportion of background clones and p represents the proportion of a given probe that returns hybridization signals.

4.6 Assessing the validity of the results

4.6.1 Identity and specificity of the source tissue

In order to isolate cells corresponding to the late and early subpopulations of neural crest cells, a culture system was employed that allowed the isolation of neural crest cells migrating away from neural tube sections of similar anteroposterior position but different age. Developmental stages were chosen based upon published observations

DISCUSSION

by Raible *et al.* (1992), with anterior trunk neural tubes from eight somite embryos being cultured to obtain early-migrating neural crest, and neural tubes from twenty-somite embryos cultured to obtain late-migrating crest cells. For early neural crest-cell culture, the developmental stage of fish used to provide neural tubes, as well as the relatively short duration of culture, makes it improbable that the cells harvested would be late-migrating cells, which should still be affixed to the neural tube at this time. It cannot, however, be established beyond question that the cells isolated from culture of older neural tube were in fact late-migrating cells—they could have been, for example, early-migrating cells that had adhered to the sides and ventral aspect of the neural tube. They were harvested with the understanding that they might at least represent a later stage (ca. 10 hrs later) in the differentiation of the neural crest.

An additional concern in choosing culture conditions was that the culture medium should contain no unidentified factors that might influence gene expression patterns. For this reason, a serum-free medium was used. Cells did not survive long in this medium, however (generally not longer than 48 hours), and it is possible that stress-associated changes in gene expression would occur even after eight hours in such culture. With this understanding, one might skeptically approach the set of genes found positive in both neural crest probes and not in ventral neural tube as the set representing cultured vs. non-cultured cells.

4.6.2 Efficacy of the subtraction

Several paradigms exist for judging the success of a differential screening approach. Following biochemical subtraction, the efficacy of subtraction (the degree of enrichment of specific clones) is traditionally assessed by Southern blot analysis. By definition, the pools of clones represented by the hybridization data sets produced here are completely subtracted by the computer. Yet given the artifactual change in cDNA abundance potentially introduced by the labeling procedure, as well as possible artifact introduced by the variability in library gridding, it remains questionable what proportion

DISCUSSION

of 'differentially scored' cDNAs really are differentially present in the original pools of amplified cDNA. Following the differential screening protocols used by Dulac (1995; 1998), selected cDNA clones are purified and re-screened by duplicate Southern blot hybridization of individual cDNAs with labeled, amplified cDNA. While such rescreening should eventually be rendered unnecessary by refined versions of the present approach, it is still perhaps the most sensitive tool for the quantitative evaluation of the current results.

4.6.3 Selection of markers as positive control

A second approach for assessing the validity of the subtracted cDNA sets is to search for known markers, both within the original gridded library and within the selected sets. As described, this approach was used with transcription factors Snail 2, AP2, and B220 to assess the cDNA sets created in this project. The output lists for the individual hybridizations, as well as the final pools of selected clones, were checked for the presence or absence of these specific markers. One marker B220 is of known position. This marker does not score positive in any of the hybridizations. It is known from *in situ* hybridizations to give a strong and immediate signal (D. Gilmour, pers. comm.); this could indicate a high abundance, in which case it might have failed to score positively due to the 'abundance filtering effect' previously described. Of the other eight positions checked, seven are scored in neural crest hybridizations, and two in neural tube hybridizations.

The presence of these genes in the neural tube pools could indicate that some neural crest cellular contamination was present in the original VNT probes. Alternatively this could indicate that low residual expression levels of these genes are maintained in neural tube up to the point at which the neural tube was harvested. AP2, for example, is expressed in a broad domain of ventral ectoderm during gastrulation before being restricted to lateral stripes during somitogenesis (Nguyen *et al.*, 1998). Perhaps this restriction, seen only by *in situ* hybridization, is not complete.

DISCUSSION

Unfortunately, few markers were selected by our subtractive parameters. While two of the three neural-crest positive markers were not selected because they scored positively in neural tube hybridizations, four others were not selected because they did not score reproducibly, each yielding only one positive result in four Late Neural Crest hybridizations. One clone was selected in the 'Both Neural Crest' clone set.

4.6.4 *In situ* hybridization

The most straightforward way to screen selected clones is by *in situ* hybridization of labeled antisense RNA. To this point, none of the *in situs* attempted has returned an acceptable signal. Problems have been encountered with background staining and appear due in some way to the RNA probe. Controls show that this background does not arise from the antibody binding or staining procedure. A potential cause of this background is a fifty base-pair sequence of multiple cloning site between the SP6 transcription promoter and the inserted cDNA sequence. Other researchers have experienced difficulties with similar plasmid cloning-site sequences (F. Weth, pers. comm.), and other groups working with this library are now PCR amplifying the gene insert using primers containing the T3 and T7 RNA polymerase promoters before transcribing labeled RNA for *in situ* probes (M. Clark; D. Gilmour, pers. comm.).

4.7 *Practical application of the hybridization data*

Though there are enough confounding factors— from the hybridization and normalization filtering alone— to make any analysis of the efficiency and reproducibility of the technique difficult, the data can still be used to create biologically relevant sets of clones. The pools isolated should have value, even if approached as only as an enriched pool to use in an *in situ* screen. The clones selected most likely represent rare cDNAs present in the neural crest, and as the filters contain fifteen to 20,000 cDNAs, with about 1500 scoring positively for a given probe, the clones of the full library could be enriched more than ten times for cDNAs of transcripts found in the neural crest. Additionally, for each of the probes, these positives have been reduced

DISCUSSION

approximately five times by subtraction of ventral neural tube cDNAs. Assuming the VNT pool is not contaminated with neural crest cells, this could represent a fifty-fold enrichment of neural-crest specific cDNAs relative to the original set of the library. Furthermore, it was the good fortune of this attempt at electronic subtraction that more densely-printed filters were used for the Ventral Neural Tube hybridizations later used as subtractive driver. This suggests that any possibly incomplete hybridization signals should not lead to incomplete enrichment of neural crest positive clones, but simply to a smaller enriched pool.

With future analysis, oligo-fingerprint clustering data could also be used to compensate for current problems of missing clones on the printed filter. If clones are missing from the filters hybridized with tissue-specific pools used as subtractive driver, the ensuing 'false negatives' will fail to subtract out shared clones from the tracer pool, leaving the tracer pool incompletely enriched. One approach to improve this situation would be to create a driver pool through extrapolation, containing all clones that are members of a cluster (putatively encoding the same transcript) in which any other member returned a positive hybridization signal. By filling out a 'virtual' driver pool in this way, one might reduce the number of such false negatives and increase the enrichment of specific clones.

This and most further theoretical analysis, however, should be postponed until the sets of clones currently selected can be assayed by *in situ* hybridization— after further sub-cloning or PCR amplification— for neural crest specificity. These identified clones might then serve as a large marker pool by which one could test the enrichment efficiency and utility of various *in silico* selection parameters and experimental approaches. Just as it is hoped that the technique employed here might further the biology of the neural crest, the results of such an *in situ* screen should themselves be used to improve the technique that generated them.

5. Citations

CITATIONS

Alberts B. Bray D. Lewis J. Raff M. Roberts K. Watson J. 1994. *Molecular Biology of the Cell*. Garland Publishing: New York, USA

Amrein H. Axel R. 1997. Genes expressed in neurons of adult male *Drosophila*. *Cell* **88**:459-469.

Baker CVH. Bronner-Fraser M. Le Douarin NM. Teillet M-A. 1997. Early- and late-migrating neural crest cell populations have equivalent developmental potential *in vivo*. *Development* **124**: 3077-3087.

Belyavsky A. Vinogradova T. Rajewsky K. 1989. PCR-based complementary DNA library construction: General complementary DNA libraries at the level of a few cells. *Nucleic Acids Research* **17**: 2919-2932.

Bonner RF. Emmert-Buck M. Cole K. Pohida T. Chuaqui R. *et al.* 1997. Laser capture microdissection: molecular analysis of tissue. *Science* **278**: 1481-1483

Bronner-Fraser M. Fraser SE. 1991. Cell lineage analysis of the avian neural crest. *Development* **113**: 17-22

Brady G. Billia F. Knox J. Hoang T. Kirsch IR. *et al.* 1995. Analysis of gene expression in a complex differentiation hierarchy by global amplification of cDNA from single cells. *Current Biology* **5**: 909-922.

Brand M. Beuchle D. Endres F. Haffter P. Hammerschmidt M. *et al.* 1995. Keeping and raising zebrafish (*Danio rerio*) in Tuebingen. *Zebrafish Science Monitor* **3**: 2-7

Chee M. Yang R. Hubbell E. Berno A. Huang X. *et al.* 1996. Accessing genetic information with high-density DNA arrays. *Science* **274**: 610-614.

Crino PB. Trojanowski JQ. Dichter MA. Eberwine J. 1996. Embryonic neuronal markers in tuberous sclerosis: Single-cell molecular pathology. *Proc Natl Acad Sci USA* **93**: 14152-14157.

Diatchenko L. Lau Y F C. Campbell A P. Chenchik A. Moqadam F. *et al.* 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* **93**: 6025-6030.

Driever W. Solnica Krezel L. Schier AF. Neuhauss SCF. Malicki J. *et al.* 1996. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**: 37-46.

Dulac C. 1998. Cloning of genes from single neurons. de Pablo F. Ferrus A. Stern CD. (Ed.). *Current Topics in Developmental Biology* **36**: 245-258. Academic Press, Inc.: San Diego, USA.

Dulac C. Axel R. 1995. A novel family of genes encoding putative pheromone receptors in mammals. *Cell* **83**: 195-206.

Eberwine J. 1996. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *Biotechniques* **20**: 584, 586, 588, 590-591.

Eberwine J. Yeh H. Miyashiro K. Cao Y. Nair S. *et al.* 1992. Analysis of gene expression in single live neurons. *Proc Natl Acad Sci USA* **89**: 3010-3014.

Eisen JS. Weston JA. 1993. Development of the neural crest in the zebrafish. *Developmental Biology* **159**: 50-59.

CITATIONS

- Emmert-Buck MR. Bonner RF. Smith PD. Chuaqui RF. Zhuang Z. 1996. Laser capture microdissection. *Science* **274**: 998-1001.
- Froussard P. 1992. A random-pcr method rPCR to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Research* **20**: 2900
- Furthauer M. Thisse C. Thisse B. 1997. A role for FGF-8 in the dorsoventral patterning of the zebrafish gastrula. *Development* **124**: 4253-4264.
- Gao D. Li Z. Murphy T. Sauerbier W. 1997. Structure and transcription of the gene for translation elongation factor 1 subunit alpha of zebrafish (*Danio rerio*). *Biochimica et Biophysica Acta* **1350**:1-5.
- Gress TM. Hoheisel JD. Lannon GG. Zehetner G. Lehrach H. 1992. Hybridization fingerprinting of high-density cDNA library arrays with cDNA pools derived from whole tissues. *Mammalian Genome* **3**: 609-619.
- Gubler U. Hoffman BJ. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* **25**: 263-9.
- Hafter P. Granato M. Brand M. Mullins MC. Hammerschmidt M. *et al.* 1996. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**: 369-89.
- Henion PD. Raible DW. Beattie CE. Stoesser KL. Weston JA. Eisen JS. 1996. Screen for mutations affecting development of zebrafish neural crest. *Developmental Genetics* **18**: 11-7.
- Hoheisel JD. Ross MT. Zehetner G. Lehrach H. 1994. Relational genome analysis using reference libraries and hybridisation fingerprinting. *Journal of Biotechnology* **35**: 121-134.
- Jesuthasan S. 1996. Contact inhibition/collapse and pathfinding of neural crest cells in the zebrafish trunk. *Development* **122**: 381-389.
- Karrer EE. Lincoln JE. Hogenhout S. Bennett AB. Bostock RM *et al.* 1995. In situ isolation of mRNA from individual plant cells— creation of cell-specific cDNA libraries. *Proc Natl Acad Sci USA* **92**: 3814-3818.
- Kelly GM. Reversade B. 1997. Characterization of a cDNA encoding a novel band 4.1-like protein in zebrafish. *Biochemistry and Cell Biology-Biochimie et Biologie Cellulaire* **75**: 623-632.
- Kelsh RN. Brand M. Jiang YJ. Heisenberg CP. Lin S. *et al.* 1996. Zebrafish pigmentation mutations and the processes of neural crest development. *Development* **123**: 369-89
- Kimmel CB. Ballard WW. Kimmel SR. Ullmann B. Schilling TF. 1995. Stages of embryonic development of the zebrafish. *Developmental Dynamics* **203**: 253-310
- Krauss S. Concordet J-P. Ingham PW. 1993. A functionally conserved homolog of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in Zebrafish embryos. *Cell* **75**: 1431-1444.
- Lander ES. 1996. The new genomics: global views of biology. *Science* **274**: 536-539.

CITATIONS

- Lanfear J. Jowett T. Holland PWH. 1991. Cloning of fish zinc-finger genes related to *krox-20* and *krox-24*. *Biochemical & Biophysical Research Communications* **179**: 1220-1224.
- Lashkari DA. Derisi JL. Mccusker JH. Namath AF. Gentile C. *et al.* 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci USA* **94**: 13057-13062.
- Le Douarin N. 1982. *The Neural Crest*. Cambridge University Press: Cambridge, UK
- Le Douarin NM. Ziller C. Couly GF. 1993. Patterning of neural crest derivatives in the avian embryo: In vivo and in vitro studies. *Developmental Biology* **159**: 24-49.
- Lennon GG. Lehrach H. 1991. Hybridization analyses of arrayed cDNA libraries. *Trends in Genetics* **7**: 314-7.
- Liang P. Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**: 967-971.
- Lisitsyn N. Lisitsyn N. Wigler M. 1993. Cloning the differences between two complex genomes. *Science* **259**: 946-951.
- Lockhart DJ. Dong H. Byrne MC. Follettie MT. Gallo MV. *et al.* 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnology* **14**: 1675-1680.
- Monyer H. Lambolez B. 1995. Molecular biology and physiology at the single-cell level. *Current Opinion in Neurobiology* **5**: 382-387.
- Neuhauss SC. Solnica-Krezel L. Schier AF. Zwartkruis F. Stemple D. *et al.* 1996. Mutations affecting craniofacial development in zebrafish. *Development* **123**: 357-67.
- Nguyen VH. Schmid B. Trout J. Connors SA. Ekker M. Mullins MC. 1998. Ventral and lateral regions of the zebrafish gastrula, including the neural crest progenitors, are established by a *bmp2b/swirl* pathway of genes. *Developmental Biology* **199**: 93-110
- Nguyen C. Rocha D. Granjeaud S. Baldit M. Bernard K. *et al.* 1995. Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics* **29**: 207-216.
- Okubo K. Hori N. Matoba R. Niiyama T. Fukushima A. *et al.* 1992. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genetics* **2**: 173-179.
- Phillips J. Eberwine JH. 1996. Antisense RNA amplification: A linear amplification method for analyzing the mRNA population from single living cells. *Methods* **10**: 283-288.
- Pietu G. Alibert O. Guichard V. Lamy B. Bois F. *et al.* 1996. Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high density cDNA array. *Genome Research* **6**: 492-503.
- Odenthal J. Nusslein-Volhard C. 1998. Fork head domain genes in zebrafish. *Development Genes and Evolution* **208**: 245-258.

CITATIONS

- Okayama H. and Berg P. 1982. High-efficiency cloning of full length cDNA. *Mol Cell Biol* **2**: 161-170.
- Raible DW. Eisen JS. 1994. Restriction of neural crest cell fate in the trunk of the embryonic zebrafish. *Development* **120**: 495-503.
- Raible DW. Eisen JS. 1996. Regulative interactions in zebrafish neural crest. *Development* **122**: 501-507.
- Raible DW. Wood A. Hodsdon W. Henion PD. Weston JA. Eisen JS. 1992. Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Developmental Dynamics* **195**: 29-42
- Sambrook J. Fritsch EF. Maniatis T. 1989. *Molecular Cloning: A laboratory manual*. Second Edition. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, USA
- Schena M. Shalon D. Davis RW. Brown PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.
- Schilling TF. Kimmel CB. 1994. Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development*. **120**: 483-494.
- Soares MB. Bonaldo MF. Jelene P. Su L. Lawton L. *et al.* 1994. Construction and characterization of a normalized cDNA library. *Proc Natl Acad Sci USA* **91**: 9228-9232.
- Strahle U. Blader P. Henrique D. Ingham PW. 1993. *Axial* a zebrafish gene expressed along the developing body axis shows altered expression in *cyclops* mutant embryos. *Genes & Development* **7**: 1436-1446.
- Sucher N J. Deitcher D. 1995. PCR and patch-clamp analysis of single neurons. *Neuron* **14**: 1095-1100.
- Tam AW. Smith MM. Fry KE. Larrick JW. 1989. Construction of complementary DNA libraries from small numbers of cells using sequence independent primers. *Nucleic Acids Research* **17**: 1269.
- Thisse C. Thisse B. Schilling TF. Postlethwait JH. 1993. Structure of the zebrafish *snail1* gene and its expression in wild-type, *spadetail* and *no tail* mutant embryos. *Development* **119**: 1203-1215.
- Thisse C. Thisse B. Postlethwait JH. 1995. Expression of *snail2*, a second member of the zebrafish *snail* family, in cephalic mesendoderm and presumptive neural crest of wild-type and *spadetail* mutant embryos. *Developmental Biology* **172**: 86-99.
- Thisse C. and Thisse B. 1998. High resolution whole-mount *in situ* hybridization. http://zfinfo.uoregon.edu/zf_info/monitor/vol5.1/vol5.1.html
- Van Gelder RN. Von Zastrow ME. Yool A. Dement WC. Barchas JD. Eberwine JH. 1990. Amplified RNA synthesized from limited quantities of heterogeneous complementary DNA. *Proc Natl Acad Sci USA* **87**: 1663-1667.
- Vos P. Hogers R. Bleeker M. Reijans M. Van De Lee T. *et al.* 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407-4414.
- Wang Z. Brown DD. 1991. A gene expression screen. *Proc Natl Acad Sci USA* **88**: 11505-11509.

CITATIONS

Westerfield M. 1993. *The Zebrafish Book*. University of Oregon Press. Eugene, OR.

Weston JA. 1991. Sequential segregation and fate of developmentally restricted intermediate cell populations in the neural crest lineage. *Curr Top Dev Biol* **25**: 133-153

Weth F. 1993. Geruchsrezeptoren des Zebrafisch: Untersuchungen zur genomischen Repräsentation und Expression. Diploma Thesis submitted to the University of Tübingen, Tübingen, Germany.